Supplementary Materials for

Social media conversations reveal large psychological shifts caused by COVID-19’s onset across U.S. cities

Ashwini Ashokkumar* and James W. Pennebaker

*Corresponding author. Email: ashwinia@utexas.edu

Published 22 September 2021, Sci. Adv. 7, eabg7843 (2021)
DOI: 10.1126/sciadv.abg7843

This PDF file includes:

Supplementary Text
Figs. S1 to S11
Tables S1 to S3
References
I. Note on ethical considerations

The survey and the Reddit project were reviewed by the authors’ institution’s Institutional Review Board and deemed as an exempt study. The analysis of social media data raises some important ethical questions given that social media posts are not originally intended for research (57). For instance, researchers have argued that even when identifying information such as names and emails are removed from the posts, it is often possible to trace the posts back to users (58). Some of these concerns are mitigated because of Reddit’s design. First, unlike Facebook and Instagram posts, which are viewable only by the audience selected by the poster, Reddit posts are public by default. Anyone with access to the internet can view all Reddit posts even without a Reddit account. All the Reddit data analyzed is already publicly viewable, which lessens concerns about privacy. Second, almost all Reddit users use user handles unrelated to their names, which means it is pretty much impossible to link posts with authors’ identities. To further mitigate privacy concerns, the Reddit dataset that has been made public includes neither people’s user handles nor the comments they posted. Only the LIWC variables corresponding to users’ comments are made publicly available. Also, all identifiable information from the survey data have been removed from the publicly available dataset. The authors are willing to share user handles and texts to university researchers with clearance from their university’s IRB.

II. Notes on the text analysis methodology

Table S1. Example words and sentences for each LIWC category

| Dimension              | Example words                                      | Example sentences from the dataset                                      |
|------------------------|----------------------------------------------------|-------------------------------------------------------------------------|
| References to COVID    | COVID, COVID19, coronavirus                        | “WHO has declared covid-19 outbreak a pandemic”                         |
| Anxiety                | anxious, worry, fearful                           | “I’m freaking out. I am panicking. Do you know what anxiety is?”        |
| Sadness                | sad, grief, crying                                | “oh no, this is tremendously sad. I’m so sorry for your loss”           |
| Anger                  | hate, annoyed, angry                              | “I hate these fucking politicians. Majority of them are fucking liars”  |
| Positive Emotion       | happy, yay, excited                               | “This made me smile and laugh :)”                                       |
| Analytic Thinking      | Note: analytic thinking is a summary variable of multiple dictionaries. | “A crumbling inflated economy created by false hopes of better healthcare, social services, and fair tax laws” |
| Cognitive Processing   | because, maybe, think                             | “It almost certainly makes you immune. It would be extremely *extremely* unusual if it doesn’t” |
| Family                 | mom, grandma, wife                                | “I know approximately 15 families who have lost some combination of parents/grandparents/aunts/uncles/spouses” |
| Friends                | buddy, friend, mate                               | “Friend and neighbor walking partner was laid off yesterday.”           |
| City                   | New York, downtown, community                     | “People from suburban Chicago and LA just say Chicago or LA so suburban Philly should just say Philly” |
| Country                | the US, United States, America                    | “Which only applies to the US Army and the US Air Force. The national guard is neither of those” |
III. Comparing 2020 with 2019

The graphs below compare the temporal patterns of social connections of 2020 versus 2019. Note that parallel graphs for the emotion and cognitive effects are included in the article (see Figures 3 and 4). Statistics comparing each phase in 2020 with baseline months (January -February 2020) can be found in the article and in the next section of this document (SI-IV). Because the 2019 means largely coincide with the 2020 baseline, separate statistics comparing 2020 and 2019 are not provided but can be made available on request.

![Graphs comparing social connections of 2020 versus 2019](image)

Figure S1. Three-day moving averages of the percentage of words pertaining to various social groups in Reddit posts from 2020 vs. 2019. The vertical bands delimit the four temporal phases (baseline, warning, isolation, and normalization). Error bands represent bootstrapped confidence intervals. The bootstrapped confidence intervals are narrow, which makes the error bands hard to see.

IV. Additional Statistics not reported in the article

Table S2. Results of ANOVA statistics comparing the focal dimensions across the four phases (baseline, warning, isolation, and normalization). All effects are statistically significant at $p < .001$ unless mentioned otherwise. Effect sizes are reported in the article.

| Dimension | ANOVA Statistic |
|-----------|-----------------|
| Anxiety   | $F(3,178837.2) = 277.10^{***}$ |
| Sadness   | $F(3,178962.5) = 11.72^{***}$ |
| Anger     | $F(3,179664.7) = 21.57^{***}$ |
| Positive Emotion      | $F(3,179359.7) = 72.54$ *** |
|----------------------|-----------------------------|
| Analytic Thinking    | $F(3,179478.2) = 139.34$ ***|
| Cognitive Processing | $F(3,179650.6) = 39.67$ *** |
| Family               | $F(3,179557.1) = 26.10$ *** |
| Friends              | $F(3,179480.1) = 6.97$ ***  |
| City                 | $F(3,179574.4) = 242.93$    |
| Impersonal references to the US (e.g., “United States”) | $F(3,178127.3) = 7.18$ ** |
| Personal references to the US (e.g., “our nation”) | $F(3,178477.8) = 2.08$, $p = .24$ |

** indicates $p < .01$. *** indicates $p < .001$

V. Within-Subjects Analyses: Analyses of only people who posted both before and during COVID

To ensure that the detected effects are not driven by changes in the cohorts of people who posted on Reddit before and during the pandemic, we re-conducted the analysis for only the people who posted at least once at baseline (i.e., before February 25, 2020) and once during any of the other three phases (i.e., between February 25 and May 25, 2020). The analysis included 571,957 posts from 49,278 people. The analysis was conducted on 500 bootstrapped stratified subsamples each of which included 10,000 randomly selected people per city in 2020. As can be seen from Figure S2, the effects look identical to the analyses reported in the main article. Statistics are reported below the figure.
Figure S2. Within-subject analysis of the psychological unfolding of the COVID outbreak. The figure represents three-day moving averages of the percentage of people referencing COVID (top-left corner), and scores corresponding to emotions (top-right corner), cognition (second row) and social connections (bottom row). The vertical bands delimit the four temporal phases: baseline, warning, isolation, and normalization. Error bands represent bootstrapped confidence intervals. Note that the bootstrapped confidence intervals are narrow, which makes the error bands hard to see.

i. **Attention to COVID.** During the baseline months (i.e., January 10, 2020 to February 24, 2020) fewer than one percent of the posts (M=0.39%) mentioned COVID. There was a sharp increase in mentions of COVID during the warning period (M = 5.93%). References to COVID peaked when it was declared a pandemic and national emergency in March (M = 12.11%), with a slight drop occurring thereafter (M=11.06%).

ii. **Emotions.** Consistent with the findings reported in the main article, natural language indicated increases in anxiety ($d_{\text{warning}} = .09$, $d_{\text{isolation}} = .14$, $d_{\text{normalization}} = .07$) and sadness ($d_{\text{warning}} = -.002$, $d_{\text{isolation}} =$...
.03, $d_{\text{normalization}} = .03$), and drops in positive emotion ($d_{\text{warning}} = -.04$, $d_{\text{isolation}} = -.11$, $d_{\text{normalization}} = -.07$) and anger ($d_{\text{warning}} = -.02$, $d_{\text{isolation}} = -.04$, $d_{\text{normalization}} = -.01$) in the days following COVID’s onset.

iii. **Cognition.** We found a drop in analytic thinking ($d_{\text{warning}} = -.04$, $d_{\text{isolation}} = -.14$, $d_{\text{normalization}} = -.09$) and an increase in cognitive processing words ($d_{\text{warning}} = .03$, $d_{\text{isolation}} = .08$, $d_{\text{normalization}} = .06$), paralleling the timelines reported in the article.

iv. **Social connections.** References to family increased in the days following the onset of the pandemic ($d_{\text{warning}} = .03$, $d_{\text{isolation}} = .05$, $d_{\text{normalization}} = .02$). There was a drop in talking about friends ($d_{\text{warning}} = -.01$, $d_{\text{isolation}} = -.03$, $d_{\text{normalization}} = -.02$; $p = .004$) and city ($d_{\text{warning}} = -.06$, $d_{\text{isolation}} = -.19$, $d_{\text{normalization}} = -.14$) but almost no change in references to country ($d_{\text{warning}} = .02$, $d_{\text{isolation}} = .03$, $d_{\text{normalization}} = .02$; $p = .01$).

The above graphs and statistics indicate that cohort effects do not drive the effects reported in the article. Rather, they are temporal changes induced by the pandemic’s onset.

VI. **Are the effects limited to posts that mention COVID?**

As depicted by Figure 2 in the article, about 10-11% of people explicitly mentioned COVID in their posts each day after the onset of the pandemic. This raises the possibility that the psychological effects reported in the paper are driven by posts about COVID. In order to test whether the reported effects extend beyond instances where people are talking about COVID, we reanalyzed the data after excluding all posts that reference words such as “COVID”, “coronavirus”, and “pandemic”. The analysis included 1,722,485 posts from 212,292 individuals. As can be seen from the figures below, the effects on people’s emotions, cognition, and social connections remain robust even in conversations that are not about COVID. Simply put, the psychological effects we found appear to be truly indicative of people’s experiences in their community during the pandemic.
Figure S3. Analysis of posts that do not reference COVID. The figure represents three-day moving averages of the effects of COVID on emotions (top-left corner), analytic thinking and cognitive processing (top-right and bottom-left corners) and social connections (bottom-right corner). The vertical bands delimit the four temporal phases (baseline, warning, isolation, and normalization). Error bands represent confidence intervals. Note that the confidence intervals are narrow, which makes the error bands hard to see.

VII. Are the effects driven by people who excessively talked about COVID?

About 20% of people explicitly mentioned COVID at least once during the pandemic. This raises the possibility that the psychological effects reported in the paper are driven by the small proportion of people who excessively discussed the pandemic. To rule out this possibility, the effects were examined separately for people who mentioned COVID at least once and those who never did. As shown in Figure S4, the effects and psychological timeline were remarkably similar across the two groups, barring some instances when a difference in magnitude of the effect was observed. Simply put, the psychological effects we found appear to are not driven by a small group of users who were disproportionately focused on the pandemic.
Figure S4. Analysis comparing people who talked about COVID versus those who never did once. The figure represents three-day moving averages of all the dimensions of interest. The vertical bands delimit the four temporal phases (baseline, warning, isolation, and normalization). Error bands represent confidence intervals.
VIII. Are the effects observable outside of city communities?

This project focused on city subreddits for two reasons. First, city-level subreddits reflect the daily issues and concerns of people living in identifiable geographic areas. Note that Reddit does not collect and/or provide GPS, IP addresses, or other location-based information about its users. Location-specific city subreddits are particularly helpful in assessing language associated with general location data on the spread of the virus (e.g., number of COVID cases in the city). Second, and more importantly, interactions in city subreddits span a range of topics relevant to the people living in the cities themselves (traffic, local politics, restaurants, weather, city celebrations, etc.). The bulk of Reddit communities are purely topic- or interest-based (e.g., gaming, relationships, pictures). Unlike topic-based subreddits, conversations that occur within city communities provide an opportunity for people to talk about shared experiences and concerns within their shared geographical space.

Because Reddit affords the opportunity for users to participate in an unlimited number of subreddits, we were able to identify all of the Reddit conversations that our city-level subreddit participants had in non-city subreddits. For example, we were able to download all the comments that people who contributed to, say, the Chicago subreddit posted on all the subreddits they participated in between January and May, 2020. Using this approach, it was possible to determine if people who expressed elevated anxiety in the Chicago subreddit also used more anxiety-related language when posting to subreddits dealing with cooking, World of Warcraft, and other groups. Such a strategy allows us to determine the degree to which the social and psychological language shifts in city subreddits generalize to others.

To address this question, we collected all the Reddit comments that 100,000 randomly selected people from our sample posted between January 10, 2020, and May 23, 2020. We obtained 20,636,872 comments. After excluding posts with fewer than 15 words, we had 11,624,050 comments (833,023 in city communities; 10,791,027 in other communities) from 100,000 people. These comments were analyzed using LIWC, and comments posted within-city were compared with those posted outside city communities (see Figure S5 below).

The timeline of attentional focus on COVID was similar across city and other communities, but COVID was more salient in city communities. Specifically, as shown in the first row of Figure S5, the percentage of people talking about the pandemic were largely similar across city and non-city communities except in the normalization phase wherein a higher percentage of people in city communities continued talking about the pandemic. At the same time, posts in the city subreddits made more references to COVID than in the non-city subreddits on average, indicating that topics relating to the pandemic were more salient in city communities. This is not surprising because a lot of the non-city communities pertained to specific topics such as gaming, sports, or jokes.

The emotional effects reported in the article were observed in non-city communities, but they were weaker than in city communities (see rows 2 and 3 in Figure S5). There was a noticeable increase in anxiety and drop in positive emotion. The trend in sadness was weak but in the same direction as in the city subreddits.

Interestingly, the effects on cognition and social connections largely did not carry over to non-city communities except for references to friends. This could mean that the shift in context or topic afforded by non-city subreddits weakens the cognitive and social effects of the pandemic. In other words, becoming immersed in specific contexts (e.g., science, sports, or gaming) provides a respite from
the cognitive and social consequences of the pandemic. However, note that our analyses are not conclusive because there are variations in the specific non-city subreddits that people visit before versus during the pandemic, which may confound our temporal analyses. For example, people flock to informational subreddits such as r/Coronavirus, r/news, r/worldnews, and r/personalfinance more after the onset of the pandemic relative to baseline. Such informational communities may put people in an analytic mindset, which may have inflated analytic thinking means during the pandemic. This reasoning is supported by the fact that informational subreddits’ language is generally higher in analytic thinking relative to more casual subreddits (e.g., advice) at baseline.
Figure S5. Psychological patterns observed over the four temporal phases across city versus non-city communities. Error bars represent confidence intervals.
IX. Weighted mean vs. Bootstrapping

The analyses reported in the main article were conducted on bootstrapped samples. This was done to account for the unequal sizes of the 18 city communities analyzed in this research. Each bootstrapped sample included randomly selected, equal-sized subsamples from each of the 18 cities. An alternate method is to weigh each city differently depending on its size. Figure S6 depicts weighted means of anxiety scores from January 10 to May 25, 2020. Comparing Figure S6 with Figure 3a in the main article (depicting bootstrapped means for anxiety) suggests that the two methods produce similar results.

![Image of Figure S6](image)

*Figure S6. Weighted means of LIWC scores measuring anxiety in Reddit posts from 2020. Error bands indicate 95% confidence intervals.*

X. Additional details on COVID growth rate analysis: Examining linear and exponential growth rates of infection

To examine the link between the spread of the virus and people’s psychological states, we tested whether daily growth rates of infections predicted the observed psychological shifts. We examined the growth rate of infections, as opposed to deaths, because only few deaths occurred in some of the cities during the early days of the pandemic (e.g., Portland). We accessed daily cumulative COVID infection counts at the county and national levels from the NYT covid case count database. City-level and national growth rates were computed in two ways following recommended practices (49).

First, we computed the exponential infection growth rate accounting for an exponential model of disease spread using the formula \( \ln(I_t/I_{t-1}) \), where \( I \) is the cumulative number of infections (in a city or in the US) and \( t \) is the time (day). Second, we computed a linear growth rate \( (I_t - I_{t-1}) / I_{t-1} \). In both cases, the growth rate was assumed to be 0 during baseline (prior to the first COVID case in a city or the US). Growth rate on the first day of recorded cases in each city or at the national level (i.e., marking an increase from 0 to an integer) was set to 0 since given our growth rate formula, growth rate on that day would be undefined. Finally, we lagged the growth rate by one day because it presumably takes a day for information regarding the number of cases to reach the public. There were three instances wherein we found a drop in the cumulative number of cases in a city between one day and the next, which likely reflect errors in reporting (two instances in Atlanta; one in Seattle). In these instances, our city growth
rate formula yielded negative numbers, which we replaced with 0s. The analyses reported in the article were conducted using the exponential growth rates, but below, we report analyses of both linear and exponential growth rates.

We computed daily means of the language dimensions across all users in each city subreddit. The resulting city-level dataset had 2430 daily means of language dimensions and daily national- and city-level growth rates. To account for the nested nature of the data (i.e., interdependence of all datapoints within cities), mixed effects models accounting for the random effect of cities’ intercepts were examined. In each model, we entered national and city level growth rates as predictors and a psychological dimension (e.g., anxiety) as the outcome. The outcome variables were standardized. In other words, the models tested the extent to which city- and national-level growth rates predicted shifts in the focal psychological dimensions.

We conducted separate analyses using the two growth rate measures (exponential and linear). Table S3 reports coefficients and total explained variance from these analyses. Both national and city level growth rates in infection amplified the psychological effects of the pandemic such that the effects were more pronounced following days when the virus spread was greater. City and national level growth rates generally accounted for independent variance, and neither was consistently more predictive than the other. (In the exponential growth rate analysis, city-level growth rate was more predictive than national growth rate, but this pattern did not fully hold up in the linear growth rate analysis.) Note that exponential growth rate generally accounted for more variance ($R^2$) than linear growth rate across the psychological dimensions.

Table S3. Coefficients and total explained variance in mixed effects models simultaneously testing the effects of city- and national-level growth rates of infection on daily means of psychological dimensions. The columns shaded in blue depict models that assumed an exponential model of infection spread, and the columns shaded in salmon depict models that assumed a linear model of infection spread. Intercepts were included as random effect to account for variations across cities’ means.

| Outcome Variable in Model | Exponential Infection Growth Rates | Linear Infection Growth Rates |
|---------------------------|------------------------------------|-------------------------------|
|                           | City level B | National level B | Total Variance Explained | City level B | National level B | Total Variance Explained |
| Attention to COVID        | 2.08***      | .59***            | 12.71%                    | .81***      | .60***            | 7.42%                    |
| Anxiety                   | 1.84***      | .1.36***          | 14.95%                    | .80***      | 1.15***          | 11.03%                   |
| Anger                     | -.84***      | -.03              | 1.70%                     | -.33***     | -.10             | 1.00%                    |
| Sadness                   | .27†         | -.15              | .15%                      | .06         | -.05             | .03%                     |
| Positive Emotion          | -1.00***     | -.66***           | 4.65%                     | -1.45***    | -.59***          | 3.26%                    |
| Analytic Thinking         | -1.70***     | -.37†             | 8.04%                     | -1.60***    | -.44**           | 4.03%                    |
| Cognitive Processing      | 1.02***      | -.04              | 2.44%                     | .37***      | .09              | 1.23%                    |
| Family                    | .83***       | .61***            | 2.99%                     | .32***      | .52***           | 1.96%                    |
| Friends                   | -.41**       | -.09              | .48%                      | -.16*       | -.12             | .30%                     |
| City                      | -1.44***     | -.34*             | 5.83%                     | -1.52**     | -.39**           | 3.06%                    |
| Country                   | .29*         | .20               | .35%                      | .09         | .17              | .18%                     |

* indicates $p < .05$. ** indicates $p < .01$. *** indicates $p < .001$. † indicates $p <= .06$

XI. Additional Reddit analysis on shifts in social connections
Social connection with country. The findings regarding people’s social connections with their country were a bit complex. As reported in the article, survey respondents reported feeling less connected to their country than pre-COVID times (see Fig. 5a in the article). At the same time, as shown in Fig 5b in the article and Fig. S2 above, people made slightly more references to the US in their conversations in the weeks following the outbreak. Additional analyses were conducted to understand how much the increase in US references reflected conversations relating to news about COVID as opposed to feelings of connectedness. Separate analyses were conducted to track formal, impersonal references to the US (“the US”, “United States”) and more warm, personal references (“our country”, “America”). As indicated in the figure below, there was an increase in impersonal references to the US ($F(3,178127.3) = 7.18, p = .008; d_{\text{warning}} = .04, d_{\text{isolation}} = .02, d_{\text{normalization}} = .01$) but not personal references ($F(3,178477.8) = 2.08, p = .24; d_{\text{warning}} = .001, d_{\text{isolation}} = .0007, d_{\text{normalization}} = .01$).

Analysis of we-words. Of particular interest was the striking increase in people’s use of first-person plural pronouns such as we, us, our (Figure S8a). Once the shutdown began, people’s use of we-words almost doubled from baseline -- a phenomenon also found in the wake of the 9/11 terrorist attacks (10). We-words often indicate an implicit shift to a collective frame of reference. When people use words like “we”, they are implicitly referring to themselves as being part of a collective, all sharing the same experiences. In fact, separate co-occurrence analyses found that the words most associated with we-words were test, people, case, virus, and death, all words used when describing the collective COVID crisis. Examples include:

“Only if we had reliable models and abundant testing”
“our healthcare system”
“the government owes us nothing”
“we should have taken action”
“We want our life back”

A theoretically interesting (and vexing) comparison with we-word usage was the relative use of affiliation-related words (e.g., friend, together, group, team). Whereas we-words spiked during the
warning and isolation phases, Figure S8b reveals that affiliation words increased only modestly during this same time. Explicitly thinking about or referring to a group, is a different process than implicitly connecting with it. Our working hypothesis is that during times of stress, people naturally orient to others (59). In fact, there is some work to suggest that merely seeing the word “we” may naturally boost self-esteem and reduce stress (60). Use of we-words can be thought of as a form of linguistic life preservers in times of anxiety.

![Graph](image)

*Figure S8a and b.* Three-day moving averages of the percentage of first-person plural pronouns (left) and non-pronoun affiliation words (right) in Reddit posts from 2020 versus 2019. The vertical bands delimit the four temporal phases: baseline, warning, isolation, and normalization. Error bands represent bootstrapped confidence intervals. The bootstrapped confidence intervals are narrow, which makes the error bands hard to see.

### XII. Examining variations across cities

The temporal patterns observed were remarkably similar across the 18 cities. Figure S9 depicts the average daily percentage of people who mentioned the pandemic in each of the 18 cities across the four temporal phases. Figure S10 depicts the temporal patterns for all the variables examined in the article. Further, the consistent patterns across the 18 cities indicate that the psychological effects of the crisis may have been relatively uniform in cities across the country, which could be due to the widespread prevalence of health and financial threats and the wide accessibility of national news.
Figure S9. Daily attentional focus on COVID in each of the 18 cities. The increase in the percentage of people who paid attention to COVID was highest in New York City and Seattle. Mirroring the timeline of the crisis in these two cities, attention afforded to COVID dropped in Seattle after March but remained elevated in New York City.
XIII. Survey methodology and analysis

Methods. The survey included questions regarding participants’ daily social, health-related, and work-related behaviors, mental health, social media use, as well as beliefs about the virus and pandemic. The survey was disseminated on a public website (http://utpsy.org/covid19/index.html) and
via Prolific. On average, the snowball (website) sample had about 1332 responses each week and the Prolific sample had 243 responses each week. The minimum number of responses per week was 27 in the snowball (website) sample and 157 in the Prolific sample. After completing the survey, respondents received feedback about their levels of anxiety, social connections, and their coping behaviors. The survey was revised three times between March and May as we added new questions to capture changes occurring in the world. Version 1 of the survey was released on March 19, Version 2 on March 28, and Version 3 on April 30. The versions largely overlapped with each other (at least 80%). The measures used in the current research were kept consistent in the three versions. The questions and scales pertaining to the measures analyzed in this research can be found in section XII of this document.

Emotion analyses. Participants rated the extent to which they felt worried about various issues on a five-point scale (1 = Not at all; 5 = A great deal). Participants reported being most anxious about the health of their family members (M = 3.74, SD = 1.12) and about inadvertently spreading the virus to strangers (M = 3.50, SD = 1.26), followed by concerns about getting sick (M = 3.09, SD = 1.11) and dying themselves (M = 2.27, SD = 1.18). Although financial concerns relating to losing jobs (M = 2.02, SD = 1.31) and paying bills (M = 1.95, SD = 1.21) were less central than COVID-related concerns to the average participant in our sample, they were more concerning to participants from lower socioeconomic backgrounds (-.2 < rs < -.05).

Further, based on work by Roxane Silver, anxiety was expected to be higher among people who had high exposure to COVID-related information. In order to measure exposure to information related to the pandemic, participants were asked how much time they spent the previous day reading or watching or communicating information related to the pandemic (M = 2.28, SD = .74). We also measured the duration of time in the previous day that they felt depressed (M = 2.09; SD =1.26). The ratings for both these items were provided on a five-point scale (1 = 0 hours; 5 = 8 hours or more). Consistent with Silver’s findings, survey respondents who reported having higher exposure to virus-related information reported feeling most depressed, r(10871) = .21. All the above reported effects were robust when in separate analysis of the snowball and Prolific samples.

Social connections analyses. Participants also rated the extent to which the pandemic “influenced how socially connected you feel” with a range of groups. They provided ratings on a five-point scale (1 = Much less connected; 3 = About the same; 5 = Much more connected). We analyzed ratings pertaining to four groups: family (M = 3.28, SD = 1.07), friends (M = 2.57, SD = 1.14), city (M = 2.43, SD = 1.11), and country (M = 2.68, SD = 1.10). As mentioned in the main article, the findings suggest that the outbreak disrupted people’s broader social ties beyond their families.
The increased connection to family is not just because most people were living with their families. The survey respondents who reported living alone or living with friends or roommates (N = 1352) also felt slightly more connected than before to family and less connected to their friends and other groups.

Further, the analysis suggested that people’s sense of community weakened as the crisis progressed. We included a measure in which participants provided absolute ratings of closeness for a range of groups including friends, city, and country, and the average rating for these items (M = 2.90, SD = .85) decreased over the course of the crisis, r(3776) = -.08. Finally, participants rated the extent to which they believed that the pandemic would “tear us apart”, and in a separate item, the extent to which it would “bring people together”. The latter was reverse scored and summed with the former to obtain a single measure that was used in the analysis (M = -.29, SD = 1.71). Consistent with the finding regarding depleting connectedness over the course of the crisis, people increasingly predicted that the crisis would tear people apart, r(11076) = .12. Both of these correlations were significant in both the snowball and Prolific samples.
REFERENCES AND NOTES

1. S. Galea, R. M. Merchant, N. Lurie, The mental health consequences of COVID-19 and physical distancing: The need for prevention and early intervention. *JAMA Intern. Med.* **180**, 817–818 (2020).

2. J. Hamblin, *Is Everyone Depressed?* (The Atlantic, 2020); www.theatlantic.com/health/archive/2020/05/depression-coronavirus/611986/.

3. F. Manjoo, *Opinion | The Hidden ‘Fourth Wave’ of the Pandemic* (N. Y. Times, 2020); www.nytimes.com/2020/12/09/opinion/coronavirus-mental-health.html.

4. R. Rossi, V. Socci, D. Talevi, S. Mensi, C. Niolu, F. Pacitti, A. Di Marco, A. Rossi, A. Siracusano, G. Di Lorenzo, COVID-19 pandemic and lockdown measures impact on mental health among the general population in Italy. *Front. Psych.* **11**, 790 (2020).

5. K. M. Fitzpatrick, C. Harris, G. Drawve, Fear of COVID-19 and the mental health consequences in America. *Psychol. Trauma Theory Res. Pract. Policy* **12**, S17–S21 (2020).

6. D. Fancourt, A. Steptoe, F. Bu, Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: A longitudinal observational study. *Lancet Psychiatry* **8**, 141–149 (2021).

7. M. Luchetti, J. H. Lee, D. Aschwanden, A. Sesker, J. E. Strickhouser, A. Terracciano, A. R. Sutin, The trajectory of loneliness in response to COVID-19. *Am. Psychol.* **75**, 897–908 (2020).

8. E. A. Holman, R. R. Thompson, D. R. Garfin, R. C. Silver, The unfolding COVID-19 pandemic: A probability-based, nationally representative study of mental health in the United States. *Sci. Adv.* **6**, eabdi390 (2020).

9. M. T. Hawes, A. K. Szenczy, T. M. Olino, B. D. Nelson, D. N. Klein, Trajectories of depression, anxiety and pandemic experiences; A longitudinal study of youth in New York during the Spring-Summer of 2020. *Psychiatry Res.* **298**, 113778 (2021).
10. M. A. Cohn, M. R. Mehl, J. W. Pennebaker, Linguistic markers of psychological change surrounding September 11, 2001. *Psychol. Sci.* **15**, 687–693 (2004).

11. D. Garcia, B. Rimé, Collective emotions and social resilience in the digital traces after a terrorist attack. *Psychol. Sci.* **30**, 617–628 (2019).

12. J. W. Pennebaker, K. D. Harber, A social stage model of collective coping: The Loma Prieta Earthquake and The Persian Gulf War. *J. Soc. Issues* **49**, 125–145 (1993).

13. R. C. Silver, E. A. Holman, D. N. McIntosh, M. Poulin, V. Gil-Rivas, Nationwide longitudinal study of psychological responses to September 11. *JAMA* **288**, 1235–1244 (2002).

14. N. M. Jones, R. C. Silver, This is not a drill: Anxiety on Twitter following the 2018 Hawaii false missile alert. *Am. Psychol.* **75**, 683–693 (2020).

15. N. M. Jones, S. P. Wojcik, J. Sweeting, R. C. Silver, Tweeting negative emotion: An investigation of Twitter data in the aftermath of violence on college campuses. *Psychol. Methods* **21**, 526–541 (2016).

16. B. Rimé, D. Páez, N. Basabe, F. Martínez, Social sharing of emotion, post-traumatic growth, and emotional climate: Follow-up of Spanish citizen’s response to the collective trauma of March 11th terrorist attacks in Madrid. *Eur. J. Soc. Psychol.* **40**, 1029–1045 (2010).

17. R. Kochhar, *Unemployment Rose Higher in Three Months of COVID-19 Than it did in Two Years of the Great Recession* (Pew Research Center, 2020); www.pewresearch.org/fact-tank/2020/06/11/unemployment-rose-higher-in-three-months-of-covid-19-than-it-did-in-two-years-of-the-great-recession/.

18. J. M. Barry, *The Great Influenza: The Epic Story of the Deadliest Plague in History* (Viking, 2004).

19. O. Dyer, Trump claims public health warnings on COVID-19 are a conspiracy against him. *BMJ* **368**, m941 (2020).
20. C. Funk, A. Tyson, *Partisan Differences Over the Pandemic Response Are Growing* (Pew Research Center Science and Society, 2020); www.pewresearch.org/science/2020/06/03/partisan-differences-over-the-pandemic-response-are-growing/.

21. J. Roozenbeek, C. R. Schneider, S. Dryhurst, J. Kerr, A. L. J. Freeman, G. Recchia, A. M. van der Bles, S. van der Linden, Susceptibility to misinformation about COVID-19 around the world. *R. Soc. Open Sci.* **7**, 201199 (2020).

22. A. D. I. Kramer, J. E. Guillory, J. T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* **111**, 8788–8790 (2014).

23. B. Fischhoff, G. Wong-Parodi, D. R. Garfin, E. A. Holman, R. C. Silver, Public understanding of Ebola risks: Mastering an unfamiliar threat. *Risk Anal.* **38**, 71–83 (2018).

24. J. W. Pennebaker, R. L. Boyd, K. Jordan, K. Blackburn, *The Development and Psychometric Properties of LIWC2015* (University of Texas at Austin, 2015).

25. A. Bruns, T. Highfield, J. Burgess, The Arab Spring and social media audiences: English and Arabic Twitter users and their networks. *Am. Behav. Sci.* **57**, 871–898 (2013).

26. H. Carter, J. Drury, G. J. Rubin, R. Williams, R. Amlôt, Applying crowd psychology to develop recommendations for the management of mass decontamination. *Health Secur.* **13**, 45–53 (2015).

27. P. C. Rosenblatt, R. M. Anderson, P. A. Johnson, The meaning of “Cabin Fever”. *J. Soc. Psychol.* **123**, 43–53 (1984).

28. L. E. Smith, B. Duffy, V. Moxham-Hall, L. Strang, S. Wessely, G. J. Rubin, Anger and confrontation during the COVID-19 pandemic: A national cross-sectional survey in the UK. *J. R. Soc. Med.* **114**, 77–90 (2021).

29. T. Schmader, M. Johns, Converging evidence that stereotype threat reduces working memory capacity. *J. Pers. Soc. Psychol.* **85**, 440–452 (2003).
30. K. Klein, A. Boals, The relationship of life event stress and working memory capacity. *Appl. Cogn. Psychol.* **15**, 565–579 (2001).

31. A. Mani, S. Mullainathan, E. Shafir, J. Zhao, Poverty impedes cognitive function. *Science* **341**, 976–980 (2013).

32. S. Seraj, K. G. Blackburn, J. W. Pennebaker, Language left behind on social media exposes the emotional and cognitive costs of a romantic breakup. *Proc. Natl. Acad. Sci.* **118**, e2017154118 (2021).

33. D. Kahneman, *Thinking, Fast and Slow* (Macmillan, 2011).

34. B. Klein, A. B. Horn, R. Kraehenmann, M. R. Mehl, A. Ehlers, Early linguistic markers of trauma-specific processing predict post-trauma adjustment. *Front. Psych.* **9**, 645 (2018).

35. A. Boals, J. B. Banks, L. M. Hathaway, D. Schuettler, Coping with stressful events: Use of cognitive words in stressful narratives and the meaning-making process. *J. Soc. Clin. Psychol.* **30**, 378–403 (2011).

36. J. W. Pennebaker, M. R. Mehl, K. G. Niederhoffer, Psychological aspects of natural language use: Our words, our selves. *Annu. Rev. Psychol.* **54**, 547–577 (2003).

37. T. Q. Phan, E. M. Airoldi, A natural experiment of social network formation and dynamics. *Proc. Natl. Acad. Sci.* **112**, 6595–6600 (2015).

38. J. Drury, C. Cocking, S. Reicher, D. Stephen, The nature of collective resilience: Survivor reactions to the 2005 London bombings. *Int. J. Mass Emerg. Disasters* **27**, 66–95 (2009).

39. J. Zaki, Catastrophe compassion: Understanding and extending prosociality under crisis. *Trends Cogn. Sci.* **24**, 587–589 (2020).

40. S. K. Cohn, Pandemics: Waves of disease, waves of hate from the Plague of Athens to A.I.D.S. *Hist. J.* **85**, 535–555 (2012).
41. S. Iyengar, G. Sood, Y. Lelkes, Affect, not ideology: A social identity perspective on polarization. *Public Opin. Q.* **76**, 405–431 (2012).

42. W. D. S. Killgore, S. A. Cloonan, E. C. Taylor, D. A. Lucas, N. S. Dailey, Loneliness during the first half-year of COVID-19 lockdowns. *Psychiatry Res.* **294**, 113551 (2020).

43. S. C. Guntuku, G. Sherman, D. C. Stokes, A. K. Agarwal, E. Seltzer, R. M. Merchant, L. H. Ungar, Tracking mental health and symptom mentions on twitter during Covid-19. *J. Gen. Intern. Med.* **35**, 2798–2800 (2020).

44. K. N. Jordan, J. Sterling, J. W. Pennebaker, R. L. Boyd, Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proc. Natl. Acad. Sci.* **116**, 3476–3481 (2019).

45. J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, D. I. Beaver, When small words foretell academic success: The case of college admissions essays. *PLOS ONE* **9**, e115844 (2014).

46. K. J. Hsu, K. N. Babeva, M. C. Feng, J. F. Hummer, G. C. Davison, Experimentally induced distraction impacts cognitive but not emotional processes in think-aloud cognitive assessment. *Front. Psychol.* **5**, 474 (2014).

47. COVID-19 - Mobility trends reports (Apple, 2020); www.apple.com/covid19/mobility.

48. M. R. Mehl, S. Vazire, N. Ramírez-Esparza, R. B. Slatcher, J. W. Pennebaker, Are women really more talkative than men? *Science* **317**, 82–82 (2007).

49. P. Sharkey, The acute effect of local homicides on children’s cognitive performance. *Proc. Natl. Acad. Sci.* **107**, 11733–11738 (2010).

50. A. Gollwitzer, C. Martel, W. J. Brady, P. Pärnamets, I. G. Freedman, E. D. Knowles, J. J. Van Bavel, Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nat. Hum. Behav.* **4**, 1186–1197 (2020).
51. M. L. Kern, G. Park, J. C. Eichstaedt, H. A. Schwartz, M. Sap, L. K. Smith, L. H. Ungar, Gaining insights from social media language: Methodologies and challenges. *Psychol. Methods* **21**, 507–525 (2016).

52. S. C. Matz, J. J. Gladstone, D. Stillwell, In a world of big data, small effects can still matter: A reply to Boyce, Daly, Hounkpatin, and Wood (2017). *Psychol. Sci.* **28**, 547–550 (2017).

53. M. Kosinski, D. Stillwell, T. Graepel, Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci.* **110**, 5802–5805 (2013).

54. G. A. Bonanno, C. R. Brewin, K. Kaniasty, A. M. L. Greca, Weighing the costs of disaster: Consequences, risks, and resilience in individuals, families, and communities. *Psychol. Sci. Public Interest* **11**, 1–49 (2010).

55. P. Whitney, C. A. Rinehart, J. M. Hinson, Framing effects under cognitive load: The role of working memory in risky decisions. *Psychon. Bull. Rev.* **15**, 1179–1184 (2008).

56. Y. Yamada, D.-B. Ćepulić, T. Coll-Martín, S. Debove, G. Gautreau, H. Han, J. Rasmussen, T. P. Tran, G. A. Travaglino; COVIDiSTRESS Global Survey Consortium, A. Lieberoth, COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak. *Sci. Data* **8**, 3 (2021).

57. S. Mannheimer, E. A. Hull, Sharing selves: Developing an ethical framework for curating social media data. *Int. J. Digit. Curation* **12**, 196–209 (2017).

58. M. Zimmer, “But the data is already public”: On the ethics of research in Facebook. *Ethics Inf. Technol.* **12**, 313–325 (2010).

59. S. E. Taylor, Tend and befriend: Biobehavioral bases of affiliation under stress. *Curr. Dir. Psychol. Sci.* **15**, 273–277 (2006).

60. W. L. Gardner, S. Gabriel, L. Hochschild, When you and I are “we”, you are not threatening: The role of self-expansion in social comparison. *J. Pers. Soc. Psychol.* **82**, 239–251 (2002).