The virus that brought science into the limelight

François Claveau, *Université de Sherbrooke and CIRST*[1]
Jean-Hugues Roy, *UQAM and CIRST*
Olivier Santerre, *Université Laval and CIRST*

Abstract

The COVID-19 pandemic has increased both journalists' use of science in their reporting and the public's interest in science-related news. This is what stands out of a computational analysis of articles and Facebook posts of French-language media in Canada. The representation of science by the media has also changed during the pandemic: science's role as a guide to governments has come into the limelight.

Introduction

The COVID-19 pandemic has put much pressure on science, forcing a mass mobilization of researchers to get a grip on the virus. The pandemic also appears as an opportunity for science, a way to prove its value to society. Early in the first wave, Quebec's Chief Scientist, Rémi Quirion, claimed that "science during this pandemic is gaining more prominence in terms of visibility in society, and in terms of relevance and importance in the eyes of decision makers and the general public." (2020)[2]

Is there evidence that such a significant change occurred in the public image of science? Using two large corpora of French-language media in Canada (more than 960,000 media articles and more than 160,000 Facebook posts), this article relies on computational methods to answer three questions related to the first one about the presence of science in the media and in the general public during the first wave of the pandemic (March to June 2020):

1. Was science more present than usual?
2. What was the representation of science by the press?
3. Did the public engage more with science?

---

[1] CIRST: [Centre interuniversitaire de recherche sur la science et la technologie](#)
[2] Our translation of: *"La science en ce temps de pandémie planétaire prend plus de relief en termes de visibilité dans la société, de pertinence et d'importance aux yeux des milieux de décision et du grand public"*.

Our results show that the COVID-19 pandemic has increased both journalists' use of science in their reporting and the public's interest in science-related news. The representation of science by the media has also changed during the pandemic: science's role as a guide to governments has come into the limelight.

Science in the press

Journalism always looked up to science to clean up its act. A century ago, Walter Lippmann deplored the sad state of American journalism, its reputation tarnished by the sensationalistic "yellow press". In *Liberty and the News*, a classic in journalism literature, he pleaded that through formal training inspired by science, the "haphazard trade" that was journalism would become a dignified profession, "for the true patterns of the journalistic apprentice are not the slick persons who scoop the news, but the patient and fearless men of science who have labored to see what the world really is" (1920, p. 82).

Half a century later, Philip Meyer argued that "we journalists would be wrong less often if we adapted to our own use some of the research tools of the social scientists" (1973, 3). Meyer demonstrated the power of those tools in 1967. After riots in Detroit claimed the lives of 43 people, he led a team of reporters who surveyed citizens and analysed their answers with a computer, a novel method at the time. The reporting earned Meyer and his team a Pulitzer Prize, the most prestigious award in American journalism, the following year and paved the way to computer-assisted reporting, a practice which evolved today into data and computational journalism.

A quantitative look

In spite of this, science seldom makes headlines. A content study of three U.S. dailies (the *New York Times*, the *Washington Post* and the *Chicago Tribune*) in a sample spanning three decades (1966-70; 1976-80; 1986-90) shows science articles make up 2% or less of the total number of articles (Pellechia 1997, 57). Metcalfe and Gascoigne found similar proportions in Australian newspapers, where "the average share of science news [...] rose from 1.3% in 1989 to 2.9% in 1993" (1995, 414). Even though Pellechia also observed an increase in the space devoted to science in the decades leading to the 1990s, scholars have noted that this trend has reversed in the decades since. "Traditional media are devoting less time and resources to monitoring [scientific] issues for the public" (Dudo 2015, 765). "As the newspaper industry responds to falling circulation with sweeping cuts, science desks are the first to suffer", lamented the editorial team of Nature (2008, 388). Furthermore, the dwindling number of science reporters "has created an opening for the publication of public relations (PR)-generated

content as journalistic content [...], thus relinquishing the traditional democratic role of the press as a watchdog that can signal misconduct, raise ethical questions and make critical observations" (Barel-Ben David, Garty, and Baram-Tsabari 2020, 2).

All of these trends are source for concern, because notwithstanding the rise in social media use during the 2010s and the fact 35% of Americans and 41% of Canadians now use Facebook for news (Newman 2020, 88, 91), most citizens still get informed about science through lay media (general interest newspapers, radio, television or websites). According to a Pew Research survey, 54% of Americans "get their science news from general sources", which is much higher than specialized sources. For example, 25% rely on science magazines or websites, 12% on science radio programs or podcasts (Funk, Gottfried, and Mitchell 2017, 5).

Few studies have yet examined how these trends played out during the COVID-19 pandemic. Most studies have focused on the spread of pseudo-scientific "news" and misinformation (Bridgman et al. 2020; Sharma et al. 2020; Vraga, Tully, and Bode 2020). The remainder have mainly focused not on scientific coverage, but on the coverage of the disease itself. Hervé, for example, found that in the third week of March, 2020, almost 80% of AFP's dispatches were related to the coronavirus (2020, 11). Ours seems to be one of the few studies examining not only if the pandemic has changed the way journalists cover science, but how.

A qualitative look

The way journalism talks about science has also been examined by scholars. There is a tendency in the press to portray science either in "a very positive, even ecstatic way" or to "present an exaggerated --often alarmistic-- negative image" of it (Schnabel 2003, 255). Those two extremes are reflected in the 2017 Pew survey. "The most common science stories people see are reports of new discoveries, followed by stories about 'strange or weird' research findings" (*op. cit.*, 23).

That's because coverage of science does not follow scientific norms, but journalistic norms, says Sharon Dunwoody. "Journalists are more likely to produce shorter stories about concrete happenings than longer, thematic stories about issues" (2008, 19). They are also looking for… well… news. Novelty and timeliness are important for journalists, who wish to "hang stories on traditional news pegs, characteristics of real-world processes that are proven attention-getters" (*op. cit.*).

This explains why most science journalism focuses on medicine and health. In her literature review of science journalism studies, Dunwoody found that the majority of science stories published by newspapers in the United States, the United Kingdom, Italy

and Canada were health-related. When looking at science, journalists seem to select more often topics with a consumer twist, "news you can use" (*op.cit.*, 18).

In the Italian study, covering half a century of science coverage in Milan daily Il Corriere della Sera, the authors found that science items are generally short, most often buried deep in the newspaper and printed on the front page in only 5.8 percent of the cases (Bucchi and Mazzolini 2003, 9). The Italian authors also found that articles on science rarely venture beyond scientific or technological issues to fray on the social, cultural economic or political aspects of a given science-related story. Less than 20 percent of the stories in their sample do so (*op. cit.*, 11). In a more recent study looking at how German science journalists select their topics, Badenschier and Wormer found that 21.4 percent of science articles "noted people or institutions with political, economic or cultural power"; the remainder noted, in a majority of cases, people or institutions with "scientific power" (2012, 75).

While our study has found a similar proportion of science articles relating to government affairs before 2020, we also found that this proportion has significantly increased during the pandemic (more details in Results, below).

Data and methods

We analysed the content of French-language Canadian general news[3] outlets (the list is in the Technical Appendix). They include dailies, weeklies, talk and community radio stations, local and national television news stations and networks, as well as news websites. From that list, we started by building two corpora.

The first corpus includes articles posted by the 139 media outlets' RSS feeds and captured by the Information Flow Observatory (IFO), a database of "information produced by Canadian media, news agencies, the public relations industry as well as by the regional and federal governments" made available to scholars and non-profit organizations (Rocheleau 2017).

The second corpus includes posts published on the media outlets' 153 Facebook pages[4]. These were gathered using CrowdTangle, a Facebook-owned social media content research tool made available to scholars through a partnership between Facebook and Harvard's Social Science One (2019).

---

[3] "General news" means we excluded media outlets specializing in any area (sports, arts, business, science).

[4] Some organizations have more than one Facebook page, this includes Radio-Canada (the French-language portion of the CBC) and TVA (a television network owned by Québecor), both of which have many regional pages such as "ICI Grand-Montréal" or "TVA Abitibi", for example.

From these two sources, we retrieved texts published in the first half (January to June inclusively) of the years 2017 to 2020. We focused on this period to isolate the first wave of the pandemic, the years 2017 to 2019 serving as contrasts to single out the peculiarities of 2020.

Each text of the two corpora (article or post) was categorized based on two dichotomous variables:

1. Whether it is related to science or not (method detailed below);
2. Whether it was published in the year of COVID-19 (2020) or not (2017 to 2019).

Table 1 presents the distribution of texts over these categories in our two corpora.

Table 1: Breakdown of texts by categories

	Articles		Facebook posts	
	Science	Not science	Science	Not science
Before COVID (2017-2019)	32,996	665,520	1046	102,166
Since COVID (2020)	17,310	247,001	562	60,100

Identifying science-related texts

In the last three decades, two main methods have emerged to perform content analysis at a great scale. The first is called the dictionary-based approach. It involves creating a list of words or expressions to look for in a corpus in order to classify a unit of text in a given category or topic. The other involves the use of machine learning to automatically identify attributes in a corpus. Topic modeling is an example of this second method.

The dictionary-based approach still "is the most widely used approach in computer-assisted content analysis." (West 2001; Guo et al. 2016, 335) Guo et al. compared both approaches. They found each have its merits and pitfalls, but that "when researchers are only looking at a specific issue or topic [...], the dictionary approach remains more 'focused'." (*op. cit.*, 351)

We thus chose a dictionary-based approach to classify our two main corpora into science-related articles or Facebook posts and non-science-related articles and Facebook posts. Our dictionary was quite focused as it contained only two strings at the beginning of words:

- ***"scien"***, which enabled us to catch terms such as *"science"* or *"scientifique"* (which, in French, is both the adjective "scientific" and the noun "scientist")

- ***"chercheu"***, which enabled us to catch both masculine, feminine and plural forms of the French word for researcher (*"chercheur", "chercheuse", "chercheurs", "chercheuses"*)

In an early iteration of our analysis, we included the term "*étude*" (study, studies). It produced too many false positives so we removed it. We still find some false positives. For instance, late in our analysis, we discovered that our indicator of science falsely matched the expression *"chercheur d'emploi"* (job seeker). Since the proportion of false positives appears very small, we did not further refine our dictionary.

We also need to add that our dictionary is certainly not identifying all instances of text with scientific content. For instance, a text could use domain-specific terms such as *"microbiologie"* (microbiology) or *"physicien"* (physicist) and thus be missed by our dictionary. The problem with adding domain-specific terms is that, on the one hand, one needs to add a lot of them to avoid biases due to underrepresentation of some domains and, on the other hand, multiplying terms generates more false positives. We thus refrained from increasing the size of our dictionary.

Finally, our method of defining science-related texts must be contrasted to a more substantive approach found in the literature. For instance, Cole (1975, 466–67) defines "science news" as "all news stories which have substantial subject matter concerning results or interpretation of empirical research in the sciences, applied science or development, technology, engineering, medicine and public health". To identify these science news, one must inspect all texts to decide whether they meet the definition. Our procedure is shallower and, consequently, we certainly identify texts that are explicitly related to science, but that would not fall under a more substantive definition of science news.

Defining the COVID period

On December 31, 2019, the World Health Organization announced that the "Wuhan Municipal Health Commission, China, reported a cluster of cases of pneumonia in Wuhan, Hubei Province [and that a] novel coronavirus was eventually identified" (WHO 2020). In Quebec, it was not until February 27, 2020, that the first case of COVID-19 was declared. Two weeks later, on March 12, Prime Minister François Legault, Minister of Health and Social Services, Danielle McCann, and the Director of Public Health, Horacio Arruda, held the first of many press conferences about the measures taken by the government to stop the spread of the disease. On March 13, the government declared a public health emergency throughout the province (2020). Figure 1 shows the daily number of new cases of COVID-19 from late February to the end of November 2020.
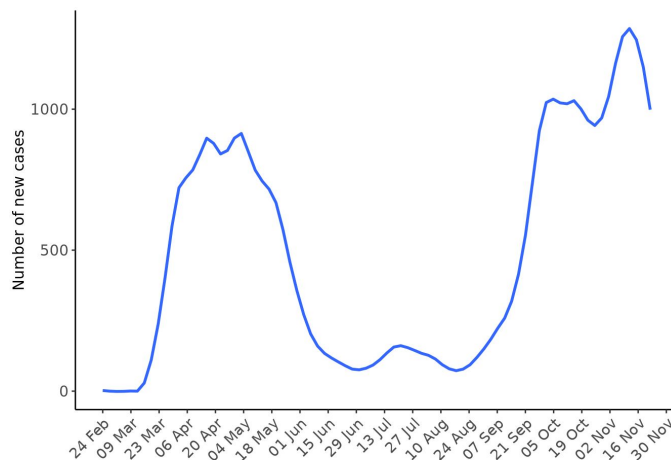
Figure 1: Number of new COVID-19 cases in Quebec
(curve smoothed using local polynomial regression) (INSPQ
2020)

To find articles mentioning the pandemic in texts published in 2020, we used another dictionary composed of four terms:

- *"virus"*, which captures both *"virus"* and *"coronavirus"*
- **"COVID"**
- **"sras-cov"** (French acronym) and
- **"sars-cov"** (English acronym, sometimes used in French).

Figure 2 shows the proportion of texts in our article corpus including at least one of these terms. This proportion exploded between March 9 and March 16 to reach a peak at around 70% of all articles before slowly trending down throughout the first wave of the pandemic.
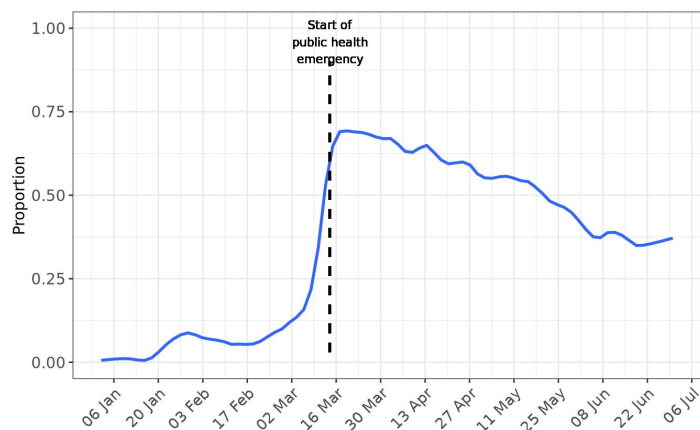
Figure 2: Proportion of articles in our corpus mentioning "virus", 'COVID', 'sras-cov' or 'sars-cov' (curve smoothed using local polynomial regression)

Figures 1 and 2 establish that March is the month when it all began in Quebec, and the same holds for Canada more generally. For much of our analysis below, we consequently focus our attention on March 1st, 2020, to June 30, 2020, as being the "COVID period."

Results

Each of the next three subsections is dedicated to answering one of the questions asked in our introduction. We present further methodological information as we proceed.

More science?

Is the pandemic responsible for an increase in scientific content in the media? Figure 3 gives the proportion of science-related articles for the full first half of each year since 2017. Three results stand out.
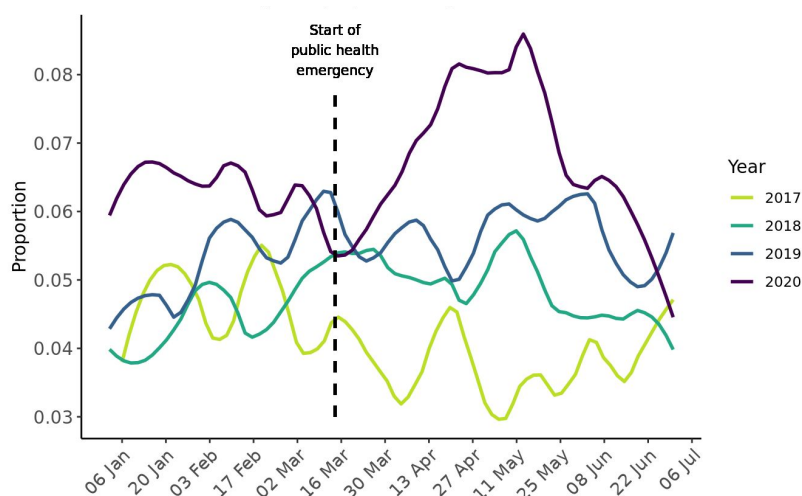
Figure 3: Proportion of science-related articles in Franco-Canadian media (curves smoothed using local polynomial regression)

First, the proportion of science-related articles in any year (between 4% and 7%) is generally higher compared to the results reported in the previous studies discussed above (less than 4%). This discrepancy is probably not due to a "Franco-Canadian difference", but rather to the distinction between "science news" (what the extant literature is about) and "news mentioning science" (what we attempt to identify).

Second, there is an upward trend in the proportion of science-related articles over our four years of data. From 4.1% in 2017, the proportion increased to 4.8% and 5.5% in

2018 and 2019 respectively, and has reached 6.5% in 2020. The growth of the proportion has been strongest from 2019 to the pandemic year, but science was already on the rise in Franco-Canadian media in prior years.[5]

Third, the proportion of science-related articles has reached unprecedented heights when the first wave of the pandemic was at its peak and it decreased as this first wave died down. There is thus a strong positive association between the pandemic context and a higher proportion of science-related articles.

Although this association is unlikely to be a coincidence, it is worth investigating whether the articles that explicitly refer to COVID-19 are more likely to be related to science according to our indicator. Figure 4 splits the proportion of science-related articles in our 2020 corpus between those explicitly referring to COVID-19 and the others. We find that our COVID sub-corpus is two times more tightly related to science (9.5%) than its complement (4.7%).
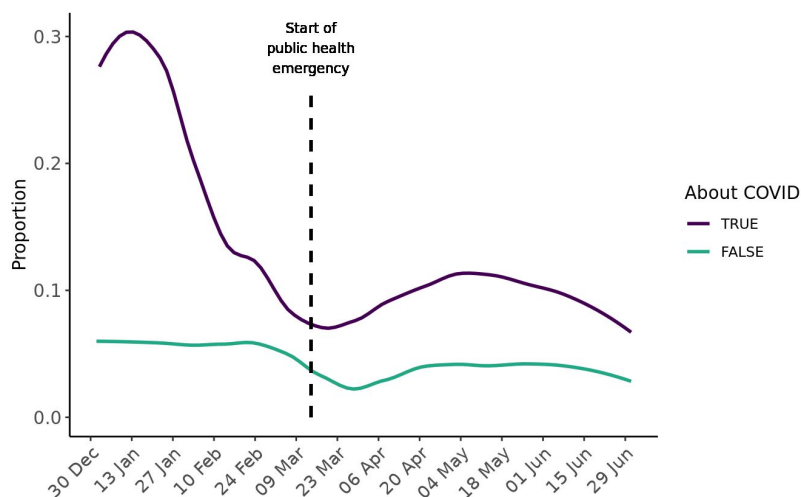
Figure 4: Proportion of science-related articles among those mentioning COVID-19 (curves smoothed using local polynomial regression)

Furthermore, we also unearth important trends. Early articles on COVID-19 are frequently detected as related to science (almost 30% in January 2020). This proportion reaches a low point when the public health emergency is announced in mid-March and then follows a pattern similar to the first wave until the end of our period. With respect to

---

[5] This result does not appear to be an artefact of some instability in our database. The trend persists (1) if we focus on the five main media outlets in the province, (2) if we drop articles below 1000 characters and (3) if we combine these two constraints. See the Technical Appendix for details.

articles not about COVID, we find that the pandemic context crowded out other scientific content: 5.8% of non-COVID articles were related to science in January and February, but this proportion drops to 3.8% in March to June 2020.

What science ?

The last result of the previous section already suggests that the science showcased in the media during the pandemic has been predominantly about the disease. This is unsurprising given the extent to which COVID-19 dominated media coverage between March and June 2020 (see Figure 2). Beyond this unprecedented focus on a topic, can we find other changes in the media representation of science in the pandemic context?

We are specifically looking for two aspects of the representation of science: the functions and the degree of uncertainty of science. Our hypotheses are that:

1. (functions) The science showcased during the pandemic is:

    a. more explicitly the "fifth branch" of government (Jasanoff 1990), **deeply involved in the policy regulation** of contemporary societies.

    b. **less presented as focusing on understanding the world**, as aiming to be a "mirror of nature" (Rorty 1979).

    c. **less presented as a source of solutions to concrete problems faced by readers**, as offering "news you can use" (Dunwoody 2008, 18).

2. (uncertainty) The science showcased during the pandemic is **more a "science in the making"** than a "ready made science" (Latour 1987, 4). We should thus find more markers of uncertainty in 2020 than in previous years in news including science as SARS-CoV-2 is a new virus about which knowledge increased daily during the first months of the pandemic.

To test these hypotheses, we need a multipronged approach. First, we propose to focus on verbs that discriminate between the media coverage of science before the pandemic (2017-2019) and during the pandemic (March to June 2020). The focus on verbs is a strategy to avoid seeing only the topical difference (pandemic science versus other sciences) because it goes without saying that names and adjectives related to public health issues are extremely more present in the 2020 corpus. Using morphosyntactic analysis, we took all verbs in the paragraphs[6] identified as related to science in the articles published between March and June from 2017 to 2020. Based on relative

---

[6] In this step, we focus on individual paragraphs instead of full articles to avoid including a large number of verbs present in a part of an article where no science is discussed.

frequencies of these verbs in 2017-2019 versus 2020, we can extract the most characteristic verbs for each period. Figure 5 presents these most characteristic verbs based on values of chi-square residuals.[7]
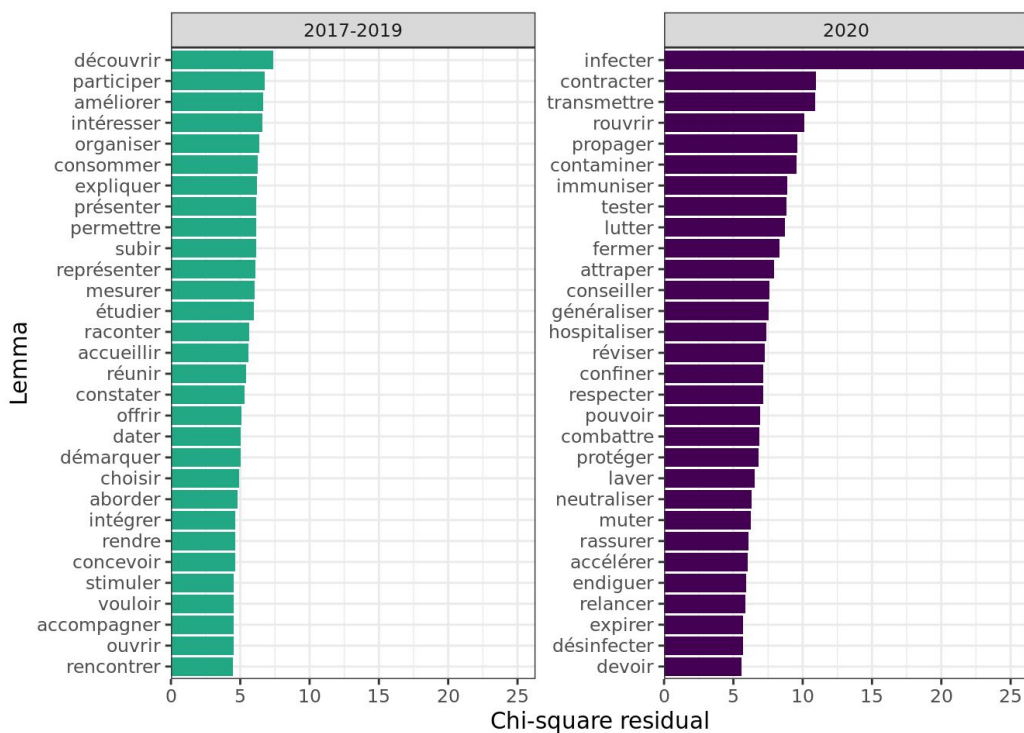
Figure 5: Top-30 most characteristic verbs in science-related paragraphs from media articles published between March and June in 2017-2019 and in 2020

If we consider only the 2020 panel of Figure 5 which gives the characteristic verbs of the pandemic period, we could conclude that our procedure does not succeed in going beyond the topic of the pandemic. Indeed, most verbs are strongly associated with public health and pandemic management. The 10 most characteristic verbs are:

- to infect,
- to contract,
- to transmit,
- to reopen*,
- to propagate,
- to contaminate,
- to immunize,
- to test*,

[7] See the Technical Appendix for examples of paragraphs most representative of the context of use of the top 10 verbs in each sub-corpus.

-   to fight,
-   to close*.

Some verbs (those with an asterisk) are clearly linked with government action, but they are so tied up with the pandemic that we should not overinterpret what they represent.

However, the characteristic verbs of 2017-2019 relate more clearly to the hypothesis of this section. By being characteristic of 2017-2019, they are simultaneously **uncharacteristic of 2020**.

The 10 verbs in science-related paragraphs that are most uncharacteristic of 2020 are thus:

-   to discover,
-   to participate (in a study),
-   to improve*,
-   to find interesting,
-   to organise,
-   to eat/drink/use* (many meanings of the French verb *"consommer"*)
-   to explain,
-   to present,
-   to allow,
-   to suffer*.

Most of these verbs (those without an asterisk) relate strongly to the representation of science as an institution to understand the world. It leaves the reader with the image of a science that should "be interesting" to the public because it aims to "discover" and to "explain". A few other verbs (those with an asterisk) are more strongly associated with science as offering health-related "news you can use": it tells readers how science can "improve" their lives, how it can help them as "consumers" or how it can relieve them from "suffering".

In line with our hypothesis, the media representation of science during the first wave of the pandemic deemphasized two images: science as a detached quest to understand the world and science as a quest that delivers practical applications to better the lives of readers. What representation took more space then? Our hypothesis is that the image of science as a guiding hand for the government gained significant ground.

To test this hypothesis, we propose to measure whether the semantic field of "government" is more strongly associated with science-related paragraphs during the pandemic than before. We again used a dictionary-based method for this assessment, relying on the UCREL semantic analysis system (Rayson et al. 2004). We used the

French-language version of its semantic lexicon[8] which we improved by removing France or Europe-specific vocabulary and adding words more related to the Canadian and Québécois contexts. This system contains the semantic field "G. Government and the Public Domain", which is further divided between "G1.1 Government etc.", "G1.2 Politics", "G2.1 Crime, law and order" and some other tags.

For our two corpora (2017-2019 and 2020), we compute the proportion of science-related paragraphs that include words associated with the semantic fields G1.1, G1.2 and G2.1. Our hypothesis implies that we should find a higher proportion of paragraphs in 2020 associated with G1.1 (Government), but the proportions of paragraphs linked to G1.2 (Politics) and G2.1 (Crime and law) are not expected to be higher during the pandemic.

Table 2: Proximity of science-related paragraphs to three semantic fields. The second and third columns give the proportion of paragraphs in 2017-2019 and in 2020, respectively, detected as associated with the semantic field. The last column gives the p-value associated to a two-sided test of equal proportions.

Semantic field	Proportion before COVID	Proportion during COVID	p-value
G1.1 Government	0.228	0.268	0.0000
G1.2 Politics	0.098	0.092	0.0241
G2.1 Crime and Law	0.105	0.093	0.0000

Table 2 strongly supports our hypothesis. We see that 26.8% of science-related paragraphs are identified as connected with the semantic field "G.1.1 Government" during the pandemic while this proportion was at 22.8% in 2017-2019. In reverse, the semantic fields G1.2 Politics and G2.1 Crime and Law were slightly more frequently associated to science-related paragraphs in 2017-2019.

We now turn to the second hypothesis investigated in this section: is science in the context of the pandemic represented as being more unstable, less certain? A first hint pointing to an affirmative answer to this question comes from the characteristic verbs in Figure 5: one characteristic verb of the 2020 corpus is *"pouvoir"*, the equivalent of "can/could/might" in English. Constructions with this verb typically convey uncertainty.

Since using conjugated forms of the verb *"pouvoir"* is only one way to express uncertainty, we need a more refined tool for uncertainty detection. Unfortunately, the extant tools are typically in English and are domain specific -- e.g., developed for health reports (for a review, see Dalloux 2017). In French, we are aware of only one corpus

---

[8] https://raw.githubusercontent.com/UCREL/Multilingual-USAS/master/French/semantic_lexicon_fr.usas (last accessed: 2020-11-25) for the French lexicon. See the Technical Appendix for details on our improvements.

annotated with uncertainty scopes, but it is composed of clinical cases in medicine (Grabar, Claveau, and Dalloux 2018).

Compared to the technical literature about uncertainty detection, our objective is relatively modest[9]: we seek an indicator to reliably determine which one of two corpora conveys uncertainty more frequently. We developed the following procedure. We coded a rule to identify sentences conveying uncertainty. This rule is a partial translation from English of the patterns used for uncertainty detection in a recent expert system, NegBio (Peng et al. 2018). It takes into account the presence of terms such as "possible" and "probable", and also the presence of verbs in the conditional tense (see the Technical Appendix for details). The rule thus created is not perfect, but we are confident that it can reliably find which corpus conveys more uncertainty: taking the proportion of sentences detected as conveying uncertainty in each corpus, we can do a simple test of equal proportions.

Table 3: Proportion of sentences conveying uncertainty in the science-related paragraphs of the article corpus. The second and third columns give the proportion of sentences in 2017-2019 and in 2020 respectively. The last column gives the p-value associated to a two-sided test of equal proportions.

	Before COVID	During COVID	p-value
Proportion of sentences conveying uncertainty	0.145	0.181	0.0000

Table 3 indicates that science-related content during the pandemic is more strongly associated to instability: the proportion of sentences detected as conveying uncertainty is 25% higher than in the previous three years. This result corroborates our hypothesis that, during the first wave of COVID-19, science was presented more as being "in the making" than "ready made" compared to the standard media coverage.

More public engagement with science?

We have established that, in Franco-Canadian lay media during the pandemic period, not only was science more present, but it was also more policy-focused and more uncertain. Now, media scrutiny does not necessarily translate into more attention in the general public. Did the pandemic context give a specific boost to general public engagement with science-related news? We turn to data from our Facebook corpus to answer this question.

Facebook's CrowdTangle returns, for each post, three main metrics:

---

[9] The technical literature attempts to reliably identify *all* markers of uncertainty in a corpora and the scope they cover in each sentence.

- number of times the post was shared,
- number of emotional reactions to the post ("likes" along with the more recently introduced "love", "wow", "haha", "sad", "mad" and "care" emoticons)
- number of comments made below the post.

The sum of these metrics is what Facebook calls "engagement".

For each post from our list of Franco-Canadian media outlets, we associate three characteristics:

1. Whether it was published during the COVID period (March to June 2020; $C = 1$) or before (March to June in 2017 to 2019, $C = 0$)
2. Whether it explicitly refers to science (based on our standard dictionary; $S = 1$) or not ($S = 0$)
3. Its total engagement E

The question of this section can thus be more precisely formulated: is the density distribution of total engagement shifted upward for science-related posts during the pandemic in contrast to the other posts (i.e., posts not explicitly about science and posts in years prior to the pandemic)?

A first pass at answering this question implies comparing the densities of total engagement for our four types of posts: before the pandemic or not, about science or not. Since the distribution of total engagement is strongly positively skewed (i.e., a small proportion of posts have extremely high values of total engagement), we take the logarithm of engagement. Figure 6 plots the four distributions.
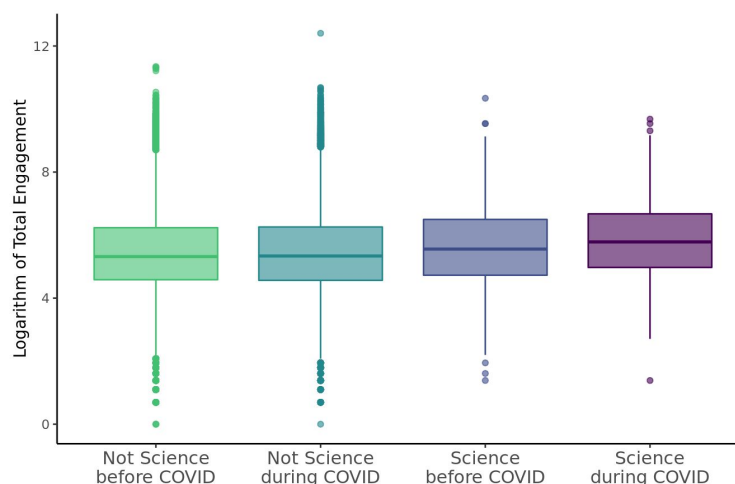
Figure 6: Engagement with Facebook posts from Franco-Canadian media according to subject (science vs non-science) and period (before vs during COVID).

It indicates that the distribution of total engagement for posts corresponding to "science during COVID" is shifted upward compared to the other types of posts, supporting the claim that there has been more attention paid to science by the general public during the pandemic. Moreover, the Figure does not suggest that posts in general have generated more engagement during the pandemic: the change in relative attention seems to be specific to science-related posts.

To establish more rigorously these conclusions, we ran the following regression model[10]:

$$log(E) = \alpha + \beta_1 S + \beta_2 C + \delta S \times C + \varepsilon, \text{ where } \varepsilon \sim NID(0, \sigma^2)$$

The parameter estimate (δ) for the interaction term ($S \times C$) is positive and significant, see Table 4. This result thus corroborates the conclusion reached by visual inspection of Figure 6: science-related posts did have a statistically significant boost with respect to public engagement during the pandemic.

Table 4: Regression results of Facebook engagement on science-related content and the pandemic context

	Dependent variable:
	log(engagement)
Is Science	0.207***
	(0.050)
Since COVID	0.021***
	(0.008)
Is Science × Since COVID	0.173**
	(0.079)
Constant	5.462***
	(0.005)
Observations	112,785
R^2	0.001
Adjusted R^2	0.001
Residual Std. Error	1.276
F Statistic	21.378***
Note:	*p<0.1; **p<0.05; ***p<0.01

---

[10] We also ran a one-way ANOVA test on the logarithm of total engagement for the four groups of posts. The hypothesis that distribution means are equal is strongly rejected (p-value = 8 x 10^{-14}).

How big was this boost in engagement? Regressions with log-transformed dependent variables give relationships between geometric means, which are reported in Table 5.

Table 5: Regression results of Facebook engagement on science-related content and the pandemic context

Parameter combination	Corresponding status	Point estimate of geometric mean
e^{α}	Not science before COVID	236
$e^{\alpha+\beta_2}$	Not science during COVID	241
$e^{\alpha+\beta_1}$	Science before COVID	290
$e^{\alpha+\beta_1+\beta_2+\delta}$	Science during COVID	352

The model thus indicates that the geometric mean of total engagement of science-related posts during the pandemic (mean = 352) is 49% higher than for pre-pandemic posts not explicitly related to science (mean = 236). Science-related posts in years before the pandemic also generated higher engagement than posts not related to science (23% more). But the pandemic context provides a significant boost in engagement for science-related posts: they have generated 46% more engagement compared to non-science posts during the pandemic. Furthermore, only science-related posts seem to benefit from this boost as the increase in engagement for non science-related posts in the COVID period is negligible (only 3%) compared to the pre-COVID period.

Science-related posts did not uniformly generate more engagement during the pandemic. To provide a more detailed picture, we first look at the terms that are characteristic of science-related posts with high engagement, using characteristic terms of posts with low engagement as a contrast. We include all adjectives, nouns and verbs.
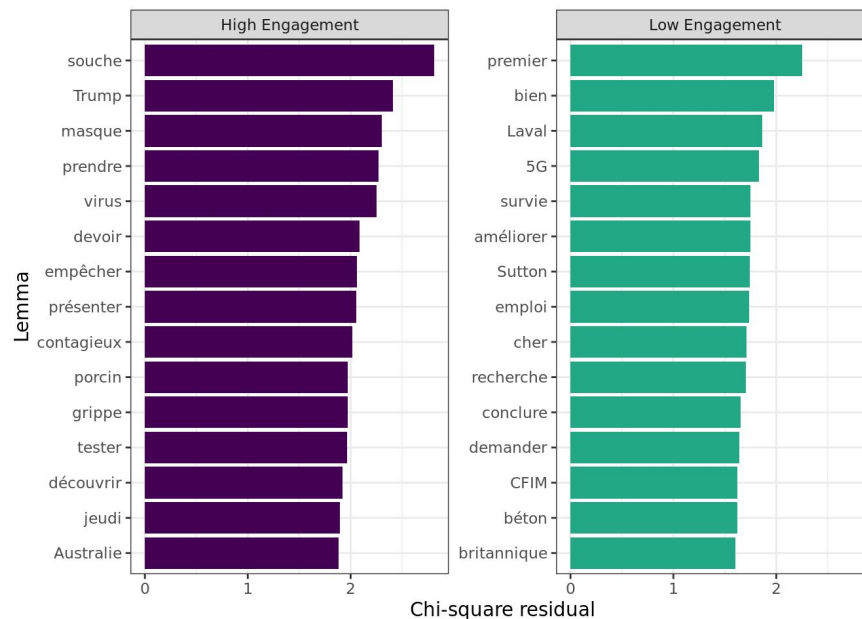
Figure 7: Top-15 characteristic terms of science-related posts generating high engagement during the pandemic (with low-engagement for contrast)

Figure 7 confirms that the science-related posts that attracted more engagement were clearly about the pandemic[11]. In contrast, the terms associated with less engaging posts are less tightly connected to the pandemic. Another telling result is the presence of "Trump": Donald Trump is the only public figure systematically generating higher-than-average engagement on science-related posts. No politician from Canada appears on the Figure.

---

[11] See the Technical Appendix for a presentation of the most characteristic contexts of each many terms, a presentation that supports our claim here that posts generating high engagement were predominantly about COVID-19.
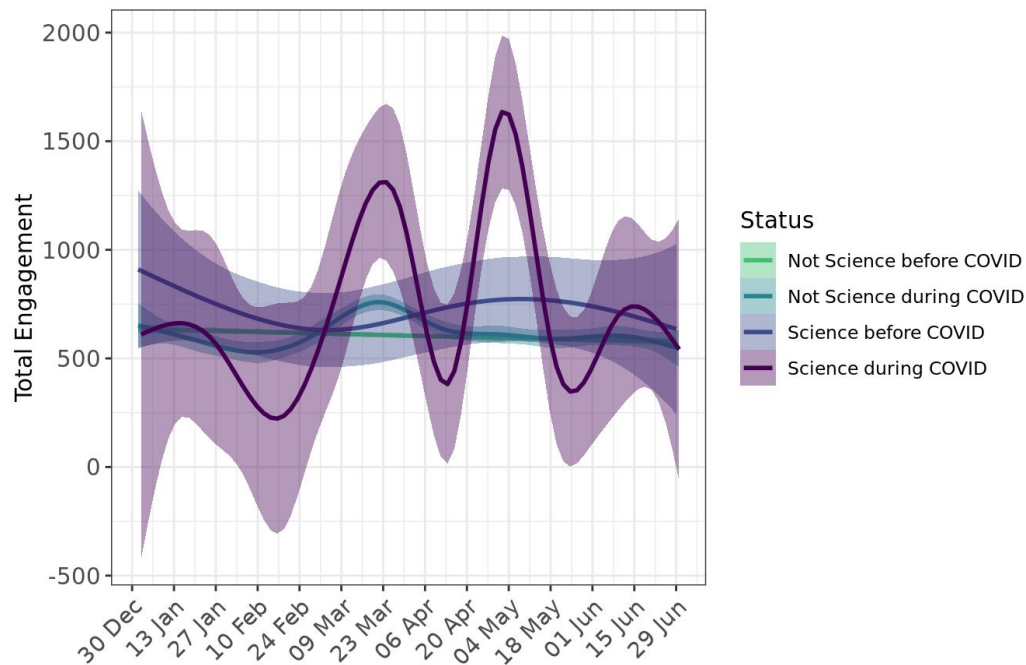
Figure 8: Engagement with Facebook posts in the first half of each year
(2017 to 2020; curves are smoothed using local polynomial regression)

To further study engagement with respect to science-related posts, our second extension is to look at the distribution of engagement over the span of the first half of the year, see Figure 8. The Figure indicates clearly that total engagement was significantly higher than normal for science-related posts during only two periods over the first wave of the pandemic. These periods correspond to:

1. Mid-March to early April

2. Late April to mid-May

The first period corresponds to the early days of public health emergency in Quebec. The general public was particularly avid of information during this period and all science-related posts generating strong engagement are about the pandemic. The most popular science-related post, published on March 23 by Radio-Canada, is about the imminent launch of a study on a possible treatment for COVID-19. The second and third most popular posts, both published by TVA Nouvelles, are respectively about a newborn being tested positive for COVID-19 and about traces of the virus having been found on the cruise ship *Diamond Princess* 17 days after the passengers and crew had left.

The second period of high engagement is well into the first wave. Almost all science-related posts with high engagement are again about COVID: the possibility of the virus also being sexually transmissible, the discovery of a new strain, the risks of

reopening schools too quickly, the suspicious death of a researcher working on COVID in the United States and the negative reactions of scientists to Trump's suggestions for self treatment. Only one post with high engagement is unrelated to the pandemic: sightings of "murder hornets" in North America.

In sum, science, and the science associated to COVID in particular, did generate significantly more engagement during the first wave of the pandemic. One must however keep in mind that the Facebook posts that generated the most engagement share many characteristics with what can be called "viral" content. They were about news items that were more sensationalistic in nature than really relevant or significant in the bulk of news about the pandemic.

Conclusion

Our study has quantified some widespread impressions about the place of science in the media during the first wave of the pandemic. While confirming these impressions, we also show that the transformation has not been as extreme as we could have thought.

Yes, science has been more present in the news in the first half of 2020. In our sample of French-language media in Canada, we find that the proportion of science-related articles grew by 20% compared to 2019. Yet, this proportion was already on the rise since 2017.

Yes, science was more frequently associated with government actions during the first wave of the pandemic (March to June 2020). The proportion of science-related paragraphs associated with the semantic field of "Government" grew by 18% compared to the three previous years.

Yes, more uncertainty has surrounded science during the first wave. The proportion of sentences conveying uncertainty in science-related paragraphs grew by 25% compared to the three previous years.

And yes, science has generated more engagement by the public. The geometric mean of engagement with science-related Facebook posts grew by 21% compared to the three previous years. Meanwhile the mean engagement with posts not related to science has barely budged (2% up). Yet, the science-related posts that have generated the most engagement exhibit the standard features of "viral" content on social media: focusing on more sensationalistic episodes than relevant news about the pandemic.

These changes are all of roughly the same magnitude. We can thus sum up our results by stating that the popular representation of science in French-language media in

Canada has leaped in the expected direction by approximately 20% during the first wave of the COVID-19 pandemic: more present, more associated with government actions, more uncertain and more engaging. We see two directions where research could build on our results: first to investigate whether they apply to journalistic discourse in different languages and countries; second to see how the patterns we measured evolve during the second wave of the pandemic, still ongoing at the time of writing.

References

Badenschier, Franziska, and Holger Wormer. 2012. "Issue Selection in Science Journalism: Towards a Special Theory of News Values for Science News?" In *The Sciences' Media Connection –Public Communication and Its Repercussions*, edited by Simone Rödder, Martina Franzen, and Peter Weingart, 28:59–85. Sociology of the Sciences Yearbook. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-2085-5_4.

Barel-Ben David, Yael, Erez S. Garty, and Ayelet Baram-Tsabari. 2020. "Can Scientists Fill the Science Journalism Void? Online Public Engagement with Science Stories Authored by Scientists." *PLoS ONE* 15 (1): 1–15. https://doi.org/10.1371/journal.pone.0222250.

Bridgman, Aengus, Eric Merkley, Peter John Loewen, Taylor Owen, Derek Ruths, Lisa Teichmann, and Oleg Zhilin. 2020. "The Causes and Consequences of COVID-19 Misperceptions: Understanding the Role of News and Social Media." *Harvard Kennedy School Misinformation Review*, June. https://doi.org/10.37016/mr-2020-028.

Bucchi, Massimiano, and Renato G. Mazzolini. 2003. "Big Science, Little News: Science Coverage in the Italian Daily Press, 1946-1997:" *Public Understanding of Science* 12 (1): 7–24. https://doi.org/10.1177/0963662503012001413.

Cole, Bruce J. 1975. "Trends in Science and Conflict Coverage in Four Metropolitan Newspapers." *Journalism Quarterly* 52 (3): 465–71. https://doi.org/10.1177/107769907505200309.

Dalloux, Clément. 2017. "Détection de l'incertitude et de La Négation : Un État de l'art." In *RECITAL 2017 - 18ème Rencontre Des Étudiants Chercheurs En Informatique En Traitement Automatique Des Langues*, 1–14. Actes de La Rencontre Des Jeunes Chercheurs En Traitement Automatique Des Langues, RECITAL, Joint à La Conférence TALN 2017. Orléans, France. https://hal.archives-ouvertes.fr/hal-01659646.

Dudo, Anthony. 2015. "Scientists, the Media, and the Public Communication of Science." *Sociology Compass* 9 (9): 761–75. https://doi.org/10.1111/soc4.12298.

Dunwoody, Sharon. 2008. "Science Journalism." In *Handbook of Public Communication of Science and Technology*, edited by Massimiano Bucchi and Brian Trench, 15–26. Routledge International Handbooks. London: Routledge.

Funk, Cary, Jeffrey Gottfried, and Amy Mitchell. 2017. "Science News and Information Today." Pew Research Center. https://www.journalism.org/wp-content/uploads/sites/8/2017/09/PJ_2017.09.20_S

cience-and-News_FINAL.pdf.

Gouvernement du Québec. 2020. "Décret Numéro 177-2020 Concernant Une Déclaration d'urgence Sanitaire Conformément à l'article 118 de La Loi Sur La Santé Publique." March 13, 2020. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/sante-services-sociaux/publ ications-adm/lois-reglements/decret-177-2020.pdf?1584224223.

Grabar, Natalia, Vincent Claveau, and Clément Dalloux. 2018. "Cas: French Corpus with Clinical Cases." In *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, 122–128.

Guo, Lei, Chris J. Vargo, Zixuan Pan, Weicong Ding, and Prakash Ishwar. 2016. "Big Social Data Analytics in Journalism and Mass Communication: Comparing Dictionary-Based Text Analysis and Unsupervised Topic Modeling." *Journalism & Mass Communication Quarterly* 93 (2): 332–59. https://doi.org/10.1177/1077699016639231.

Hervé, Nicolas. 2020. "Coronavirus - Étude de l'intensité médiatique." http://www.herve.name/research/ina_nherve_mediatisation_coronavirus_v1_7.pd f.

INSPQ. 2020. "Données COVID-19 Au Québec." *Institut National de La Santé Publique Du Québec* (blog). 2020. https://www.inspq.qc.ca/covid-19/donnees.

Jasanoff, Sheila. 1990. *The Fifth Branch: Science Advisers as Policymakers*. Cambridge, MA: Harvard University Press.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

Metcalfe, Jenni, and Toss Gascoigne. 1995. "Science Journalism in Australia." *Public Understanding of Science* 4 (4): 411–28. https://doi.org/10.1088/0963-6625/4/4/005.

Meyer, Philip. 1973. *Precision Journalism. A Reporter's Introduction to Social Science Methods.* Bloomington & London: Indiana University Press.

Nature. 2008. "Critical Journalism. Science Coverage Is on the Wane When Public Scrutiny of Science Is More Important than Ever." *Nature* 452 (7186): 387–88. https://doi.org/10.1038/452387b.

Newman, Nic. 2020. "Reuters Institute Digital News Report 2020." Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FIN AL.pdf.

Pellechia, Marianne G. 1997. "Trends in Science Coverage: A Content Analysis of Three US Newspapers." *Public Understanding of Science* 6 (1): 49–68. https://doi.org/10.1088/0963-6625/6/1/004.

Peng, Yifan, Xiaosong Wang, Le Lu, Mohammadhadi Bagheri, Ronald Summers, and Zhiyong Lu. 2018. "NegBio: A High-Performance Tool for Negation and Uncertainty Detection in Radiology Reports." *AMIA Summits on Translational Science Proceedings* 2018 (May): 188–96.

Quirion, Rémi. 2020. "La science en temps de pandémie." *La Presse*, April 25, 2020, sec. Opinions. https://www.lapresse.ca/debats/opinions/2020-04-25/la-science-en-temps-de-pan demie.

Rayson, P., Dawn Archer, S. Piao, and A. M. McEnery. 2004. "The UCREL Semantic Analysis System." In *Proceedings of the beyond Named Entity Recognition Semantic Labelling for NLP Tasks Workshop, Lisbon, Portugal, 2004*, 7–12. PRT. https://eprints.lancs.ac.uk/id/eprint/1783/.

Rocheleau, Sylvain. 2017. *Observatoire de La Circulation de l'information*. https://oci-ifo.org/en/home/.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton, N.J: Princeton University Press.

Schnabel, Ulrich. 2003. "God's Formula and Devil's Contribution: Science in the Press." *Public Understanding of Science* 12 (3): 255–59. https://doi.org/10.1177/0963662503123004.

Sharma, D.C., Abhishek Pathak, Rameshwar Nath Chaurasia, Deepika Joshi, Rajesh Kumar Singh, and Vijay Nath Mishra. 2020. "Fighting Infodemic: Need for Robust Health Journalism in India." *Diabetes & Metabolic Syndrome* 14 (5): 1445–47. https://doi.org/10.1016/j.dsx.2020.07.039.

Social Science One. 2019. "CrowdTangle Data at Facebook." *Social Science One* (blog). 2019. https://socialscience.one/rfp-crowdtangle.

Vraga, Emily K., Melissa Tully, and Leticia Bode. 2020. "Empowering Users to Respond to Misinformation about Covid-19." *Media and Communication* 8 (2): 475–79. https://doi.org/10.17645/mac.v8i2.3200.

West, Mark D. 2001. *Theory, Method, and Practice in Computer Content Analysis*. Greenwood Publishing Group.

WHO. 2020. "Archived: WHO Timeline - COVID-19." April 27, 2020. https://www.who.int/news/item/27-04-2020-who-timeline---covid-19.