

Methodology

This report uses two different research components and methodologies: a nationally representative survey of U.S. adults conducted through Pew Research Center’s American Trends Panel (ATP) and a content analysis of tweets.

This is the latest report in Pew Research Center’s ongoing investigation of the state of news, information and journalism in the digital age, a research program funded by The Pew Charitable Trusts, with generous support from the John S. and James L. Knight Foundation.

The American Trends Panel survey methodology

Overview

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 May 17 to May 31, 2021. The sample is comprised of panelists who indicated that they use Twitter on the Wave 85 survey conducted on the ATP in March 2021. A total of 2,548 panelists responded out of 2,643 who were sampled, for a response rate of 96%. The cumulative response rate accounting for nonresponse to the recruitment surveys and attrition is 4%. The break-off rate among panelists who logged on to the survey and completed at least one item is less than 1%. The margin of sampling error for the full sample of 2,548 respondents is plus or minus 3.4 percentage points.

American Trends Panel recruitment surveys

Recruitment dates	Mode	Invited	Joined	Active panelists remaining
Jan. 23 to March 16, 2014	Landline/ cell RDD	9,809	5,338	2,181
Aug. 27 to Oct. 4, 2015	Landline/ cell RDD	6,004	2,976	1,241
April 25 to June 4, 2017	Landline/ cell RDD	3,905	1,628	620
Aug. 8 to Oct. 31, 2018	ABS	9,396	8,778	5,893
Aug. 19 to Nov. 30, 2019	ABS	5,900	4,720	2,323
June 1 to July 19, 2020; Feb. 10 to March 31, 2021	ABS	3,197	2,812	2,442
	Total	38,211	26,252	14,700

Note: Approximately once per year, panelists who have not participated in multiple consecutive waves or who did not complete an annual profiling survey are removed from the panel. Panelists also become inactive if they ask to be removed from the panel.

PEW RESEARCH CENTER

Panel recruitment

The ATP was created in 2014, with the first cohort of panelists invited to join the panel at the end of a large, national, landline and cellphone random-digit-dial survey that was conducted in both English and Spanish. Two additional recruitments were conducted using the same method in 2015 and 2017, respectively. Across these three surveys, a total of 19,718 adults were invited to join the ATP, of whom 9,942 (50%) agreed to participate.

In August 2018, the ATP switched from telephone to address-based recruitment. Invitations were sent to a random, address-based sample of households selected from the U.S. Postal Service's Delivery Sequence File. Two additional recruitments were conducted using the same method in 2019 and 2020, respectively. Across these three address-based recruitments, a total of 18,493 adults were invited to join the ATP, of whom 16,310 (88%) agreed to join the panel and completed an initial profile survey. In each household, the adult with the next birthday was asked to go online to complete a survey, at the end of which they were invited to join the panel. Of the 26,252 individuals who have ever joined the ATP, 14,700 remained active panelists and continued to receive survey invitations at the time this survey was conducted.

The U.S. Postal Service's Delivery Sequence File has been estimated to cover as much as 98% of the population, although some studies suggest that the coverage could be in the low 90% range.¹ The American Trends Panel never uses breakout routers or chains that direct respondents to additional surveys.

Sample design

The overall target population for this survey was Twitter users ages 18 and older, living in the U.S., including Alaska and Hawaii. The sample consisted of 2,643 panelists who indicated that they use Twitter on the Wave 85 survey conducted on the ATP in March 2021.

Questionnaire development and testing

The questionnaire was developed by Pew Research Center in consultation with Ipsos. The web program was rigorously tested on both PC and mobile devices by the Ipsos project management team and Pew Research Center researchers. The Ipsos project management team also populated test data which was analyzed in SPSS to ensure the logic and randomizations were working as intended before launching the survey.

¹ AAPOR Task Force on Address-based Sampling. 2016. "[AAPOR Report: Address-based Sampling](#)."

Incentives

All respondents were offered a post-paid incentive for their participation. Respondents could choose to receive the post-paid incentive in the form of a check or a gift code to Amazon.com or could choose to decline the incentive. Incentive amounts ranged from \$5 to \$20 depending on whether the respondent belongs to a part of the population that is harder or easier to reach. Differential incentive amounts were designed to increase panel survey participation among groups that traditionally have low survey response propensities.

Data collection protocol

The data collection field period for this survey was May 17 to May 31, 2021. Postcard notifications were mailed to all ATP panelists with a known residential address on May 17.

Invitations were sent out in two separate launches: Soft Launch and Full Launch. Sixty panelists were included in the soft launch, which began with an initial invitation sent on May 17, 2021. The ATP panelists chosen for the initial soft launch were known responders who had completed previous ATP surveys within one day of receiving their invitation. All remaining English- and Spanish-speaking panelists were included in the full launch and were sent an invitation on May 18.

All panelists with an email address received an email invitation and up to four email reminders if they did not respond to the survey. All ATP panelists that consented to SMS messages received an SMS invitation and up to four SMS reminders.

Invitation and reminder dates

	Soft Launch	Full Launch
Initial invitation	May 17, 2021	May 18, 2021
1st reminder	May 21, 2021	May 21, 2021
2nd reminder	May 24, 2021	May 24, 2021
3rd reminder	May 26, 2021	May 26, 2021
Final reminder	May 28, 2021	May 28, 2021

Data quality checks

To ensure high-quality data, the Center's researchers performed data quality checks to identify any respondents showing clear patterns of satisficing. This includes checking for very high rates of leaving questions blank, as well as always selecting the first or last answer presented. As a result of

this checking, one ATP respondent was removed from the survey dataset prior to weighting and analysis.

Weighting

The ATP data was weighted in a multistep process that accounts for multiple stages of sampling and nonresponse that occur at different points in the survey process. First, each panelist begins with a base weight that reflects their probability of selection for

their initial recruitment survey (and the probability of being invited to participate in the panel in cases where only a subsample of respondents were invited). The base weights for panelists recruited in different years are scaled to be proportionate to the effective sample size for all active panelists in their cohort. To correct for nonresponse to the initial recruitment surveys and gradual panel attrition, the base weights for all active panelists are calibrated to align with the population benchmarks identified in the accompanying table to create a full-panel weight.

For ATP waves in which only a subsample of panelists are invited to participate, a wave-specific base weight is created by adjusting the full-panel weights for subsampled panelists to account for any differential probabilities of selection for the particular panel wave. For waves in which all active panelists are invited to participate, the wave-specific base weight is identical to the full-panel weight.

In the final weighting step, the wave-specific base weights for panelists who completed the survey are again calibrated to match the population benchmarks specified above. These weights are trimmed (typically at about the 1st and 99th percentiles) to reduce the loss in precision stemming from variance in the weights. Sampling errors and test of statistical significance take into account the effect of weighting.

For this wave, the sample was weighted to align with benchmarks that were estimated among all Twitter users who responded to Wave 85.

Weighting dimensions

Variable	Benchmark source
Age x Gender	Twitter users from ATP Wave 85
Education x Gender	
Education x Age	
Race/Ethnicity x Education	
Born inside vs. outside the U.S. among Hispanics and Asian Americans	
Years lived in the U.S.	
Census region x Metro/Non-metro	
Volunteerism	
Voter registration	
Party affiliation	
Frequency of Internet use	
Religious affiliation	

Note: These dimensions were also used to construct the Wave 85 weight. See the Wave 85 methodology statement for more details.

PEW RESEARCH CENTER

The following table shows the unweighted sample sizes and the error attributable to sampling that would be expected at the 95% level of confidence for different groups in the survey.

Group	Unweighted sample size	Plus or minus ...
Total sample	2,548	3.4 percentage points
Gets news on Twitter	1,783	4.0 percentage points
Does not get news on Twitter	751	6.1 percentage points

Sample sizes and sampling errors for other subgroups are available upon request. In addition to sampling error, one should bear in mind that question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of opinion polls.

Dispositions and response rates

Final dispositions	AAPOR code	Total
Completed interview	1.1	2,548
Logged onto survey; broke off	2.12	14
Logged onto survey; did not complete any items	2.1121	8
Never logged on (implicit refusal)	2.11	72
Survey completed after close of the field period	2.27	0
Completed interview but was removed for data quality		1
Screened out		0
Total panelists in the survey		2,643
Completed interviews	I	2,548
Partial interviews	P	0
Refusals	R	95
Non-contact	NC	0
Other	O	0
Unknown household	UH	0
Unknown other	UO	0
Not eligible	NE	0
Total		2,643
AAPOR RR1 = $I / (I+P+R+NC+O+UH+UO)$		96%

Cumulative response rate	Total
Weighted response rate to recruitment surveys	12%
% of recruitment survey respondents who agreed to join the panel, among those invited	69%
% of those agreeing to join who were active panelists at start of Wave W90	56%
Response rate to Wave W90 survey	96%
Cumulative response rate	4%

Analysis of Twitter users' tweets

The analysis of Twitter users' behaviors and activities on the site is based on a subsample of 1,026 respondents to the main survey who had previously agreed to share their Twitter handle for research purposes and provided a valid handle. Researchers from the Center used these handles to collect information about their on-site behaviors using the Twitter API.

The handles for these respondents were initially collected in a previous wave of the ATP fielded March 8 to March 14, 2021. These respondents were then included in the subsequent survey of all Twitter users that the main analysis in this report is based on.

Of the 2,548 adults who completed this survey, 1,099 (43%) had previously agreed to provide their Twitter handle. After the survey was fielded, researchers reviewed each account individually and removed any accounts that were suspended, invalid, or that belonged to institutions, products or international entities.

This final sample of 1,026 U.S. adult Twitter users with valid, active handles was weighted using an iterative technique that matches gender, age, race, place of birth among Hispanics and Asian Americans, years lived in the U.S., education, region, party identification, volunteerism, voter registration, and metropolitan area to American Trends Panel March 2021 (Wave 85) survey respondents who indicated in that survey that they use Twitter. The margin of error for the full sample is plus or minus 6 percentage points.

Content analysis of users' tweets

The findings in this post, examining the content of users' tweets, are based on tweets produced by respondents with public accounts during the time period of June 12, 2021 to Aug. 31, 2021. Of the 1,026 respondents that agreed to share their handle and provided a valid one, 620 had their accounts set to public (so researchers were able to access their tweets) and tweeted during the time period studied.

Throughout this time, there were over 140,000 tweets from the respondents with public accounts. Researchers sampled by selecting 40 random tweets from respondents who had tweeted more than 40 times, and all of the tweets from respondents who had tweeted 40 times or less during the time period (e.g., if a user tweeted 10 times, all 10 of those tweets would be sampled and coded; if a user had 100 tweets, 40 tweets would be randomly selected for coding). This allowed researchers to decrease the number of tweets analyzed by human coders, while still keeping a robust sample of tweets for each respondent. In the end, more than 12,000 tweets from 620 respondents were sampled, coded, and analyzed. Among the (unweighted) respondents, one-third tweeted five or

fewer times, one-third tweeted between six and 39 times, and 34% tweeted 40 or more times (and thus 40 of their tweets were sampled for coding).

The measures in this post are the average percent across respondents. For example, if Respondent A tweeted 10 times, and three of those tweets (30%) were about news, and Respondent B tweeted 40 times, and four of those tweets (10%) were about news, the average % of tweets about news between the two respondents would be 20%.

Human coders coded these tweets for four variables. First, they coded whether the tweet was about current news. If it was, they then coded the tweets for what topic the tweet was about, whether it referenced COVID-19, and whether it contained the author's opinion. Agreement was tested using Krippendorff's alpha.

Below are the main measures and their alphas:

- Tweet is about news: 0.70
- Tweet topic: 0.70
- Tweet contains a reference to COVID-19: 0.92
- Tweet includes the author's opinion: 0.84

Upon reaching an acceptable level of intercoder reliability, tweets were coded by four coders and a lead researcher. Throughout the coding process, staff discussed questions as they arose and arrived at decisions under supervision of the content analysis team leader. In addition, the lead researcher checked coders' accuracy throughout the process.

In addition to human coding, researchers used data from the Twitter API to indicate whether the tweet was original, a retweet, a reply or a quote tweet; and whether the tweet included a photo, video, link or hashtag.

Question wording and respondent consent language for collection of Twitter handles

Respondents were asked to provide their Twitter handles for research purposes using the following language:

Tweet frequency for content analysis sample

Number of tweets during time period studied	# of respondents	% of respondents
1 to 5	202	33%
6 to 10	65	10%
11 to 20	61	10%
21 to 39	80	13%
40 or more	212	34%
Total	620	100

Source: Analysis of tweets by 620 members of the American Trends Panel, June 12-Aug. 31, 2021.

PEW RESEARCH CENTER

In many of our surveys, we ask about how people like you use social media. But there's a lot we cannot learn unless you allow us to connect information about your public Twitter behavior to your answers.

No matter how often you use Twitter, would you be willing to share your Twitter handle to help us with this research?

- 1 Yes
- 2 No

This project will ONLY collect information that an ordinary user would see when they visit your profile. We will treat this data just like we treat your survey responses and will NEVER share anything that can be linked back to you.

How we will use your Twitter handle

If you provide us with your Twitter handle, we will only be able to collect what an ordinary user would see when visiting your profile. The specific information we will have access to depends on the privacy settings of your Twitter account.

If your account is set to PUBLIC, we will be able to see the accounts you follow; the accounts that follow you; the text of any tweets you have posted; and basic information in your profile, such as your description and the age of your account. We will NOT have access to any private messages you have sent or received.

If your account is set to PRIVATE, we will be able to see how many accounts you follow or how many times you have tweeted – but will NOT be able to see any private messages you have sent or received; the actual content of your tweets; or the names of the accounts you follow or that follow you.

We expect that this research will be concluded in approximately 18 months. At that point we will ask you again whether or not you would like to provide your handle. If you provided your Twitter handle to us previously, that is why you are seeing this question again.

If at any point you wish to opt out of research related to your Twitter account, contact info@americantrendspanel.org. For more information about our privacy practices, please see our Privacy Policy.