# **ECONSTOR** Make Your Publications Visible.

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Xia, Ziqian; Ye, Jinquan; Debnath, Ramit

## Working Paper A Comment on "A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy" by Lorenz-Spreen et al., (2023)

I4R Discussion Paper Series, No. 206

**Provided in Cooperation with:** The Institute for Replication (I4R)

*Suggested Citation:* Xia, Ziqian; Ye, Jinquan; Debnath, Ramit (2025) : A Comment on "A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy" by Lorenz-Spreen et al., (2023), I4R Discussion Paper Series, No. 206, Institute for Replication (I4R), s.l.

This Version is available at: https://hdl.handle.net/10419/312257

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# **INSTITUTE** for **REPLICATION**

No. 206 I4R DISCUSSION PAPER SERIES

# A Comment on "A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy" by Lorenz-Spreen et al., (2023)

Ziqian Xia Jinquan Ye Ramit Debnath

February 2025



# **I4R DISCUSSION PAPER SERIES**

I4R DP No. 206

## A Comment on "A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy" by Lorenz-Spreen et al., (2023)

Ziqian Xia<sup>1,3</sup>, Jinquan Ye<sup>2,3</sup>, Ramit Debnath<sup>3</sup>

<sup>1</sup>Tongji University, Shanghai/China <sup>2</sup>Duke University, Durham/USA <sup>3</sup>University of Cambridge/Great Britain

#### FEBRUARY 2025

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and metascientific work in the social sciences. Provided in cooperation with EconStor, a service of the <u>ZBW – Leibniz Information Centre for Economics</u>, and <u>RWI – Leibniz Institute for Economic Research</u>, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur University of Ottawa Anna Dreber Stockholm School of Economics Jörg Ankel-Peters *RWI – Leibniz Institute for Economic Research* 

E-Mail: joerg.peters@rwi-essen.de RWI – Leibniz Institute for Economic Research Hohenzollernstraße 1-3 45128 Essen/Germany www.i4replication.org

### A comment on "A systematic review of worldwide causal and correlational evidence on digital media and democracy" by Lorenz-Spreen et al., (2023)<sup>\*</sup>

Ziqian Xia<sup>1,2</sup>, Jinquan Ye<sup>2,3,#</sup>, Ramit Debnath<sup>2</sup>

1. School of Economics and Management, Tongji University

2. Collective Intelligence & Design Group, University of Cambridge

3. Nicholas School of the Environment, Duke University

#### Abstract

In this study, we conduct a direct replication of Lorenz-Spreen et al. (2023)'s systematic review on digital media and democracy to verify and enhance the robustness of their findings. We successfully reproduce the main claims of the original paper, uncovering a few minor coding errors that do not affect the study's primary results. Our replication updates the dataset with recent publications up to March 23, 2024. After screening 2,567 papers, we included 110 that met the original study's criteria. We then replicated the results from original paper and employed topic modeling and word frequency analysis to identify key themes and trends. This replication confirms the original findings and provides an updated perspective on the relationship between digital media and democracy.

<sup>\*</sup> Authors: Ziqian Xia: Tongji University and University of Cambridge. E-mail: <u>ziqian.research@gmail.com</u>. Jinquan Ye: Duke University. Email: <u>Jinquan.ye@duke.edu</u>. Ramit Debnath: University of Cambridge. Email: <u>rd545@cam.ac.uk</u>. # Corresponding author.

I4R DP No. 206

#### 1. Introduction

This report focuses on the replication and reanalysis of the study titled "A Systematic Review of Causal and Correlational Evidence on Digital Media and Democracy" by Lorenz-Spreen, Oswald, Lewandowsky & Hertwig, (2023) (hereafter referred to as the "original study")<sup>1</sup>. The original study systematically reviewed 496 empirical articles to investigate the impact of digital media on various political variables. The primary data sources for the review were the Scopus and Web of Science databases, which include articles examining the effects of digital media on democracy, with a focus on variables such as political participation, trust, populism, and polarization. The review process adhered to the MOOSE guidelines (Brooke, Schwartz & Pawlik. 2021), and the detailed review protocol was preregistered on the Open Science Framework (OSF)<sup>†</sup>.

The original study employed a comprehensive coding and screening process to ensure the inclusion of relevant empirical evidence. It covered articles published up to September 15, 2021, excluding theoretical works, small-scale studies, and articles without empirical data. The methodological framework involved coding various dimensions of digital media and political outcomes, assessing study quality, and identifying potential biases.

The original study concluded that while digital media can enhance political participation and information consumption—beneficial aspects often observed in autocracies and emerging democracies—it is also associated with negative outcomes such as declining political trust, increasing populism, and growing polarization, particularly in established democracies. Our replication of this study, incorporating an expanded dataset, confirms that these conclusions remain consistent. The results demonstrate that the impact of digital media on democracy varies by political context and specific variables, aligning with the original findings and underscoring the need for continued research and regulatory efforts.

In this replication effort conducted for the Institute for Replication in collaboration with Nature Human Behaviour (Brodeur et al. 2024), we used the methods and code

<sup>†</sup> https://osf.io/7ry4a/

provided by the original authors to reanalyze the data and validate their findings. The authors addressed our questions about their classification principles and methodologies, allowing us to closely follow the approach outlined in their original study. We successfully replicated their results and applied the same statistical techniques to regenerate all figures. Our updated dataset included publications up to March 23, 2024, and we re-ran the MOOSE checklist to ensure methodological rigor. Despite encountering minor issues, such as small discrepancies in data reporting, these were easily resolved and did not affect the overall findings. By adhering to these rigorous standards and incorporating updated data, this replication validated the original findings and contributed to a more comprehensive understanding of the interplay between digital media and democracy.

#### 2. Computational Reproducibility

During our reproduction of the study, we identified several minor issues that, while they required some attention, did not impact our ability to accurately reproduce the original results. These issues were easily resolved and had no effect on the study's findings. We successfully computationally reproduced all the main results (Figure 1, Figure 2, Figure 4 and Figure 4) from the raw data.

	Fully	Partial	No
Raw data provided	Х		
Cleaning code provided		Х	
Analysis data provided	х		
Analysis code provided	х		
Reproducible from raw data	х		
Reproducible from analysis data	х		

Table 1: Summary of Reproducibility Assessment and Availability of Data and Code

For example, although the "quality\_label\_preprocessing.R" file was missing from the replication package, its absence did not hinder the replication process, as it only served to streamline labels for outcome measures and methods. Both we and the original authors confirmed that this omission had no effect on the results. We also addressed the overlooked installation of the *readxl* package, which was necessary for data importation, by adding the appropriate commands to our script. Additionally, minor discrepancies between the counts reported in the original study and those produced by our replication code, such as a difference of one count in certain variable combinations, were observed. However, these discrepancies were negligible and did not alter the study's overall conclusions. Specifically, for the combination of variables denoted as A-C-B, the original paper reported a count of 40, while the replicate code provided by the authors yielded 39. Similarly, for the A-B combination, the paper indicated a count of 301, whereas our code produced 302.

In summary, despite these minor challenges, we were able to reproduce the study's results accurately and confirm that the original findings hold true. The adjustments we made were straightforward and did not affect the study's main conclusions, demonstrating the robustness of the original analysis.

#### 2.1 Discrepancies Between Pre-analysis Plan and Article

The replication was preregistered with a pre-analysis plan (PAP). However, upon collecting the updated dataset, we encountered certain limitations that necessitated deviations from the original plan.

Initially, the PAP specified the use of bibliometric analysis to map the research landscape, identify key themes and trends, and assess the influence of publications within the dataset. The plan included conducting citation analysis, co-authorship network analysis, and keyword co-occurrence analysis using bibliometric software tools. However, due to the limited number of data points in the final updated dataset, the citation network was highly parsed and scattered, making it unsuitable for a meaningful network analysis. Consequently, we deviated from the original plan and instead employed topic modeling (Robledo, S. & Zuluaga, M. 2022 & Bickel, M. W. 2019) and word frequency analysis on the titles and abstracts to identify trends.

Furthermore, the PAP outlined the use of Qualitative Comparative Analysis (QCA) to examine in-depth case studies and investigate how digital media influences democratic processes. The intent was to identify patterns and conditions under which digital media

6

positively or negatively impacts democracy. However, during the analysis, we discovered that the limited number of explanatory variables in our dataset hindered our ability to find effective causal combinations that could explain the relationship between democracy and digital media use. As a result, the QCA did not yield the anticipated insights, and we were unable to produce the detailed, context-rich analysis initially planned.

#### 3. Robustness Reproduction and Replication with New Data

#### [Figure 1 should be inserted here]

By updating the dataset of the articles, we conducted a classification of research topics and variables following the original paper: Figure 1a shows the combinations of variables involved in the research articles, mainly focusing on digital media (such as social media, online news), political variables (such as trust, participation), and characteristics of the information ecology (such as misinformation, selective exposure). This is consistent with the conclusion in the original article that the research sample primarily focuses on the two variables of digital media and political outcomes, with an increase in the number (89 new cases added).

Geographically, the number of related studies in the United States still significantly leads other countries, and although some African countries have made breakthroughs from zero, there is still considerable geographic variation overall.

Regarding the distribution of methods used in the research and the combinations of political and digital media variables, the original paper covers a wide range of politically relevant variables, which have been grouped into broader categories. The updated data indicates that Participation remains the most frequently studied topic. This demonstrates that the research primarily focuses on variables related to digital media use, such as political information, polarization, and participation, while variables like instant messaging platforms and search platforms are relatively less studied.

[Figure 2 should be inserted here]

7

The updated dataset further highlights the trends observed in the original research, Fig. 2 emphasizing the complex and multifaceted impact of digital media on democracy. Specifically, digital media usage has shown a generally positive effect on political participation, particularly in increasing voter turnout and political mobilization. Studies across established democracies and other regions consistently indicate that digital media can effectively stimulate political engagement. However, when it comes to trust, digital media use is predominantly associated with a decline, particularly in trust in government, media, social institutions, and society at large. While some studies suggest that digital media may have a positive impact on trust in science or democratic trust in specific regions, the overall trend points toward a weakening of trust. The relationship between digital media use and political knowledge, as well as news exposure diversity, is generally positive, though the results are somewhat nuanced and not universally beneficial. Additionally, digital media fosters political expression in ways that can support democratic values. On the other hand, the data also underscores associations that are detrimental to liberal democracy. Digital media usage is closely linked to increases in political polarization, hate speech, and populism, with most studies indicating that these factors have a negative impact on democracy. Furthermore, the trend of declining trust is reiterated, particularly in relation to government, media, and societal institutions, reflecting a consistent and broad negative influence. Lastly, the emergence of homophilic network structures within digital media environments has been shown to harm democratic processes by fostering echo chambers and reducing exposure to diverse viewpoints.

[Figure 3 should be inserted here]

We also confirm the previous conclusions regarding the impact of digital media on democracy, as illustrated in Figs. 3a and 3b. Specifically, beneficial effects are predominantly observed in emerging democracies in South America, Africa, and parts of Asia and South Asia, indicating that digital media usage in these regions generally promotes positive democratic outcomes such as trust, knowledge, participation, exposure, and expression. Mixed effects are noted in Europe, the United States, Russia, and China, where digital media has a varied impact without a clear positive or negative trend. Detrimental effects are mainly found in Europe (e.g., Germany, Finland, Sweden) and partially in Russia, associated with issues such as hate, polarization, populism, homophily, and misinformation, highlighting the more complex challenges in these regions.

[Figure 4 should be inserted here]

In Fig. 4 A and B, there is a significant difference in sample sizes associated with different sampling methods, such as behavioral data, non-probabilistic samples, probabilistic samples, and randomized control trials. Notably, behavioral data typically have larger sample sizes, whereas randomized control trials tend to have smaller ones. Fig. 4A shows that studies investigating positive outcomes like trust, knowledge, participation, diversity exposure, and expression more frequently use probabilistic samples. These studies tend to report positive impacts on liberal democracy (indicated by green markers), although some may show no significant association (indicated by gray markers). In Fig. 4B, when examining negative outcomes such as hate, polarization, populism, network homophily, and misinformation, behavioral data are used more frequently than in positive outcome studies. These studies generally report negative impacts on liberal democracy (indicated by orange markers). The expanded dataset reveals a slight discrepancy with the original text on page 7, which stated, "If anything, large probabilistic samples report relatively less beneficial associations for both types of outcomes." Due to the inclusion of additional large-sample studies using probabilistic sampling, there is an increased representation of positive impacts on liberal democracy within positive outcomes, making the positive results no longer dominated by negative impacts and non-significant associations. The expansion of the sample size has altered this observational result.

[Figure 5 should be inserted here]

Institute for Replication

I4R DP No. 206

The extended bibliometric analysis reveals several important patterns in the research landscape of digital media and democracy. The frequent appearance of terms such as "participation," "information," and "exposure" indicates a strong emphasis on how digital media influences civic engagement and the dissemination of political information. Additionally, terms like "polarization," "misinformation," and "trust" suggest that research is also keenly focused on the negative aspects of digital media, particularly concerning its role in exacerbating political divides and spreading false information.

The topic model visualization (Figure 5B) further supports these observations by categorizing key themes within the literature. Latent Dirichlet Allocation (LDA) is a probabilistic model used to identify patterns within large text datasets. By employing this model, analysis can uncover hidden structures within newly introduced literature, enabling a detailed classification of potential research topics. Many top studies also utilize LDA for topic modeling. For example, LDA has been applied to analyze interdisciplinary research trends in the scientific field (Sietsma, Ford, & Minx. 2024) and to uncover the thematic structure of political discourse and misinformation propagation on social media (Wang et. al., 2022; ). These examples demonstrate LDA's broad applicability in extracting latent topics from text corpora. The identified topics include issues related to public trust and misinformation (Topic 1), the role of social media in political participation and offline influence (Topic 2), and the impact of selective exposure and polarization on political attitudes (Topic 3). Topics 4 and 5 delve into the relationship between users' engagement with platforms like Twitter and Facebook and their influence on public discourse and democratic participation.

#### 4. Conclusion

In this replication study, we successfully reproduced the main findings of the original study by Lorenz-Spreen et al. (2023). Our work confirmed that digital media's impact on democracy varies significantly depending on the political context and specific variables. We found that while digital media can foster political participation and information dissemination, it can also contribute to negative outcomes such as declining

political trust, rising populism, and increased polarization, especially in established democracies. Our replication effort highlighted the robustness of the original study's conclusions. The application of additional analytical techniques, such as topic modeling and word frequency analysis, provided further insights into the evolving themes within the literature on digital media and democracy.

However, due to limitations in our dataset, we had to deviate from the originally planned Qualitative Comparative Analysis (QCA) intended to explore the factors underlying the causal association between digital media and democracy. Future replicators could attempt this method with a more extensive set of causal studies. Additionally, conducting a meta-analysis on specific subtopics could provide valuable insights, though ensuring inter-comparability across studies would be crucial.

Given the rapid growth in this field, it is essential to update the evidence base every few years to incorporate the latest research findings. As more studies emerge, continuous reassessment will be necessary to maintain an accurate and comprehensive understanding of how digital media influences democratic processes.

#### References

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S. & Hertwig, R. A systematic review of worldwide causal and correlational evidence on digital media and democracy. Nat Hum Behav 7, 74–101 (2023).

Brodeur, A., Dreber, A., Hoces de la Guardia, F. et al. Reproduction and replication at scale. Nat Hum Behav 8, 2–3 (2024). https://doi.org/10.1038/s41562-023-01807-2.

Brooke, B. S., Schwartz, T. A. & Pawlik, T. M. MOOSE reporting guidelines for metaanalyses of observational studies. JAMA surgery 156, 787–788 (2021).

Robledo, S. & Zuluaga, M. Topic modeling: Perspectives from a literature review. IEEE Access 11, 4066–4078 (2022).

- Bickel, M. W. Reflecting trends in the academic landscape of sustainable energy using probabilistic topic modeling. Energ Sustain Soc 9, 49 (2019).
- Sietsma, A. J., Ford, J. D. & Minx, J. C. The next generation of machine learning for tracking adaptation texts. Nat. Clim. Chang. 14, 31–39 (2024).

- Wang, X., Zhang, M., Fan, W., & Zhao, K. Understanding the spread of COVID-19 misinformation on social media: The effects of topics and a political leader's nudge. Journal of the Association for Information Science and Technology, 73(5), 726-737 (2022).
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 19(1), 22-36. (2017)

#### Acknowledgement

We acknowledge the financial support provided by the Bill and Melinda Gates Foundation [OPP1144], the Cambridge Humanities Research Grant (CHRG), and ai@cam. We would like to express our gratitude to Sourjyamoy Barman for his valuable assistance with screening and data collection. We also state that we have no personal or professional relationship with the original authors of the study, and there are no conflicts of interest in this replication effort.

#### **Figures**



**Figure 1 (Updated). A**, Combinations of variables in the sample: digital media (A), political variables (B) and content features such as selective exposure or misinformation (C). Numbers in brackets count articles in our sample that measure an association between variables. **B**, Geographic distribution of articles that reported site of data collection. c,d, Distribution of measurements (counted separately whenever one article reported several variables) over combinations of outcome variables and methods (**C**) and over combinations of outcome variables and methods (**D**).



**Figure 2 (Updated). A.** Directions of associations are reported for various political variables (see Fig. 1d for a breakdown). Sub-item **B-E** show examples of the distribution of associations with trust, news exposure, polarization and network homophily over the different digital media variables with which they were associated



Trust + knowledge + participation + diversity exposure + expression

Hate + polarization + populism + network homophily + misinformation

**Figure 3 (Updated). A**, Geographical distribution of reported associations for the variables trust, knowledge, participation, exposure and expression. Pie charts show the composition of directions for each country studied. **B**, Geographic representation of reported associations for the variables hate, polarization, populism, homophily and misinformation. **C**, Data and variables in a, in absolute numbers of reported associations and sorted along the Liberal Democracy Index. **D**, Data and variables in b, in absolute numbers of reported associations and sorted along the Liberal Democracy Index.



**Figure 4 (Updated). A**, Sample size vs sampling methods for variables of trust, knowledge, participation, exposure and expression. Each dot represents one measurement, colour coded according to the direction of the reported association. **B**, Sample size vs sampling method for variables of hate, polarization, populism, network homophily and misinformation



**Figure 5. A**. Word cloud representing the most frequently occurring terms in the updated dataset. **B**. Top terms in each of the five topics identified through topic modeling analysis.