Methodology

The American Trends Panel survey methodology

The American Trends Panel (ATP), created by Pew Research Center, is a nationally representative panel of randomly selected U.S. adults. Panelists participate via self-administered web surveys. Panelists who do not have internet access at home are provided with a tablet and wireless internet connection. Interviews are conducted in both English and Spanish. The panel is being managed by Ipsos.

Data in this report is drawn from the panel wave conducted Aug. 31 to Sept. 7, 2020. A total of 9,220 panelists responded out of 9,810 who were sampled, for a response rate of 94%. This does not include two panelists who were removed from the data due to extremely high rates of refusal or straightlining. The cumulative response rate accounting for nonresponse to the recruitment surveys and attrition is 5%. The break-off rate among panelists who logged on to the survey and completed at least one item is 1%. The margin of sampling error for the full sample of 9,220 respondents is plus or minus 1.7 percentage points.

More information about this wave of the American Trends Panel can be found here.

Keyword analysis

Researchers conducted their keyword analysis on a collection of 9,220 raw responses to the following open-ended survey question:

Thinking about how things have been going in your life since the beginning of the coronavirus outbreak, in what ways, if any, have things been difficult or challenging for you personally?

Overall, 77% of respondents provided one or more examples of difficulties or challenges resulting from the pandemic, 16% skipped the question, and 6% said that the coronavirus pandemic has had no negative impact on their lives. Roughly equal shares of Democrats and Republicans provided a substantive response to the question.

To identify and remove the 6% of respondents who indicated they have not experienced any coronavirus-related difficulties, researchers developed a series of exact pattern matches (e.g., “none,” “no challenges”) and coded a sample of 250 responses to compare their own decisions with those made via pattern-matching. Researchers determined that the patterns successfully removed nonresponses with a recall of 99% and 100% precision.

The remaining responses that contained descriptions of negative impacts were then cleaned using a series of standard text pre-processing methods, including supervised spelling correction, removal of 339 common English "stop words," contraction expansion, and lemmatization (reducing words to their root form). Researchers identified 149 words and phrases (one-, two- and three-word phrases) that appeared in at least 100 responses. Researchers calculated how many responses a word or phrase appeared in as a weighted statistic.

Word usage was then examined across a variety of demographic and socioeconomic factors, including income, gender and political affiliation. A handful of words were used frequently by a particular demographic group (at least 5% of the group's responses) and used more frequently than by other groups (at least 50% more often).

In addition to the notable difference in the use of the word "mask" between Republicans and Democrats, Americans in the upper-income tier (with household incomes above roughly $119,400) were more likely to mention "travel" (13% did so, compared to 6% of Americans with lower incomes), "friend" (20% vs. 12%) and "work [from] home" (6% vs. 2%). Meanwhile, members of Generation Z were more likely to mention terms such as "online" (12% vs. 2%), "school" (14% vs. 7%) and "vacation" (6% vs. 2%) relative to older adults.

These demographic keyword differences were confirmed for statistical significance using logistic regressions that predicted the presence of each keyword while controlling for gender, employment, income, geographic region, urbanicity, political affiliation, generation, marital status, ethnicity, response language and logged word count.

Categorizing mask responses

Researchers also manually examined the 943 responses (an unweighted number) that mentioned the word "mask" or one of its variations (e.g., "masking") and coded these responses into two additional categories:

1. Respondent expresses frustration with other people:

    a. Not following safety precautions like wearing masks,
    b. Complaining about masks,
    c. Wearing masks incorrectly,
    d. Politicizing or being skeptical of safety precautions like wearing masks, and/or
    e. Not taking the virus seriously.

2. Respondent says that they:

    a. Choose not to follow safety precautions like wearing masks,
    b. View safety precautions like wearing masks as unnecessary, ineffective, oppressive or unfair,
    c. Are skeptical about COVID-19 in general, or feel that it's a hoax/conspiracy or is exaggerated/overblown, and/or
    d. Believe that COVID-19 is being used for political gain to manipulate people or strip them of their freedom.

A sample of 100 responses were double-coded to ensure acceptable inter-rater reliability. Cohen's Kappa was 0.94 for the first category and 0.89 for the second category.