Sudden Attention Shifts on Wikipedia During the COVID-19 Crisis

Manoel Horta Ribeiro,[1*] **Kristina Gligorić,**[1*] **Maxime Peyrard,**[1*]
Florian Lemmerich,[2] **Markus Strohmaier,**[3] **Robert West**[1]

[1]EPFL, [2]University of Passau, [3]RWTH Aachen & GESIS

{manoel.hortaribeiro, kristina.gligoric, maxime.peyrard, robert.west}@epfl.ch
florian.lemmerich@uni-passau.de, markus.strohmaier@cssh.rwth-aachen.de

Abstract

We study how the COVID-19 pandemic, alongside the severe mobility restrictions that ensued, has impacted information access on Wikipedia, the world's largest online encyclopedia. A longitudinal analysis that combines pageview statistics for 12 Wikipedia language editions with mobility reports published by Apple and Google reveals massive shifts in the volume and nature of information seeking patterns during the pandemic. Interestingly, while we observe a transient increase in Wikipedia's pageview volume following mobility restrictions, the nature of information sought was impacted more permanently. These changes are most pronounced for language editions associated with countries where the most severe mobility restrictions were implemented. We also find that articles belonging to different topics behaved differently; e.g., attention towards entertainment-related topics is lingering and even increasing, while the interest in health- and biology-related topics was either small or transient. Our results highlight the utility of Wikipedia for studying how the pandemic is affecting people's needs, interests, and concerns.

1 Introduction

The coronavirus disease 2019 (COVID-19) pandemic has led to the implementation of unprecedented non-pharmaceutical interventions ranging from case isolation to national lockdowns (Flaxman et al. 2020). These interventions, along with the disease itself, have created massive shifts in people's lives. For instance, in mid-May 2020, more than a third of the global population was under lockdown (Liverpool et al. 2020), and millions have since lost their jobs or have moved to work-from-home arrangements (Hale et al. 2020).

This scenario has fueled an unprecedented effort to better understand the disease and its spread (Cohen 2020), as well as to develop clinical treatments and vaccines (WHO 2020; Collins and Stoffels 2020). Yet, recent work suggests that the impact of the pandemic transcends health-related issues (Ryan et al. 2020). Thus, research should also identify how the pandemic has impacted human needs, interests, and concerns. After all, many challenges of the COVID-19 crisis did not originate directly from the virus, but from the social, economic, and psychological implications of the measures taken to prevent its spread (Bonaccorsi et al. 2020).
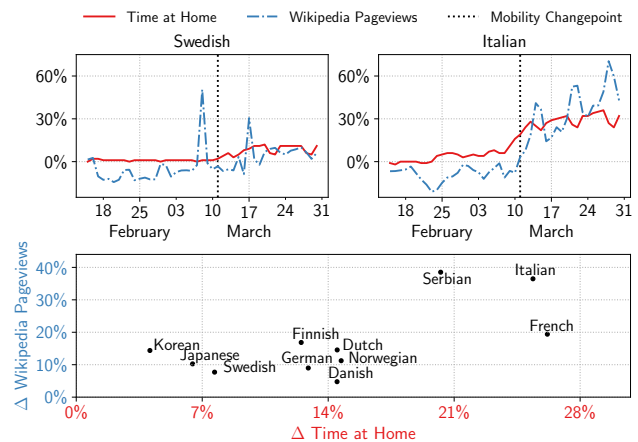
*These authors contributed equally.

Figure 1: Wikipedia access vs. mobility — Association between increase in time spent at home (from Google mobility reports; red) and increase in Wikipedia access volume (from Wikipedia pageview statistics; blue), both in terms of relative change over a five-week baseline period in early 2020. Top: time series for Sweden/Swedish Wikipedia and Italy/Italian Wikipedia; dotted vertical lines: changepoints in mobility time series. Bottom: summary for 11 of the 12 languages studied (excluding English); x-axis: post-minus-pre-changepoint difference in relative mobility change; y-axis: post-minus-pre-changepoint difference in total Wikipedia access volume (Pearson's $r = 0.63$, $p = 0.03$).

Quantifying the impact of COVID-19 is particularly challenging due to its global nature: traditional survey-based studies face the difficulty of *scale*, since participants ought to be spread throughout the globe, and of *immediacy*, since the pandemic is an ongoing crisis where needs, interests, and concerns are dynamic (Diaz et al. 2016; Salganik 2019).

Unlike most previous events that directly impacted so many lives around the world, the COVID-19 pandemic developed in a time of widespread Internet access. This digitization enables researchers to explore the impact of the pandemic across society by analyzing how it has impacted users' digital traces. For instance, Wikipedia, the world's largest encyclopedia and one of the most visited sites on the Web, captures rich digital traces from readers and makes them publicly available in aggregated form. Wikipedia is

used hundreds of millions of times each day to address a wide spectrum of information needs, ranging from getting information to make personal decisions to reading up on what was discussed in the media (Singer et al. 2017).

By analyzing Wikipedia access logs, we may thus study how the shifts in needs, interests, and concerns brought by COVID-19 have affected the readership of a key information source on the Web, while addressing the challenges of *scale* (since Wikipedia has versions for languages spoken all over the globe) and *immediacy* (since the "always-on" nature of usage logs allows us to analyze information seeking patterns at any time).

Present work. This work aims to elucidate how information access patterns across 12 Wikipedia languages editions have shifted during the current COVID-19 pandemic, in response both to the disease itself and to the massive mobility restrictions imposed by governments. More specifically, we ask:

- **RQ1** How has the COVID-19 crisis impacted information seeking behavior in Wikipedia?

- **RQ2** Have changes in information seeking behavior lasted after mobility restrictions were lifted?

We argue that a longitudinal understanding of information seeking through the lens of Wikipedia is a starting point to measure the impact of the pandemic and its associated interventions on information seeking behavior. Examining the differences in the *volume* (i.e., how much?) and in the *nature* (i.e., what?) of information seeking may shed light not only on the needs, interests, behaviors, and concerns of people throughout the crisis, but also on how the pandemic is shaping society in the long run. Also, observing what changes remained after mobility returned to normality can prove useful to disentangle effects due to the lockdowns from the drifts created by the pandemic itself.

Methods and data. Our analyses combine Wikipedia access logs, mobility reports made available by Apple and Google, and information on the dates of non-pharmaceutical interventions. We enrich the Wikipedia access logs in different ways, categorizing articles as COVID-19-related or not, attributing articles to a series of predefined "topics", and calculating how distant from normal is the distribution of views over articles. Also, we calculate relevant points in time related to mobility change, capturing, for different countries, when mobility suddenly decreased and when it returned to normality. Methodologically, drawing meaningful conclusions from the longitudinal Wikipedia access logs is challenging due to the presence of trends and seasonalities. We overcome these hurdles by performing careful difference-in-differences analyses in a regression framework. We detail our methods and the data employed in Sec. 3.

Findings. Our analysis of the number of pageviews over time (Sec. 4.1) shows that there was a sharp increase in the *volume* of Wikipedia usage, by up to 40%, following the sudden decrease in mobility induced by non-pharmaceutical interventions (cf. Fig. 1). This increase, however, was transient: our careful differences-in-differences study suggests that in the long run, there was either a decrease or no significant increase for 11 out of the 12 languages.

Considering not the volume, but the *nature* of information seeking in Wikipedia (Sec. 4.2), as measured through our "distance from normality" metric, we again observe the same patterns: a sharp increase following mobility restrictions, followed by a decrease after mobility returns to normality. Yet, here, our differences-in-differences study suggests that part of the increase remained, suggesting that the impact of the crisis on information seeking behavior persisted beyond mobility restrictions.

Lastly, we zoom in on what kinds of articles received more or less attention throughout the pandemic (Sec. 4.3). We find that some topics (e.g., BIOLOGY) saw transient increases or decreases in volume, while others saw persistent increases (e.g., VIDEO GAMES) or decreases (e.g., SPORT). Overall, surprisingly, many of the topics with persistent increases relate not to basic needs related to the pandemic, but to entertainment and self-actualization.

We conclude by discussing our findings in the light of the research questions and in the overall context of assessing the social impact of the pandemic (Sec. 5). We make available all the code and the data necessary to reproduce our results at: www.github.com/epfl-dlab/wiki_pageviews_covid.

2 Related Work

We review related work that focused on the COVID-19 infodemic, work that measured changes in user behavior during the pandemic, and, lastly, other research that studied Wikipedia in the general context of crises and pandemics.

2.1 The COVID-19 Infodemic

COVID-19 was the first event of its magnitude to take place in the era of social media and of user-generated content. The important role that platforms such as Twitter and Facebook have in today's society has prompted researchers to study patterns of virality and of information sharing in social media during the pandemic.

Kouzy et al. (2020) studied the spread of COVID-19-related misinformation on Twitter by analyzing a sample of tweets ($n = 673$), collected in February 2020, with trending hashtags and keywords related to the disease, finding a high prevalence of mis- or unverifiable information. Other studies that followed have found similar and complementary results. Yang, Torres-Lugo, and Menczer (2020) also found a high prevalence of low-credibility information, which is disproportionally spread by bots. Ahmed et al. (2020) identified and analyzed the drivers behind one of the main COVID-19-related conspiracies, which postulates the spread of the infection to be related to the 5G standard for cellular networks. Analyzing data containing conspiracy-related hashtags, they found that a handful of users were driving the conspiratory content and that many of those using the hashtag were denouncing the conspiracy theory.

Whereas these studies detected and measured content that may impact behaviors and beliefs (e.g., conspiracy theories), we focus on the inverse question: *Can we uncover changes in behavior via digital traces?* We argue that both these directions are important to the overarching goal of understanding the social dynamics of the spread of the virus (Depoux

et al. 2020), since to understand the impact of misinformation, one must be capable of measuring people's needs, interests, and concerns.

2.2 The Web in Times of COVID-19

Recent work has also more broadly characterized how COVID-19 has altered people's online behavior, and how digital traces can be used to better understand the pandemic.

A first and broader theme in this direction concerns how COVID-19 has increased Internet traffic. Feldmann et al. (2021) showed that, as a result of COVID-19-induced lockdowns, Internet traffic of residential users increased by 15–20%. Traffic increases were observed in applications that people use when at home, such as Web conferencing, VPN, gaming, and messaging. Results in the same direction were also found in survey-based studies analyzing Internet time (Colley, Bushnik, and Langlois 2020), and by a smaller-scale study measuring the increase in the stress on a campus network (Favale et al. 2020).

Second, and more related to the work at hand, are works leveraging digital traces to understand the impact of COVID-19 on mental health, economy, society, and human needs (Abay, Tafere, and Woldemichael 2020; Gupta et al. 2020; Tubadji, Boy, and Webber 2020). We highlight two recent papers based on search data. Lin, Liu, and Chiu (2020) used Google search data on COVID-19-specific keywords to predict the speed of the spread of the disease. They found, e.g., that searches for "*wash hands*" are correlated with a lower spreading speed of the disease. Suh et al. (2020) measured changes in human needs using Bing search logs. They found that, for a variety of different "need categories", there was elevated increase in search activity, and that subcategories related to the most basic needs received the largest boost.

The outlined related work is complementary to the study at hand. Having signals from multiple distinct digital traces such as network usage, social media activity, and search logs may allow stakeholders to paint a more comprehensive picture of how the pandemic has impacted society.

2.3 Wikipedia in Times of Crisis

Finally, our work extends a rich literature studying behavioral changes of Wikipedia readers during unexpected events, crises, and catastrophes (García-Gavilanes et al. 2017; Zhang et al. 2019). In work that is most closely related, researchers used Wikipedia pageviews in order to monitor and forecast diseases at a global scale (Generous et al 2014a,b) and to study anxiety and information seeking about infectious diseases, such as influenza (McIver and Brownstein 2014), H1N1 (Tausczik et al. 2012), and Zika (Tizzoni et al. 2020).

Resonating with existing work, our results highlight the integral role played by Wikipedia in times of crisis and its usefulness to understand how these events impacted the behavior of its readers. Moreover, it is worth noting that the difference-in-differences methodology we employ here may also be of use for future research aimed at understanding how a given crisis has impacted user behavior while accounting for seasonal effects.

Language	# articles	2019 pageviews	2020 pageviews
English	6,047,509	83,566,105,101	51,911,047,562
Japanese	1,197,856	12,335,323,771	7,829,206,874
German	2,415,136	10,090,208,904	6,083,708,597
French	2,195,949	7,663,315,198	4,867,880,748
Italian	1,594,039	5,996,763,417	3,945,321,500
Dutch	2,003,807	1,678,509,656	1,010,487,782
Swedish	3,735,720	1,019,647,051	591,716,448
Korean	490,314	837,989,910	499,911,756
Finnish	481,854	665,104,786	399,430,747
Norwegian	531,478	359,423,070	206,080,079
Danish	258,063	309,196,927	177,158,005
Serbian	632,128	274,704,611	206,425,442

Table 1: Basic dataset statistics: the number of articles in each Wikipedia version, and the number pageviews received in 2019 (full year) and 2020 (January through July).

3 Materials and Methods

3.1 Data

The analyses of this paper combine information about the content and usage of Wikipedia with information about the progression of the pandemic.

Wikipedia. We selected 12 languages that have Wikipedia editions of various sizes. When choosing language editions, we considered (1) the size of the edition, (2) whether the language was spoken in relatively few countries, (3) the kinds of non-pharmaceutical interventions imposed in those countries. The overarching goal in selecting the languages was to have relevant language editions representing different attitudes towards the crisis, preferably from languages easily tied to one country. For each language edition, we obtained publicly available pageview statistics,[1] which specify how frequently each article was accessed daily between 1 January 2018 and 31 July 2020, aggregating both the desktop and the mobile versions of the site. The languages selected as well as basic statistics for their Wikipedia editions are listed in Table 1.

Enriching Wikipedia articles. To gain further insight into which kinds of articles people are reading, we enriched our pageview data with two additional sources. First, we labeled each article with one of 57 topics based on the ORES `article topic` model.[2] We collected the topic predictions (one page can belong to multiple topics) from the model for each of the articles in English (the language for which the model was developed) and linked articles across the 12 languages studied. Second, we determined for each article if it was related to the pandemic.[3] Analyzing such articles separately allows us to disentangle the changes in Wikipedia usage due to seeking information about the COVID-19 pandemic from other, perhaps less obvious patterns.

---

[1]https://dumps.wikimedia.org/other/pageviews/

[2]https://www.mediawiki.org/wiki/ORES/Articletopic

[3]Based on the article list in https://covid-data.wmflabs.org. The number of COVID-19-related articles varied substantially across language editions; e.g., English had over 300, Swedish had only 9.
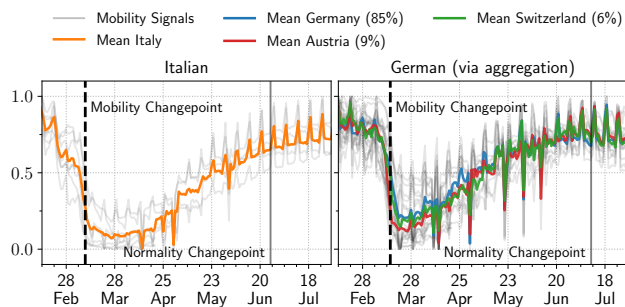
Figure 2: Robust detection of mobility and normality changepoints from Google and Apple mobility reports — The reports specify by what percentage the time spent at various location types has changed, compared to a baseline period in January 2020. We use time series for all location types, with multiple changepoint detection algorithms and varying hyperparameters. Given the abruptness of the decrease in most countries, different runs largely agree; we use the average as a robust changepoint. For languages spoken in multiple languages, e.g., German (right), we average mobility time series from the main countries where the language is spoken (here three), weighted according to the proportion of native speakers (details in Appendix A).

Pandemic timeline. Nine of the 12 languages are primarily spoken in a single country. For each country, we used Wikipedia to manually determine five days of particular interest in the context of COVID-19: first reported case, first death, ban of public events, school closure, lockdown. Two problems with employing the aforementioned days of interest in statistical analyses are (1) that it is not guaranteed that they would impact movement patterns across different countries homogeneously (e.g., it could be that for some of the countries people stayed more at home even before the lockdown was enacted), and (2) that not all days of interest were observed for the countries of interest (e.g., in Sweden there has been no country-wide lockdown by 31 July 2020).

Mobility reports. To address these limitations, we turn to daily mobility reports published by Apple and Google,[4] which capture population-wide movement patterns based on cellphone location signals. The mobility reports specify, for each day, by what percentage the time spent in various location types (e.g., residential areas, workplaces, retail and recreation, etc.) differed from a pre-pandemic baseline period in early 2020. Both government-mandated lockdowns as well as self-motivated social distancing measures manifest themselves as sharp changes in the mobility time series, which we detect automatically using changepoint detection algorithms (Fig. 2, left; details in Appendix A). We henceforth refer to these points as *mobility changepoints.* We use mobility changepoints as heuristic dates for when people started spending substantially more time in their homes.

---

[4]https://www.(apple|google).com/covid19/mobility/. We use country-wise mobility data from January to the end of July for all countries that have the languages chosen as official languages (see Appendix A).

Unlike choosing one of the days of interest, this leads to a meaningful "treatment" across different countries.

Three of the 12 languages are spoken more widely than in one predominant country: English, German, and French. As Wikipedia pageview statistics are available only at the language level, not at the country level, we determined a mobility changepoint for these language editions by aggregating mobility reports for the countries in which the language is official (Fig. 2, right; details in Appendix A).

We emphasize that, since any Wikipedia edition can be accessed from anywhere (barring censorship), the link between Wikipedia language editions and countries is merely approximate, even for languages that are official in only a single country. This should be kept in mind when interpreting our results, especially for the English edition, which is read widely across the globe.

Additionally, we calculated another important set of "changepoints", which do not represent abrupt changes in mobility patterns, but rather corresponds to the times when mobility returned to normal. We refer to these as *normality changepoints* (also depicted in Fig. 2). These times were also calculated with mobility data, by identifying the time when the future average mobility remained within a 10% band around the respective baseline level (defined as the pre-pandemic mobility levels by Google and Apple). In cases of languages spoken in multiple countries, we maintained the same aggregation scheme as for *mobility changepoints*.

3.2 Difference-in-differences

We now introduce our study design. Difference-in-differences regression is a "quasi-experimental" technique that mimics an experimental design with observational data by studying the effect of a treatment (e.g., mobility changes) on a treatment group vs. a control group (Angrist and Imbens 1995). The specific regression models are discussed as we present results, whereas here we focus on explaining and justifying our setup, illustrated on the left-hand side of Fig. 4.

The difference-in-differences method aims at separating the true treatment effect from simultaneous (e.g., seasonal) changes that would have occurred even without the treatment. To do so, we calculate the post-minus-pre-treatment difference (in 2020) and compare it to the difference between the corresponding time periods in the previous year (2019). Subtracting the 2019 difference from the 2020 difference (yielding the "difference in differences") thus removes changes that would have occurred even without the intervention (assuming 2020 would otherwise have looked like 2019) and gives a better estimate of the treatment effect.

In the first scenario, Fig. 4(a), we compare the difference in activity five weeks *before* vs. five weeks *after the mobility changepoint.* The idea here is to capture what changes were introduced by the sudden halts in human mobility induced by non-phamaceutical interventions. In the second scenario, Fig. 4(b), we compare the difference five weeks *after the mobility changepoint* vs. five weeks *after the normality changepoint.* Here, we capture how information seeking patterns changed as mobility restrictions waned. The third scenario, Fig. 4(c), compares the difference in activity between two scenarios where mobility was similar in 2020: five weeks

before the mobility changepoint and five weeks *after the normality changepoint.* Lastly the fourth scenario, Fig. 4(d), is an adaptation of the first for events other than mobility (e.g., school closure or official lockdown declaration).

Notice that for some languages there were less than five weeks in the data after the normality changepoint. In these cases we used all data up to 31 July 2020 (and the same period of 2019).

3.3 Distance from Normality

We introduce a notion of "distance from normality" as follows. On each day, the pageviews in a given language edition form a distribution over articles, characterizing how users' attention was distributed. We represent each daily distribution as an "attention vector" with one entry per article and entries summing to 1.

With over 6 million Wikipedia articles, many of which are rarely visited, attention vectors are large and noisy. Therefore, we first applied principal component analysis (PCA) in order to project attention vectors into a low-dimensional subspace. In the subspace, two attention vectors are naturally compared via their Euclidean distance.

The notion of "normal" attention is captured by the average attention vector over all days of 2019, i.e., well before the pandemic; and for each subsequent day, the distance from normality is given by the Euclidean distance of that day's attention vector from the average attention vector. Notice that we calculated the attention vector separately for each Wikipedia language edition using all its articles and then performed the dimensionality reduction.

We use this metric in addition to the aforementioned topics to understand overall changes in the information seeking patterns of Wikipedia users.

4 Results

4.1 Shifts in Overall Pageview Volume

We begin by examining the *volume* of information seeking throughout the pandemic. To do so, we leverage the count of pageviews across all articles for each of the 12 languages. To distinguish between an increase in information seeking about COVID-19-related articles vs. the remainder of Wikipedia, we analyze the two sets of articles separately.

Pageview trends in COVID-19-related articles. Wikipedia has been shown to be an accurate and up-to-date source of COVID-19-related information,[5] and we start by investigating how heavily this information was accessed by users. Fig. 3 (column 1) tracks the popularity of COVID-19-related articles (as a fraction of all pageviews) in all 12 languages over the course of the pandemic, from 14 January (the earliest time for which mobility reports are available) to 31 July 2020. Vertical lines in the figure represent the dates associated with noteworthy events (e.g., first case), as well as non-pharmaceutical interventions (e.g., lockdowns).

In nearly all languages, the share of COVID-19-related pageviews increased up to the mobility changepoint (dashed vertical line), from where their daily pageviews tended to

slightly decrease.[6] The eventual return to normal levels of mobility did not seem to have any sharp impact on the pageview share of these articles. We emphasize that these time series are plotted on logarithmic *y*-scales, such that a linear slope in the plots, even if small, corresponds to an exponential rate of change.

COVID-19-related articles were generally among the most popular during the period of study; e.g., 12 of the 15 most accessed articles in the English version were related to the pandemic. In some languages, the fraction of pageviews going to COVID-19-related articles surpassed 2% on some days, a considerable share of Wikipedia's overall volume, considering that all Wikipedia editions considered here have over half a million articles (Table 1).

Pageview trends in non-COVID-19-related articles. Next, we focus on Wikipedia articles *not* related to COVID-19, the vast majority. Their popularity during the period of study, in terms of the daily total number of pageviews, is shown as solid lines in Fig. 3 (column 2; linear *y*-scales). The dotted lines correspond to the same period in 2019, precisely one year earlier.[7] Notice that, in this plot, the *y*-scales vary per language edition, since their sizes differ substantially.

A clear pattern emerges in many language editions. Access volumes in 2020 closely mirrored those in 2019 up to the mobility changepoint associated with the respective language. Thereafter, the access volumes of 2020 began to rise up and above those of 2019 for nearly all languages. The trend was particularly strong for languages spoken in the countries with the most severe lockdown measures (e.g., Serbia, Italy, France), and weakest for languages spoken where mobility-related interventions were weaker than elsewhere (e.g., Japan, Scandinavia). Finally, once normality was approximately restored (as indicated by the normality changepoint), we see that the pageview volume again aligned itself with the quantities observed in the previous year. Again, for countries with weaker lockdowns, the "back to normality" effect is less noticeable since the pageview counts never deviated substantially from the baseline pattern to begin with.

We further explore this relationship in Fig. 3 (column 3) via cumulative plots, where the daily 2020-minus-2019 difference is accumulated from 14 January onward. Since languages have vastly different access volumes overall (cf. column 2), and to have a common scale across languages, we express the cumulative difference in terms of multiples of the average daily pageview volume attained in 2019 by the respective language. The increase is dramatic for some languages; e.g., the Serbian Wikipedia edition experienced about 57 days' worth of surplus access volume between mid-March and late July 2020, corresponding roughly to a

---

[5]https://wikimediafoundation.org/covid19/

[6]Some languages saw extreme upticks on certain days. Most of them can be linked to the creation of important COVID-19-related articles; e.g., the Swedish article CORONAVIRUSUTBROTTET 2020 I SVERIGE (English: COVID-19 PANDEMIC IN SWEDEN) was created on 22 March 2020 and immediately received wide attention.

[7]Since Wikipedia access volumes tend to follow a weekly periodicity, our alignment was manually adjusted to ascertain that matched days correspond to the same day of the week.
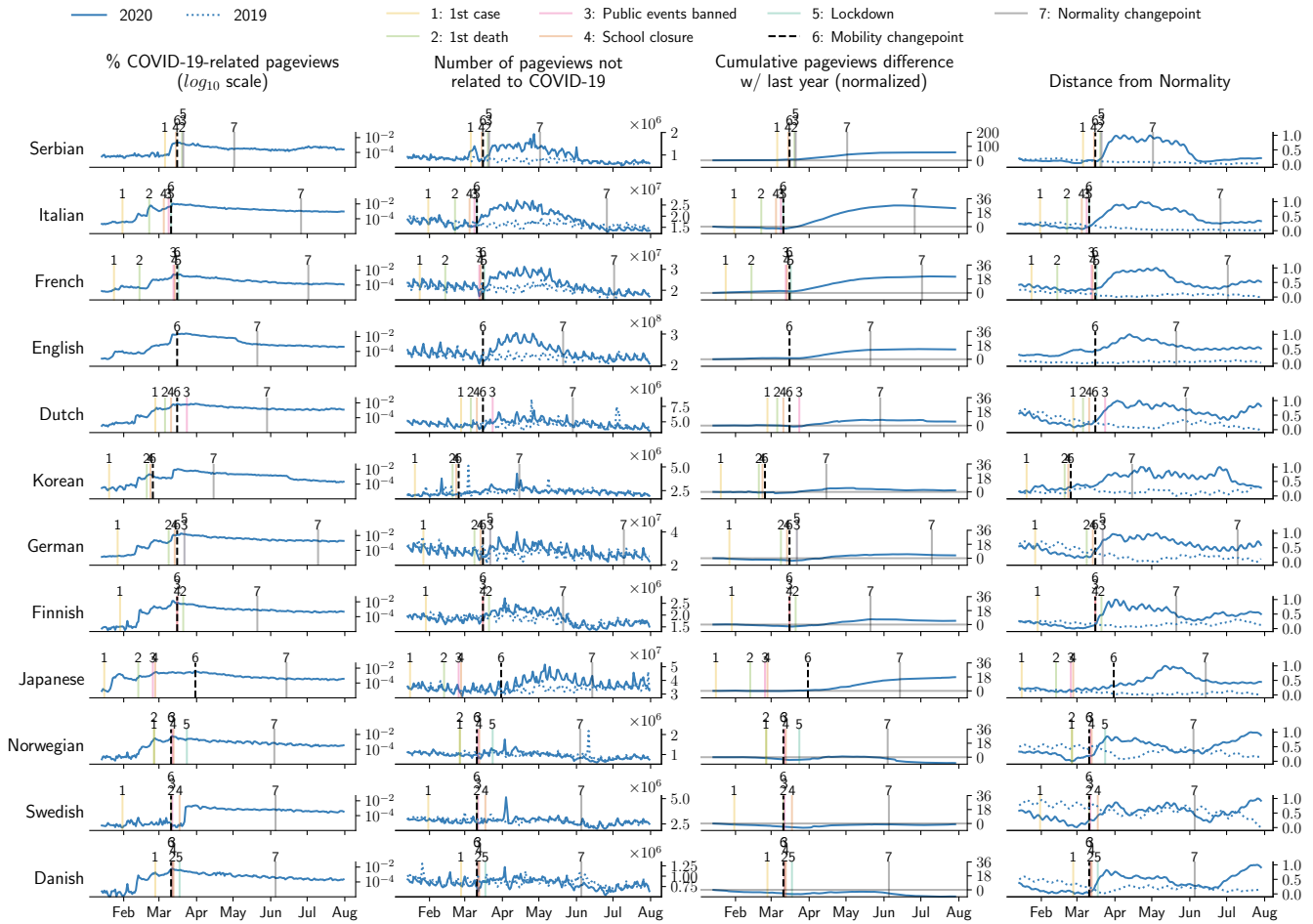
Figure 3: Evolution of Wikipedia access for 12 language — Column 1: percentage of pageviews spent on COVID-19-related pages (logarithmic scale). Column 2: total number of pageviews for all other pages (linear scales); solid: 14 January to 31 July 2020; dotted: corresponding period in 2019. Column 3: cumulative difference in total number of pageviews between study period of 2020 and corresponding period of 2019, reported in terms of multiples of average daily number of pageviews in 2019 (y-axes identical across languages, except for Serbian, where cumulative volume is much higher). Column 4: "distance from normality" with respect to topical attention (cf. Sec. 4.2) for 2020 (solid) and 2019 (dotted). All plots show, as vertical lines, mobility changepoints and pandemic-related events.

tripling in access volume, compared to the previous year. Strong effects are also observed for Italian (27 surplus days at its peak), French (21 days), Japanese (17 days), English (13 days), Dutch (7 days), and Finnish (7 days). As mobility returned to normality, the cumulative gain stabilized (e.g., Serbian, French) or even decreased (e.g., Dutch, Italian).

Note that eight of the 12 languages (Italian, Japanese, Finnish, Dutch, German, Korean, Danish, Serbian) initially ran a deficit, with Wikipedia being visited *less* in 2020 compared to the corresponding days in 2019 (reflected as negative values in the plots of column 3 of Fig. 3), but all except three languages (Norwegian, Swedish, Danish) recovered and eventually ran a surplus by the end of the study period.

Difference-in-differences regression. In order to go beyond visual inspection and to precisely quantify the shifts in pageview volume, we take a regression-based difference-in-differences approach, as described in Sec. 3.2.

In this setup, we consider, for each language, a time window of 10 weeks (70 days) split around either the mobility or the normality changepoint in 2020, as well as the corresponding time window in 2019. Each of these 140 days contributes one data point per language, for a total of $140 \times 12 = 1,680$ data points. As the dependent variable y, we use the logarithm of the number of pageviews, and as independent variables, the following three factors: year (2019 or 2020), period (before or after calendar day of mobility changepoint), language. We now model y as a linear function of these three factors and all their two- and three-way interactions. In R formula notation,

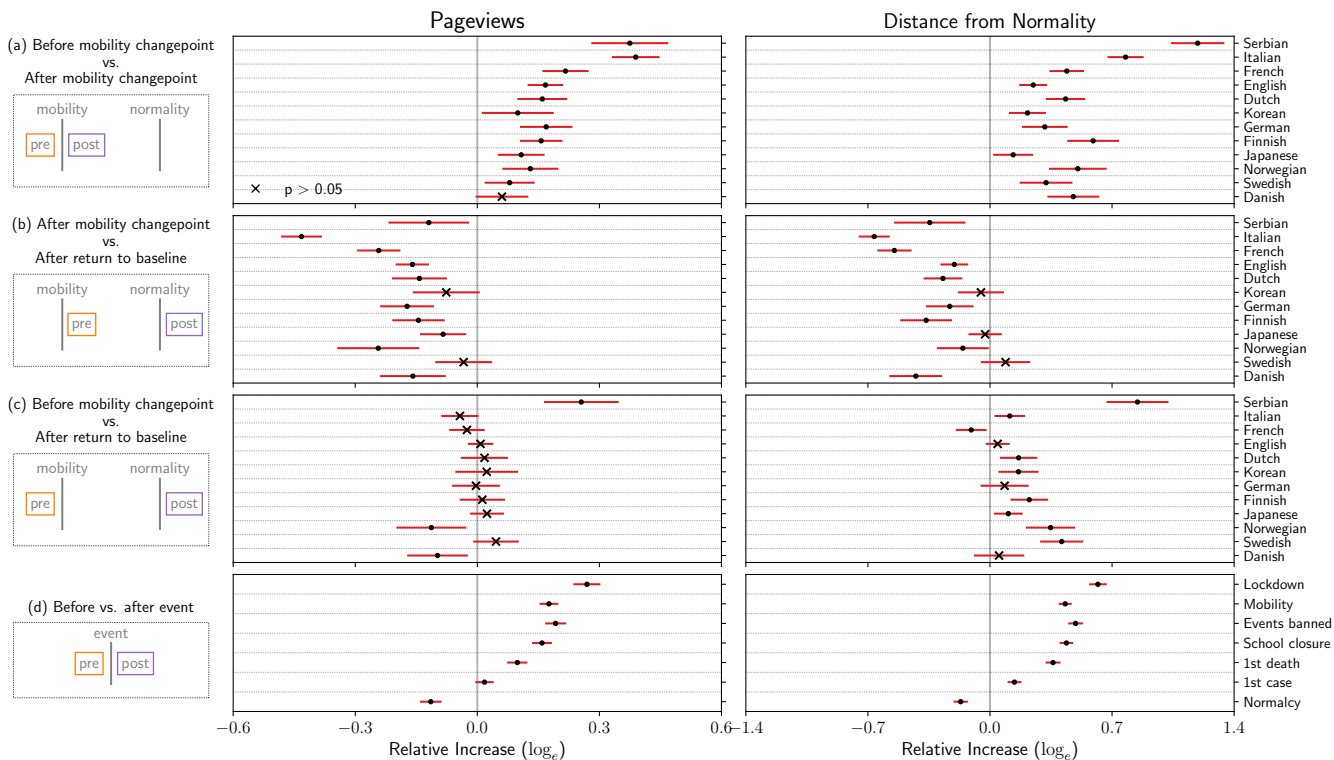$$y \sim \texttt{year} * \texttt{period} * \texttt{language}. \tag{1}$$

213

Figure 4: Estimated effects of restricted mobility in pageviews and in attention — For the 12 studied language editions, we depict the results of difference-in-differences for estimating effects of changepoints non-pharmaceutical intervention in the total number of pageviews (on the left) and in the distance from normality metric (on the right). Effects are shown for different diffs-in-diffs setups illustrated in diagrams in the left-hand side (see Sec. 3.2 for details). Error bars represent 95% CIs approximated as 2 standard errors. Dependent variables are used in logarithmic form, such that exponentiated coefficients capture the multiplicative increase due to the treatment. ($R^2 > 0.95$ for all models.)

Here, $a * b$ is shorthand notation for $a + b + a : b$, where in turn $a : b$ stands for the interaction of a and b.

Pageview volumes were considered in logarithmic form for two reasons: first, because raw pageview counts are far from normally distributed, with numerous large outliers, and second, because the logarithm makes the model multiplicative, thus implicitly normalizing the estimated effects and making it possible to compare languages with different pageview volumes: if b is the coefficient of the three-way interaction $year : period : language$, then e^b captures the multiplicative factor by which pageview volumes increased when mobility dropped, after accounting for differences stemming from the year alone or the period alone,[8] which are already captured by the coefficients of $year : language$ and $period : language$, respectively.

The estimated logarithmic effects are plotted for all languages in Fig. 4(a) (column 1). The results confirm the effects observed visually in Fig. 3 (column 3); e.g., the logarithmic pre-vs.-post mobility changepoint effect on the Ital-

ian version is around 0.38 (corresponding to an increase in pageviews to $e^{0.38} \approx 146\%$). For language editions where the effect was less pronounced in the visual analysis, we also find smaller (e.g., for Swedish the increase was to approximately 108%) or insignificant (e.g., Danish) effects. Another set of interesting cases are languages for which the increase in pageviews was substantial but more gradual, e.g., Japanese. Although Japanese saw a larger cumulative pageview difference than English (Fig. 3 column 3), here its effect is smaller (0.10 for Japanese vs. 0.16 for English).

In a similar fashion, Fig. 4(b) (column 1) shows the same analysis, but now comparing the five-week period *after the mobility changepoint* with the (up-to-)five-week period *after the normality change point*. Interestingly, the results of this analysis are, to a large extent, diametrically opposite of those of the above analysis. Whereas there (around the mobility changepoint), Serbian and Italian saw the biggest increases in pageview volume, here (around the normality changepoint), they saw the biggest decreases. For instance, the logarithmic effect is -0.43 for Italian (a multiplicative decrease to $e^{-0.43} \approx 64\%$ of the original value).

In Fig. 4(c) (column 1), we fit yet another model that shares the pre-intervention with the setup in Fig. 4(a) and the post-intervention period with Fig. 4(b). Its purpose is

---

[8]For instance (cf. Fig. 3, column 2), Swedish pre-mid-March pageviews were consistently lower in 2020 than in 2019; and Finnish post-mid-March pageviews were higher than pre-mid-March pageviews even in 2019, without the pandemic.

to determine long-term effects, i.e., whether there were any changes between the periods before the mobility changepoint and after the normality changepoint. From Fig. 4(c) (column 1), we observe that most languages saw no long-term effects, whereas two (Norwegian and Danish) saw a significant negative long-term effect, and only Italian saw a significant positive long-term effect.

Lastly, in the fourth and last setup, we fit a slightly different model:

$$y \sim \texttt{year} * \texttt{period} + \texttt{language}, \qquad (2)$$

where the dependent variable y is again the logarithm of the number of pageviews, `language` now merely serves as a language-specific baseline, and `year : period` is the only interaction term which, captures the estimated effect in a language-independent way. We fit this model not only for the case where `period` is defined by the mobility changepoint, but also where it is defined by the five other pandemic-related events. This way, we may compare effect sizes for the various events when considered as "treatments".

The estimated logarithmic effects are plotted for each event in Fig. 4(d) (column 1). We find that events that are more tightly related to decreased mobility are associated with the largest increases in pageviews. For example, the pre-vs.-post first-death effect is only around 0.09 ($e^{0.10} \approx$ 110%); for school closure, it is around 0.16 ($e^{0.16} \approx$ 117%); and for the actual lockdown, it is largest, at 0.27 ($e^{0.27} \approx$ 131%). Moreover, as expected from the previous results, we also find that the effect around the normality changepoint is significantly negative, at -0.11 ($e^{-0.11} \approx$ 89%).

This corroborates the previous results on mobility and Wikipedia access from yet another angle. Earlier, in Fig. 3 and Fig. 4(a), we showed that languages spoken in countries with a stricter lockdown saw a larger pageview increase, whereas here we showed that not only does this hold when considering all languages together, but also that events associated with a mobility decrease are significantly more associated with an increase in pageviews than other pandemic-related events.

4.2 Shifts in Information Seeking Patterns

So far, we observed an overall increase in pageviews following mobility restrictions. Then, once mobility went back to normal, so did pageviews. Next, we explore whether these shifts affected pageviews evenly across articles or whether users' attention shifted from certain topics to others.

To gain a better understanding, we analyze the temporal trends for the distance-from-normality metric (defined in Sec. 3.3) and perform a difference-in-differences analysis in the same setup as in Sec. 4.1.

Distance from normality. Distances from normality are plotted as time series in Fig. 3 (column 4). We also plot a baseline time series computed on data from exactly one year earlier (dotted line), where normality is defined as the average attention vector of 2018, and distance from normality was computed for mid-January to the end of September 2019. We observe that the (dotted) 2019 baseline curves stay

flat around the mobility changepoints. If additional pageviews gained after the changepoints mirrored the average attention distribution, one would expect equally flat curves for 2020. Quite on the contrary, however, the curves for 2020 increase sharply at the mobility changepoints, a clear indication of large topical shifts in overall attention. Interestingly, we find that these changes occur even for languages for which there was no substantial pageview increase, such as Danish, Swedish, and Norwegian. This suggests that, although these languages were not reading more, their information seeking patterns still differed overall.

As mobility returned to normality, we find that Wikipedia readers' attention also shifted back towards normality. That is, the distance from normality became smaller after the effect of mobility restrictions waned. Yet, it is important to notice that the level at which the distance from normality settled thereafter was generally larger than in the previous year (e.g., 0.3 vs. 0.03 for Italian). Also, for some language editions (Dutch and Scandinavian languages), we see a second sudden increase in the metric around mid-July, for which we found no obvious explanation.

Difference-in-differences regression. To quantify the attention shifts more objectively, we perform a difference-in-differences regression analogous to the one of Sec. 4.1 (Eq. 1), but this time with the outcome y being the logarithm of the distance from normality, rather than of the pageview volume. We again employ the setups previously described and show the interaction terms on the right-hand side of Fig. 4. Overall, in the two first setups (Fig. 4(a-b), column 2), we see results similar to those obtained in the pageview analysis. For all languages, there was a sharp increase in the distance from normality when comparing the periods before vs. after the mobility changepoint. Comparing the period after the mobility changepoint with the period after the normality changepoint, we observe a significant decrease. Results are also similar to what we observed for pageviews in the fourth scenario, (Fig. 4(d)), where we compare "before vs. after" for several pandemic-related events. Here again, we see a gradation, where the regressions for events more closely associated with mobility decrease (e.g., lockdowns) are associated with stronger effects.

Interestingly, the picture in the third scenario (Fig. 4(c)) differs from the pageview analysis: here the distance from normality remains higher than baseline for eight of the 12 languages. This scenario compares the period before the mobility changepoint with the period after the normality changepoint. This indicates that attention shifted massively during the mobility restrictions but settled in a new state after mobility went back to normal.

This suggests that a portion of the effect was transient and clearly associated with the mobility decrease, but that there was also another part that lingered months after mobility decreases waned. A plausible explanation for this phenomenon could be that the societal impact of the pandemic and its associated non-pharmaceutical interventions have altered the interests, needs, and concerns of people for good.
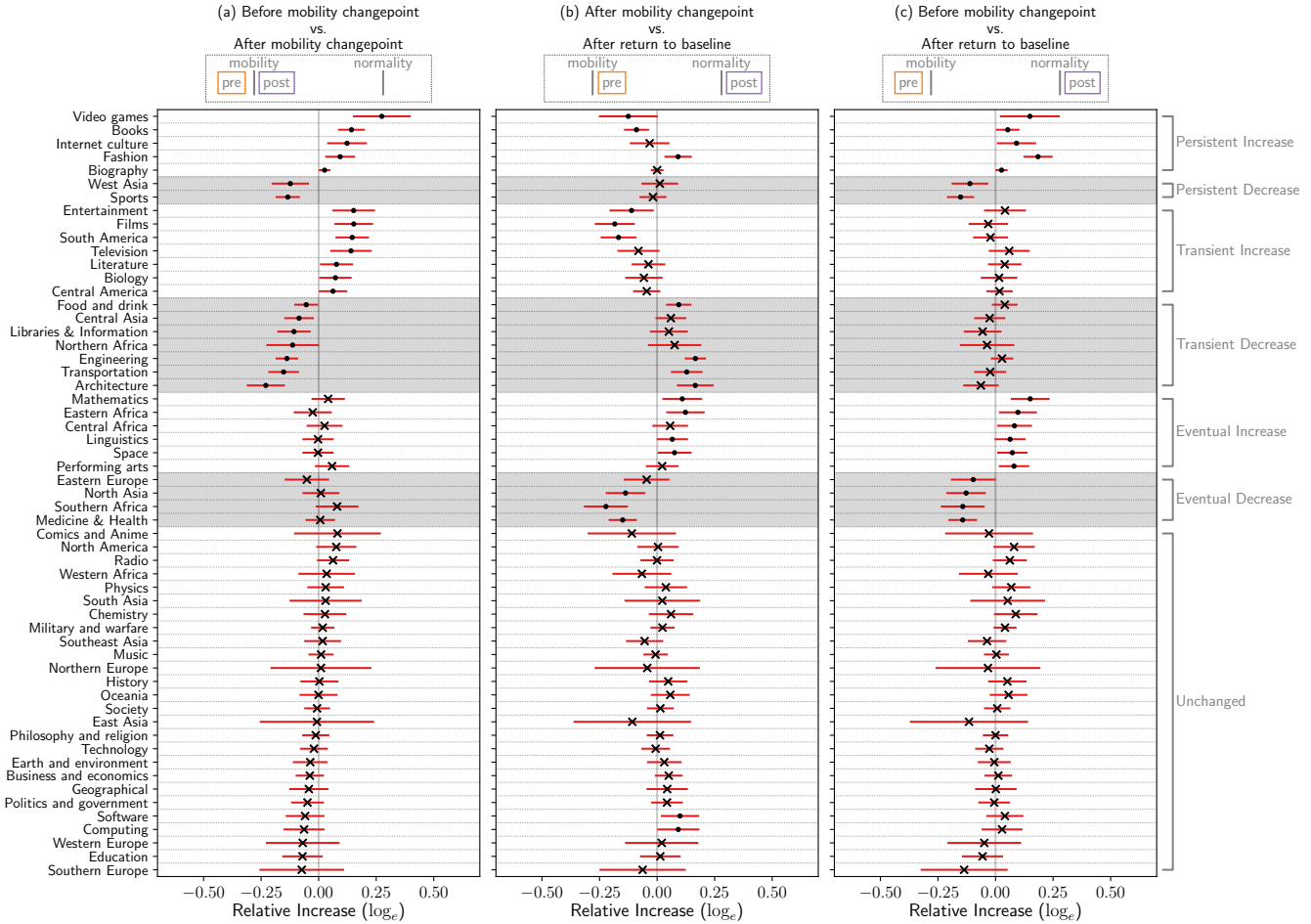
Figure 5: Shifts in topic-specific pageview volumes — Effect of mobility decrease (and eventual return to normalcy) on relative pageview share of 57 topics, estimated via difference-in-differences regression, pooled across 12 languages. Again we use three different setups in order to capture how things changed through different moments of the pandemic (see Sec. 3.2 for details). Topics are grouped according to their behaved given changes in human mobility, as shown in the brackets in the right-handside. Error bars represent 95% CIs. ($R^2 > 0.85$ for all models.)

4.3 Shifts in Topic-specific Pageview Volume

So far, we observed a massive increase in pageview volume following the sudden mobility change induced by COVID-19-related interventions, which was not explained by a simple proportional increase according to the prior attention distribution (Sec. 4.1). Rather, the increase in pageview volume was accompanied by major shifts in attention (Sec. 4.2). Moreover, we have also found that, although the shifts in pageview volume were reverted once mobility restrictions waned, this was not the same for the shifts in attention: information seeking patterns seem to have been impacted in the long run. Next, we investigate the shift further, with the goal of identifying which topics gained, and which lost, attention, and also which of these changes remained after mobility reverted to normality.

Again, we account for trends and seasonality by formulating difference-in-differences regression models analogous to that of Eq. 1, but we now use a model without a language

term and with an added topic term (for notation, cf. Eq. 1):

$$y \sim \texttt{year} * \texttt{period} * \texttt{topic}. \qquad (3)$$

We code the 57 topics as 56 dummy variables, with one arbitrary topic serving as a baseline topic. This amounts to 95,760 data points, one for each combination of year, calendar day (35 days before as well as after), topic, and language. The outcome variable y of interest is the logarithm of the fraction of pageviews going to articles of the respective topic on the respective day. We again fit models for different setups (here we consider only the first three scenarios discussed in Sec. 3.2). Summing the coefficient of $\texttt{year}:\texttt{period}:\texttt{topic}$ with the coefficient of $\texttt{year}:\texttt{period}$ (corresponding to the baseline topic) reveals changes in topical interest around the changepoints of interest. Given the three distinct difference-in-differences estimations, we are able to categorize different topics according to how they behaved given changes in human mobility: Did they change? Did these changes endure once mobility went

back to normal? The 57 topics are annotated with these different "behavioral groups" on the right-hand side of Fig. 5.

In the first two sectors of the figure (the first two colored blocks), we have topics for which a persistent increase or decrease was observed. These topics experienced an increase following the mobility changepoint, as shown in column (a), which was sustained once the pandemic ended, as shown in column (c). Among the topics for which there was a persistent *increase* in pageview volumes, we see entertainment-related subjects such as VIDEO GAMES, BOOKS, and INTERNET CULTURE. Meanwhile, topics that experienced a persistent *decrease* include a geographical location (e.g., WEST ASIA) as well as SPORTS, a topic clearly related to outdoor activities. These findings suggest that people shifted information seeking needs from outdoor (e.g., SPORTS) to indoor activities (e.g., VIDEO GAMES).

In the third and fourth sections of the figure, we have topics for which the increase or decrease in pageview volume was transient. A noteworthy topic here is BIOLOGY, which may have received a boost given the sudden importance of all things virus-related, but that eventually, as people got used to the pandemic, returned to the normal levels of volume compared to the previous year. After the mobility changepoint, as shown in column (a), pages related to biology had an increase to around 109%, followed by a decrease to around 89% after things went back to normal, as shown in column (b). Overall, comparing the five-week period before the mobility changepoint with the five-week period after the normality changepoint, the effect is not significant.

Lastly, the last five sections of the figure contain either topics that observed a "reversed" increase or decrease,[9] shown in the sixth and seventh sections, or that experienced no change following the mobility changepoint at all, in the eighth, nineth and tenth sections. For the latter kind, some topics also saw an eventual increase or decrease in the difference-in-differences scenarios depicted in columns (b) or (c). Of these last sections, a remarkable topic is MEDICINE & HEALTH, which presents a substantial decrease overall, as shown in column (c), but which did not experience an increase following the mobility changepoint, as shown in column (a). We further investigated the time series for this topic and found that this happened because the shift in interest for the topic happened *before* the mobility changepoint for many countries. This is aligned with previous suggestions that the "panic" of the disease preceded its actual spread (Depoux et al. 2020).

All in all, our results suggest that, while COVID-19-related topics saw transient increases (e.g., BIOLOGY) or no increase at all (e.g., MEDICINE & HEALTH) following mobility restrictions, several topics related to entertainment saw long-lasting increases. Similarly, SPORTS, a topic related to the outdoors, saw persistent decreases. Importantly, although mobility returned to normal in many places, the pandemic is far from ending. It will be interesting to see whether these shifts are medium-term changes that will re-

vert once the pandemic is over or long-term changes that will remain for years to come.

5   Discussion and Conclusions

In summary, our findings suggest that the sharp decrease in human mobility induced by COVID-19-related non-pharmaceutical interventions has boosted the *volume* of information seeking on Wikipedia and has changed the *nature* of the information sought, even considering non-COVID-19-related articles (RQ1). Once mobility returned to its normal levels, the volume of information seeking also returned to its prior levels, but the kinds of content sought did not (RQ2). Lastly, touching upon both research questions, we zoom in on what articles received more or less attention during this period, finding that some topics, e.g., BIOLOGY, saw transient increases or decreases in volume, while others saw persistent increases (e.g., VIDEO GAMES) or decreases (e.g., SPORT).

These findings may be help in characterizing the societal impact of the pandemic. First, the trends in the *volume* of pageviews may themselves be interpreted as a trace of the social impact of the pandemic. The increase in information seeking that closely followed changes in mobility patterns suggests that, once we are deprived of activities halted by non-pharmaceutical interventions, we resort even more to information seeking on the Web, an activity related to entertainment and self-actualization.

Second, our analysis of the *nature* of what the information sought suggests that the societal impact of the crisis has persisted beyond the associated mobility restrictions. This acts as evidence that long-term changes induced by the pandemic are not only geopolitical or economic in nature, but also related to human needs, interests, and concerns.

Third, our results suggest that most of the topics that experienced persistent increases (e.g., VIDEO GAMES and INTERNET CULTURE) are not related to needs often framed as "basic", such as safety and physiological needs (Maslow 1943). Although this does not imply that the latter were not increased,[10] it nevertheless reinforces that there was an increase in people's need for entertainment and arguably self-actualization.

An alternate angle in which we can frame the aforementioned discussion is to consider not how Wikipedia logs shed light on the societal impact of the pandemic, but how this specific event sheds light on Wikipedia's role in times of crisis. While previous work has highlighted how Wikipedia is able to provide crucial information to millions of people during epidemics (McIver and Brownstein 2014), we show that things do not stop there: Wikipedia can also play the role of providing information related to non-essential needs, and browsing Wikipedia becomes more widespread once other activities are hindered by an unexpected crisis.

Limitations. In this paper, we performed a careful analysis of access patterns on Wikipedia. As our findings are observational, we cannot conclusively establish causal effects of pandemic-related events on user behavior. Indeed, lock-

---

[9]That is, they changed in one direction with the decrease in mobility, but then experienced a "bounce back" after mobility went back to normal.

[10]And indeed previous work has found evidence of increases on queries related to such basic needs (Suh et al. 2020).

down measures and the ensuing decrease in mobility coincided with numerous other life changes inflicted by the pandemic, and any of these concomitant changes—rather than the increase in time spent at home—might in principle be the true cause of the changes in Wikipedia access patterns. For instance, one might argue that an increased concern about health issues might drive people to access more encyclopedic information. Nevertheless, our work presents circumstantial indications that there is indeed a causal link between mobility and the volume and nature of information seeking, as we find that Wikipedia -derived signals follow mobility patterns closely (Fig. 4a-b analyzed jointly) and that there exists a dose–response relationship between mobility interventions and Wikipedia -derived signals (Fig. 4d).

Another limitation of this work has to do with the biases of Wikipedia readership. Given the popularity of the world's largest online encyclopedia, it is tempting to try to generalize our findings to large populations. Yet, we must recall that this may not be true due to uneven access to the Internet and varying levels of digital literacy. It could be the case, for example, that the groups with the largest increase in basic needs lack the capability to access Wikipedia, and that we therefore lose the signals associated with this important population.

Future work. A further interesting causal question asks in what activities people engage when forced to spend time at home. Answering this would require drawing conclusions from the particular lockdown that we are observing in the context of COVID-19 to general situations of restricted mobility. One idea would be to treat COVID-19 as an instrumental variable (Angrist and Imbens 1995), i.e., a haphazard event that systematically nudges people to stay at home, while affecting people's interests only via the lockdown measures. For instance, it would be interesting to explore whether COVID-19 can be used as an instrumental variable to estimate what books people read, or what dishes they cook, when they are forced to, or—let the glass be half-full!—when they are given the chance to, spend more time at home.

Acknowledgments

A   Obtaining Mobility Changepoints

Mobility data. To estimate the effective lockdown dates, we use the mobility reports made available by Google and Apple. Google released community-level reports (Aktay et al. 2020) indicating the daily percentage change in visits to predefined categories of places: *Retail and Recreation* aggregates places like restaurants, cafés, shopping centers, *Grocery and Pharmacy*, *Parks*, *Transit Stations* for pub-

Language	Mean	SD	Language	Mean	SD
English	03/16	0.14	Dutch	03/16	0.97
French	03/16	0.00	Norwegian	03/11	0.93
German	03/16	0.07	Danish	03/11	0.06
Korean	02/25	1.01	Swedish	03/11	1.02
Japanese	03/31	1.07	Serbian	03/16	0.89
Finnish	03/16	0.14	Italian	03/11	0.00

Table 2: Changepoint detection results averaged over different parameter choices. Standard deviations (SD) in days.

lic transport hubs, *Workplaces*, and *Residential* which estimates stay-at-home changes. The changes, for a given place on a given day, are reported in comparison to a baseline value, i.e., the median volume for the same day of the week computed across a five-week period between 3 January and 6 February 2020. Similarly, Apple reports relative changes compared to a baseline volume measured on 13 January 2020 along 3 dimensions of mobility: *Driving*, *Walking*, and *Transit*. Combined, these represent nine types of time series capturing mobility behavior.

Handling the language-vs.-country mismatch. Some languages (e.g., English, French, German) cannot easily be matched to one particular geographical area. For them, we collect the largest countries in terms of native speakers for which mobility data was available. We then produce an aggregate of the mobility data across countries weighted by the percentage of native speakers of the language in each country. For English, we aggregated: United States (68.9%), United Kingdom (16.1%), Canada (5.8%), Australia (5.4%), South Africa (1.5%), Ireland (1.2%), and New Zealand (1.1%). For French, we aggregated: France (62.9%), Canada (10.1%), Cameroon (9.2%), Belgium (7.9%), Senegal (4.3%), Benin (3.8%), and Switzerland (1.8%). For German, we aggregated: Germany (87%), Austria (8.7%), and Switzerland (6.3%).

Changepoint detection. Changepoint detection is the task of identifying state changes in time series. We can benefit from the large literature on offline changepoint detection to identify the effective time of lockdown based on changes in mobility time series. These techniques usually rely on two components: (1) a cost function assessing the quality of a particular signal segmentation and (2) a technique to search the space of possible segmentations, guided by the cost function (Aminikhanghahi and Cook 2017). For more details, we refer to the review of (Truong, Oudre, and Vayatis 2020). In this work, we considered the cost functions and search algorithms available as part of the `ruptures` package.[11]

Several different design choices can be made: (1) different subsets of mobility time series can be selected and (2) different changepoint algorithms can be employed. We ran the pipeline with many different parameters and report the results in Table 2. The standard deviations are due to changing parameters and are small, often below one day. Indeed, for most countries, the changes in mobility are very clear, and different methods largely agree.

---

[11] https://github.com/deepcharles/ruptures

References

Abay, K. A.; Tafere, K.; and Woldemichael, A. 2020. Winners and Losers from COVID-19: Global Evidence from Google Search. In *World Bank Policy Research*.

Ahmed, W.; Vidal-Alaball, J.; Downing, J.; and Seguí, F. L. 2020. COVID-19 and the 5G conspiracy theory. In *Journal of Medical Internet Research*.

Aktay et al., A. 2020. Google COVID-19 Community Mobility Reports: Anonymization Process Description. *arXiv* 1–6.

Aminikhanghahi, S.; and Cook, D. J. 2017. A Survey of Methods for Time Series Change Point Detection. In *Knowledge and information systems*.

Angrist, J. D.; and Imbens, G. W. 1995. Identification and estimation of local average treatment effects. Technical report, National Bureau of Economic Research.

Bonaccorsi et al., G. 2020. Economic and social consequences of human mobility restrictions under COVID-19. In *Proceedings of the National Academy of Sciences*.

Cohen, J. 2020. Unprecedented nationwide blood studies seek to track U.S. coronavirus spread. In *ScienceMag*.

Colley, R. C.; Bushnik, T.; and Langlois, K. 2020. Exercise and screen time during the COVID-19 pandemic. In *Health Reports*.

Collins, F. S.; and Stoffels, P. 2020. Accelerating COVID-19 Therapeutic Interventions and Vaccines: An Unprecedented Partnership for Unprecedented Times. In *JAMA*.

Depoux, A.; Martin, S.; Karafillakis, E.; Preet, R.; Wilder-Smith, A.; and Larson, H. 2020. The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine* 1–2.

Diaz, F.; Gamon, M.; Hofman, J. M.; Kıcıman, E.; and Rothschild, D. 2016. Online and social media data as an imperfect continuous panel survey. In *PloS one*.

Favale, T.; Soro, F.; Trevisan, M.; Drago, I.; and Mellia, M. 2020. Campus traffic and e-Learning during COVID-19 pandemic. In *Computer Networks*.

Feldmann et al., A. 2021. Implications of the COVID-19 Pandemic on the Internet Traffic. In *Broadband Coverage in Germany; 15th ITG-Symposium*, 1–5.

Flaxman et al., S. 2020. Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Technical report, Imperial College London.

García-Gavilanes, R.; Mollgaard, A.; Tsvetkova, M.; and Yasseri, T. 2017. The memory remains: Understanding collective memory in the digital age. In *Science advances*.

Generous et al, N. 2014a. Global disease monitoring and forecasting with Wikipedia. In *PLoS computational biology*.

Generous et al, N. 2014b. Global disease monitoring and forecasting with Wikipedia. In *PLoS computational biology*.

Gupta et al., S. 2020. Effects of social distancing policy on labor market outcomes. Technical report, National Bureau of Economic Research.

Hale, T.; Webster, S.; Petherick, A.; Phillips, T.; and Kira, B. 2020. Oxford COVID-19 Government Response tracker. https://bit.ly/2ZF3Q7h. Acessed: 2020-06-04.

Kouzy et al., R. 2020. Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter. In *Cureus*. Cureus Inc.

Lin, Y.-H.; Liu, C.-H.; and Chiu, Y.-C. 2020. Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries. In *Brain, Behavior, and Immunity*.

Liverpool et al., A. 2020. Covid-19 News: Coronavirus restrictions to ease slightly in England. https://www.newscientist.com/article/2244516-uk-plans-to-further-ease-lockdown-as-new-case-rate-remains-high/. Acessed: 2020-06-04.

Maslow, A. 1943. A Theory of Human Motivation. In *Psychological Review*.

McIver, D. J.; and Brownstein, J. S. 2014. Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time. In *PLOS Computational Biology*. Public Library of Science.

Ryan, B. J.; Coppola, D.; Canyon, D. V.; Brickhouse, M.; and Swienton, R. 2020. COVID-19 Community Stabilization and Sustainability Framework. In *Disaster Medicine and Public Health Preparedness*.

Salganik, M. J. 2019. *Bit by bit: Social research in the digital age*. Princeton University Press.

Singer, P.; Lemmerich, F.; West, R.; Zia, L.; Wulczyn, E.; Strohmaier, M.; and Leskovec, J. 2017. Why We Read Wikipedia. In *WWW*.

Suh, J.; Horvitz, E.; White, R. W.; and Althoff, T. 2020. Population-Scale Study of Human Needs During the COVID-19 Pandemic: Analysis and Implications. *arXiv* 1–18.

Tausczik, Y. R.; Faasse, K.; Pennebaker, J. W.; and Petrie, K. J. 2012. Public Anxiety and Information Seeking Following the H1N1 Outbreak: Blogs, Newspaper Articles, and Wikipedia Visits. In *Health Communication*.

Tizzoni, M.; Panisson, A.; Paolotti, D.; and Cattuto, C. 2020. The impact of news exposure on collective attention in the United States during the 2016 Zika epidemic. In *PLOS Computational Biology*.

Truong, C.; Oudre, L.; and Vayatis, N. 2020. Selective Review of Offline Change Point Detection Methods. In *Signal Processing*. Elsevier.

Tubadji, A.; Boy, F.; and Webber, D. J. 2020. Narrative economics, public policy and mental health. In *Covid Economics*.

WHO. 2020. 172 countries and multiple candidate vaccines engaged in COVID-19 vaccine Global Access Facility. https://www.who.int/news/item/24-08-2020-172-countries-and-multiple-candidate-vaccines-engaged-in-covid-19-vaccine-global-access-facility. Acessed: 2020-06-04.

Yang, K.-C.; Torres-Lugo, C.; and Menczer, F. 2020. Prevalence of low-credibility information on twitter during the covid-19 outbreak. In *arXiv*, 1–5.

Zhang, A. F.; Wang, R.; Blohm, E.; Budak, C.; Robert Jr, L. P.; and Romero, D. M. 2019. Participation of New Editors after Times of Shock on Wikipedia. In *AAAI ICWSM*.