Methodology

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 17-31, 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.

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

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

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.

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.

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

Data collection protocol

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

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, 2021.

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
First reminder	May 21, 2021	May 21, 2021
Second reminder	May 24, 2021	May 24, 2021
Third 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.

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

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.

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

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 on to survey; broke off	2.12	14
Logged on to 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 behavioral data

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 8-14, 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. The analysis of Twitter behavioral data in this report is based on the 1,026 respondents who both completed the survey and had a valid, active handle at the time of the study.

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 United States, education, region, party identification, volunteerism, voter registration, and metropolitan area to American Trends Panel March 2021 (Wave 85) survey respondents who provided their Twitter handles for research purposes. The margin of error for the full sample is plus or minus 6 percentage points.

Collecting tweets from survey respondents

The findings in this report that examine users' patterns of political posting and engagement are based on tweets produced by respondents whose accounts were set to "public" during the time period of May 1, 2020, to May 31, 2021. All tweets posted by these public accounts during this time frame were collected using the Twitter API, resulting in a total of 959,254 original tweets, replies, quote tweets and retweets from 942 users.

Identifying political tweets from our sample of users

To identify which tweets featured political content, researchers from the Center used a custom-trained machine learning classifier. This model was fine-tuned on a training set of 6,413 tweets that had been hand-coded according to whether or not they referred to political content such as officials and activists, social issues, or news and current events.

To increase the model's generalizability, the training data used was posted by users not in the American Trends Panel. This training set was selected by randomly sampling tweets from the Twitter Powertrack historical search API, filtered to posts originating from within the U.S. between May 2020 and May 2021. These 6,413 sampled tweets were then labeled by a team of seven annotators (Krippendorff's alpha = 0.76) for the presence or absence of recognizable political content, using the following codebook:

> Political content on Twitter is varied, and can be about officials and activists, social issues, or news and current events.
>
> Looking at the tweet displayed here, would you categorize it as POLITICAL or NOT POLTICAL content?
>
> Pick what you think is the best option, even if you're not entirely sure.

Where multiple coders disagreed on the appropriate label for a tweet, the ultimate label was decided using a modified Dawid-Skene aggregation model.[5]

Using this training set, researchers then fine-tuned a machine learning model, based on the DistilRoBERTa architecture in the Transformers library for Python. While the training data came from outside the American Trends Panel, the trained model was validated against a test set of 1,082 tweets sampled from the 959,254 tweets that were posted by the respondents who volunteered a handle for this study. The tweets in this test set were each coded for political content by three human annotators (Krippendorff's alpha = 0.82), with labels decided by majority rule in the case of disagreement. Evaluated against this test set, the trained model achieved an accuracy score of 0.88 and an F1 score of 0.91.

Once trained, this model was used to categorize all remaining uncoded tweets for the presence or absence of political content. The study excluded non-English tweets.

Distinct keywords of political tweets

Researchers also conducted a distinct keyword analysis using the complete set of 959,254 tweets posted by the respondents who volunteered a handle for this study.

Text from each document (tweet) was converted into a set of features representing words and phrases by applying a series of pre-processing functions to the text of the tweets. First, "stop words," including common English words, were removed. Then, the text of each post was

converted to lowercase, and common contractions were expanded into their constituent words. Punctuation was removed, and each sentence was tokenized using the resulting white space. Finally, words were lemmatized (reduced to their semantic root form) and filtered to those containing three or more characters. Terms were then grouped into one-, two- and three-word phrases.

Distinctive keywords and phrases used in tweets with political and nonpolitical content were identified using pointwise mutual information. Researchers then calculated the proportion of political tweets referencing each distinctive term (phrase) as well as common variations.