

Arno Simons and Alexander Schniedermann

3. Preprints in the German news media before and during the COVID-19 pandemic. A comparative mixed-method analysis

Abstract: Mainstream media widely references scientific publications for claims of factuality and authority. But how did science journalism deal with the sudden surge in preprint publications that provided rapid but often uncertain knowledge during the COVID-19 pandemic? While several studies have investigated various aspects of preprint-based science communication, only a few have focused on the public discourse in Germany, albeit with substantial challenges and controversies. In this mixed-method study, we identified the usage of preprints for 1,006 in about 390,000 German news stories, qualitatively analyzed the contexts of these preprints, and developed codes that reflect the epistemic sentiments. We further compared the code compositions of news stories that cover the pandemic with those about other topics. We found that the amount of news stories which used preprints increased with the pandemic. Frequent framings of preprints include accessibility, timeliness, and uncertainty, where the latter was more prominent in corona-related than corona-unrelated news stories. Beside using preprints as sources for claims, some news stories referred to them as a publication genre to turn scientific publishing into a story itself. Based on our findings we argue that journalists have to be transparent about their usage of preprints as well as reflect the benefits and drawbacks of using them.

Keywords: health communication, preprints, German press, scientific uncertainty, mixed-methods

1 Introduction and research question

Over the last decade, journalism in general and science journalism in particular have been changing in response to digitalization as well as to new demands for immediacy and transparency (Dunwoody, 2021; Allan, 2009). While journalism has always been profoundly about “getting it right,” truth-telling strategies and standards are currently being re-negotiated (Craig, 2016; Le Masurier, 2015; Karlsson, 2011). The outbreak of the COVID-19 pandemic has only propelled this development, and it has put the truth-telling role of journalism to a new test

(Dunwoody, 2020). On the one hand, the pandemic has revealed the difficulties involved in formulating and legitimating wide-ranging policy responses to a problem whose nature is inherently scientific while our scientific knowledge of it is highly uncertain (Bicchieri et al., 2021; Parviainen et al., 2021; Kreps & Kriner, 2020). On the other hand, the pandemic has spurred an unprecedented growth and acceleration of scientific research and publishing in the biomedical field and beyond (Wang & Tian, 2021; Horbach, 2020; Torres-Salinas, 2020), resulting in a number of highly appreciated discoveries and inventions such as in the content of new mRNA-vaccines (Dolgin, 2021).

In this context, an interesting phenomenon related to both the general transformation of journalism and the specific development of pandemic science and its media coverage is the dissemination and uptake of academic preprints. Preprints can be defined in several ways depending on their stage in the conventional publication process (Till, 2001). We use the term to denote academic papers that are published online (“e-preprints”), on dedicated preprint servers or other openly accessible outlets, at a time when they have not (yet) been peer-reviewed in a process typically organized by academic journals.

For journalists, preprints are both tempting and risky. On the “pro-side,” preprints provide fast and free access to the latest scientific findings, which became especially relevant during the pandemic. Of the 125,000+ COVID-19-related scientific papers released within 10 months of the first confirmed case more than 30,000 were hosted by open-access preprint servers before they were published by peer-reviewed journals (Fraser et al., 2021; cf. Colavizza et al., 2021). The dissemination of corona-related preprints not only accelerated the scholarly discussion of scientific results but it also allowed for faster journalism and a more timely information of pandemic policymaking (Fleerackers et al., 2021; Fraser et al., 2021; Horbach, 2020). On the “con-side,” preprints lack academic peer review, and so their findings have to be taken with extra care, especially when interpreted by non-scientists, including journalists (Fleerackers et al., 2021; Chiarelli et al., 2019). In this light, preprints may “change the rules of the expertise game” (Heimstädt, 2020) with an effect also on the way science is reported in the media.

How do journalists deal with preprints, especially when they become more important within science as well as for informing policy? In particular, how do journalists frame preprints with regard to their advantages and disadvantages as discussed by scientists, policymakers, and journalists? Here we present results from a systematic qualitative and quantitative case study on the coverage and framing of preprints in 1,006 German news stories identified in 390,942 stories issued by seven major German news outlets between 2018 and 2021. We

compare the use and reporting of preprints before and after the beginning of the pandemic and in news context both related and unrelated to COVID-19.

2 Preprints in science and in the media

While the surge in and public visibility of preprints during corona is unprecedented, preprints have been around in science for a very long time (Moore, 1965). The first large-scale exploration of preprinting occurred in the 1960s during a six-year experiment with so-called Information Exchange Groups (IEGs), the members of which exchanged hard-copies of preprinted manuscripts via mail, financed by the US National Institutes of Health (NIH). The organizers and participants of the IEGs, as well as of similar initiatives in the physical sciences around the same time, were generally happy with these services, but strong opposition came from academic publishers and leading journals, such as *Science* and *Nature*, who feared that their business model was in danger (Cobb, 2017; Till, 2001). Only much later, with the development of new editing software, the advent of the World Wide Web (WWW), and the creation of the preprint server arXiv in 1991 supported by the National Science Foundation (NSF), did preprinting really take off as an alternative publishing practice in physics and increasingly in other disciplines.

Today, preprinting is not only “a device for quick scientist-to-scientist communication” (Moore, 1965, p. 127), but more generally a tool for accelerated science communication to both internal and external audiences, including journalists.¹ The two key advantages of preprinting are easy and fast dissemination of research findings and better accessibility (Chiarelli et al., 2019). This combination of timeliness and accessibility makes preprints attractive also to journalists looking for new and accessible information and sources.

The biggest caveat with preprinting is the lack of formal peer review. The peer-review process organized by academic journals is meant to ensure quality standards, improve performance, and provide credibility of a manuscript. Peer review thus functions as a filter for quality and relevance, especially for external audiences like journalists and policymakers, who are not generally capable of assessing the value of a scientific manuscript. Since preprints have not gone

¹ The timeliness aspect of preprinting also plays a role in the establishment of priority claims. In several cases, scientists have used preprints as a tool to outpace competitors during otherwise lengthy publication processes, sometimes linked to the sharing of preprints with the media as a social amplifier (Weingart, 1998; Lewenstein, 1995; Nelkin, 1995).

through this filter, they are typically associated with greater scientific uncertainties than peer-reviewed publications in academic journals (Chiarelli et al., 2019).

Science journalists are increasingly aware of these issues. Since 2019, the stylebook of the Associated Press (2019), a leading reference for journalists around the world, emphasizes that any “research that has not been peer-reviewed, including articles posted on preprint servers, should be reported with extreme care,” and links this to the remark that science reporting “comes [with] unique responsibility. A misleading or incorrect story could lead someone to make unwise, harmful choices.”²

To sum up, preprints are both tempting and risky for journalists: tempting because of their timeliness and accessibility, risky because of the scientific uncertainties due to the lack of peer review. According to Heimstädt (2020),

it is up to journalists and policymakers to familiarise themselves with the most important preprint servers and their specific moderation techniques (e.g. sanity checks of uploaded preprints by a small editorial team). Only when understanding the governance of such new and more open scientific practices will they be able to leverage the benefits of fast science while avoiding the threat of disinformation.

3 Sources, factuality, and framing

The use of preprints as sources in news stories is (a new) part of a set of journalistic practices called “truth-telling,” aimed at generating a sense of “factuality” (Pan & Kosicki, 1993; van Dijk, 1988; Tuchman, 1980). According to Pan and Kosicki (1993, p. 61), news stories are often characterized by a “hypothesis-testing (or research finding) aspect,” which is especially true when they report about research-related topics such as COVID-19. But often there is “no clear distinction between factuality and persuasion. The rhetorical claim of news being factual and impartial helps establish the epistemological status of news as a source of factual information and the authority of news as a mirror of reality” (p. 62).

² That preprints “should not be reported in news media as established information” has also been demanded by leading preprint servers during the pandemic. Such warnings have been posted, for example, on medRxiv’s website since their launch in June 2019 (<https://web.archive.org/web/20190630063933/https://www.medrxiv.org/>) and on bioRxiv’s epidemiology page since February 2020 (<https://web.archive.org/web/20200202105940/https://www.biorxiv.org/collection/epidemiology>).

As van Dijk (1988, p. 84) highlights in his book *News as Discourse*, “If propositions are to be accepted as true or plausible, there must be special means to enhance their appearance of truth and plausibility. News discourse has a number of standard strategies to promote the persuasive process for assertions.” Such strategies of persuasion—which resemble practices in science publication (Latour, 1987; Fleck, 1980)—include the use of authoritative sources, direct quotations, and discursive markers indicating precision and exactness, such as numbers or particular adjectives (Pan & Kosicki, 1993; van Dijk, 1988; Tuchman, 1980).

The construction of “factuality” in media reporting (as well as in science) can be analyzed in terms of framing (Pan & Kosicki, 1993; Entman, 1993; van Dijk, 1988; Latour, 1987). Pan and Kosicki (1993, p. 59) define a news story’s “frame” as “an idea that connects different semantic elements of a story (e.g., descriptions of an action or an actor, quotes of sources, and background information) into a coherent whole.” Whereas the frame intended by the journalist never perfectly matches the one comprehended by the reader (van Dijk, 1988), on the writing side, the choice of conventionally understandable structural and lexical “framing devices” opens “a cognitive ‘window’ through which a news story is ‘seen’” (Pan & Kosicki, 1993, p. 59).

While truth-telling likely “is the most essential component of journalism... the means by which this can be accomplished... change radically in the digital environment” (Karlsson, 2011, p. 279). One such change, according to Karlsson, is the trend toward immediacy in an ever accelerating news cycle. Speed and timeliness generally provide a comparative advantage in journalism (Craig, 2016). But with the increased acceleration of online journalism, timeliness comes with a caveat. The trend toward immediacy increases the likelihood of provisory, incomplete or dubious news reporting (Craig, 2016; Le Masurier, 2015; Karlsson, 2011). This speed-accuracy tradeoff is one of the factors that has led to a restructuring of journalistic authority. To maintain journalism as an authoritative source of information, many authors have called for transparency as a new publicly communicated standard for establishing accountability and legitimacy (Karlsson, 2011; Allen, 2008).

These developments have been reflected in recent research on media framing of science in general and of scientific publications in particular. Studies find, for example, that news outlets increasingly point their readers to academic sources via hyperlinks to demonstrate credibility and transparency (Stroobant & Raeymaeckers, 2019; Karlsson & Sjøvaag, 2018). Research on the framing of scientific uncertainties reveals a mixed picture. In a study of 149 news stories in nine major US and Canadian online news outlets, Matthias et al. (2020, p. 1) find

that academic sources are mostly framed as “certain,” only sometimes as “controversial” and least often “uncertain.” Dumas-Mallet et al. (2018) find that most of 426 news stories covering 40 initial biomedical studies frame these studies as “initial” but only 21% mention that they should be “confirmed by replication.” A systematic, quantitative content analysis by Guenther et al. (2019) of 128 science stories published in seven major German media outlets reveals that while scientific findings are predominantly depicted as “certain,” uncertainty framings are more common within specifically dedicated science sections or in stories with natural scientific or medical scientific content. We found only one study on the media framing of preprints.

Fleerackers et al. (2021) studied the mentions of 100 preprints in 457 news stories and found that about half of the references to preprints in their sample contained one or more of four “uncertainty framing devices”: formulations that the cited study was 1) a “preprint”; 2) “unreviewed”; 3) “preliminary”; and/or 4) “in need of verification.”

Our analysis complements current research by combining traditional science media studies and more recent approaches that utilize large-scale quantitative data. In the latter case, researchers select scientific publications and search for their media coverage, often by utilizing large-scale databases such as Altmetric (Fleerackers et al., 2021; Matthias et al., 2020). Although this data-driven approach provides access to a huge number of publications and media items, it is mostly based on standardized links via URLs/DOIs or via bibliographic information such as author names, publication year or journal title which can be matched with other bibliographic sources.³ In addition, it has been found that such databases are often biased towards recency, English language, and particular disciplines. In contrast, more traditional science media studies start their analysis by sampling news outlets (Guenther et al., 2019; Riles et al., 2015) or journalistic databases (Dumas-Mallet et al., 2018). Such approaches can identify situations in which scientific publications cover other formats, for example expert interviews. But in contrast to larger datasets, they are usually limited in their timeframe or selection of news outlets, which makes them less suited to cover longer and broader events like the pandemic in a longitudinal manner.

³ For some of their sources (news, policy, and patents), Altmetric additionally uses text mining: <https://help.altmetric.com/support/solutions/articles/6000240263-text-mining>

4 Data and coding

Our data pipeline consisted of four steps. First, we selected the time span and the sources. We picked the 48 months long period between January 2018 and December 2021 inclusive, to be able to compare the coverage of preprints before and after the outbreak of COVID-19 in December 2019/ January 2020. The selection of sources was based on three criteria: their type, their circulation, and their accessibility. We wanted to include both major daily and weekly newspapers as well as major online news providers. Accessibility was an issue because we had no institutional access to any of the targeted sources, most of which contain at least some paid content. In several cases, we could leverage trial subscriptions to get past a paywall, but in other cases there were simply no affordable pricing options.

Table 1: Retrieval strategy and coverage per source (sorted by number of retrieved stories).

Source	Type	Retrieval	Coverage/ Confidence	Stories
Spiegel Online (SPON)	daily/ online	per day	complete/ high	184,636
Süddeutsche Zeitung/ SZ.de (SUED)	mixed	keyword search	subpops 1&2/ high	72,908
Bild/Bild.de (BILD)	mixed	keyword search	subpops 1&2/ medium	41,187
FAZ.net (FAZNET)	daily/ online	category search	subpops 1&2/ high	36,708
Die Zeit (ZEIT)	weekly/ print	per issue	complete/ high	21,224
Zeit Online (ZON)	daily/ online	keyword search	subpops 1&2/ high	20,341
Der Spiegel (SPIEGEL)	weekly/ print	per issue	complete/ high	13,938
			Total	390,942

Second, we identified and downloaded retrievable news stories of interest. By retrievable we mean stories for which we could get hold of their metadata, including most importantly their URL but also the title and the date of publication. Generally, we aimed at retrieving the total population of stories published by each source in the study period. However, this was only possible for three sources: SPIEGEL, SPON, ZEIT. For all other sources, in which the total population was unknown to us, we developed strategies for searching and retrieving all stories mentioning preprints and other types of scientific publications (subpop-1) and, given that preprints have increased dramatically during the pandemic, all stories related to COVID-19 (subpop-2). Depending on a source's spe-

cific search capability and limitations, we defined multiple and sometimes complex queries to retrieve these two subpopulations. For all but one source (BILD), we are highly confident to have retrieved these subpopulations. Since BILD limits its search results to 1,000 articles, we used a large number of connected keywords including “preprint AND corona,” “preprints AND physics,” “corona AND masks,” “corona AND vaccines,” etc. to narrow down the search space, hoping to retrieve as many news stories as possible. In the end we are still only medium confident to have achieved our goal of retrieving the two subpopulations.

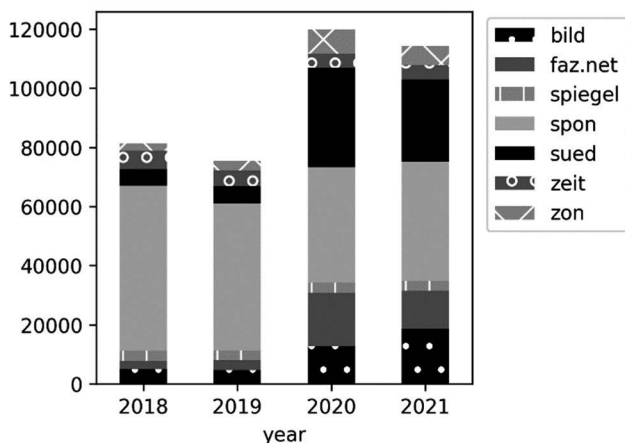


Figure 1: Number of news stories per source and year ($N = 390,942$, all sources).

We downloaded all retrieved stories and compiled a dataset containing the metadata, the html data, and the extracted plain text data. All this was done with scripts that we coded in Python, using standard libraries as well as external libraries like *Requests* and *BeautifulSoup* for html parsing, *Pandas* for data management and analysis, and *Matplotlib* for visualization. Our final dataset contains 390,942 news stories from seven major German news outlets. Table 1 lists the overall retrieval strategies and coverage per source. Figure 1 plots the number of retrieved stories per source and year.

Third, we labeled all stories according to two categories of interest. First, we call stories “p-stories” if they relate to preprints in one of the following ways: either a) their plain text contains any of the terms: “preprint,” the German syn-

onyms “Vorveröffentlichung,” “Vorab-Veröffentlichung” or the name of one of over 50 preprint servers⁴; or b) their html includes at least one hyperlink to a URL that contains the substring “preprint” or, again, the name of one of over 50 preprint servers. Second, we made a distinction between “corona stories” and “non-corona stories.” All stories whose plain text contains any of the terms “corona,” “COVID-19,” “2019-nCov” or “cov-2” went into the former category, or all stories went into the latter.

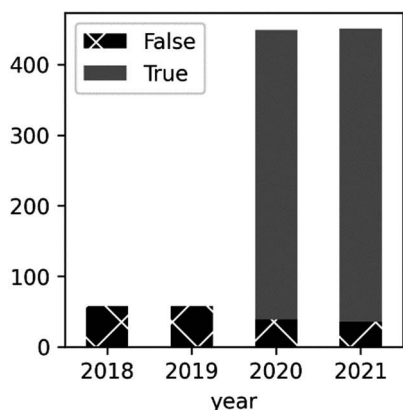


Figure 2: Number of p-stories (N = 1,006, all sources).

Figure 2 shows the distribution of 1,006 identified p-stories over time. The four bars represent absolute numbers of p-stories per given year of our four-year study period. The colors indicate how many of the p-stories were related (orange) or unrelated (blue) to corona reporting. As expected, we find that the number of p-stories rose steeply with the start of the pandemic. The increase from 2018 (58) and 2019 (58) to 2020 (449) and 2021 (451) amounts to a multiplication close to factor 8. As indicated by the coloring, the absolute majority of p-stories in 2020 (91%) and 2021 (92%) were focused on corona.

⁴ To be more precise, all of our dictionaries contained regular expressions. For example, to cover the names of over 50 preprint servers, we used shortcut expressions such as `r"[rin]xiv(?:!m)"` to include servers that contain “xiv” while excluding false positive matches such as “NXIVM,” an American cult, which engaged in sex trafficking, forced labor, and racketeering, and which was covered in the German press during our study period.

Finally, we qualitatively analyzed all 1,006 p-stories on how they framed preprints. While our analysis was guided by the methodology and results of previous studies, especially Fleerackers et al. (2021), we started with a few rounds of open coding, allowing us to “get a feel” for the ways in which preprints are framed in our sample. We then gradually moved from open codes to more defined and settled codes by iteratively comparing, refining, splitting, and merging our codes until we reached a good level of intercoder agreement and overall saturation. Eventually, we defined 12 codes of interest relating to the following five concepts: 1) framing, 2) naming, 3) referencing, 4) genre, and 5) focussing. In the following sections we present these concepts and associated codes both qualitatively and quantitatively as findings, since that is what they really are, the results of an iterative, theory generating coding process (cf. Glaser & Strauss, 1967).

5 Framing and naming

Given the controversiality of preprints among researchers and journalists, we analyzed to what extent the discussed pros and cons of preprints can also be found as framing devices in our sample set of p-stories. On the con-side we found uncertainty framing devices quite similar to those reported by Fleerackers et al. (2021). Several p-stories in our sample frame preprints as “unreviewed,” “preliminary,” “premature,” or “in need of verification,” sometimes linked to warnings that preprints “may lead to the circulation of unscientific claims” or “fuel bad policy decisions.” We coded such stories as uncertainty. On the pro-side we found p-stories associating preprints with notions of timeliness or access, which we also used as codes.⁵ We especially noticed emphases on the recency of preprint studies, expressions like “a preprint study published on Wednesday” (2021, SUEd) or “the results have just been published on a preprint server” (2021, SPON). In one case, a brand new preprint was added to a story via an update to demonstrate a concern for immediacy: “Update: In the story it says the results of the recovery-study... are not yet available. This has changed since the publication of our story. In the meantime, there is a preprint: MedRxiv: Horby et al., 2020” (2020, ZON).

⁵ To be able to analyze “pro-framing” in comparison to “con-framing,” we will sometimes use a combined code `timeliness_or_access` for the former. The latter then is equal to the uncertainty code.

When highlighting access, p-stories contain formulations like “The complete paper is available on the preprint server PsyarXiv” (2021, ZON). The following example shows a strong double framing of both timeliness and access: “In their publication, which has been freely accessible to everyone on the preprint server ‘arXiv’ since Thursday, the researchers provide the standard deviation of 4.8 sigma” (2018, FAZ).

Table 2a: Corona p-stories. Annual distribution of codes.

	2018	2019	2020	2021	Total
p-Stories			410	405	815
references			484	538	1022
Code					
name_de			10%	6%	8%
name_en			42%	46%	44%
any_name			47%	49%	48%
no_name			53%	51%	52%
timeliness			26%	16%	21%
access			5%	4%	4%
timeliness_or_access			27%	18%	22%
uncertainty			37%	40%	38%
any_framing			49%	48%	48%
min_1_ref			92%	98%	95%
p_in_focus			28%	18%	23%
genre			22%	13%	17%

In terms of numbers, we see striking differences between corona p-stories and non-corona p-stories. Tables 2a and 2b, which show the distribution of codes over time for these two categories, as well as Tables 3a and 3b, which show the cross-tabulation of codes, reveal that uncertainty framings occur more often in corona p-stories than in non-corona p-stories. Comparing only the total columns in 2a and 2b, we see that the percentages of timeliness (19% in 2a vs. 20% in 2b) and access (4% in 2a vs. 5% in 2b) do not change much in relation to corona-coverage, while uncertainty framings are more than twice as frequent in corona articles (38% in 2a vs. 14% in 2b). When paying attention to time, we find that uncertainty framings have doubled in non-corona p-stories from 2020 (10%) to

2021 (22%), which we interpret as a result of a more intense discussion of the scientific uncertainties surrounding preprints related to corona preprints.

Tables 3a and 3b reveal that of all non-corona p-stories that use either of the three framings (uncertainty, timeliness, or access) (`any_framing`) 76% include pro-framings (`timeliness_or_access`) and only 44% include uncertainty framings, whereas this relationship is reversed in corona p-stories (44% `timeliness_or_access` vs. 81% uncertainty).

Table 2b: Non-corona p-stories. Annual distribution of codes.

	2018	2019	2020	2021	Total
p-Stories	58	58	39	36	191
references	67	69	43	43	222
Code					
<code>name_de</code>	5%	3%	8%	0%	4%
<code>name_en</code>	21%	22%	28%	28%	24%
<code>any_name</code>	22%	26%	33%	28%	27%
<code>no_name</code>	78%	74%	67%	72%	73%
<code>timeliness</code>	29%	14%	21%	19%	21%
<code>access</code>	9%	3%	0%	8%	5%
<code>timeliness_or_access</code>	31%	17%	21%	25%	24%
<code>uncertainty</code>	12%	12%	10%	22%	14%
<code>any_framing</code>	34%	24%	28%	39%	31%
<code>min_1_ref</code>	90%	88%	100%	97%	93%
<code>p_in_focus</code>	59%	72%	64%	50%	62%
<code>genre</code>	10%	12%	5%	6%	9%

When analyzing how preprints were framed in terms of uncertainty, timeliness, and access, we felt a need to differentiate between framing and naming. While we interpret the word “preprint” or its direct German translations as names that can be used for any particular preprint or for the preprint genre in general, we did not automatically interpret such names as framing devices. In particular, we did not follow the suggestion by Fleerackers and colleagues to interpret the presence of the word “preprint” as an uncertainty device: unlike “unreviewed,” “preliminary,” etc., the term “preprint” does not directly point to uncertainty or any aspect of it. Thus a link between the term “preprint” and uncertainty can only be evoked in the minds of readers already associating preprints with uncer-

tainty. However, by the same token readers may also associate preprints with timeliness or access or any other attribute. In other words, associating “preprint” with uncertainty requires not only a knowledge of what preprints are, but also of the pros-and-cons discussions related to the preprint genre. What is more, if any of the three framings under investigation could be said to be included in the name “preprint,” it arguably is timeliness, because the prefix “pre” means something like “before” or “prior to.”

The German language context adds even more complexity to the matter. When explicitly naming preprints, German journalists mostly use the English expression “Preprint,” which we coded as `name_en`. But sometimes they use a direct German translation. In such cases, which we coded `name_de`, “pre” is translated as either “vor” or “vorab” while “print” is typically translated as “Veröffentlichung,” “Publikation,” “Studie” or “Druck.” Like “Preprint,” none of these translated names invoke aspects of uncertainty directly, which is why we did not treat them as uncertainty framing devices. However, in contrast to the rather general prefix “vor,” the prefix “vorab” has a more distinct meaning, which can be translated as “in advance.” Ironically, the key example sentence for the use of “vorab” provided by the *Duden*—Germany’s authoritative dictionary of the Standard High German language—reads “Die Presse wurde vorab informiert” (“The press has been informed in advance”). We therefore decided to interpret the use of “vorab” (but not “vor”) in German names for preprint as a timeliness framing device and coded such instances accordingly. We hesitated to also code “Vorveröffentlichung,” “vorveröffentlicht” or “Preprint” as timeliness, because, as stated above, we believe that the meaning of the prefix “pre/vor” is less straightforward.

We assigned a combined code `any_name` to all p-stories that use either an English or a German name or both. P-stories that don’t include any proper name for preprints were coded `no_name`. Note that such stories can still include descriptions and framings of preprints or of some of their aspects, e.g. in formulations like “the un-reviewed publication” or “a study, which has been released in advance via an online platform.”

Naming practices have evolved over time. Naming is twice as common in corona p-stories compared to non-corona p-stories (Tables 2a vs. 2b: `name_en`: 44% vs. 24%, `name_de`: 8% vs. 4%, `any_name`: 48% vs. 27%). The proportion of English names to German names fluctuates a bit over the years and across categories, but it has always been in favor of the English “Preprint,” and it seems to have increased especially since 2021, possibly indicating that “Preprint” is about to become the dominant name also in the German language context.

Table 3a: Cross-tabulation of codes assigned to corona p-stories. Relative portions of code co-occurrences against absolute codes (diagonal). Read horizontal from row to column.

	name_de	name_en	any_name	no_name	timeliness	access	timeliness_or_access	uncertainty	any_framing	min_1_ref	p_in_focus	genre
name_de	66	50%	⊆	↖	47%	12%	52%	56%	76%	86%	39%	50%
name_en	9%	358	⊆	↖	27%	8%	29%	54%	64%	89%	31%	35%
any_name	17%	92%	391	↖	27%	7%	30%	53%	64%	90%	31%	35%
no_name	↖	↖	↖	424	15%	1%	16%	25%	34%	100%	16%	1%
timeliness	18%	57%	63%	37%	169	13%	⊆	55%	⊆	88%	33%	34%
access	23%	77%	83%	17%	63%	35	⊆	63%	⊆	77%	34%	54%
timeliness_or_access	19%	58%	64%	36%	93%	19%	182	55%	⊆	88%	32%	34%
uncertainty	12%	62%	66%	34%	30%	7%	32%	313	⊆	91%	29%	36%
any_framing	13%	58%	63%	37%	43%	9%	46%	79%	395	92%	29%	31%
min_1_ref	7%	41%	45%	55%	19%	3%	21%	37%	47%	773	23%	13%
p_in_focus	14%	59%	64%	36%	29%	6%	31%	48%	61%	93%	188	32%
genre	23%	90%	97%	3%	40%	13%	44%	80%	87%	72%	43%	141

Our qualitative analysis of naming practices similarly points to the negotiation of how to best name and refer to preprints. For example, several p-stories contain expressions like “a so-called preprint,” as if to educate the readership that preprints exist in the first place and that it can be legitimate to use them as sources in news stories. One context explains preprint naming by referring to a famous German corona-news podcast: “They are also known by the names ‘working paper’ or ‘preprint’... this formulation has been popularized by the Drostén-Podcast” (2021, SÜED). We interpret such educational interventions also as indications for a broader struggle within journalism of how to deal with preprints as sources, given their timeliness and accessibility but also their uncertainty.

Table 3b: Cross-tabulation of codes assigned to non-corona p-stories. Relative portions of code co-occurrences against absolute codes (diagonal). Read horizontally from row to column.

	name_de	name_en	any_name	no_name	timeliness	access	timeliness_or_access	uncertainty	any_framing	min_1_ref	p_in_focus	genre
name_de	8	38%	⊆	↯	75%	0%	75%	50%	88%	75%	75%	63%
name_en	7%	46	⊆	↯	26%	9%	30%	22%	43%	80%	67%	24%
any_name	16%	90%	51	↯	31%	8%	35%	25%	49%	80%	69%	27%
no_name	↯	↯	↯	140	17%	4%	19%	9%	24%	97%	60%	2%
timeliness	15%	30%	40%	60%	40	13%	⊆	20%	⊆	90%	68%	18%
access	0%	40%	40%	60%	50%	10	⊆	60%	⊆	60%	90%	40%
timeliness_or_access	13%	31%	40%	60%	89%	22%	45	27%	⊆	84%	69%	22%
uncertainty	15%	38%	50%	50%	31%	23%	46%	26	⊆	73%	65%	35%
any_framing	12%	34%	42%	58%	68%	17%	76%	44%	59	85%	66%	22%
min_1_ref	3%	21%	23%	77%	20%	3%	21%	11%	28%	177	60%	3%
p_in_focus	5%	26%	29%	71%	23%	8%	26%	14%	33%	90%	119	13%
genre	29%	65%	82%	18%	41%	24%	59%	53%	76%	35%	94%	17

5.1 Preprints as traceable sources

In the absolute majority of p-stories, preprints were used as sources to state and/or support claims or arguments. For each p-story we coded how many unique preprints it refers to, resulting in a total of 1,244 unique references distributed across p-stories as shown in Figure 3.⁶ The number of references per p-story is not correlated with corona reporting and remains similar over time. Overall, close to 80% of p-stories contain exactly one reference, between 10% and 20% of p-stories contain two references, and very few p-stories contain

⁶ Unique reference means that if a preprint was mentioned more than once in a story it was counted only once. We did not match references across p-stories, and thus cannot quantify the number of unique preprints cited in our sample. In Figure 3 we make use of lines merely to improve the readability of the plot, not to suggest that our numbers were continuous between the years.

three, four or more references. Most preprint references are traceable to their original source. We coded the traceability of preprints by checking if one of the following four minimal combinations of metadata was present in the text: 1) a DOI; 2) a hyperlink to the source; 3) a standardized scientific reference; or 4) authors' name(s) plus publication date plus either a complete title or at least a fair description of the content of the study. P-stories that contain at least one reference to a concrete preprint (min_1_ref) almost always enable the tracing of at least one of these preprints (min_1_trace), regardless of whether these p-stories relate to corona or not (97% in Table 3a vs. 98% in Table 3b).

References to preprints can take different forms. First of all, such references can be explicit: “According to a preprint study from India, the transmissibility is 50% more than in the British variant” (2021, BILD). In rare cases, found especially in ZON, references to preprints even take the form of standardized citation styles, e.g., “Large studies from Israel (The Lancet: Haas et al., 2021), England and Scotland (Lancet Preprint: Vasileio et al., 2021) show that the vaccines prevent almost all symptomatic cases and even most infections” (2021, ZON).

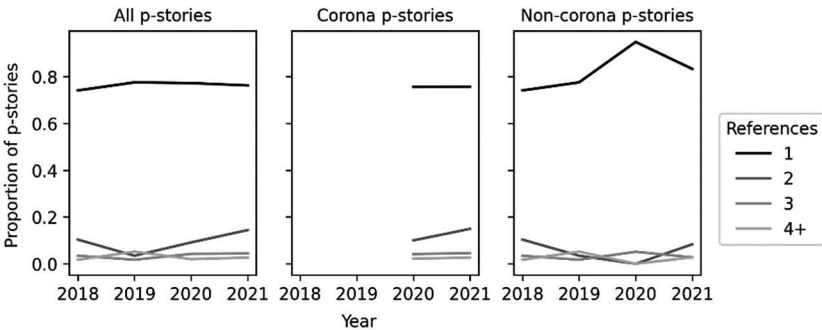


Figure 3: Number of references to preprints per p-story over time and across categories.

But in the majority of cases preprints are more implicitly referenced, without being named. Less than half of the corona p-stories that reference at least one preprint and less than a quarter of the corona p-stories that reference at least one preprint also name preprints. This can be read off the cross-tabulation of min_1_ref and min_1_ref in Tables 3a (45%) and 3b (23%).

In cases where preprints are cited but not named, preprints are often simply called “studies,” as in “Around 15 percent of all Twitter users are bots, or nearly 50 million! This is the result of a study” (2018, BILD). In other cases preprints are added via hyperlinks but the anchor text does not mention that the link goes

to a preprint, as in “the universe expands accelerated” (2019, FAZNET). Then there are cases where preprints are not named but some of their typical aspects are described. For example, the context “A not independently reviewed analysis from... Guangzhou estimates a value around 19 percent” (2020, SPON), which contains a hyperlink to a preprint, does not call its source a preprint but highlights that the source has not undergone peer review, invoking an uncertainty framing. Sometimes, p-stories also hyperlink to preprint versions of already published studies. We interpret such cases as indirect indications of the importance of public accessibility of sources for journalists.

6 Preprints as a genre in science communication

Preprints are not only used as (traceable) sources, they are also discussed as a distinct category, or genre, of the academic literature. In a broad sense, the preprint genre becomes invoked when a referenced study is called a preprint (any_name). But our code for genre (genre) is more restricted in that it applies only when the genre is discussed more directly. A minimum case was the formulation “a so-called preprint,” where the term “so-called” emphasizes that the cited preprint is but a “token” of a more general “type,” i.e., the preprint genre. But more generally, we applied the code to p-stories that discuss general aspects of preprints, such as that their lack of peer review must be taken with caution or that preprints are becoming more frequent in science publishing. Illustratively, p-stories have covered the scientific publication system in general, featuring titles such as “Researchers have to publish constantly - this harms science” (2018, SUED).

Genre talk often seems to be linked to educational interventions. Already simple expressions like “a so-called preprint” inform the readership about a new name and concept. Other contexts further explicate that preprints are “a kind of discussion material” (2020, SPON), that they are “currently common in science” (2020, SUED) or that they are “publications, which have not yet taken the high hurdle of being positively reviewed in a disciplinary journal” (2021, ZON). More extensive examples are given below.

In terms of overall numbers, we again see differences between corona and non-corona p-stories. The genre code is more present in the former group than in the latter, and in both groups occurrences decrease over time. A total of 22% of corona p-stories include genre talk in 2020, compared to 13% in 2021. Around 10% of non-corona p-stories addressed the preprint genre in 2018 and 2019, but only around 5% did in the following two years.

In p-stories that describe preprints as a genre, the balance between “pro-” and “con-framings” changed with the pandemic. In non-corona p-stories, genre coincides roughly equally with both uncertainty (53%) and timeliness_or_access (59%), whereas in corona stories, genre coincides much more often with uncertainty (80%) than with timeliness_or_access (44%). This supports our general finding that uncertainty framings occur more often in corona p-stories than in non-corona p-stories.

In some p-stories the preprint genre is discussed as a distinct topic of its own, unrelated to concrete examples or a referenced source. Cases of genre without any reference to a preprint (`min_1_ref = false`) are found more often in non-corona p-stories (100% - 35% = 65% in Table 3b) than in corona ones (100% - 72% = 28% in Table 3a). This result is mainly due to a number of p-stories in 2018 and 2019 discussing overall changes in pre-pandemic science publishing practices, especially in physics and biomedical sciences. Ten of these stories—amounting to almost 12% of all p-stories in these two years—debate the pros and cons of preprints without citing a single concrete preprint. One story speculates that preprints could be the “ideal future” (FAZNET, 2018) of science publication, stressing the pioneering role of physics and biosciences in this regard and quoting a Nobel Laureate as envisioning a world in which all publications are free and post-publication peer-reviewed. Another story speaks of preprint publishing as a form of “basic democracy,” where peers collectively and openly decide upon the quality of a study, but it also notes that through open preprint servers “journalists, policy makers, and lobbyists alike get hold of premature and unreviewed publications, and possibly spread or mobilize half-cooked, false or interest-driven results” (SUED, 2018). This discussion indicates that, already before the pandemic, the emerging role of preprints as a genre in science communication had become a newsworthy topic in the German media.

In later corona p-stories, we observe a profound change of tone in genre debates. Due to the overall explosion of preprints and their new importance for informing policymaking, journalists discussed preprint publishing much more critically. Whereas the earlier genre debates positively highlight free and fast access and collective post-publication review, pandemic genre debates stress the downsides of preprints when accessed by non-scientists and especially when political decisions are based on their results. One p-story highlights, for example, that “especially in relation to the new coronavirus this intensive exchange has a downside” (2020, SUED). Another one explicitly juxtaposes the different epistemic risks associated with physics preprints in comparison to medical ones, which public health decisions are based on:

As long as preprint servers were relevant only to academia, this did not mean a problem: with some delay, the self-control of science eventually worked. Meanwhile the instrument has begun to suffer from its own popularity... In the case of a sloppy study from physics, media sensationalism rarely has negative consequences for the general public... but this can be different in the life sciences. This became clear several times in the corona crisis, for example when a study overestimated the number of unreported infections in the USA or when politicians jumped to conclusions from poorly done experiments (2021, ZEIT).

7 Preprints in focus

In some p-stories the focus is so much on a particular preprint or on preprints in general that they are, in fact, stories about preprints, rather than stories merely mentioning or using preprints. We coded such stories as *p_in_focus*.

P-in-focus-stories typically contain some sort of reference to preprints in their headlines. “Studies present first results on the danger of the British variant B.1.1.7” (2020, SPON) and “Study shows why all hipsters look alike” (2019, ZON) are two example headlines of p-stories focussing on particular preprints. Example headlines of p-in-focus-stories targeting the genre as a whole are “Scientific Results Become Freely Available” (2019, SUEDE) or “Studies on the Coronavirus: Stress Test for Science” (2020, SUEDE).

A special case in this regard is the German debate about a particular preprint by Christian Drosten and colleagues, which not only employed various framings of preprints in general, but also led to a broader discussion about science journalism during the pandemic. The Drosten preprint argued that children were as infectious as adults and cautioned “against an unlimited re-opening of schools and kindergartens” (Jones et al., 2020, p. 6), thereby providing substantial fuel for political conflict. One month after its publication, BILD wrote that the Drosten study was “grossly false,” allegedly based on expert opinions uttered on various platforms. Although this led to a criticism of the journalistic methods used by BILD, it nevertheless triggered a public controversy about the credibility of Christian Drosten. In September 2020 the German press council reprimanded the BILD story, deciding that the wording “grossly false” was not covered by the expert opinions uttered on the Internet. In our dataset, seven p-stories in 2020 and one p-story in 2021 treat the Drosten preprint, its implications or the scandal around it as *p_in_focus* without evoking any reference to a broader, more abstract preprint genre. In 2021 a story about “Science Journalism” in ZON critically picked up on this example:

The pseudo debate launched by Bild... exemplifies how a yellow press paper wants to misunderstand the heart of research, and how it can scandalize a fake controversy with the help of alleged counter-expertise, because lay people can hardly judge such cases, given the high degree of complexity.

More than half of the non-corona p-stories focus on preprints (62% total in Table 2a), compared to only 23% of corona p-stories (total in Table 3b). Stories focusing on preprints more often contain uncertainty framings when also covering corona topics (48% in Table 3a) compared to when they cover non-corona topics (14% in Table 3b). No such difference exists for timeliness or access, confirming the overall observation that preprints are generally framed more controversially in corona p-stories. That 94% of non-corona p-stories marked as genre are also marked as `p_in_focus` (Table 3b) is a result of the abovementioned pre-pandemic genre discussion in 2018 and 2019, in which preprints were discussed as part of a changing academic publishing system. In contrast, the finding that only 36% of non-corona p-stories coded `p_in_focus` are also coded `any_name` is linked to the already discussed finding that the majority of non-corona p-stories that use preprints as sources (`min_ref_1`) do not flag their sources as preprints.

The existence of p-in-focus stories is again an indication that preprints are a newsworthy topic in the German media, both as individual publications (sources) and as an increasingly important genre in science publishing.

8 Conclusion and discussion

Guided by the general question of how preprints are covered and discussed in the media, we used a novel approach that is not only based on a large-scale dataset, but also aggregates the results of an in-depth qualitative analysis. In more concrete terms, we analyzed 1,006 news stories mentioning preprints (p-stories), identified in a sample of 390,942 stories issued by seven major German news outlets from 2018 to 2021. Our study exceeds previous quantitative analyses by using a text-mining approach that is based on various concepts and synonyms for preprints and thereby is not dependent on the availability of hyperlinks or URLs to preprint servers. As the outbreak of COVID-19 in January 2020 occurred in the middle of our study period, we were able to compare the framing of preprints in two groups, corona p-stories and non-corona p-stories. Our qualitative analysis produced 12 codes of interest relating to the five core concepts: 1) framing, 2) naming, 3) referencing, 4) genre, and 5) focussing.

A first and expected finding was that the number of p-stories rose steeply with the beginning of the pandemic, mirroring the increase of preprints in the

dissemination of COVID-19 research. The absolute majority of p-stories in our sample relate to corona reporting. The number of non-corona p-stories remains constant over time.

Second, preprints are framed both in terms of the pros and cons ascribed to them in the broader discourse: scientific uncertainty on the “con-side” and timeliness and access on the “pro-side.” Among the two pro-framings, timeliness was always much more salient than access for any cross-comparison of categories/codes. While the rates of combined pro-framings were subject to annual variations between 17% (in 2019) and 31% (in 2018), we found a more distinctive increase of uncertainty framings in corona p-stories compared to non-corona p-stories. Although uncertainty framings generally increased in both groups over time, they more than doubled in non-corona p-stories from the year 2020 to the year 2021, suggesting that the pandemic led to an overall more cautious attitude towards preprints, even in the case of non-pandemic related topics.

Third, we found a variety of naming practices in relation to preprints, and suggested an analytical distinction between naming and framing. Unlike Fleerackers et al. (2021), we did not interpret the name “preprint” as an uncertainty framing device, arguing that the name itself does not directly invoke uncertainty. We did, however, interpret the German translation “Vorab-Veröffentlichung” as a timeliness framing device, because the prefix “vorab” very specifically means “in advance” and is, ironically, commonly defined by the example sentence “the press has been informed in advance.”⁶ Both our quantitative and qualitative analysis revealed that naming practices are still evolving in the German media. The English name “Preprint” has been used more often over time and seems to gradually replace its German translations.

Fourth, we find that preprints are reported and discussed both as sources and as a genre of academic publishing. In the majority of cases, p-stories reference concrete preprints to state and/or support claims or arguments, and almost always p-stories then contain enough information so that the interested reader can, in principle, trace the original preprint, for example via a hyperlink. When addressing preprints as a genre, p-stories often educate the reader about the nature of preprints as well as negotiate the pros and cons associated with preprints. During the pandemic we observe a profound change of tone in genre debates. In comparison to pre-pandemic times, uncertainty framings of preprints doubled while framings of timeliness and access remained constant.

⁶ <https://www.duden.de/rechtschreibung/vorab>

Finally, we showed that preprints sometimes become a key focus of news stories. Such stories do not merely mention preprints or use them as sources to make an argument, rather they are stories about preprints. P-in-focus stories are much more common in areas unrelated to corona. In fact, more than half of all non-corona p-stories are also p-in-focus stories, and most of them were published in 2018 and 2019, covering for example the general role of preprints in physics or other disciplines. We interpret the existence of p-in-focus stories as an indication that preprints have become a newsworthy topic in the German media, both as individual publications (sources) and as an increasingly important genre in science publishing.

On the basis of these findings, we'd like to discuss the following two points. First, we believe that the use and coverage of preprints present a real challenge to (science) journalism. In science, early access to unreviewed results can be important to drive cutting-edge scientific debates, especially debates in which the involved scientists are able to judge the quality of these results on their own. Journalists, in contrast, are not generally capable of assessing the quality of unreviewed publications, but at the same time they are competing with each other for the latest news and sources. This creates a situation where journalists may become tempted to cite preprints even when their results are of poor quality and thus potentially dangerous when used to inform policy and practice.

Our data shows that some journalists frame preprints in terms of uncertainty, timeliness, and access, and also that uncertainty framings seem to become more salient over time. However, many journalists do not discuss the pros and cons of preprints, and an alarmingly large share of journalists does not even mention that they use or talk about preprints. Describing preprints merely as “studies” or “publications” is dangerous, because it suggests a higher level of scientific certainty as should be associated with preprints. We therefore prompt journalists to become more aware of the differences between peer-reviewed journal articles and preprints. More generally, journalists should acknowledge that there is a palette of scientific genres—both within the realm of peer-reviewed publications and within the realm of unreviewed academic publications—each associated with different (epistemic) advantages and disadvantages. Understanding these differences is a prerequisite for a responsible and transparent communication of research to non-scientific audiences.

Second, in using preprints as sources and in publicly negotiating the pros and cons of preprints (or failing to do so), journalists may influence scientific preprint practices in direct and indirect ways. The politicization of preprints, especially when resulting in personal attacks of their authors, as in the case of the Drosten study, can intimidate and de-incentivise scientists to make their

findings publicly available prior to official peer review. For example, a statistician who critically commented Drosten's first manuscript on a preprint server confessed: "Had I known that Bild reads this sentence, I had definitely not written it [sic]" (2020, SPON). When, in addition, policymakers argue and make decisions on the basis of preprinted results, scientists increasingly fear that others may try to hold them responsible for political consequences of such decisions. In this regard, it is no surprise that we increasingly see warnings issued by preprint servers directly aimed at journalists and policymakers.

To better understand the changing role of preprints in journalism, especially since the explosion of preprints during the pandemic, further research is needed. One direction would be to interview journalists and scientists on the paradoxes and interrelationships that we just elaborated on. How do journalists reflect their use of preprints? How do scientists view the impact of media attention to preprint publishing? Another direction could be to link citation context analysis in the media with bibliometric data from databases like Web of Science or Scopus. This could enable the analysis of citation latency as well as, potentially, the causal modeling of preprint reporting in the media on citation behavior of scientists. In addition, empirical analysis of how different framings impact on the understanding by readers can enrich the literature on narrative framing.

9 References

- Allan, S. (2009). The future of science journalism. *Journalism*, 10(3), 280-282.
<https://doi.org/10.1177/1464884909102570>
- Allen, D.S. (2008). The trouble with transparency. *Journalism Studies*, 9(3), 323-340.
<https://doi.org/10.1080/14616700801997224>
- Beck, J., Ferguson, C.A., Funk, K., Hanson, B., Harrison, M., Ide-Smith, M., Lammey, R., Levchenko, M., Mendonça, A., Parkin, M., Penfold, N., Pfeiffer, N., Polka, J., Puebla, I., Rieger, O.Y., Rittman, M., Sever, R., & Swaminathan, S. (2020). *Building trust in preprints: Recommendations for servers and other stakeholders*. OSF Preprints.
<https://doi.org/10.31219/osf.io/8dn4w>
- Bicchieri, C., Fatas, E., Aldama, A., Casas, A., Deshpande, I., Lauro, M., Parilli, C., Spohn, M., Pereira, P., & Wen, R. (2021). In science we (should) trust: Expectations and compliance across nine countries during the COVID-19 pandemic. *PLOS ONE*, 16(6), e0252892.
<https://doi.org/10.1371/journal.pone.0252892>
- Boetto, E., Golinelli, D., Carullo, G., & Fantini, M.P. (2021). Frauds in scientific research and how to possibly overcome them. *Journal of Medical Ethics*, 47(12), e19-e19.
<https://doi.org/10.1136/medethics-2020-106639>
- Brondi, S., Pellegrini, G., Guran, P., Fero, M., & Rubin, A. (2021). Dimensions of trust in different forms of science communication: The role of information sources and channels used

- to acquire science knowledge. *Journal of Science Communication*, 20(3), A08.
<https://doi.org/10.22323/2.20030208>
- Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). *Preprints and scholarly communication: An exploratory qualitative study of adoption, practices, drivers and barriers* (8:971). F1000Research. <https://doi.org/10.12688/f1000research.19619.2>
- Cobb, M. (2017). The prehistory of biology preprints: A forgotten experiment from the 1960s. *PLOS Biology*, 15(11), e2003995. <https://doi.org/10.1371/journal.pbio.2003995>
- Colavizza, G., Costas, R., Traag, V.A., Eck, N.J. van, Leeuwen, T. van, & Waltman, L. (2021). A scientometric overview of COR-19. *PLOS ONE*, 16(1), e0244839.
<https://doi.org/10.1371/journal.pone.0244839>
- Craig, G. (2016). Reclaiming slowness in journalism. *Journalism Practice*, 10(4), 461-475.
<https://doi.org/10.1080/17512786.2015.1100521>
- Dijk, T.A. van (1988). *News as discourse*. Routledge. <https://doi.org/10.4324/9780203062784>
- Dolgin, E. (2021). The tangled history of mRNA vaccines. *Nature*, 597(7876), 318-324.
<https://doi.org/10.1038/d41586-021-02483-w>
- 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.
<https://doi.org/10.1177/1075547017752166>
- Dunwoody, S. (2014). Science journalism: Prospects in the digital age. In *Routledge handbook of public communication of science and technology*. Routledge.
- Eisen, M.B., Akhmanova, A., Behrens, T.E., & Weigel, D. (2020). Publishing in the time of COVID-19. *eLife*, 9, e57162. <https://doi.org/10.7554/eLife.57162>
- 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>
- Fleck, L. (1980). *Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv* (T. Schnelle & L. Schäfer, Eds.; 13. Edition). Suhrkamp Verlag.
- Fleerackers, A., Riedlinger, M., Moorhead, L., Ahmed, R., & Alperin, J.P. (2022). Communicating scientific uncertainty in an age of COVID-19: An investigation into the use of preprints by digital media outlets. *Health Communication*, 37(6), 726-738.
<https://doi.org/10.1080/10410236.2020.1864892>
- Fraser, N., Brierley, L., Dey, G., Polka, J.K., Pálffy, M., Nanni, F., & Coates, J.A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLOS Biology*, 19(4), e3000959.
<https://doi.org/10.1371/journal.pbio.3000959>
- Guenther, L., Bischoff, J., Löwe, A., Marzinkowski, H., & Voigt, M. (2019). Scientific evidence and science journalism. *Journalism Studies*, 20(1), 40-59.
<https://doi.org/10.1080/1461670X.2017.1353432>
- Heimstädt, M. (2020, April 3). Between fast science and fake news: Preprint servers are political. *Impact of Social Sciences*.
<https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/03/between-fast-science-and-fake-news-preprint-servers-are-political/>
- Horbach, S.P.J.M. (2020). Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. *Quantitative Science Studies*, 1(3), 1056-1067.
https://doi.org/10.1162/qss_a_00076

- Jones, T.C., Mühlemann, B., Veith, T., Biele, G., Zuchowski, M., Hofmann, J., Stein, A., Edelman, A., Corman, V.M., & Drosten, C. (2020). *An analysis of SARS-CoV-2 viral load by patient age*. medRxiv. <https://doi.org/10.1101/2020.06.08.20125484>
- Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism*, 12(3), 279-295. <https://doi.org/10.1177/1464884910388223>
- Karlsson, M., & Sjøvaag, H. (2018). Hyperlinks and linking practice. In *The international encyclopedia of journalism studies* (S. 1-5). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118841570.iejs0231>
- Kreps, S.E., & Kriner, D.L. (2020). Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic. *Science Advances*, 6(43), eabd4563. <https://doi.org/10.1126/sciadv.abd4563>
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLOS ONE*, 10(6), e0127502. <https://doi.org/10.1371/journal.pone.0127502>
- Latour, B. (1988). *Science in action: How to follow scientists and engineers through society* (Reprint Edition). Harvard University Press.
- Le Masurier, M. (2015). What is slow journalism? *Journalism Practice*, 9(2), 138-152. <https://doi.org/10.1080/17512786.2014.916471>
- Lewenstein, B.V. (1995). From fax to facts: Communication in the cold fusion saga. *Social Studies of Science*, 25(3), 403-436. <https://doi.org/10.1177/030631295025003001>
- Matthias, L., Fleerackers, A., & Alperin, J.P. (2020). Framing science: How opioid research is presented in online news media. *Frontiers in Communication*, 5. <https://www.frontiersin.org/articles/10.3389/fcomm.2020.00064>
- Mede, N.G. (2022). Legacy media as inhibitors and drivers of public reservations against science: Global survey evidence on the link between media use and anti-science attitudes. *Humanities and Social Sciences Communications*, 9(1), Art. 1. <https://doi.org/10.1057/s41599-022-01058-y>
- Moore, C.A. (1965). Preprints. An old information device with new outlooks. *Journal of Chemical Documentation*, 5(3), 126-128. <https://doi.org/10.1021/c160018a003>
- Nelkin, D. (1995). *Selling science: How the press covers science and technology*. W.H. Freeman & Co Ltd.
- Pan, Z., & Kosicki, G.M. (1993). Framing analysis: An approach to news discourse. *Political Communication*, 10(1), 55-75. <https://doi.org/10.1080/10584609.1993.9962963>
- Parviainen, J. (1998). *Bodies moving and moved: A phenomenological analysis of the dancing subject and the cognitive and ethical values of dance art*. Tampere University Press.
- Riles, J.M., Sangalang, A., Hurley, R.J., & Tewksbury, D. (2015). Framing cancer for online news: Implications for popular perceptions of cancer. *Journal of Communication*, 65(6), 1018-1040. <https://doi.org/10.1111/jcom.12183>
- Soderberg, C.K., Errington, T.M., & Nosek, B.A. (2020). Credibility of preprints: An interdisciplinary survey of researchers. *Royal Society Open Science*, 7(10), 201520. <https://doi.org/10.1098/rsos.201520>
- 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
- The Associated Press. (2019). *The associated press stylebook 2019: And briefing on media law* (New Edition). Basic Books.

- Till, J.E. (2001). Predecessors of preprint servers. *Learned Publishing*, 14(1), 7-13.
<https://doi.org/10.1087/09531510125100214>
- Torres-Salinas, D., Robinson-Garcia, N., & Castillo-Valdivieso, P.A. (2020). *Open access and altmetrics in the pandemic age: Forecast analysis on COVID-19 literature* (S. 2020.04.23.057307). bioRxiv. <https://doi.org/10.1101/2020.04.23.057307>
- Tuchman, G. (1980). *Making news*. Free Press.
- Wang, P., & Tian, D. (2021). Bibliometric analysis of global scientific research on COVID-19. *Journal of Biosafety and Biosecurity*, 3(1), 4-9.
<https://doi.org/10.1016/j.jobbb.2020.12.002>
- Weingart, P. (1998). Science and the media. *Research Policy*, 27(8), 869-879.
[https://doi.org/10.1016/S0048-7333\(98\)00096-1](https://doi.org/10.1016/S0048-7333(98)00096-1)
- Weingart, P., & Guenther, L. (2016). Science communication and the issue of trust | JCOM. *Journal of Science Communication*, 15(5), C01. <https://doi.org/10.22323/2.15050301>