

Title: Codebook appendix for “Vaccine hesitancy in online spaces: A scoping review of the research literature, 2000-2020”

Authors: Timothy Neff (1), Jonas Kaiser (2), Irene Pasquetto (3), Dariusz Jemielniak (4), Dimitra Dimitrakopoulou (5), Siobhan Grayson (6), Natalie Gyenes (7), Paola Ricaurte (8), Javier Ruiz-Soler (9), Amy Zhang (10)

Date: October 11th, 2021

Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

Appendix: Codebook

Variables should receive one code only (not multiple codes), indicating assessment of the primary characteristic relevant for each variable. If an article passes the filter variable (V11), every other cell in that article’s row in the codesheet must include a single numeral as coding (in other words, do not leave any cell in that article’s row empty). The only exception is that article’s cell under V17, which is a text variable and can be left empty.

V1. Article ID

(already filled out)

V2. Article title

(already filled out)¹

V3. Coder

(fill out with name)

V4. Journal name

(already filled out)

V5. Academic field

(already filled out)

V6. Publication year

(already filled out)

V7. Author names

(already filled out)

V8. Author affiliations

(already filled out)

V9. Keywords

(already filled out)

V10. [DELETED]

V11. Relevant for our research interest?

This variable asks for the relevance of the article within the scope of our research interest. Relevance means these conditions are met: 1) online, vaccination-related communication (internet, social media,

¹ Based on the output from the Web of Science search, metadata was already filled out.

blogs, etc.) is referenced in the article's abstract; 2) one paragraph in the introduction, one paragraph in the methods section, and one paragraph in the conclusion mentions online vaccination-related communication; 3) the article must directly research anti-vaccination (hesitancy, opposition, etc.) communication expressed online. If the article only conducts online surveys, it should be coded (0) no.

Important: Anti-vaccination needs to be addressed multiple times in the paper, i.e., papers that just mention the issue once are coded as 0.

0 = No (if "no," stop coding)

1 = Yes

V12. [DELETED]

V13. Open access

Is this article openly accessible?

0 = No

1 = Yes

V14. Social media platforms (code one)

Are social media platforms the focus of direct analysis (not simply mentioned)? If so, which ones?

(Note: Analysis of websites or blogs is not meant here; code 0)

0 = No

1 = Instagram

2 = Twitter

3 = Facebook

4 = YouTube

5 = Reddit

6 = Other social media platform (e.g., TikTok, VKontakte, etc.)

7 = Multiple platforms (multiple, specific social media platforms specified as included in article's analysis)

8 = Other (Only in case a different social media platform is analyzed)

V15. Other digital media (code one)

If not, what other digital platforms are the focus of direct analysis (not simply mentioned)?

0 = No (if purely social media, code 0)

1 = Blog

2 = News/Online media

3 = Wikipedia

4 = Forums or other online communities

5 = Websites

6 = Other (e.g., 4chan)

7 = Multiple platforms (multiple, specific platforms specified as included in article's analysis)

V16. Country of focus given by Web of Science

(already filled out)

V17. Country of focus different from Web of Science classification

If the country or countries of focus are different than the Web of Science classification, type in full country name(s), separated by comma; if there are four or more countries type "International;" if not leave the field empty.

V18. Study methodology (code one; based on whether the study uses these key terms)

Search article for specific terms. We do NOT code what we think the researchers are doing, we code what the scholars are describing their own study as. If they are not specifying their method, code unspecified. If they both say qualitative and quantitative or quantitative and computational, etc. code mixed.

- 1 = Theoretical (e.g., literature reviews)
- 2 = Qualitative and descriptive statistics
- 3 = Quantitative and/or statistical significance of results assessed (look for p values)
- 4 = Computational (e.g., classifiers, machine learning, topic modeling, network analysis)
- 5 = Mixed
- 6 = Unspecified/other

V19. Causal inference (code one; based on whether the study uses these key terms)

This variable is aimed at understanding if the authors are claiming to have identified a statistically significant causal relationship between variables (e.g., misinformation -> anti-vaccination attitudes).

Signal words to look out for: causal, causality, direct relationship.

- 0 = No
- 1 = Yes

V20. Type of journal (code one)

- 1 = Public health/medicine
- 2 = Media and communications
- 3 = Anthropology
- 4 = Computer science
- 5 = Social sciences/humanities (e.g., economics, political science, linguistics, etc.)
- 6 = General (*Science*, *Nature*, *PLOS One*, *PNAS*)
- 7 = Other

V21. What is the main issue?

(automatically filled out)

V22. Title & abstract language used (code one)

- 1 = Anti-vaccine/anti-vaccination
- 2 = Vaccine hesitancy
- 3 = Vaccine skepticism
- 4 = Vaccine denial
- 5 = Vaccine critical
- 6 = Uses at least two terms interchangeably
- 7 = Other

V23. Stakeholders

The following variables assess which stakeholders are named by the authors in the article. Code one variable per stakeholder. The stakeholders have to be clearly identified by the authors. For example, if the article mentions that has findings that could be helpful for “policymakers and health professionals,” code V23a as 3, and then code V23b as 4. Note that this also follows the order in which stakeholders are mentioned (policymakers mentioned first, so coded in V23a; health professionals mentioned second, so coded in V23b). Please place a code in each variable, even if that code is 8 for the first three and 0 for the last, indicating no stakeholders are mentioned.

V23a. Are stakeholders named in the article, and which ones? (code one)

- 1 = Journalists/the media
- 2 = Patients and their relatives
- 3 = Policymakers and legislators
- 4 = Medical professionals and the public health sector
- 5 = Technology companies/platforms
- 6 = Academic researchers (“more research needed”)
- 7 = Other
- 8 = None

V23b. Are stakeholders named in the article, and which ones? (code one)

- 1 = Journalists/the media
- 2 = Patients and their relatives
- 3 = Policymakers and legislators
- 4 = Medical professionals and the public health sector
- 5 = Technology companies/platforms
- 6 = Academic researchers (“more research needed”)
- 7 = Other
- 8 = None

V23c. Are stakeholders named in the article, and which ones? (code one)

- 1 = Journalists/the media
- 2 = Patients and their relatives
- 3 = Policymakers and legislators
- 4 = Medical professionals and the public health sector
- 5 = Technology companies/platforms
- 6 = Academic researchers (“more research needed”)
- 7 = Other
- 8 = None

V23d. More than 3 stakeholders named

- 0 = No
- 1 = Yes

V24. Demographic variables of research subjects

Focus of the article. This means: Is the article targeted at a specific demographic (e.g., mothers in Berlin, or in Boulder, Colorado, etc.). If there’s no such focus or if the article only mentions genders, ethnicities, or ages, (e.g., 51% of survey respondents were male, 49% female; 15% were 18 or younger, 50% were 19 to 45, 35% were 46 or older), then code 0.

V24a. Gender

- 0 = No
- 1 = Yes

V24b. Ethnicity

- 0 = No
- 1 = Yes

V24c. Age

0 = No

1 = Yes

V25. Research focus (code one)

This variable asks for the unit of analysis as specified in the article (usually in the methods section).

1 = Identity: Who is vaccine hesitant? (e.g., surveys, polls, papers on collective identity and beliefs, etc.)

2 = Communication: How do people talk about vaccination? What do people say about vaccination? (e.g., analysis of anti-vaccination forums/subreddits, anti-vaccination hashtags; analysis of discourse on vaccinations as a whole)

3 = Information behavior: How do people search for information about vaccination? (information seeking and retrieval practices about vaccination)

4 = Conspiracy theories (e.g., vaccine hesitant as part of bigger conspiracy theory communities)

5 = Computational identification of anti-vaccination sentiments

6 = Other