

## Communicating Scientific Uncertainty in an Age of COVID-19: An Investigation into the Use of Preprints by Digital Media Outlets

Alice Fleerackers , Michelle Riedlinger , Laura Moorhead , Rukhsana Ahmed & Juan Pablo Alperin

To cite this article: Alice Fleerackers , Michelle Riedlinger , Laura Moorhead , Rukhsana Ahmed & Juan Pablo Alperin (2021): Communicating Scientific Uncertainty in an Age of COVID-19: An Investigation into the Use of Preprints by Digital Media Outlets, Health Communication, DOI: [10.1080/10410236.2020.1864892](https://doi.org/10.1080/10410236.2020.1864892)

To link to this article: <https://doi.org/10.1080/10410236.2020.1864892>



Published online: 03 Jan 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



# Communicating Scientific Uncertainty in an Age of COVID-19: An Investigation into the Use of Preprints by Digital Media Outlets

Alice Fleerackers <sup>a</sup>, Michelle Riedlinger <sup>b</sup>, Laura Moorhead <sup>c</sup>, Rukhsana Ahmed <sup>d</sup>, and Juan Pablo Alperin <sup>e</sup>

<sup>a</sup>Interdisciplinary Studies, Simon Fraser University; <sup>b</sup>School of Communication, Queensland University of Technology; <sup>c</sup>Journalism, College of Liberal and Creative Arts, San Francisco State University; <sup>d</sup>Department of Communication, University at Albany, State University of New York; <sup>e</sup>School of Publishing, Simon Fraser University

## ABSTRACT

In this article, we investigate the surge in use of COVID-19-related preprints by media outlets. Journalists are a main source of reliable public health information during crises and, until recently, journalists have been reluctant to cover preprints because of the associated scientific uncertainty. Yet, uploads of COVID-19 preprints and their uptake by online media have outstripped that of preprints about any other topic. Using an innovative approach combining altmetrics methods with content analysis, we identified a diversity of outlets covering COVID-19-related preprints during the early months of the pandemic, including specialist medical news outlets, traditional news media outlets, and aggregators. We found a ubiquity of hyperlinks as citations and a multiplicity of framing devices for highlighting the scientific uncertainty associated with COVID-19 preprints. These devices were rarely used consistently (e.g., mentioning that the study was a preprint, unreviewed, preliminary, and/or in need of verification). Less than half of the stories we analyzed contained framing devices emphasizing uncertainty. Outlets in our sample were much less likely to identify the research they mentioned as preprint research, compared to identifying it as simply “research.” This work has significant implications for public health communication within the changing media landscape. While current best practices in public health risk communication promote identifying and promoting trustworthy sources of information, the uptake of preprint research by online media presents new challenges. At the same time, it provides new opportunities for fostering greater awareness of the scientific uncertainty associated with health research findings.

The public expectation and need for credible health information in the midst of the COVID-19 pandemic have put a renewed focus on science and its internal processes, while simultaneously challenging traditional journalistic sourcing practices in the absence of relevant peer-reviewed research. As researchers respond by publishing research as so-called preprints, journalistic reporting on research that has yet to be peer reviewed is filling this gap and driving public discourse (Majumder & Mandl, 2020). While this surge in preprint media coverage could benefit publics by connecting them with timely and relevant public health information, it could prove problematic if the uncertainties associated with the research are not made transparent. In this study, we address this tension by analyzing the framing devices used by digital media outlets to emphasize the scientific uncertainty of COVID-19-related preprints in the early stages of the pandemic.

## Literature Review

### *Preprints and scientific uncertainty*

Preprints are generally recognized by the scientific community as unvalidated and uncertain science, and journalists have been reluctant to report on them (Associated Press Stylebook 2020: And Briefing on Media Law, 2020; Haelle,

2020; Kille, 2015). Yet, this reluctance eased during the early months of the COVID-19 crisis, with online media coverage of COVID-19-related preprints outstripping that of preprints about any other topic (Fraser et al., 2020). While timely reporting of this emerging research is important for information-seeking publics, it can also mislead if findings are reported too early, without validation from the research community (Kharasch, 2020). While the potential for results to be invalidated through subsequent studies is an inherent aspect of science (see S. Schneider, 2016), this ever-present “scientific uncertainty” (Gustafson & Rice, 2019) may be further amplified in the case of preprints. Without peer review, results that are not supported by a wider scientific community can spread; indeed, a widely circulated COVID-19 preprint linking the SARS-CoV-2 spike protein to HIV-1 glycoproteins was later withdrawn by the authors because of criticism from other scientists about methodological flaws and faulty interpretation of results (Fraser et al., 2020).

The scientific uncertainty inherent in COVID-19 preprints presents challenges when communicating research findings. Journalists may ignore uncertainty when sharing research findings, particularly when dealing with risk communication (Peters & Dunwoody, 2016). Media have historically been found to gloss over unknowns and uncertainties when covering health issues (Dan & Raupp, 2018; Hove et al., 2015; Jung Oh

et al., 2012), perhaps to reduce the risk of alienating audiences with limited understanding of scientific work (S. Schneider, 2016; Stroobant & Raeymaeckers, 2019). Yet, communicating scientific uncertainty can be beneficial. For example, Jensen (2008) found that college students viewed both journalists and scientists as more trustworthy “when news coverage of cancer research was hedged (e.g., study limitations were reported)” and “when the hedging was attributed to the scientists responsible for the research” (p. 347).

Importantly, the choice to highlight scientific uncertainty varies by topic; media coverage of controversial research areas may exaggerate scientific uncertainties and disputes to appear “balanced” (Corbett & Durfee, 2004; Dixon & Clarke, 2012; Zehr, 1999); to add conflict to stories (e.g., J. Schneider, 2010); or to involve researchers debating scientific uncertainties among themselves (Dunwoody, 1999). Journalists are also guided by audience expectations and influenced by the practices of their colleagues, editors, and competitors when reporting scientific uncertainty (Guenther et al., 2015; Guenther & Ruhrmann, 2016). Media outlets’ portrayal of uncertainty becomes all the more important in times of crisis, when individuals look to them for timely guidance (Austin et al., 2012). Best practices in public health risk communication are grounded in core communication values associated with fostering public understanding of risks (Sellnow et al., 2009), which include transparency and credibility (Covello et al., 1989). To help publics navigate the risks associated with health crises like COVID-19, communicators should strive to be honest, frank, and open—clearly addressing unknowns and uncertainties—and coordinating and collaborating with trustworthy sources (Covello & Allen, 1988).

### **Framing uncertainty through hyperlinks**

Definitions and theoretical perspectives of framing vary widely (Entman, 1993; Scheufele, 1999; Scheufele & Scheufele, 2010); however, analyzing *emphasis frames* (Chong & Druckman, 2007), which “select some aspects of a perceived reality and make them more salient” (Entman, 1993, p. 52), can be useful for understanding what journalistic content communicates (e.g., Guenther et al., 2019; Semetko & Valkenburg, 2000). However, a review of the literature by Guenther et al. (2020) points to the paucity of emphasis framing studies in health communication that use content analyses to investigate media reporting. We aim to help fill this gap with a timely study investigating the framing devices that emphasize the (un)certainty of preprint research in digital media stories.

According to Coddington (2012), textual references associated with hyperlinks are particularly important for framing researchers to investigate because the language associated with hyperlinks “work together to frame the content and context of the hyperlink” (p. 2018). Yet, hyperlinks do more than frame. Online health media stories often use them to cite authoritative sources, such as academic research sites or government resources (Karlsson & Sjøvaag, 2018; Stroobant & Raeymaeckers, 2019). In theory, these “hyperlinks as citations” (Karlsson & Sjøvaag, 2018; p. 1) act as credibility markers (Coddington, 2012; Luzón, 2009; Stroobant & Raeymaeckers, 2019), influencing how audiences perceive and trust media messages (Borah, 2014). However,

hyperlinks as citations do not always fulfill this role in online health stories. In a study of opioid-related research media coverage in the US and Canada, Matthias et al. (2020) found that journalists incorporate research into their published work via hyperlinks to peer reviewed articles, but provide little context or information to help readers evaluate the validity of these studies and the certainty of claims. Importantly, this tendency seems to be more common in media stories that report on research in the context of some larger issue than in stories that focus on the research itself (Matthias et al., 2020).

### **Scientific uncertainty in a changing media landscape**

Researchers have found that the way media stories represent scientific (un)certainty depends on the reporting context (Peters & Dunwoody, 2016). According to others, these contextual differences may only amplify as the media landscape diversifies, and as blogs, aggregators, and other “digital-native” media outlets join traditional journalistic news sources (Barthel, 2019; Berkowitz, 2009; Bruns, 2018; Hermida, 2019; Stocking, 2019). For example, Hurley and Tewksbury (2012) found that focused providers (e.g., *New York Times*, MSNBC) differed from news aggregators both in terms of the frames they used and the degree of (un)certainty they incorporated into their coverage. Similarly, an analysis comparing an “independent” and a “mainstream” media outlet in New Zealand found notable differences in the frames the outlets used when communicating about the relationship between climate change and health, with the mainstream outlet favoring negative and sensationalist framing (Harrison et al., 2020).

Despite the growing popularity of digital-native news outlets (Stocking, 2019), digital news startups (Carlson & Usher, 2016), and native in-platform publishing (Bruns, 2018), these content providers have been largely overlooked in media scholarship (Hurley & Tewksbury, 2012; Lee & Chyi, 2015). When they have been studied, they have often been conceptualized as “periphery” to “core” legacy outlets—amassed into catch-all categories like “hybrid” journalism or dismissed as low-quality—rather than examined as integral, interconnected components of a diverse, ever-changing media ecosystem (Bakker, 2012; Deuze & Witschge, 2018; Witschge et al., 2019). As the boundaries between newsmakers, reporters, consumers, and distributors blur, drawing distinctions between content providers and curators has become more difficult—and, arguably—less valid (Berkowitz, 2009; Hermida, 2019; Jenkins & Deuze, 2008). For example, while so-called “periphery” outlets may rely more heavily on nontraditional practices such as republishing stories produced by other outlets or relying on publicity materials for content, these practices have increasingly been adopted by “core” outlets as well (Bakker, 2012). Similarly, while news bloggers and aggregators can be viewed as “parasitic” competitors to mainstream news outlets, they can also complement their work by amplifying their stories and increasing their web traffic (Bruns, 2018; Lee & Chyi, 2015). Additionally, these publishers often mimic the norms and values of professional journalism. As Coddington (2019) argues, an “amalgam of standards and practices shapes

aggregation as a hybrid practice that is built on professional journalism yet marginal within it” (p. 1).

In response to these ongoing transformations to the media landscape, researchers have called for “scholarship to address the dance between stability and change, to capture the diversity in the field” (Witschge et al., 2019, p. 655). This article responds to that call by exploring how a diversity of online media outlets represent—or frame—scientific (un)certainly in preprints about COVID-19. We answer three interconnected research questions:

RQ1. What content producers and curators (outlets) in the media ecosystem are communicating about COVID-19-related preprints?

RQ2. How are outlets using hyperlinking practices when communicating about COVID-19-related preprints?

RQ3. How are outlets using framing devices that emphasize uncertainty when communicating about COVID-19-related preprints?

## Method

### Sample selection and collection

To understand how preprints on topics related to COVID-19 were reported in online media, we focused our analysis on preprints posted on medRxiv and bioRxiv—the two top-ranked preprint servers for publishing studies related to COVID-19 (Kwon, 2020). These two servers noted a rapid uptake of COVID-19-related submissions in the early months of 2020 (Fraser et al., 2020) and are among the most widely used preprint servers for biomedical research (Polka & Penfold, 2020). We relied on the dataset from Fraser et al. (2020) that includes all the submissions published in both of these servers between January 1 and April 30, 2020 that were available through the bioRxiv Application Programming Interface (API<sup>1</sup>) on May 1, 2020. From the original set of 14,812 preprints, we used the 2,527 (17.06%) preprints that Fraser et al. (2020) identified as COVID-19-related through the presence of relevant terms in the preprints’ title or abstract.

On June 1, 2020, we searched for these 2,527 COVID-19-related preprints in the Altmetric database by querying the Altmetric Explorer<sup>2</sup> with their Digital Object Identifiers (DOIs). Altmetric tracks online activity of research, including references or “mentions” in media stories, by identifying links to publications (i.e., a hyperlink or a publication identifier, such as a DOI) and by regularly scanning the text of thousands of media stories and using natural language processing techniques to identify study details such as author names, journal titles, and study timeframes (Altmetric.com, 2018). This search yielded 14,717 total media mentions across 801 (31.7%) of the 2,527 COVID-19-related preprints. We noted that some outlets published multiple stories mentioning the same preprint. To avoid double counting and to understand how outlets first introduced these preprints to their readers, we kept only the first mention of each preprint by each outlet. This restricted the set to 10,572 mentions across the same 801 preprints. We further limited our study to the 8,270 mentions with titles identified as being in English via the langdetect<sup>3</sup> Python library.

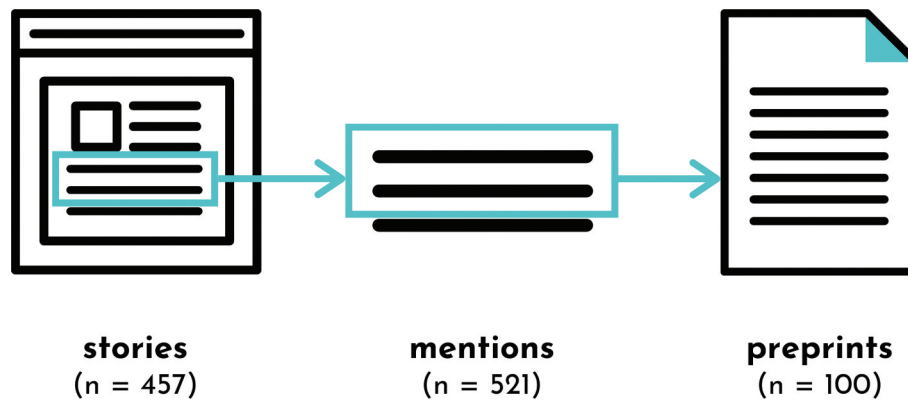
To identify which outlets in the media ecosystem were most actively communicating about COVID-19-related preprints, we restricted our sample to the 15 media outlets most prevalent in this subset of English-language mentions (after eliminating two sources: 1) *Infosurhoy* because all URLs Altmetric had collected for this source were invalid, and 2) *Google News* because URLs attributed to this source redirected to stories posted on other media outlets). Importantly, Altmetric applies a broad conception of media outlets—one that does not filter by history, audience size, or influence. As such, the 15 outlets we analyzed were not necessarily those with the greatest readership or level of public recognition, but rather the outlets incorporating the greatest number of COVID-19-related preprints into their coverage during the study time period.

After the restrictions listed above, these 15 outlets accounted for 1,117 mentions. To focus our analysis on preprints that were circulated widely, we further restricted our sample to those mentions that were about the 100 most mentioned preprints in the original dataset.<sup>4</sup> This final dataset comprised 590 mentions in 457 stories across the 15 outlets (a media story can contain multiple mentions of different preprints; see Figure 1, below). We removed 69 mentions because of broken story URLs or because the stated preprint was not actually mentioned in the story. The remaining 521 mentions were quantitatively analyzed.

### Codebook development

Following Evensen and Clarke (2012), our codebook was developed deductively, drawing on both relevant scholarly literature and professional guidelines. We adapted codes from previous studies examining (un)certain or initial scientific evidence in media stories (Dumas-Mallet et al., 2018; Matthias et al., 2020) and informed by relevant work on media framing of scientific (un)certainly (Dan & Raupp, 2018; Gustafson & Rice, 2019; Jung Oh et al., 2012; Nisbet et al., 2003). We referenced professional journalism resources (e.g., tip sheets, blog posts, stylebooks) describing best practices for reporting on COVID-19 preprints to complement the scholarly literature (Associated Press Stylebook 2020: And Briefing on Media Law, 2020; Helmuth, 2020; Jaklevic, 2020; Khamsi, 2020; Ordway, 2020).

Coding was binary (cf. Hart & Feldman, 2014), with variables of interest broken down into multiple questions to be coded as either 0 = *no/false* or 1 = *yes/true* (cf. Semetko & Valkenburg, 2000). Although the codebook emphasized manifest content, it allowed for implicit or latent interpretations (Evensen & Clarke, 2012), as frames may be interpreted somewhat differently by different individuals (Entman, 1993; Gamson & Modigliani, 1989; Price & Tewksbury, 1997). Coding assessed both the story overall (e.g., “Is this story a published press release?”) as well as the specific mention of the preprint in question (e.g., “Does the story mention that the study is a preprint?”). Given that transparency is key for valid, reliable framing content analysis (Matthes & Kohring, 2008), the complete codebook is available in the supplementary materials. Brief definitions and examples of codes are available in Table 1, below.



**Figure 1.** Conceptualization and sample size of preprints, mentions, and stories. In this study, we define “stories” as media articles that implicitly or explicitly refer to one or more COVID-19-related preprints and “mentions” as those parts of the story that relate to the preprint in question.

**Table 1.** Overview of codes.

Code	Description	Examples
science communication story	The primary focus of the story is to communicate the results and/or implications of the preprint.	<ul style="list-style-type: none"> <li>• Study of twins reveals genetic effect on COVID-19 symptoms [headline]</li> <li>• New MIT machine learning model shows relaxing quarantine rules will spike COVID-19 cases [headline]</li> </ul>
reposted story	The story was first published by another source.	<ul style="list-style-type: none"> <li>• This article by Joseph Eisenberg Professor and Chair of Epidemiology at the University of Michigan, first appeared in The Conversation on February 5, 2020. [attribution line]</li> <li>• This story is auto-aggregated by a computer program and has not been created or edited by Dailyhunt. Publisher: News Karnataka [disclaimer]</li> </ul>
press release	The story is a published press release.	<ul style="list-style-type: none"> <li>• Provided by Leiden University [attribution line]</li> <li>• SALINAS, CA – Congressman Jimmy Panetta (D-Carmel Valley) joined 63 Representatives in calling on Administration officials for improved testing ... [opening line]</li> </ul>
defines preprint	The story defines preprints in some way.	<ul style="list-style-type: none"> <li>• ... in one preprint study, meaning it is currently under peer review ...</li> <li>• Pre-prints are a way of getting research out quickly to get rapid responses, without waiting for peer-review, but they have some really important limitations.</li> </ul>
mentions “preprint”	The story mentions that the study is a preprint.	<ul style="list-style-type: none"> <li>• Two new preprints about the likely prevalence of the novel coronavirus ...</li> <li>• Another preprint study of outbreaks in Japan suggests ...</li> </ul>
mentions work is unreviewed	The story explains that the study has not been peer reviewed.	<ul style="list-style-type: none"> <li>• Their results, published Friday in a study that has yet to be peer-reviewed ...</li> <li>• However, a recent study under review shows ...</li> </ul>
mentions work is preliminary	The story suggests that the study is preliminary.	<ul style="list-style-type: none"> <li>• On 20 February the researchers posted a preliminary version of the study ...</li> <li>• The research is still early.</li> </ul>
mentions verification is needed	The story suggests that the study results are inconclusive (i.e., should be replicated or verified).	<ul style="list-style-type: none"> <li>• ... clearly further scientific research is required to substantiate these claims.</li> <li>• The researchers called for the “immediate validation” of the results.</li> </ul>
indicates mention is research	The story refers to the preprint as scientific research.	<ul style="list-style-type: none"> <li>• A new study suggests ...</li> <li>• A startling paper by a team of French scientists ...</li> </ul>
includes a hyperlink to preprint	The story contains a hyperlink to the study it cites.	<ul style="list-style-type: none"> <li>• A small <u>study</u>, done in macaque monkeys, shows ... [underline = hyperlink]</li> <li>• A separate <u>multi-center comparative clinical trial in China</u> indicated ...</li> </ul>
link does not indicate it is research/preprint	The story hyperlinks to the study, but does not make it clear that it is a research study or a preprint.	<ul style="list-style-type: none"> <li>• Brazil has many advantages over its neighbors for an effective pandemic response: <u>universal health coverage</u>, a large community-based primary care delivery system ...</li> <li>• The infection risk is especially high among <u>household contacts</u>.</li> </ul>

### Intercoder reliability

Following best practices for mass communication research (Lacy et al., 2015), coding was performed by one researcher (AF) and

a second, independent coder who was not involved in developing the codebook (cf. Strekalova, 2015). The lead author tested the codebook by coding 69 representative stories that were not part of



**Table 2.** Intercode reliability scores.

Code	Krippendorff's alpha
science communication story	.97
reposted story	.97
press release	1.00
defines preprint	.76
mentions "preprint"	.81
mentions work is unreviewed	.96
mentions work is preliminary	.88
mentions verification is needed	.91
indicates mention is research	.92
includes a hyperlink to preprint	.75
link does not indicate it is research/preprint	.75

the main coding (i.e., they mentioned COVID-19-related medRxiv and bioRxiv preprints in the top-15 outlets, but mentioned preprints not among the top 100 used in our final sample) and refined the codebook as needed (Lombard et al., 2002). A second coder was provided with the codebook and a demonstration of the method. The second coder then independently coded the same 69 stories. All coding was performed using Excel (cf. Evensen & Clarke, 2012). Krippendorff's alpha reliability scores were calculated using Python's krippendorff library.<sup>5</sup> Given that this measure of intercode reliability is conservative and that the study is exploratory, we set the minimum acceptable level of reliability at .70 (Lacy et al., 2015; Lombard et al., 2002). This level was met or exceeded for all codes, as shown in Table 2 (below).

### Intercode reliability scores

After intercode testing, both coders met to discuss discrepancies, particularly regarding the codes with lower levels of agreement: *defines preprint*, *includes a hyperlink to preprint*, and *link does not indicate it is research/preprint*. Sources of disagreements were identified, and appropriate coding approaches were reviewed and clarified in the codebook (e.g., preprint definitions do not have to be correct to be coded as *defines preprint*; a hyperlink to a different version of a preprint than the one in our data should be coded as *includes a hyperlink to a preprint*). Finally, the main data set (n = 521 stories, none of which were used for the intercode reliability test) was divided for coding by the two coders. Although coding was largely performed individually, coders consulted with one another to resolve difficult cases, discussing possibilities until they reached consensus (cf. Evensen & Clarke, 2012).

### Statistical methods

Binary logistic regressions were performed using the Python statsmodels<sup>6</sup> package.

### Results

In the following, we present the findings of our study alongside the three research questions.

#### RQ1. What content producers and curators (outlets) in the media ecosystem are communicating about COVID-19-related preprints?

The data collection and analysis yielded a diverse set of outlets that mentioned COVID-19-related preprints most frequently in the study period (Table 3). These outlets included legacy media (e.g., *The Guardian*, *New York Times*), digital-native news outlets (e.g., *Inverse*), medical-niche publications (e.g., *News Medical*, *MedicalXpress*, *Medscape*), technology-niche publications (e.g., *Business Insider*, *Wired*), Web portals (e.g., MSN, Yahoo! News), a native in-platform publisher (Medium), several news aggregators (e.g., *Dailyhunt*, *National Interest*), and *The Conversation*, a nonprofit outlet that "sources its content exclusively from university scholars and provides journalistic editing services to its authors" (Bruns, 2018, p. 52-53). Regardless of categorization of these outlets, all but one (*Inverse*) showed a clear tendency to either publish reposted content (i.e., stories aggregated from other media outlets or press releases) or publish original content. We loosely categorized these outlets as "aggregators," defined, for the purposes of this study, as media outlets for which at least two thirds of stories were originally published by another source (i.e., were coded as reposted stories). Although we did not include press releases in this analysis, we note that doing so does not change the categorization of any outlet. *News Medical* published the greatest proportion of stories that we categorized as science communication stories (n = 15, 65.2%), where the primary focus was to communicate the research results and/or implications of the preprint. *MedicalXpress* and *Medscape* also published a large proportion of science communication stories (30.2% and 33.3%, respectively).

**Table 3.** Number and type of stories mentioning COVID-19-related preprints published by top 15 outlets.

Outlet	Total Stories	Press Releases		Reposted Stories		Science Communication (scicomm) Stories	
		Number	Percent	Number	Percent	Number	Percent
Business Insider	31	0	0	6	19.4	3	9.7
Dailyhunt*	27	1	3.7	26	96.3	8	29.6
Foreign Affairs New Zealand*	25	6	24	18	72.0	5	20.0
Inverse	22	0	0	10	45.5	3	13.6
MedicalXpress*	43	10	23.3	29	67.4	13	30.2
Medium	25	0	0	0	0.0	3	12.0
Medscape	15	0	0	1	6.7	5	33.3
MSN*	36	0	0	35	97.2	7	19.4
New York Times	29	0	0	0	0.0	4	13.8
The Conversation	41	0	0	0	0.0	2	4.9
The Guardian	24	0	0	0	0.0	5	20.8
News Medical	23	2	8.7	4	17.4	15	65.2
National Interest*	32	1	3.1	26	81.2	1	3.1
Wired	17	0	0	2	11.8	3	17.6
Yahoo! News*	67	6	9	61	91.0	16	23.9

NB: Media outlets categorized as aggregators are demarcated with an asterisk.

**Table 4.** Number and percent of stories by hyperlinking practice and outlet.

Outlet	Includes a hyperlink to preprint		Indicates mention is research		Link does not indicate it is research/preprint	
	Number	Percent	Number	Percent	Number	Percent
Business Insider	30	96.8	28	90.3	3	9.7
Dailyhunt	9	33.3	21	77.8	2	7.4
Foreign Affairs New Zealand	24	96.0	17	68.0	9	36.0
Inverse	22	100.0	17	77.3	5	22.7
MedicalXpress	39	90.7	33	76.7	11	25.6
Medium	24	96.0	19	76.0	8	32.0
Medscape	14	93.3	14	93.3	1	6.7
MSN	30	83.3	30	83.3	2	5.6
New York Times	28	96.6	26	89.7	2	6.9
The Conversation	41	100.0	25	61.0	18	43.9
The Guardian	24	100.0	22	91.7	2	8.3
News Medical	20	87.0	22	95.7	0	0.0
National Interest	32	100.0	20	62.5	16	50.0
Wired	17	100.0	17	100.0	0	0.0
Yahoo! News	63	94.0	57	85.1	9	13.4
Total	417	91.2	368	80.5	88	19.3

NB: Because some stories mention more than one COVID-19-related preprint, a story may be counted as both having only a hyperlink and as indicating the mention pertains to research (i.e., if a story cites two different preprints, one may be described as research and the other included with only a hyperlink).

## RQ2. How are outlets using hyperlinking practices when communicating about COVID-19-related preprints?

As there were no notable differences in the hyperlinking practices of outlets we categorized as aggregators and those that mostly posted original content, we do not draw distinctions in the rest of our results. Similarly, our findings remained unchanged whether or not press releases were included in the analyses; because of this—and because there were so few press releases in our sample ( $n = 26$ )—we do not make comparisons with this group in the following results. The vast majority of stories mentioned a single preprint ( $n = 419$ , 91.7%) or two of the top 100 COVID-19-related preprints that comprised our sample ( $n = 28$ , 6.1%). The remaining stories mentioned between three and five preprints. We also noted that many stories mentioned preprints beyond our sample, but these were not systematically studied.

For the remaining analysis, we considered a story to include a particular practice if it was used when mentioning at least one of the preprints associated with that story. In most cases, stories included a hyperlink to a preprint ( $n = 417$ , 91.2%). Similarly, the majority of stories indicated that what was being mentioned was research ( $n = 368$ , 80.5%), for example, by referring to the preprint as “a study” or “new research.” Nearly 20% of stories hyperlinked to a preprint without any indication of what the hyperlink pointed to ( $n = 88$ ).

However, these practices varied by outlet (Table 4). While most outlets in our sample included a hyperlink to the COVID-19 preprints in their stories, only 27 (33.3%) of Dailyhunt’s stories had at least one hyperlink to a preprint (all other outlets included hyperlinks in over 80% of their stories; five outlets included them in 100%). Two outlets stand out for the infrequency of references to research; *The Conversation* and the *National Interest* only described the preprints they mentioned as research research in approximately 60% of their stories. In many of these instances, the

two outlets simply hyperlinked to a preprint without further indication that the hyperlink led to research (i.e., there was no mention of words such as “study” or “findings”). For example, a story from *The Conversation* titled “Predicting COVID-19: what applying a model in Kenya would look like” (Nanyingi, 2020) stated (hyperlink to preprint underlined):

There is an urgent need for serological tests. These find antibodies in the blood—molecules made by the immune system in response to a pathogen’s attack—and would measure how much the virus spread and how many people recovered.

While these outlets had the highest number of stories that included hyperlinks without indication of research (43.9% and 50%, respectively), two additional outlets published over 30% of their stories with uncontextualized hyperlinks (*Foreign Affairs New Zealand* and *Medium*) and two more had over 20% of stories with such hyperlinks (*Inverse* and *MedicalXpress*).

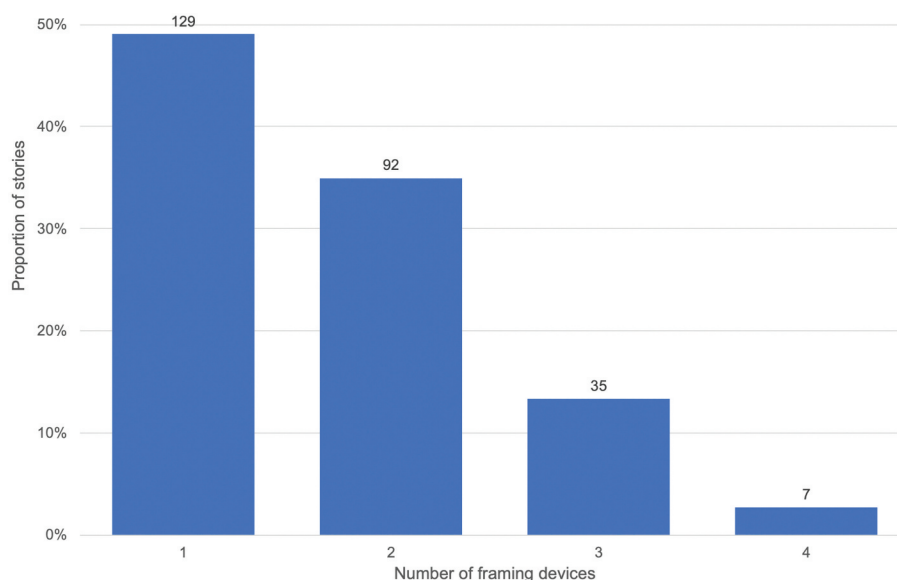
## RQ3. How are outlets using framing devices that emphasize uncertainty when communicating about COVID-19-related preprints?

Regardless of the practices used to identify or hyperlink to a preprint posted on medRxiv and bioRxiv, more than half of all stories made use of one or more framing devices to emphasize scientific uncertainty (i.e., they mentioned that the study was a preprint, unreviewed, preliminary, and/or in need of verification;  $n = 263$ , 57.5% of stories). In nearly half of these instances, the stories included a single framing device ( $n = 129$ , 49%), whereas 92 stories (35%) included two framing devices, 35 stories (13.3%) included three devices, and the remaining seven (2.7%) had all four devices (Figure 2).

Noting that content was unreviewed was the most common uncertainty framing device, appearing in 172 (37.6%) of stories, followed in equal numbers by identifying the content as a preprint and noting that further verification was needed ( $n = 99$ , 21.7% for both devices). Indicating that the work was preliminary was the least common uncertainty framing device, appearing in only 76 (16.6%) of the stories.

The devices used to frame uncertainty about COVID-19-related preprints varied by outlet. While some outlets indicated some form of uncertainty framing in over 80% of their stories (i.e., *News Medical*, *Medscape*, *Wired*), others did so less than half the time (*The Conversation*, *Foreign Affairs New Zealand*, *Medium*, *New York Times*) (Table 5). To test the significance of these differences, we calculated a logistic regression that examined whether the probability of finding an uncertainty framing device varied depending on the publication outlet. More formally, we calculated a model in the form  $P(Y = 1) = \beta_0 + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{15} x_{15})$ , where  $Y$  is a binary outcome variable coded as 1 if at least one uncertainty framing device was used and 0 otherwise, and  $x_1 x_2 \dots x_{15}$  are a set of predictor variables corresponding to each of the 15 outlets. A Wald Test easily rejects the null hypothesis that the outlets are equally likely to use at least one uncertainty device ( $F = 39.32$ ,  $p < .001$ ).

Some outlets, such as *News Medical*, *MedicalXpress*, and *Wired*, included specific definitions of preprints within their stories. We identified a variety of definitions in 46 stories



**Figure 2.** Proportion of stories using different numbers of framing devices that emphasize scientific uncertainty

**Table 5.** Percentage of stories by type of uncertainty and outlet.

Outlet	Mentions “preprint”	Mentions work is unreviewed	Mentions work is preliminary	Mentions verification is needed	Any of four devices
Business Insider	6.5	54.8	19.4	9.7	67.7
Dailyhunt	22.2	48.1	7.4	22.2	59.3
Foreign Affairs New Zealand	24.0	28.0	12.0	16.0	44.0
Inverse	27.3	54.5	31.8	27.3	72.7
MedicalXpress	30.2	34.9	4.7	23.3	53.5
Medium	12.0	16.0	8.0	16.0	32.0
Medscape	73.3	40.0	33.3	33.3	86.7
MSN	13.9	44.4	22.2	25.0	61.1
New York Times	3.4	37.9	17.2	20.7	44.8
The Conversation	12.2	24.4	9.8	22.0	41.5
The Guardian	12.5	45.8	29.2	29.2	62.5
News Medical	65.2	39.1	13.0	21.7	82.6
National Interest	12.5	28.1	21.9	15.6	53.1
Wired	64.7	52.9	41.2	23.5	88.2
Yahoo! News	11.9	34.3	11.9	23.9	55.2

(10.1%). For example, some stories highlighted the uncertain nature of preprints in their definitions, such as this one published by *News Medical* (Mandal, 2020):

The preprint paper is a version of a scholarly or scientific paper that precedes formal peer review and publication in a peer-reviewed scholarly or scientific journal ... medRxiv publishes preliminary scientific reports that are not peer-reviewed and, therefore, not be regarded as conclusive, guide clinical practice/health-related behavior, or treated as established information.

Other outlets offered a more positive spin in their preprint definitions. For example, a *Wired* story titled “Blood From Covid-19 Survivors May Point the Way to a Cure” (Foreign Affairs New Zealand, 2020) described them as “the 21st century way to report data almost in real time.”

We identified many of these preprint definitions in stories that specifically indicated that the study mentioned was a preprint (29 of 99 stories, 29.3%). Outlets, such as the *New York Times*, *Medscape*, and *Wired*, had previously published stories specifically about peer review, preprints, and COVID-19, which they occasionally hyperlinked to in their COVID-19 coverage as a shortcut for defining preprints. For example:

“The paper, which has not yet undergone peer review, appeared on the [Medrxiv preprint server](#).” (from *Wired* [Molteni, 2020]; the bolded hyperlink leads to the story “Biology’s Roiling Debate Over Publishing Research Early,” which provides an overview of what preprints are, why they can be beneficial, as well as why they can be detrimental)

“The study, however, was published on a [preprint server](#), [medRxiv](#), where, as *Medscape* readers know, researchers publish early versions of a manuscript before they are peer-reviewed.” (from *Medscape* [Coffey & Oransky, 2020]; hyperlink leads to the article “To Maintain Trust in Science, Lose the Peer Review,” which explores the pitfalls of peer review and the barriers media face when covering research that is not open access)

“The research was posted on [MedRxiv](#), a website where scientists have been posting articles [submitted for publication elsewhere that have not yet been through peer review](#).” (from the *New York Times* [Yan, 2020]; hyperlink leads to a story on preprints, including their strengths and weaknesses, titled “Coronavirus Tests Science’s Need for Speed Limits”)

All outlets in our sample, particularly those focused on medical issues (*News Medical*, *MedicalXpress*, and *Medscape*), published stories specifically about the COVID-19-related preprints in our study (Table 3). We coded these stories as



“science communication” if the results and/or implications of the preprint were the primary focus. We estimate the probability of whether a story will include an uncertainty framing device if the story is coded as a science communication story from the logistic model in the form  $P(Y = 1) = \beta_0 + \beta_1 scicomm$ , where  $Y$  is a binary outcome variable coded as 1 if at least one uncertainty framing device was used and 0 otherwise, and where  $scicomm$  is a predictor variable coded as 1 if the preprint was the primary focus of the story and 0 otherwise. We found that such stories were statistically more likely to include some uncertainty framing device when compared to other stories (odds ratio = 9.64,  $p < .001$ ). We ran a similar model using the 26 stories that were coded as press releases and found an increased likelihood that these contained an uncertainty framing device (odds ratio = 2.58,  $p = .046$ ). However, the significance of this effect disappears when both variables are considered in the same model ( $p = .337$ ), while the increased likelihood of science communication stories containing an uncertainty framing device remains statistically different from zero ( $p < .001$ ).

## Discussion

This study aimed to identify what digital content providers are communicating about COVID-19-related preprint research and the hyperlinking practices these outlets use when mentioning preprints in media stories. Because preprint research is characterized by a high level of scientific uncertainty (Berg et al., 2016; Chiarelli et al., 2019; Fry et al., 2019), we were particularly interested in how these outlets framed this uncertainty.

To inform this study, we drew on research about scientific (un)certainty in health reporting (Hove et al., 2015; Jung Oh et al., 2012; Matthias et al., 2020) and work documenting the changing media landscape (Bakker, 2012; Bruns, 2018; Hermida, 2019; Lowrey, 2012). We found that a diverse range of legacy and digital-native content providers and curators are using the 100 most-mentioned COVID-19-related preprints in stories, and each media organization had a unique approach to covering them. We loosely categorized these outlets as legacy, aggregators, medical-niche, digital-native news, technology-niche, and outlets such as Medium and *The Conversation*, which resist categorizations found in existing research.

We recognize that digital outlets will continue to diversify and innovate (Bruns, 2018; Lowrey, 2012; Witschge et al., 2019), but we also note trends in how the outlets identified in our study used preprints in their stories. Hyperlinking was a ubiquitous practice, with over 90% of stories we analyzed including a hyperlink to at least one preprint. Identifying those preprints and hyperlinks as pointing to research was also common. This standardization or routinization of practices points to some stability in the media landscape (Lowrey, 2012)—at least among the 15 outlets that were most active in covering COVID-19-related preprints. Existing research has similarly documented a tendency among digital communicators to hyperlink to academic sources, often to demonstrate credibility and transparency (Coddington, 2012; Karlsson & Sjøvaag, 2018; Stroobant & Raeymaeckers, 2019). However,

our findings extend the literature by examining how hyperlinks to COVID-19-related preprints were used as citations by a diversity of media outlets, and how even uncertain science may be leveraged as a credibility marker—especially when described as “research” rather than as a preprint. Indeed, outlets in our sample were much less likely to identify the research they mentioned as preprints—perhaps to maintain credibility, but perhaps also to avoid alienating readers with limited knowledge of scientific methods. Although more research is needed to understand the motivations behind media outlets’ use of different uncertainty framing devices, avoiding terms like “preprint” may be a strategic editorial approach adopted by media professionals who are known to pay close attention to audience preferences when making editorial decisions (Arenberg & Lowrey, 2019; McKenzie et al., 2011; Tandoc, 2015; Vu, 2014).

Some outlets (e.g., *Medscape*, *Wired*) framed the preprints they mentioned as uncertain in almost every story; others (e.g., *The Conversation*, *New York Times*) did so in less than half of the stories we analyzed. Over 40% of stories in this study did not frame the preprint as uncertain at all; of those that did, most included just a single framing device—typically a statement that the research had not been peer reviewed. This may be because peer review is a feature that obviously distinguishes preprints from other research articles (Fraser et al., 2020; da Silva, 2020), but it may also suggest that some communicators are wary of emphasizing uncertainties that are more directly tied to the quality of the research, such as its preliminary nature or the need for verification, as these could potentially undermine credibility and trust (Frewer et al., 2003; van der Bles et al., 2020). These findings align with previous studies examining editorial framing of health issues (e.g., Dan & Raupp, 2018; Hove et al., 2015; Jung Oh et al., 2012; Matthias et al., 2020) that find science is seldom framed as uncertain. In our study, less than half of the stories we analyzed contained framing devices emphasizing uncertainty; yet, this was still far more frequent than past studies have reported, perhaps because of the nature of the topic of our study. Previous studies have largely focused on media coverage of peer reviewed health research, while our work focused specifically on preprints, which are scientifically uncertain research. In the context of risk communication where transparency is deemed essential (Bourrier, 2018; Covello & Allen, 1988; O’Malley et al., 2009; Sellnow et al., 2009), some media outlets may recognize the importance of framing them as such. Indeed, although research examining media coverage of unreviewed health science is limited, available evidence from beyond the COVID-19 context suggests media outlets may frame scientific uncertainty differently when communicating during a public health crisis. These studies—which focused on media coverage of other forms of preliminary health research (e.g., findings from initial biomedical studies or medical conference proceedings)—found that only about one in five media stories mentioned the uncertain or unverified nature of the findings they communicated (Dumas-Mallet et al., 2018; Lai & Lane, 2009). While more research is needed in this area—particularly comparing media coverage of peer reviewed and unreviewed research findings, as well as the portrayal of preprints across different topics,

communication contexts, and outlets—it is encouraging to consider that media may be more attentive to addressing scientific uncertainties when such transparency has important implications for public health, such as the COVID-19 pandemic.

In comparing the hyperlinking and framing practices of these diverse outlets that mention preprints frequently, this study revealed greater similarities between outlets from different categories than previous research (Harrison et al., 2020; Hurley & Tewksbury, 2012; Stroobant, 2019)—at least in their coverage of COVID-19-related preprints. For example, both MSN and Dailyhunt can be considered aggregators, but MSN's communication practices had more in common with niche outlets such as *Wired* or *News Medical*, which were among the most likely to hyperlink to preprints and identify them as such. Similarly, while *The Guardian* and the *New York Times* shared similar hyperlinking practices (i.e., they almost always hyperlinked to the COVID-19-related preprints they mentioned, and rarely did so without identifying them as research), the *New York Times* was less likely to include a framing device emphasizing uncertainty in their stories. In this instance, the *New York Times* had more in common with outlets like *The Conversation* or *Foreign Affairs New Zealand*, both of which used fewer uncertainty framing devices than other outlets. These similarities across “categories” could suggest, as other scholars have argued, that drawing divides between legacy or “core” journalism and alternative, “peripheral” outlets may no longer make sense; the boundaries between them appear to have become blurred (Deuze & Witschge, 2018; see also Bruns, 2018; Chadwick, 2017; Hermida, 2019). Future studies examining the use of uncertainty framing devices among a larger sample of media outlets could provide important insights into the degree to which such blurring is, indeed, taking place.

Within our sample, the digital-native outlets (e.g., *The Conversation*, *MedicalXpress*, Yahoo! News) published the most stories citing preprints. This may be explained in part by the resources required to cover scientific research: when contributors get “no-pay” or “low-pay” for original content outlets have fewer financial barriers to providing content (Bakker, 2012; Coddington, 2019). The dominance of *The Conversation* in the lack of framing devices emphasizing uncertainty, however, is surprising, given its official terms and conditions specify that “Research, as a general principle, should not be reported before it has been subjected to a recognized process of peer review” (Terms and Conditions, 2020, para. 10). The urgent nature of a crisis situation like the COVID-19 pandemic may have prompted revisions to media policies by some outlets and provides an area for future research.

Professional journalism resources, such as tip sheets, blog posts, and stylebooks, that describe best practices for reporting on COVID-19 preprints (e.g., Associated Press Stylebook 2020: And Briefing on Media Law, 2020; Helmuth, 2020; Jaklevic, 2020; Khamsi, 2020; Ordway, 2020) recommend using the uncertainty framing devices that we investigate in this paper. Interestingly, aggregators like *MedicalXpress* and Yahoo! News were more likely to follow these guidelines when mentioning the most reported on COVID-19-related preprints, compared with some of the more “traditional” outlets in our sample. These findings depart from research suggesting that aggregators are

less likely to include uncertainty in their coverage (Hurley & Tewksbury, 2012) and more likely to provide shallow, sensational content (Coddington, 2019). While more research is needed in this area, one possible explanation is that aggregators are increasingly heterogeneous, with some licensing high-quality journalistic content. Aggregators, at least those identified in this study, may also have new content options outside traditional news media. For instance, *The Conversation*, with articles commonly republished by aggregators within the dataset, represents a relatively new form of content available for free and from researchers at a mix of academic institutions. Finally, several of the aggregators in our sample also occasionally reposted press releases, the majority of which were science communication stories and thus more likely to include at least one uncertainty framing device.

*The Conversation* also stands apart, as does Medium, for having the fewest stories hyperlinking to a highly mentioned COVID-19-related preprint without framing it as uncertain. Both outlets were relatively unlikely to identify those preprints as preprints. This may be expected in the case of Medium, which allows anyone to publish content with little or no editorial oversight<sup>7</sup>, and hence is likely to feature stories by authors with limited awareness or concern of what preprints are or the scientific uncertainty surrounding them. The limited uncertainty in *The Conversation* stories is more surprising, given academics' typical reliance on uncertainty framing devices to communicate scientific findings (Zehr, 1999). *The Conversation's* readership, while largely working in non-academic settings (Our Audience, 2020), is mostly university educated and it may be assumed that this audience is already knowledgeable about preprints. This raises issues around article sharing by outlets with readership beyond *The Conversation's* initial or intended audience. As such, *The Conversation's* distribution and editorial approach—what it calls “academic rigor, journalistic flair”<sup>8</sup>—present a tension deserving of further research.

Across the 15 outlets we analyzed, “science communication” stories—that is, stories focused on communicating the results or implications of a particular COVID-19-related preprint—were more likely to portray that preprint as uncertain compared to stories using preprints for other purposes (e.g., to cover a wider issue, to support an argument). This finding supports recent research by Matthias et al. (2020) which similarly found that scientific uncertainty was more likely to be conveyed in science communication rather than issue-focused stories. These findings are not surprising considering media preferences for novelty and significance; framing research as scientifically uncertain does not enhance a story unless the implications for the audience necessitate it (Fahnestock, 1986). This tendency may also come down to a question of word count and reader experience. Discussing the uncertainty associated with a preprint takes up valuable space—even in online publications that typically work to keep articles brief—and could disrupt the “flow” of a story (Van Leuven et al., 2018). While outlets and their editorial staff may deem an explanation of a preprint as warranted in a story focused on that research, they may be less likely to do so when a preprint is mentioned only in passing.

There are practical implications from these research findings that we wish to highlight. Media are a key source

of public health information during times of crisis (Austin et al., 2012); yet many of the outlets we analyzed do not appear to follow public health risk communication best practices when it comes to the portrayal of uncertainty surrounding COVID-19 preprint research—at least not consistently. We recognize that covering this unvalidated science poses challenges for public health risk communicators, as doing so requires balancing the public's need for timely, relevant information with risk communication best practices of communicating with transparency and openness regarding unknowns and uncertainties (Covello & Allen, 1988; Covello et al., 1989). Still, our findings suggest that achieving these seemingly conflicting aims is possible; *Wired* and the *New York Times* have both produced explanations about the uncertainty inherent in science, peer review, and the scientific method, and each outlet includes hyperlinks to these “meta”-science stories when mentioning preprint research. These practices could be seen as a first step toward a standardized industry practice, although they offer no guarantee that readers will click through for additional information (Yaros, 2011). Evaluating the certainty of scientific findings can be challenging for readers without a science background; however, readers can at least understand whether research is established or preliminary with the help of editorial framing devices such as those analyzed in this study. This seems especially important for global issues with such local and personal relevance as COVID-19.

These issues warrant further study and could be explored by building on the research method we employed, which is innovative in two ways. First, our approach builds on the emerging field of altmetrics (Erdt et al., 2016), which seeks new ways to capture how and when research is shared and communicated online, by analyzing the contexts into which research is mentioned. Second, while much previous research has analyzed uncertainty frames using a holistic approach (i.e., is the story uncertain?), often with nontransparent coding schemes (Matthes & Kohring, 2008), we offer our full coding scheme and break down uncertainty frames into four distinct framing devices, allowing the identification of the relative prevalence of each device overall, as well as for each outlet.

This study comes with limitations. First, we focused on the framing of uncertainty specific to COVID-19-related preprints, rather than uncertainty in these stories more generally. We see this focus as a strength, as, to our knowledge, no existing research has examined how preprints are portrayed in online media coverage. However, by restricting the unit of analysis to the preprint mention, we might have missed some of the context surrounding the preprint that could influence how the preprint is perceived by readers (Corbett & Durfee, 2004; Tewksbury & Riles, 2018). Second, we further restricted our sample to mentions of the 100 most mentioned COVID-19-related preprints, which may not be representative of coverage of less popular preprints. Indeed, media outlets sometimes take their lead from other outlets when deciding how and whether to cover issues (Golan, 2006; Wang & Guo, 2018); seeing a preprint mentioned by multiple outlets—especially without the inclusion of an uncertainty framing device—may have encouraged the outlets in our sample to see the research as sound and verified,

and portray it accordingly in their own coverage. Third, we found that a large proportion of our stories used preprints with a hyperlink, but this may be an artefact of how Altmetric tracks research mentions and could bias the data to include hyperlinked preprints over text-based mentions of preprints. While we can be certain that Altmetric's natural language processing is successful in identifying text-based mentions of preprints at least some of the time, the exact precision and recall of this approach remains unknown. We encourage scholars to complement our findings using other data sources and methodologies—as well as systematically evaluate the effectiveness of Altmetric's text-based data collection. Fourth, we restricted our sample to English-language stories from the 15 media outlets that mentioned the most COVID-19-related preprints, but these may not be the most influential or most read outlets covering the pandemic, nor representative of preprint coverage in other languages. Future research could examine whether our findings apply to international outlets, as well as those chosen based on the size and influence of their audiences or the reach of their stories. Finally, interviews with content providers and curators could advance our understanding of how they approach the communication of scientific uncertainty surrounding preprint research in their media stories and help us develop a more complete explanation of media communication of preprint research.

## Notes

1. <https://api.biorxiv.org>
2. <https://www.altmetric.com/explorer/>
3. <https://pypi.org/project/langdetect/>
4. Like media coverage of other scholarly research, mentions of COVID-19 preprints follow a skewed distribution, with a small number of highly influential preprints receiving the bulk of the coverage, and a long tail of other preprints that received only one or two mentions during the study period.
5. <https://pypi.org/project/krippendorff/>
6. <https://www.statsmodels.org/>
7. Medium offers guidelines and best practices, but does not enforce them <https://medium.com/creators>
8. <https://theconversation.com/>

## Acknowledgements

We gratefully acknowledge Altmetric.com for access to data on news media mentions.

## Funding

This work was supported by the Social Sciences and Humanities Research Council of Canada [435-2020-0401].

## ORCID

Alice Fleerackers  <http://orcid.org/0000-0002-7182-4061>  
 Michelle Riedlinger  <http://orcid.org/0000-0003-4402-4824>  
 Laura Moorhead  <http://orcid.org/0000-0001-9185-6290>  
 Rukhsana Ahmed  <http://orcid.org/0000-0003-0381-4491>  
 Juan Pablo Alperin  <http://orcid.org/0000-0002-9344-7439>



## References

- Altmetric.com. (2018). *How does Altmetric's text mining work?* Altmetric. <https://help.altmetric.com/support/solutions/articles/6000196551-how-does-altmetric-s-text-mining-work>
- Arenberg, T., & Lowrey, W. (2019). The impact of web metrics on community news decisions: A resource dependence perspective. *Journalism & Mass Communication Quarterly*, 96(1), 131–149. <https://doi.org/10.1177/1077699018801318>
- Associated Press Stylebook 2020: And Briefing on Media Law. (2020). New York: Basic Books. [https://www.apstylebook.com/ap\\_stylebook/health-and-science](https://www.apstylebook.com/ap_stylebook/health-and-science)
- Austin, L., Fisher Liu, B., & Jin, Y. (2012). How audiences seek out crisis information: Exploring the social-mediated crisis communication model. *Journal of Applied Communication Research*, 40(2), 188–207. <https://doi.org/10.1080/00909882.2012.654498>
- Bakker, P. (2012). Aggregation, content farms and huffinization. *Journalism Practice*, 6(5–6), 627–637. <https://doi.org/10.1080/17512786.2012.667266>
- Barthel, M. (2019). *Newspaper fact sheet: State of the news media*. Pew Research Center. <https://www.journalism.org/fact-sheet/digital-news/>
- Berg, J. M., Bhalla, N., Bourne, P. E., Chalfie, M., Drubin, D. G., Fraser, J. S., ... Wolberger, C. (2016). Preprints for the life sciences. *Science*, 352(6288), 899–901. <https://doi.org/10.1126/science.aaf9133>
- Berkowitz, D. (2009). Journalism in the broader cultural mediascape. *Journalism: Theory, Practice & Criticism*, 10(3), 290–292. <https://doi.org/10.1177/1464884909102587>
- Borah, P. (2014). The hyperlinked world: A look at how the interactions of news frames and hyperlinks influence news credibility and willingness to seek information. *Journal of Computer-Mediated Communication*, 19(3), 576–590. <https://doi.org/10.1111/jcc4.12060>
- Bourrier, M. (2018). Risk communication 101: A few benchmarks. In M. Bourrier & C. Bieder (Eds.), *Risk communication for the future* (pp. pp. 1–14). Springer.
- Bruns, A. (2018). *Gatewatching and news curation: Journalism, social media, and the public sphere*. Peter Lang.
- Carlson, M., & Usher, N. (2016). News startups as agents of innovation: For-profit digital news startup manifestos as metajournalistic discourse. *Digital Journalism*, 4(5), 563–581. <https://doi.org/10.1080/21670811.2015.1076344>
- Chadwick, A. (2017). *The hybrid media system: Politics and power*. Oxford University Press.
- Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). Preprints and scholarly communication: An exploratory qualitative study of adoption, practices, drivers and barriers. *F1000Research*, 8. <https://doi.org/10.12688/f1000research.19619.2>
- Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, 10(1), 103–126. <https://doi.org/10.1146/annurev.polisci.10.072805.103054>
- Coddington, M. (2012). Building frames link by link: The linking practices of blogs and news sites. *International Journal of Communication*, 6(2012), 2007–2026. <https://ijoc.org/index.php/ijoc/article/view/1476>
- Coddington, M. (2019). Aggregation and journalism. In *Oxford Research Encyclopedia of Communication*. <https://doi.org/10.1093/acrefore/9780190228613.013.778>
- Coffey, D., & Oransky, I. (2020, March 20). The week that wasn't in COVID-19: Centenarian survives, high-risk blood types. *Medscape*. <https://www.medscape.com/viewarticle/927299?src=rss>
- Corbett, J. B., & Durfee, J. L. (2004). Testing public (un)certainly of science: Media representations of global warming. *Science Communication*, 26(2), 129–151. <https://doi.org/10.1177/1075547004270234>
- Covello, V. T., McCallum, D. B., & Pavlova, M. (1989). Principles and guidelines for improving risk communication. In V. T. Covello, D. B. McCallum, & M. T. Pavlova (Eds.), *Effective risk communication* (pp. 3–16). Springer US. [https://doi.org/10.1007/978-1-4613-1569-8\\_1](https://doi.org/10.1007/978-1-4613-1569-8_1)
- Covello, V. T., & Allen, F. (1988). Seven cardinal rules of risk communication. US Environmental Protection Agency. Office of Policy Analysis.
- da Silva, J. A. T. (2020). An alert to COVID-19 literature in predatory publishing venues. *The Journal of Academic Librarianship*, 46(5), 1–2. <https://doi.org/10.1016/j.acalib.2020.102187>
- Dan, V., & Raupp, J. (2018). A systematic review of frames in news reporting of health risks: Characteristics, construct consistency vs. name diversity, and the relationship of frames to framing functions. *Health, Risk & Society*, 20(5–6), 203–226. <https://doi.org/10.1080/13698575.2018.1522422>
- Deuze, M., & Witschge, T. (2018). Beyond journalism: Theorizing the transformation of journalism. *Journalism: Theory, Practice & Criticism*, 19(2), 165–181. <https://doi.org/10.1177/1464884916688550>
- Dixon, C. E., & Clarke, G. N. (2012). Heightening uncertainty around certain science: Media coverage, false balance, and the autism-vaccine-controversy. *Science Communication*, 35(3), 358–382. <https://doi.org/10.1177/1075547012458290>
- Dumas-Mallet, E., Smith, A., Boraud, T., & Gonon, F. (2018). Scientific uncertainty in the press: How newspapers describe initial biomedical findings. *Science Communication*, 40(1), 124–141. <https://doi.org/10.1177/1075547017752166>
- Dunwoody, S. (1999). Scientists, journalists, and the meaning of uncertainty. In S. M. Friedman, S. Dunwoody, & C. L. Rogers (Eds.), *Communicating uncertainty: Media coverage of new and controversial science* (pp. 59–79). Routledge.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Erdt, M., Nagarajan, A., Sin, S.-C. J., & Theng, Y.-L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics*, 109(2), 1117–1166. <https://doi.org/10.1007/s11192-016-2077-0>
- Evensen, D. T., & Clarke, C. E. (2012). Efficacy information in media coverage of infectious disease risks: An ill predicament?. *Science Communication*, 34(3), 392–418. <https://doi.org/10.1177/1075547011421020>
- Fahnestock, J. (1986). Accommodating science. The rhetorical life of scientific facts. *Written Communication*, 3(3), 275–296. <https://doi.org/10.1177/0741088386003003001>
- Foreign Affairs New Zealand (2020, March 18). *MIL-OSI United Kingdom: Expert comment on what more do we now know about the virus causing the COVID-19 outbreak, and what can scientists do to best respond?*. <https://foreignaffairs.co.nz/2020/03/19/mil-osi-united-kingdom-expert-comment-on-what-more-do-we-now-know-about-the-virus-causing-the-covid-19-outbreak-and-what-can-scientists-do-to-best-respond/>
- Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálffy, M., & Alexis, J. (2020). Preprinting a pandemic: The role of preprints in the COVID-19 pandemic. *BioRxiv*, 36. <https://doi.org/10.1101/2020.05.22.111294>
- Frewer, L., Hunt, S., Brennan, M., Kuznesof, S., Ness, M., & Ritson, C. (2003). The views of scientific experts on how the public conceptualize uncertainty. *Journal of Risk Research*, 6(1), 75–85. <https://doi.org/10.1080/136987032000047815>
- Fry, N. K., Marshall, H., & Mellins-Cohen, T. (2019). In praise of preprints. *Microbial Genomics*, 5(4), 1–3. <https://doi.org/10.1099/mgen.0.000259>
- Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95(1), 1–37. <https://doi.org/10.1086/229213>
- Golan, G. (2006). Inter-media agenda setting and global news coverage. *Journalism Studies*, 7(2), 323–333. <https://doi.org/10.1080/146161670050533643>
- Guenther, L., Bischoff, J., Löwe, A., Marzinkowski, H., & Voigt, M. (2019). Scientific evidence and science journalism: Analysing the representation of (un)certainly in German print and online media. *Journalism Studies*, 20(1), 40–59. <https://doi.org/10.1080/1461670X.2017.1353432>
- Guenther, L., Froehlich, K., & Ruhrmann, G. (2015). Un)certainly in the news: Journalists' decisions on communicating the scientific evidence of nanotechnology. *Journalism & Mass Communication Quarterly*, 92(1), 199–220. <https://doi.org/10.1177/1077699014559500>
- Guenther, L., Gaertner, M., & Zeitz, J. (2020). Framing as a concept for health communication: A systematic review. *Health Communication*, 1–9. Advance online publication. <https://doi.org/10.1080/10410236.2020.1723048>

- Guenther, L., & Ruhrmann, G. (2016). Scientific evidence and mass media: Investigating the journalistic intention to represent scientific uncertainty. *Public Understanding of Science*, 25(8), 927–943. <https://doi.org/10.1177/0963662515625479>
- Gustafson, A., & Rice, R. E. (2019). The effects of uncertainty frames in three science communication topics. *Science Communication*, 41(6), 679–706. <https://doi.org/10.1177/1075547019870811>
- Haelle, T. (2020). Tips on covering preprints about coronavirus research. [Blog post]. *Association of Health Care Journalists*. [https://healthjournalism.org/resources-tips-details.php?id=1113#.Xu\\_FTy2ZPxU](https://healthjournalism.org/resources-tips-details.php?id=1113#.Xu_FTy2ZPxU)
- Harrison, S., Macmillan, A., & Rudd, C. (2020). Framing climate change and health: New Zealand's online news media. *Health Promotion International*, 1–11. <https://doi.org/10.1093/heapro/daz130>
- Hart, P. S., & Feldman, L. (2014). Threat without efficacy? Climate change on U.S. network news. *Science Communication*, 36(3), 325–351. <https://doi.org/10.1177/1075547013520239>
- Helmuth, L. (2020). Tipsheet: Covering the coronavirus epidemic effectively without spreading misinformation. [Blog post]. *The Open Notebook*. <https://www.theopennotebook.com/2020/03/02/tipsheet-covering-the-coronavirus-epidemic-effectively-without-spreading-misinformation/#>
- Hermida, A. (2019). The existential predicament when journalism moves beyond journalism. *Journalism*, 20(1), 177–180. <https://doi.org/10.1177/1464884918807367>
- Hove, T., Paek, H.-J., Yun, M., & Jwa, B. (2015). How newspapers represent environmental risk: The case of carcinogenic hazards in South Korea. *Journal of Risk Research*, 18(10), 1320–1336. <https://doi.org/10.1080/13669877.2014.923025>
- Hurley, R. J., & Tewksbury, D. (2012). News aggregation and content differences in online cancer news. *Journal of Broadcasting & Electronic Media*, 56(1), 132–149. <https://doi.org/10.1080/08838151.2011.648681>
- Jaklevic, M. (2020). Strong caveats are lacking as news stories trumpet preliminary COVID-19 research. [Blog post]. *Health News Review*. <https://www.healthnewsreview.org/2020/04/strong-caveats-are-lacking-as-news-stories-trumpet-preliminary-covid-19-research/>
- Jenkins, H., & Deuze, M. (2008). Editorial: Convergence culture. *Convergence: The International Journal of Research into New Media Technologies*, 14(1), 5–12. <https://doi.org/10.1177/1354856507084415>
- Jensen, J. D. (2008). Scientific uncertainty in news coverage of cancer research: Effects of hedging on scientists' and journalists' credibility. *Human Communication Research*, 34(3), 347–369. <https://doi.org/10.1111/j.1468-2958.2008.00324.x>
- Jung Oh, H., Hove, T., Paek, H.-J., Lee, B., Lee, H., & Kyu Song, S. (2012). Attention cycles and the H1N1 pandemic: A cross-national study of US and Korean newspaper coverage. *Asian Journal of Communication*, 22(2), 214–232. <https://doi.org/10.1080/01292986.2011.642395>
- Karlsson, M., & Sjøvaag, H. (2018). Hyperlinks and linking practice. In T. Vos, & F. Hanusch (Eds.), *The international encyclopedia of journalism studies* (pp. 1–5). American Cancer Society. <https://doi.org/10.1002/9781118841570.iej0231>
- Khamsi, R. (2020). What best practices are you following in covering preprints during the pandemic? [Blog post]. *Health Journalism*. [https://healthjournalism.org/core-topic.php?id=4&page=shared\\_wisdom](https://healthjournalism.org/core-topic.php?id=4&page=shared_wisdom)
- Kharasch, E. D. (2020). Uncertainty and certainty. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 133(1), 1–4. <https://doi.org/10.1097/ALN.0000000000003388>
- Kille, L. W. (2015). The growing problem of Internet “link rot” and best practices for media and online publishers. *Journalist's Resource*, Harvard Shorenstein Center on Media Politics and Public Policy. [Blog post]. <https://journalistsresource.org/studies/society/internet/website-linking-best-practices-media-online-publishers/>
- Kwon, D. (2020). How swamped preprint servers are blocking bad coronavirus research. *Nature*, 581(7807), 130–131. <https://doi.org/10.1038/d41586-020-01394-6>
- 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>
- Lai, W. Y. Y., & Lane, T. (2009). Characteristics of medical research news reported on front pages of newspapers. *PLoS ONE*, 4(7), e6103. <https://doi.org/10.1371/journal.pone.0006103>
- Lee, A. M., & Chyi, H. I. (2015). The rise of online news aggregators: Consumption and competition. *International Journal on Media Management*, 17(1), 3–24. <https://doi.org/10.1080/14241277.2014.997383>
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- Lowrey, W. (2012). Journalism innovation and the ecology of news production: Institutional tendencies. *Journalism & Communication Monographs*, 14(4), 214–287. <https://doi.org/10.1177/1522637912463207>
- Luzón, M. J. (2009). Scholarly hyperwriting: The function of hyperlinks in academic weblogs. *Journal of the American Society for Information Science and Technology*, 60(1), 75–89. <https://doi.org/10.1002/asi.20937>
- Majumder, M. S., & Mandl, K. D. (2020). Early in the epidemic: Impact of preprints on global discourse about COVID-19 transmissibility. *The Lancet Global Health*, 8(5), e627–e630. [https://doi.org/10.1016/S2214-109X\(20\)30113-3](https://doi.org/10.1016/S2214-109X(20)30113-3)
- Mandal, A. (2020, April 8). Air pollution and PM2.5 raise risk of COVID-19 death. *Medical News*. <https://www.news-medical.net/news/20200408/Air-pollution-and-PM25-raise-risk-of-COVID-19-death.aspx>
- Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, 58(2), 258–279. <https://doi.org/10.1111/j.1460-2466.2008.00384.x>
- Matthias, L., Fleerackers, A., & Alperin, J. P. (2020). Framing science: How opioid research is presented in online news media. *Frontiers in Communication*, 5(64). <https://doi.org/10.3389/fcomm.2020.00064>
- McKenzie, C. T., Lowrey, W., Hays, H., Chung, J. Y., & Woo, C. W. (2011). Listening to news audiences: The impact of community structure and economic factors. *Mass Communication and Society*, 14(3), 375–395. <https://doi.org/10.1080/15205436.2010.491934>
- Molteni, M. (2020, January 24). Scientists predict Wuhan's outbreak will get much worse. *Wired*. <https://www.wired.com/story/scientists-predict-wuhans-outbreak-will-get-much-worse/>
- Nanyingi, M. (2020, May 31). Predicting COVID-19: What applying a model in Kenya would look like. <https://theconversation.com/predicting-covid-19-what-applying-a-model-in-kenya-would-look-like-134675>
- Nisbet, M. C., Brossard, D., & Kroepsch, A. (2003). Framing science: The stem cell controversy in an age of press/politics. *Harvard International Journal of Press/Politics*, 8(2), 36–70. <https://doi.org/10.1177/1081180X02251047>
- O'Malley, P., Rainford, J., & Thompson, A. (2009). Transparency during public health emergencies: From rhetoric to reality. *Bulletin of the World Health Organization*, 87(8), 614–618. <https://doi.org/10.2471/BLT.08.056689>
- Ordway, D. (2020). Covering biomedical research preprints amid the coronavirus: 6 things to know [Blog post]. <https://journalistsresource.org/tip-sheets/research/medical-research-preprints-coronavirus/>
- Our Audience. (2020). *The conversation Australia*. <https://theconversation.com/au/audience>
- Peters, H. P., & Dunwoody, S. (2016). Scientific uncertainty in media content: Introduction to this special issue. *Public Understanding of Science*, 25(8), 893–908. <https://doi.org/10.1177/0963662516670765>
- Polka, J. K., & Penfold, N. C. (2020). *Biomedical preprints per month, by source and as a fraction of total literature* (Version 3.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.3819276>
- Price, V., & Tewksbury, D. (1997). News values and public opinion: A theoretical account of media priming and framing. In G. A. Barnett & F. J. Boster (Eds.), *Progress in the communication sciences* (Vol. 13, pp. 173–212). Ablex.
- Scheufele, B., & Scheufele, D. (2010). Of spreading activation, applicability, and schemas: Conceptual distinctions and their operational implications for measuring frames and framing effects. In P. D'Angelo &



- J. A. Kuypers (Eds.), *Doing news framing analysis: Empirical and theoretical perspectives* (1st ed., pp. 110–134). Routledge.
- Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, 49(1), 103–122. <https://doi.org/10.1111/j.1460-2466.1999.tb02784.x>
- Schneider, J. (2010). Making space for the “nuances of truth”: Communication and uncertainty at an environmental journalists’ workshop. *Science Communication*, 32(2), 171–201. <https://doi.org/10.1177/1075547009340344>
- Schneider, S. (2016). Communicating uncertainty: A challenge for science communication. In J. L. Drake, Y. Y. Kontar, J. C. Eichelberger, T. S. Rupp, & K. M. Taylor (Eds.), *Communicating climate-change and natural hazard risk and cultivating resilience: Case studies for a multi-disciplinary approach* (pp. 267–278). Springer International Publishing. [https://doi.org/10.1007/978-3-319-20161-0\\_17](https://doi.org/10.1007/978-3-319-20161-0_17)
- Sellnow, T. L., Ulmer, R. R., Seeger, M. W., & Littlefield, R. S. (2009). *Effective risk communication—A message-centered approach*. Springer Science & Business Media.
- Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication*, 50(2), 93–109. <https://doi.org/10.1111/j.1460-2466.2000.tb02843.x>
- Stocking, G. (2019). *Trends and facts on online news: State of the news media*. Pew Research Center. <https://www.journalism.org/fact-sheet/digital-news/>
- Strekalova, Y. A. (2015). Informing dissemination research: A content analysis of U.S. newspaper coverage of medical nanotechnology news. *Science Communication*, 37(2), 151–172. <https://doi.org/10.1177/1075547014555025>
- Stroobant, J. (2019). Finding the news and mapping the links: A case study of hypertextuality in Dutch-language health news websites. *Information, Communication & Society*, 22(14), 2138–2155. <https://doi.org/10.1080/1369118X.2018.1477971>
- Stroobant, J., & Raeymaeckers, K. (2019). Hypertextuality in net-native health news: A quantitative content analysis of hyperlinks and where they lead to. *Journal of Applied Journalism & Media Studies*, 8(3), 367–385. [https://doi.org/10.1386/ajms\\_00007\\_1](https://doi.org/10.1386/ajms_00007_1)
- Tandoc, E. C. (2015). Why web analytics click. *Journalism Studies*, 16(6), 782–799. <https://doi.org/10.1080/1461670X.2014.946309>
- Terms and Conditions. (2020). *The conversation*. <https://theconversation.com/ca/terms-and-conditions>
- Tewksbury, D., & Riles, J. M. (2018). Framing in an interactive news environment. In P. D’Angelo (Ed.), *Doing news framing analysis II: Empirical and theoretical perspectives* (pp. 138–162). Routledge.
- van der Bles, A. M., van der Linden, S., Freeman, A. L. J., & Spiegelhalter, D. J. (2020). The effects of communicating uncertainty on public trust in facts and numbers. *Proceedings of the National Academy of Sciences*, 117(14), 7672–7683. <https://doi.org/10.1073/pnas.1913678117>
- Van Leuven, S., Kruikemeier, S., Lecheler, S., & Hermans, L. (2018). Online and newsworthy: Have online sources changed journalism?. *Digital Journalism*, 6(7), 798–806. <https://doi.org/10.1080/21670811.2018.1498747>
- Vu, H. T. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism*, 15(8), 1094–1110. <https://doi.org/10.1177/1464884913504259>
- Wang, W., & Guo, L. (2018). Framing genetically modified mosquitoes in the online news and Twitter: Intermedia frame setting in the issue-attention cycle. *Public Understanding of Science*, 27(8), 937–951. <https://doi.org/10.1177/0963662518799564>
- Witschge, T., Anderson, C., Domingo, D., & Hermida, A. (2019). Dealing with the mess (we made): Unraveling hybridity, normativity, and complexity in journalism studies. *Journalism*, 20(5), 651–659. <https://doi.org/10.1177/1464884918760669>
- Yan, W. (2020, April 14). Coronavirus tests science’s need for speed limits. *New York Times*. <https://www.nytimes.com/2020/04/14/science/coronavirus-disinformation.html?searchResultPosition=1>
- Yaros, R. A. (2011). Effects of text and hypertext structures on user interest and understanding of science and technology. *Science Communication*, 33(3), 275–308. <https://doi.org/10.1177/1075547010386803>
- Zehr, S. (1999). Scientists’ representations of uncertainty. In S. M. Friedman, S. Dunwoody, & C. L. Rogers (Eds.), *Communicating uncertainty: Media coverage of new and controversial science* (pp. 59–79). Routledge.