The Harvard Kennedy School Misinformation Review[1]
June 2020, Volume 1, Special Issue on COVID-19 and Misinformation
Attribution 4.0 International (CC BY 4.0)
Reprints and permissions: misinforeview@hks.harvard.edu
DOI: https://doi.org/10.37016/mr-2020-028
Website: misinforeview.hks.harvard.edu

Research Article

The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media

We investigate the relationship between media consumption, misinformation, and important attitudes and behaviours during the coronavirus disease 2019 (COVID-19) pandemic. We find that comparatively more misinformation circulates on Twitter, while news media tends to reinforce public health recommendations like social distancing. We find that exposure to social media is associated with misperceptions regarding basic facts about COVID-19 while the inverse is true for news media. These misperceptions are in turn associated with lower compliance with social distancing measures. We thus draw a clear link from misinformation circulating on social media, notably Twitter, to behaviours and attitudes that potentially magnify the scale and lethality of COVID-19.

Authors: Aengus Bridgman (1), Eric Merkley (2), Peter John Loewen (2), Taylor Owen (3), Derek Ruths (4), Lisa Teichmann (5), Oleg Zhilin (4)
Affiliations: (1) Political Science, McGill University, (2) Munk School of Global Affairs and Public Policy, University of Toronto, (3) Max Bell School of Public Policy, McGill University, (4) Computer Science, McGill University, (5) Languages, Literatures, and Cultures, McGill University
How to cite: Bridgman, A.; Merkley, E.; Loewen, PJ; Owen, T; Ruths, D.; Teichmann, L.; Zhilin, O. (2020) The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media, *The Harvard Kennedy School (HKS) Misinformation Review*, Volume 1, Special Issue on COVID-19 and Misinformation
Received: April 27th, 2020 Accepted: June 10th, 2020 Published: June 18th, 2020.

Research questions

- How prevalent is misinformation surrounding COVID-19 on Twitter, and how does this compare to Canadian news media?
- Does the type of media one is exposed to influence social distancing behaviours and beliefs about COVID-19?
- Is there a link between COVID-19 misinformation and perceptions of the pandemic's severity and compliance with social distancing recommendations?

Essay summary

- We evaluate the presence of misinformation and public health recommendations regarding COVID-19 in a massive corpus of tweets as well as all articles published on nineteen Canadian

[1] *A publication of the Shorenstein Center for Media, Politics, and Public Policy, at Harvard University, John F. Kennedy School of Government.*

news sites. Using these data, we show that preventative measures are more encouraged and covered on traditional news media, while misinformation appears more frequently on Twitter.

- To evaluate the impact of this greater level of misinformation, we conducted a nationally representative survey that included questions about common misperceptions regarding COVID-19, risk perceptions, social distancing compliance, and exposure to traditional news and social media. We find that being exposed to news media is associated with fewer misperceptions and more social distancing compliance while conversely, social media exposure is associated with more misperceptions and less social distancing compliance.
- Misperceptions regarding the virus are in turn associated with less compliance with social distancing measures, even when controlling for a broad range of other attitudes and characteristics.
- Association between social media exposure and social distancing non-compliance is eliminated when accounting for effect of misperceptions, providing evidence that social media is associated with non-compliance through increasing misperceptions about the virus.

Argument and Implications

The COVID-19 pandemic has been accompanied by a so-called "infodemic" – a global spread of misinformation that poses a serious problem for public health. Infodemics are concerning because the spread of false or misleading information has the capacity to change transmission patterns (Kim et al., 2019) and consequently the scale and lethality of a pandemic. This information can be shared by any media, but there is reason to be particularly concerned about the role that social media, such as Facebook and Twitter, play in incidentally boosting misperceptions. These platforms are increasingly relied upon as primary sources of news (Mitchell et al., 2016) and misinformation has been heavily documented on them (Garrett, 2019; Vicario et al., 2016). Scholars have found medical and health misinformation on the platforms, including that related to vaccines (Radzikowski et al., 2016) and other virus epidemics such as Ebola (Fung et al., 2016) and Zika (Sharma et al., 2017).

However, misinformation content typically makes up a low percentage of overall discussion of a topic (e.g. Fung et al., 2016) and mere exposure to misinformation does not guarantee belief in that misinformation. More research is thus needed to understand the extent and consequences of misinformation surrounding COVID-19 on social media. During the COVID-19 pandemic, Twitter, Facebook and other platforms have engaged in efforts to combat misinformation but they have continued to receive widespread criticism that misinformation is still appearing on prominent pages and groups (Kouzy et al., 2020; NewsGuard, 2020). The extent to which misinformation continues to circulate on these platforms and influence people's attitudes and behaviours is still very much an open question.

Here, we draw on three data sets and a sequential mixed method approach to better understand the consequences of online misinformation for important behaviours and attitudes. First, we collected nearly 2.5 million tweets explicitly referring to COVID-19 in the Canadian context. Second, we collected just over 9 thousand articles from nineteen Canadian English-language news sites from the same time period. We coded both of these media sets for misinformation and public health recommendations. Third, we conducted a nationally representative survey that included questions related to media consumption habits, COVID-19 perceptions and misperceptions, and social distancing compliance. As our outcome variables are continuous, we use Ordinary Least Squares (OLS) regression to identify relationships between news and social media exposure, misperceptions, compliance with social distancing measures,

and risk perceptions. We use these data to illustrate: 1) the relative prevalence of misinformation on Twitter; and 2) a powerful association between social media usage and misperceptions, on the one hand, and social distancing non-compliance on the other.

Misinformation and compliance with social distancing

We first compare the presence of misinformation on Twitter with that on news media and find, consistent with the other country cases (Chadwick & Vaccari, 2019; Vicario et al., 2016), comparatively higher levels of misinformation circulating on the social media platform. We also found that recommendations for safe practices during the pandemic (e.g. washing hands, social distancing) appeared much more frequently in the Canadian news media. These findings are in line with literature examining fake news which finds a large difference in information quality across media (Al-Rawi, 2019; Guess & Nyhan, 2018).

Spending time in a media environment that contains misinformation is likely to change attitudes and behaviours. Even if users are not nested in networks that propagate misinformation, they are likely to be incidentally exposed to information from a variety of perspectives (Feezell, 2018; Fletcher & Nielsen, 2018; Weeks et al., 2017). Even a highly curated social media feed is thus still likely to contain misinformation. As cumulative exposure to misinformation increases, users are likely to experience a reinforcement effect whereby familiarity leads to stronger belief (Dechêne et al., 2010).

To evaluate this empirically, we conducted a national survey that included questions on information consumption habits and a battery of COVID-19 misperceptions that could be the result of exposure to misinformation. We find that those who self-report exposure to the misinformation-rich social media environment do tend to have more misperceptions regarding COVID-19. These findings are consistent with others that link exposure to misinformation and misperceptions (Garrett et al., 2016; Jamieson & Albarracín, 2020). Social media users also self-report less compliance with social distancing.

Misperceptions are most meaningful when they impact behaviours in dangerous ways. During a pandemic, misperceptions can be fatal. In this case, we find that misperceptions are associated with reduced COVID-19 risk perceptions and with lower compliance with social distancing measures. We continue to find strong effects after controlling for socio-economic characteristics as well as scientific literacy. After accounting for the effect of misperceptions on social distancing non-compliance, social media usage no longer has a significant association with non-compliance, providing evidence that social media may lead to less social distancing compliance through its effect on COVID-19 misperceptions.

Implications

While some social media companies have made efforts to suppress misinformation on their platforms, there continues to be a high level of misinformation relative to news media. Highly polarized political environments and media ecosystems can lead to the spread of misinformation, such as in the United States during the COVID-19 pandemic (Allcott et al., 2020; Motta et al., 2020). But even in healthy media ecosystems with less partisan news (Owen et al., 2020), social media can continue to facilitate the spread of misinformation. There is a real danger that without concerted efforts to reduce the amount of misinformation shared on social media, the large-scale social efforts required to combat COVID-19 will be undermined.

We contribute to a growing base of evidence that misinformation circulating on social media poses public health risks and join others in calling for social media companies to put greater focus on flattening the curve of misinformation (Donovan, 2020). These findings also provide governments with stronger evidence that the misinformation circulating on social media can be directly linked to misperceptions and public health risks. Such evidence is essential for them to chart an effective policy course. Finally, the

methods and approach developed in this paper can be fruitfully applied to study other waves of misinformation and the research community can build upon the link clearly drawn between misinformation exposure, misperceptions, and downstream attitudes and behaviours.

Next steps

We found use of social media platforms broadly contributes to misperceptions but were unable to precise the overall level of misinformation circulating on non-Twitter social media. Data access for researchers to platforms such as Facebook, YouTube, and Instagram is limited and virtually non-existent for SnapChat, WhatsApp, and WeChat. Cross-platform content comparisons are an important ingredient for a rich understand of the social media environment and these social media companies must better open their platforms to research in the public interest.

Findings

Finding 1: Misinformation about COVID-19 is circulated more on Twitter as compared to traditional media.

We find large differences between the quality of information shared about COVID-19 on traditional news and Twitter. Figure 1 shows the percentage of COVID-19 related content that contains information linked to a particular theme. The plot reports the prevalence of information on both social and news media for: 1) three specific pieces of misinformation; 2) a general set of content that describes the pandemic itself as a conspiracy or a hoax; and 3) advice about hygiene and social distancing during the pandemic. We differentiate content that shared misinformation (red in the plot) from content that debunked misinformation (green in the plot).

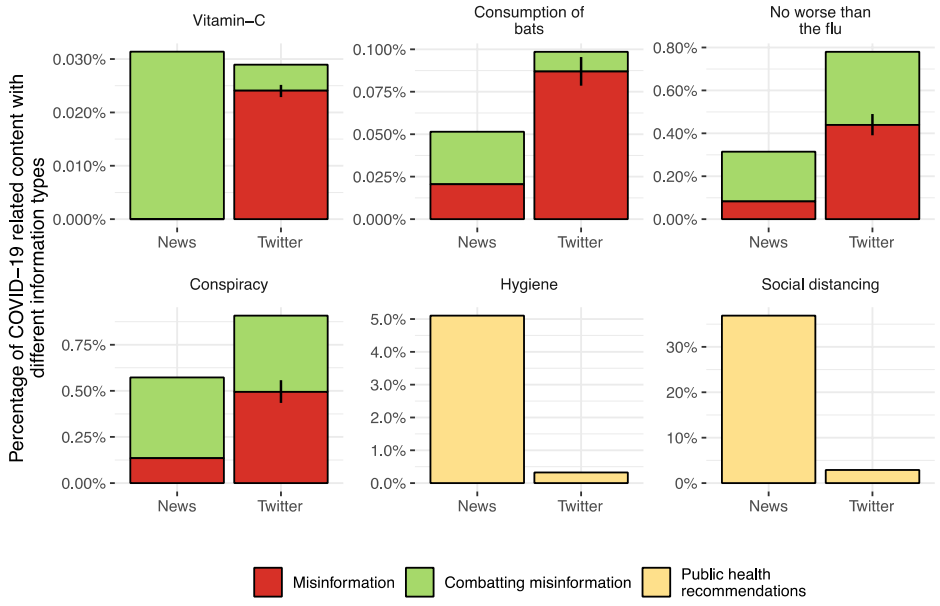


Figure 1. Relative frequencies of COVID-19 related information. *Figure 1 plots the percentage of all COVID-19 related content that includes reference to a particular theme. Confidence intervals are shown only for the tweets where a random sample was manually categorized and the results extrapolated to all tweets in the corpus.*

There are large differences between the levels of misinformation on Twitter and news media. Misinformation was comparatively more common on Twitter across all four categories, while debunking was relatively more common in traditional news. Meanwhile, advice on hygiene and social distancing appeared much more frequently in news media. Note that higher percentages are to be expected for longer format news articles since we rely on keyword searches for identification. This makes the misinformation findings even starker – despite much higher average word counts, far fewer news articles propagate misinformation.

Finding 2: There is a strong association between social media exposure and misperceptions about COVID-19. The inverse is true for exposure to traditional news.

Among our survey respondents we find a corresponding strong association between social media exposure and misperceptions about COVID-19. These results are plotted in Figure 2, with controls included for both socioeconomic characteristics and demographics. Moving from no social media exposure to its maximum is expected to increase one's misperceptions of COVID-19 by 0.22 on the 0-1 scale and decreased self-reported social distancing compliance by 0.12 on that same scale.

This result stands in stark contrast with the observed relationship between traditional news exposure and our outcome measures. Traditional news exposure is *positively* associated with correct perceptions regarding COVID-19. Moving from no news exposure to its highest level is expected to reduce misperceptions by 0.12 on the 0-1 scale and to increase social distancing compliance by 0.28 on that same scale. The effects are plotted in Figure 2. Social media usage appears to be correlated with COVID-19 misperceptions, suggesting these misperceptions are partially a result of misinformation on social media. The same cannot be said of traditional news exposure.

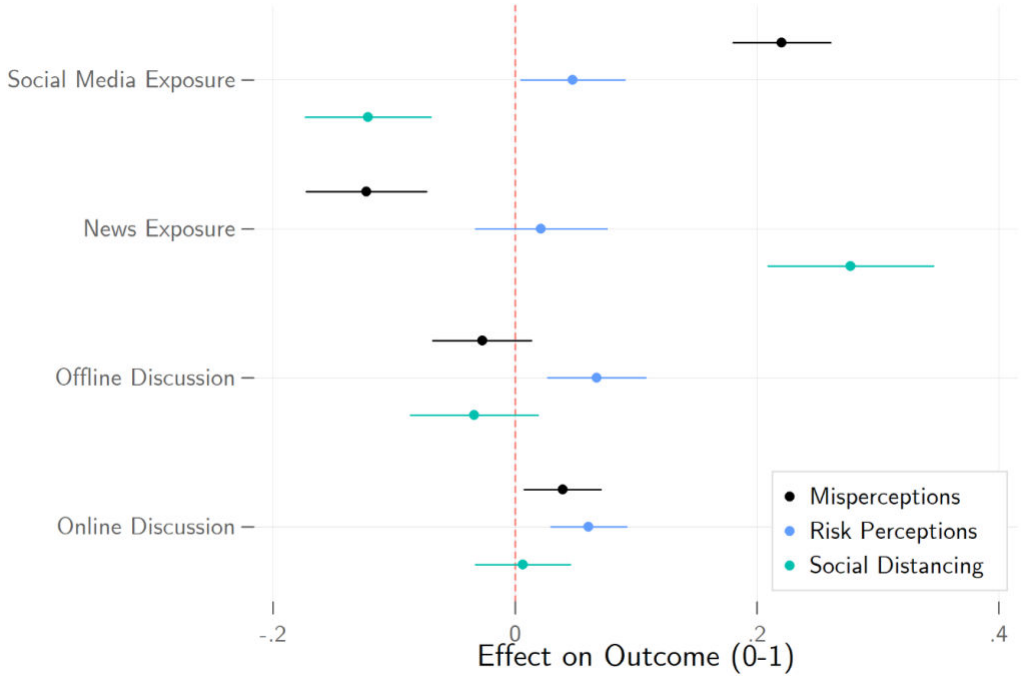


Figure 2. Effects of Media and Discussion on COVID-19 Misperceptions, Risk Perceptions, and Social Distancing Practice.
Figure 2 plots the coefficients from three Ordinary Least Squares regression models. All measures scaled from 0-1, so effects can be interpreted as going from the minimum to the maximum of explanatory variables on the 0-1 scales of each outcome. Additional controls for income, education, age, gender, religiosity, language, and region not plotted. 95% confidence intervals plotted.

Finding 3: Misperceptions about the pandemic are associated with lower levels of risk perceptions and social distancing compliance.

COVID-19 misperceptions are also powerfully associated with *lower* levels of social distancing compliance. Moving from the lowest level of COVID-19 misperceptions to its maximum is associated with a reduction of one's social distancing by 0.39 on the 0-1 scale. The previously observed relationship between social media exposure and misperceptions disappears, suggestive of a mediated relationship. That is, social media exposure increases misperceptions, which in turn reduces social distancing compliance. Misperceptions is also weakly associated with lower COVID-19 risk perceptions. Estimates from our models using COVID-19 concern as the outcome can be found in the left panel of Figure 3, while social distancing can be found in the right panel.

Finally, we also see that the relationship between misinformation and both social distancing compliance and COVID-19 concern hold when including controls for science literacy and a number of fundamental predispositions that are likely associated with both misperceptions and following the advice of scientific experts, such as anti-intellectualism, pseudoscientific beliefs, and left-right ideology. These estimates can similarly be found in Figure 3.

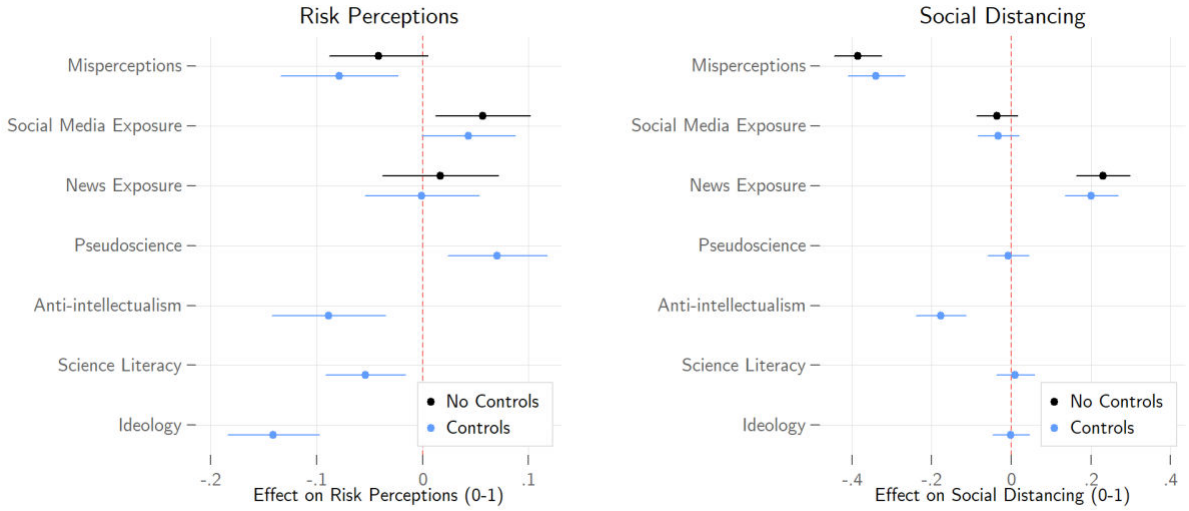


Figure 3. Effects of COVID-19 Misperceptions on Risk Perceptions (Left) and Social Distancing (Right). *Figure 3 plots the coefficients from Ordinary Least Squares regression models. Left panel plots the effects of misperceptions on COVID-19 risk perceptions with and without science literacy and predispositions as controls. Right panel does the same for social distancing. All measures scaled from 0-1, so effects can be interpreted as going from the minimum to the maximum of explanatory variables on the 0-1 scales of each outcome. Additional controls for online and offline discussion, income, education, age, gender, religiosity, language, and region not plotted. 95% confidence intervals plotted.*

Methods

Canadian Twitter and news data were collected from March 26th to April 6th, 2020. We collected all English-language tweets from a set of 620,000 users that have been determined to be likely Canadians. For inclusion, a given user must self-identify as Canadian-based, follow a large number of Canadian political elite accounts, or frequently use Canadian-specific hashtags. News media was collected from nineteen prominent Canadian news sites with active RSS feeds. These tweets and news articles were searched for "covid" or "coronavirus", leaving a sample of 2.25 million tweets and 8,857 news articles.

Of the COVID-19 related content, we searched for terms associated with four instances of misinformation that circulated during the COVID-19 pandemic: that COVID-19 was no more serious than the flu, that vitamin C or other supplements will prevent contraction of the virus, that the initial animal-to-human transfer of the virus was the direct result of eating bats, or that COVID-19 was a hoax or conspiracy. Given that we used keyword searches to identify content, we manually reviewed a random sample of 500 tweets from each instance of misinformation. Each tweet was coded as one of four categories: propagating misinformation, combatting misinformation, content with the relevant keywords but unrelated to misinformation, or content that refers to the misinformation but does not offer comment.

We then calculated the overall level of misinformation for that instance on Twitter by multiplying the overall volume of tweets by the proportion of hand-coded content where misinformation was identified. Each news article that included relevant keywords was similarly coded. The volume of the news mentioning these terms was sufficiently low that all news articles were hand coded. To identify health recommendations, we used a similar keyword search for terms associated with particular recommendations: 1) social distancing including staying at home, staying at least 6 feet or 2 meters away and avoiding gatherings; and 2) washing hands and not touching any part of your face.[2]

For survey data, we used a sample of nearly 2,500 Canadian citizens 18 years or older drawn from a probability-based online national panel fielded from April 2-6, 2020. Quotas we set on age, gender, region, and language to ensure sample representativeness, and data was further weighted within region by gender and age based on the 2016 Canadian census.

We measure levels of COVID-19 misperceptions by asking respondents to rate the truthfulness of a series of nine false claims, such as the coronavirus being no worse than the seasonal flu or that it can be warded off with Vitamin C. Each was asked on a scale from definitely false (0) to definitely true (5). We use Cronbach's Alpha as an indicator of scale reliability. Cronbach's Alpha ranges from 0-1, with scores above 0.8 indicating the reliability is "good." These items score 0.88, so we can safely construct a 0-1 scale of misperceptions from them.

We evaluate COVID-19 risk perceptions with a pair of questions asking respondents how serious of a threat they believe the pandemic to be for themselves and for Canadians, respectively. Each question was asked on a scale from not at all (0) to very (4). We construct a continuous index with these items.

We quantify social distancing by asking respondents to indicate which of a series of behaviours they had undertaken in response to the pandemic, such as working from home or avoiding in-person contact with friends, family, and acquaintances. We use principal component analysis (PCA) to reduce the number of dimensions in these data while minimizing information loss. The analysis revealed 2 distinct dimensions in our questions. One dimension includes factors strongly determined by occupation, such as working from home and switching to online meetings. The other dimension contains more inclusive behaviours such as avoiding contact, travel, and crowded places. We generate predictions from the PCA for this latter dimension to use in our analyses. The factor loadings can be found in Table A1 of the supporting materials.

We gauge news and social media consumption by asking respondents to identify news outlets and social media platforms they have used over the past week for political news. The list of news outlets included 17 organizations such as mainstream sources like CBC and Global, and partisan outlets like Rebel Media and National Observer. The list of social media platforms included 10 options such as Facebook, Twitter, YouTube, and Instagram. We sum the total number of outlets/platforms respondents

[2] Further details on the media collection strategy and hand-coding schema are available in the supporting materials.

report using and take the log to adjust for extreme values. We measure offline political discussion with an index based on questions asking how often respondents have discussed politics with family, friends, and acquaintances over the past week. Descriptions of our primary variables can be found in Table A2 of the supporting materials.

We evaluate our hypotheses using a standard design that evaluates the association between our explanatory and outcome variables controlling for other observable factors we measured. In practice, randomly assigning social media exposure is impractical, while randomly assigning misinformation is unethical. This approach allows us to describe these relationships, though we cannot make definite claims to causality.

We hypothesize that social media exposure is associated with misinformation on COVID-19. Figure 2 presents the coefficients of models predicting the effects of news exposure, social media exposure, and political discussion on COVID-19 misinformation, risk perceptions, and social distancing. Socio-economic and demographic control estimates are not displayed. Full estimation results can be found in the Table A3 of the supporting materials.

We further hypothesize that COVID-19 misinformation is associated with lower COVID-19 risk perceptions and less social distancing compliance. Figure 3 presents the coefficients for models predicting the effects of misinformation, news exposure, and social media exposure on severity perceptions and social distancing. We show models with and without controls for science literacy and other predispositions. Full estimation results can be found in the Table A4 of the supporting materials.

Limitations and Robustness

A study such as this comes with clear limitations. First, we have evaluated information coming from only a section of the overall media ecosystem and during a specific time-period. The level of misinformation differs across platforms and online news sites and a more granular investigation into these dynamics would be valuable. Our analysis suggests that similar dynamics exist across social media platforms, however. In the supplementary materials we show that associations between misperceptions and social media usage are even higher for other social media platforms, suggesting that our analysis of Twitter content may underrepresent the prevalence of misinformation on social media writ large. As noted above, existing limitations on data access make such cross-platform research difficult.

Second, our data is drawn from a single country and language case study and other countries may have different media environments and levels of misinformation circulating on social media. We anticipate the underlying dynamics found in this paper to hold across these contexts, however. Those who consume information from platforms where misinformation is more prevalent will have greater misperceptions and that these misperceptions will be linked to lower compliance with social distancing and lower risk perceptions.

Third, an ecological problem is present wherein we do not link survey respondents directly to their social media consumption (and evaluation of the misinformation they are exposed to) and lack the ability to randomly assign social media exposure to make a strong causal argument. We cannot and do not make a causal argument here but argue instead that there is strong evidence for a misinformation to misperceptions to lower social distancing compliance link.

Bibliography

Allcott, H., Boxell, L., Conway, J. C., Gentzkow, M., Thaler, M., & Yang, D. Y. (2020). *Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic* (Working Paper No. 26946; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w26946

Al-Rawi, A. (2019). Gatekeeping Fake News Discourses on Mainstream Media Versus Social Media. *Social Science Computer Review*, *37*(6), 687–704. https://doi.org/10.1177/0894439318795849

Chadwick, A., & Vaccari, C. (2019). *News sharing on UK social media: Misinformation, disinformation, and correction* [Report]. Loughborough University. https://repository.lboro.ac.uk/articles/News_sharing_on_UK_social_media_misinformation_disinformation_and_correction/9471269

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review*, *14*(2), 238–257. https://doi.org/10.1177/1088868309352251

Donovan, J. (2020). Social-media companies must flatten the curve of misinformation. *Nature*. https://doi.org/10.1038/d41586-020-01107-z

Feezell, J. T. (2018). Agenda Setting through Social Media: The Importance of Incidental News Exposure and Social Filtering in the Digital Era. *Political Research Quarterly*, *71*(2), 482–494. https://doi.org/10.1177/1065912917744895

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society*, *20*(7), 2450–2468. https://doi.org/10.1177/1461444817724170

Fung, I. C.-H., Fu, K.-W., Chan, C.-H., Chan, B. S. B., Cheung, C.-N., Abraham, T., & Tse, Z. T. H. (2016). Social Media's Initial Reaction to Information and Misinformation on Ebola, August 2014: Facts and Rumors. *Public Health Reports*, *131*(3), 461–473. https://doi.org/10.1177/003335491613100312

Garrett, R. K. (2019). Social media's contribution to political misperceptions in U.S. Presidential elections. *PLoS ONE*, *14*(3). https://doi.org/10.1371/journal.pone.0213500

Garrett, R. K., Weeks, B. E., & Neo, R. L. (2016). Driving a Wedge Between Evidence and Beliefs: How Online Ideological News Exposure Promotes Political Misperceptions. *Journal of Computer-Mediated Communication*, *21*(5), 331–348. https://doi.org/10.1111/jcc4.12164

Guess, A., & Nyhan, B. (2018). Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign. *European Research Council*, 49.

Jamieson, K. H., & Albarracín, D. (2020). The Relation between Media Consumption and Misinformation at the Outset of the SARS-CoV-2 Pandemic in the US. *Harvard Kennedy School Misinformation Review*, *2*. https://doi.org/10.37016/mr-2020-012

Kim, L., Fast, S. M., & Markuzon, N. (2019). Incorporating media data into a model of infectious disease transmission. *PLOS ONE*, *14*(2), e0197646. https://doi.org/10.1371/journal.pone.0197646

Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E. W., & Baddour, K. (2020). Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. *Cureus*, *12*(3). https://doi.org/10.7759/cureus.7255

Mitchell, A., Gottfried, J., Barthel, M., & Shearer, E. (2016, July 7). The Modern News Consumer. *Pew Research Center's Journalism Project*. https://www.journalism.org/2016/07/07/the-modern-news-consumer/

Motta, M., Stecula, D., & Farhart, C. E. (2020). *How Right-Leaning Media Coverage of COVID-19 Facilitated the Spread of Misinformation in the Early Stages of the Pandemic* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/a8r3p

NewsGuard. (2020). *Superspreaders*. https://www.newsguardtech.com/superspreaders/

Owen, T., Loewen, P., Ruths, D., Bridgman, A., Gorwa, R., MacLellan, S., Merkley, E., & Zhilin, O. (2020). *Lessons in Resilience: Canada's Digital Media Ecosystem and the 2019 Election*. Public Policy Forum. https://ppforum.ca/articles/lessons-in-resilience-canadas-digital-media-ecosystem-and-the-2019-election/

Radzikowski, J., Stefanidis, A., Jacobsen, K. H., Croitoru, A., Crooks, A., & Delamater, P. L. (2016). The Measles Vaccination Narrative in Twitter: A Quantitative Analysis. *JMIR Public Health and Surveillance*, *2*(1), e1. https://doi.org/10.2196/publichealth.5059

Sharma, M., Yadav, K., Yadav, N., & Ferdinand, K. C. (2017). Zika virus pandemic—Analysis of Facebook as a social media health information platform. *American Journal of Infection Control*, *45*(3), 301–302. https://doi.org/10.1016/j.ajic.2016.08.022

Shin, J., Jian, L., Driscoll, K., & Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, *83*, 278–287. https://doi.org/10.1016/j.chb.2018.02.008

Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental Exposure, Selective Exposure, and Political Information Sharing: Integrating Online Exposure Patterns and Expression on Social Media. *Journal of Computer-Mediated Communication*, *22*(6), 363–379. https://doi.org/10.1111/jcc4.12199

Funding
The project was funded through the Department of Canadian Heritage's Digital Citizens Initiative.

Competing interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics
The research protocol was approved by the institutional review board at University of Toronto. Human subjects gave informed consent before participating and were debriefed at the end of the study.

Copyright
This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are properly credited.

Data Availability
All materials needed to replicate this study are available via the Harvard Dataverse: https://doi.org/10.7910/DVN/5QS2XP.

Supporting materials

1. Methodology Details

1.1. Twitter collection strategy

We drew upon a large-scale Twitter data collection effort in the Canadian context. 620, 000 Twitter users were identified as relevant to the Canadian conversation on the platform. For inclusion, a user needed to: a) have been manually identified as a politician, journalist, academic, or public personality (N = 6000); b) identify themselves as Canadian or as located in Canada in their Twitter profile; c) have recently tweeted with Canadian-specific hashtags (e.g. #toronto) AND have followed at least 10 Canadian politicians and news personalities; d) be in the top percentile for following Canadian politicians and news personalities. All tweets for this population have been collected since January 1, 2020, with the Tweets used in this study running from 1 week before the survey was in the field to the last day the survey was in the field.

1.2. Media collection strategy

The html of Canadian news sites with active RSS feeds was collected then processed using newspaper3k in Python. Data was collected for the following English-language publications: Global News, CBC, Toronto Star, Financial Post, National Post, Winnipeg Free Press, Toronto Sun, Montreal Gazette, Huffington Post, Ottawa Citizen, Calgary Herald, Edmonton Journal, National Observer, Macleans, iPolitics, Winnipeg Sun, Vancouver Sun, Vancouver Prince, and Edmonton Sun. All articles posted between 1 week before the survey was put into the field and the last day the survey was in the field were included.

1.3. Coding schema

We used the following dictionary schema to identify COVID-19 related content. The hygiene and social distancing language was taken from official health recommendations from: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/guidance-documents.html (accessed April 20, 2020). Each document was tokenized and then token-matching was effected for the following terms.

COVID-19	`('coronavirus','covid','covid-19')`
Flu	`('flu')`
Vitamin-C	`('vitamin-c','vitamin c')`
Hygiene	`('hand wash', 'wash hands', 'hand washing', '30 seconds', 'thirty seconds', 'with soap', 'wash your hands', 'hand sanitizer', 'avoid touching your eyes', 'avoid touching your face', 'avoid touching your mouth')`
Social distancing	`('work from home', 'stay home', 'stay at home', 'avoid all non-essential trips', 'gather in groups', 'avoid places', 'avoid public', 'avoid crowds', 'avoid gatherings', 'grocery delivery', 'avoid large gatherings', 'avoid small gatherings', 'and small gatherings', 'and large gatherings', 'limit events', 'limit meetings', 'self-isolate', 'isolation', 'must isolate', 'social distance', 'social distancing', 'six feet', '6 ft', '2 meter', '2 meters', 'maintain distance')`
Conspiracy	`('hoax', 'fraud', 'deception', 'swindle', 'dupe', 'con', 'trick', 'deceive', 'scam', 'scheme', 'racket', 'overblown', 'exaggerated', 'overdone', 'inflated', 'embellished', 'hyperbolic', 'conspiracy', 'hyperbole', 'harmless')`

2. Supporting tables for survey data

Table A1 shows factor loadings for social distancing measures. Factor 1 generally captures compliance with physical distancing recommendations and is used throughout the paper. Factor 2 contains behaviours strongly determined by occupation, such as working from home and switching to online meetings. Table A2 provides the descriptions for the explanatory and outcome variables.

Table A1. Factor loadings for social distancing measures

Variable	Factor 1	Factor 2	Uniqueness
Worked from home	-0.164	0.733	0.437
Avoided bars, restaurants, and crowds	0.767	-0.001	0.412
Avoided grocery stores at peak times	0.601	0.210	0.595
Avoided in-person contact	0.707	0.068	0.495
Stocked up on provisions	0.258	0.404	0.770
Kept distance of two metres	0.767	-0.001	0.412
Switched to virtual meetings	0.266	0.662	0.492
Switched to online shopping	0.238	0.614	0.566
Avoided domestic travel	0.715	0.158	0.465
Avoided public transit	0.705	0.187	0.468

Table A2. Variable Descriptions

Measure	Description	Mean	SD
Risk Perceptions	0-1; 1) How serious of a threat do you think the coronavirus is to yourself? (very, somewhat, not very, not at all); 2) How serious of a threat do you think the coronavirus is to Canadians? (very, somewhat, not very, not at all)	0.78	0.21
Social Distancing	0-1; First dimension of PCA. More details in Table A1	0.75	0.26
Misperceptions	0-1; Please rate the truthfulness of the following claims: 1) The coronavirus is no worse than the seasonal flu; 2) Drinking water every 15 minutes will help prevent the coronavirus; 3) The Chinese government developed the coronavirus as a bioweapon; 4) Homeopathy and home remedies can help manage and prevent the coronavirus; 5) The coronavirus was caused by the consumption of bats in China; 6) The coronavirus will go away by the summer; 7) Vitamin C can ward off the coronavirus; 8) There is a vaccine for the coronavirus that national governments and pharmaceutical companies won't release; 9) High temperatures, such as from saunas and hair dryers, can kill the coronavirus (definitely true, probably true, probably false, definitely false, unsure)	0.31	0.21
Social Media Exposure	Logged, 0-1; Over the past week, which of the following social media applications did you use to watch, read, or listen to news about politics? Please select all that apply. 1) Twitter 2) Facebook; 3) Instagram; 4) Youtube; 5) Reddit; 6) LinkedIn; 7) Tumblr; 8) WhatsApp; 9) Snapchat; 10) WeChat; 11) Another application	0.31	0.25
News Exposure	0-1; Over the past week, which of the following news media	0.42	0.19

Measure	Description	Mean	SD
	outlets did you watch, read, or listen to for news about politics? Please select all that apply. 1) CBC; 2) CTV; 3) Global; 4) CityNews; 5) Globe and Mail; 6) National Post; 7) Toronto Star; 8) Local newspaper; 9) TVA (QC only); 10) TV5 (QC only); 11) La Presse (QC only); 12) Journal de Montreal (QC only); 13) Journal de Quebec (QC only); 14) Le Devoir (QC only); 15) Radio-Canada (QC only); 16) Rebel Media; 17) National Observer; 18) Toronto Sun; 19) The Tyee; 20) Post Millennial; 21) APTN; 22) True North News; 23) Press Progress; 24) Huffington Post; 25) Another outlet		
Offline Discussion	0-1; How often in the past week did you talk about politics or public affairs with the following people? 1) Your family; 2) Your friends; 3) Your co-workers (never, once, a few times, almost every day, daily, don't know)	0.40	0.26
Online Discussion	0-1; How many days in the past week did you talk about politics or public affairs with anyone via the Internet, including social media, e-mail, chat rooms, and micro-blogging sites? (never, once, a few times, almost every day, daily, don't know)	0.29	0.32
Pseudoscience Beliefs	0-1; 1) Some numbers are especially lucky for people; 2) Faith healing is a valid alternative to conventional medicine; 3) Homeopathy is a valid alternative to conventional medicine; 4) Some ancient civilizations were visited by extraterrestrials; 5) Some people possess psychic powers; 6) The position of the planets have an influence on the events of every day life (strongly, somewhat, neither agree/disagree)	0.41	0.24
Anti-intellectualism	0-1; Below is a list of groups and institutions in society. Please tell us the degree to which you trust or distrust members of these groups or institutions: 1) Experts; 2) Economists; 3) Scientists; 4) Doctors and medical professionals; 5) University professors (distrust a lot, distrust somewhat, neither, trust somewhat, trust a lot, don't know)	0.27	0.19
Science Literacy	0-1; 1) The center of the Earth is very hot; 2) The continents have been moving their location for millions of years and will continue to move; 3) All radio-activity is man-made; 4) Electrons are smaller than atoms; 5) Lasers work by focusing sound waves; 6) It is the father's gene that decides whether the baby is a boy or a girl; 7) Antibiotics kill viruses as well as bacteria.	0.61	0.25
Ideology	0-1; 1) The government should take measures to reduce differences in income levels; 2) Protecting the environment is more important than creating jobs; 3) Canada should increase the number of immigrants it admits each year; 4) People who don't get ahead should blame themselves, not the system; 5) The government should see to it that everyone has a decent standard of living (Strongly, somewhat, neither agree/disagree). Each item coded in left-wing (-1) and right-wing (1) direction. Don't knows and neither coded as neutral (0)	0.42	0.22

The two following regression tables provide the full coefficient list for Figures 2 and 3 in the main body of the paper.

Table A3. OLS Estimates for Figure 2

	Misperceptions	Risk	Compliance
	1	2	3
Social Media	0.22***	0.05**	-0.12***
	(0.02)	(0.02)	(0.03)
News Exposure	-0.12***	0.02	0.28***
	(0.03)	(0.03)	(0.03)
Offline Discussion	-0.03	0.07***	-0.03
	(0.02)	(0.02)	(0.03)
Online Discussion	0.04**	0.06***	0.01
	(0.02)	(0.02)	(0.02)
Income	0.01***	0.00	-0.01***
	(0.00)	(0.00)	(0.00)
Education	-0.02***	-0.00*	0.02***
	(0.00)	(0.00)	(0.00)
Age	-0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)
Religiosity	0.06***	0.02***	-0.02***
	(0.00)	(0.00)	(0.00)
Urban Density	0.00	0.00	-0.01***
	(0.00)	(0.00)	(0.00)
French	0.02	-0.08***	-0.10***
	(0.02)	(0.03)	(0.03)
Female	-0.04***	0.00	0.10***
	(0.01)	(0.01)	(0.01)
Constant	0.22***	0.05**	-0.12***
R^2	0.31	0.11	0.22
N	2022	2022	2022

Note: Standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01. Controls for region (Atlantic, Quebec, Ontario, West)

Table A4. OLS Estimates for Figure 3

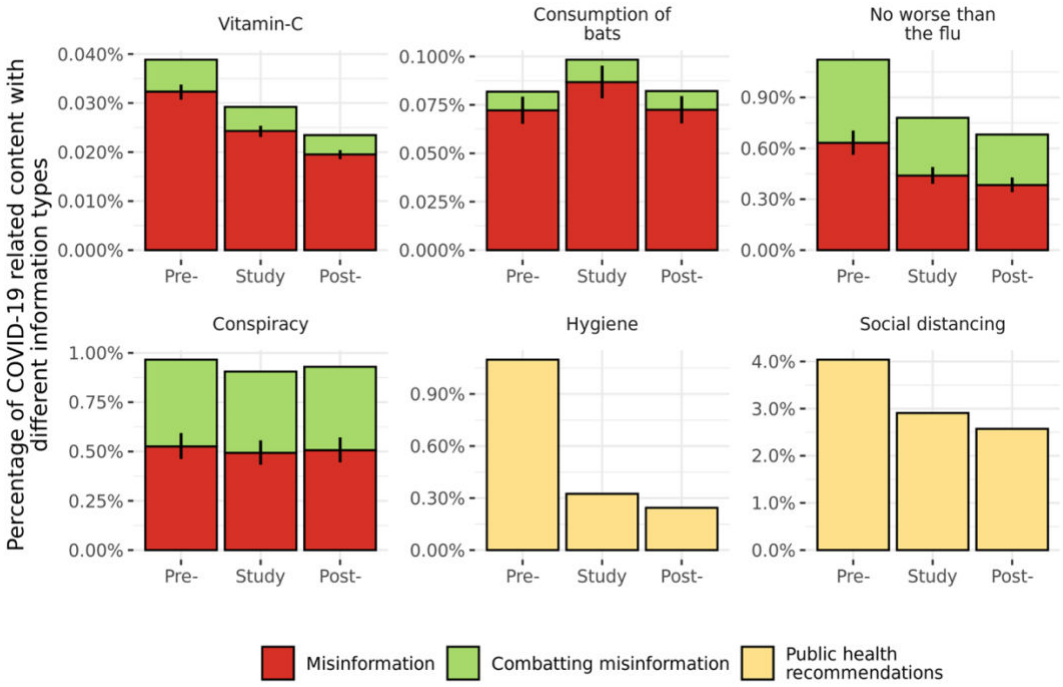
	Risk	Social Distancing	Risk	Social Distancing
	1	2	3	4
Misperceptions	-0.04*	-0.39***	-0.08***	-0.34***
	(0.02)	(0.03)	(0.03)	(0.04)
Social Media	0.06**	-0.04	0.04*	-0.03
	(0.02)	(0.03)	(0.02)	(0.03)
News Exposure	0.02	0.23***	-0.00	0.20***
	(0.03)	(0.03)	(0.03)	(0.03)
Offline Discussion	0.07***	-0.04*	0.06***	-0.06**
	(0.02)	(0.03)	(0.02)	(0.03)
Online Discussion	0.06***	0.02	0.05***	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Income	0.00	-0.01*	0.00	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Education	-0.01**	0.01***	-0.01**	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Age	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Religiosity	0.02***	0.00	0.02***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Urban Density	0.00	-0.01***	0.00	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
French	-0.08***	-0.09***	-0.09***	-0.08***
	(0.03)	(0.03)	(0.02)	(0.03)
Female	0.00	0.09***	-0.01	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)
Pseudoscience			0.07***	-0.01
			(0.02)	(0.03)
Anti-intellectualism				-0.09***
			(0.03)	(0.03)
Science Literacy			-0.05***	0.01
			(0.02)	(0.02)
Ideology			-0.14***	-0.00
			(0.02)	(0.02)
Constant	0.77***	0.63***	0.88***	0.69***
R^2	0.11	0.29	0.16	0.30
N	2022	2022	2022	2022

Note: Standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01. Controls for region (Atlantic, Quebec, Ontario, West)

3. Supplementary analysis

3.1. Time window of social media analysis

The misinformation research community is always shooting at a moving target as misinformation evolves quickly (Shin et al., 2018). A key challenge our design faces is that by the time a survey is designed and fielded, any misperception questions may be somewhat less relevant. To evaluate this, we looked at a similar corpus of tweets that included 6.9 million COVID-19 tweets collected from January 1st to April 16th. The results shown in Supporting Figure 1 indicate that while there are some differences across periods under examination, the misinformation themes examined during the March 26 to April 3 period were not vastly different than in pre- or post-periods.



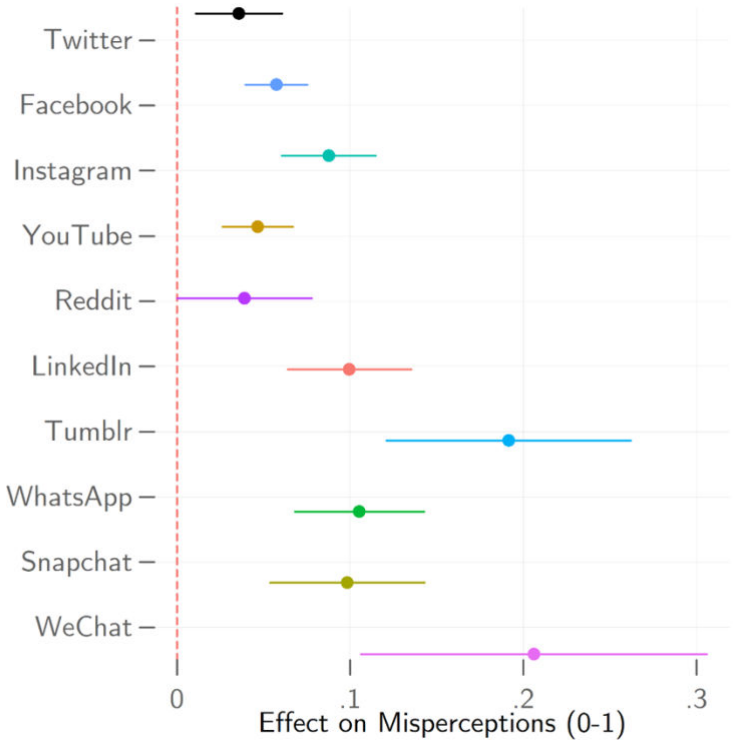
Supporting Figure 1. Relative frequencies of COVID-19 related information. *The percentage of all COVID-19 related content that includes reference to a particular theme in the pre-March 26, March 26 – April 3, and post-April 3 periods. Confidence intervals are shown only for the tweets where a random sample was manually categorized, and the results extrapolated to all tweets in the corpus. Note that the manual coding was done only during the Study period.*

3.2. Individual social media platforms and misperceptions

We are interested in the link between social media exposure and COVID-19 related attitudes and behaviours broadly, not just exposure to Twitter. To the extent that misinformation is circulating on other platforms, classifying respondents who are exposed to other social media platforms as untreated (i.e. 0) will bias our coefficients toward zero. We do not believe this is theoretically or methodologically defensible. We have some reason to believe the prevalence of misinformation on Twitter (and Facebook) is lower than other social media platforms. These platforms have been the most visible ones in the misinformation conversation, and they have been the ones to take measures to try to limit misinformation.

We also provide evidence below of an association between exposure to *each individual* social media platform and COVID-19 misperceptions after controlling for traditional news exposure, offline and

online political discussion, as well as controls for socio-economic status and demographics. The association between exposure and misperceptions is actually weaker for Twitter than for other platforms. If these results provide any indication of the amount of misinformation circulating on social media, it suggests our Twitter analysis is a conservative estimate.



Supporting Figure 2. *Estimated effects of social media application exposure on COVID-19 misperceptions. Note: controls for traditional news exposure, offline and online political discussion, income, education, age, gender, religiosity, language, and region. 95% confidence intervals.*