Analysis and Results of Time Reference Web Probes on the RANDS COVID-19 Survey

Kristen Cibelli Hibben, Valerie Ryan, Travis Hoppe, and Paul Scanlon

Summary

How do people interpret questions about when the Coronavirus pandemic began? Using the RANDS during COVID-19 Round 1 survey, respondents were asked when they thought the Coronavirus pandemic began; respondents did not generally agree on a date. Most respondents thought the pandemic started sometime between November 2019 and March 2020, while others gave more than one date. Some respondents thought this question was about when the first cases appeared, while others thought it was about when the government declared a pandemic. Some respondents provided different dates for different countries.

Respondents were also asked when the pandemic first affected their life. Most people said it first affected them in March 2020. These responses were similar across age, education, and race and ethnicity.

Results indicate that adding a month and year (March 2020) to survey questions that use the words "since the Coronavirus pandemic began" may be useful. This is suggested because respondents provided a wide range of answers for when they thought the pandemic began. For survey questions to work well, respondents have to understand the question in the same way.

Background

The Division of Research and Methodology (DRM) of the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention has been conducting a series of surveys using commercially recruited panels termed the Research and Development Survey (RANDS)[1]. RANDS is an ongoing series of surveys designed to explore the feasibility of using these panels to collect information on health outcomes and to augment NCHS' question evaluation and research program with mixed and novel methodologies for detecting and understanding sources of measurement error. In response to the coronavirus disease 2019 (COVID-19) pandemic, two additional rounds of RANDS were implemented during June – August, 2020. These surveys were used to collect information on COVID-19-related health issues and other pandemic-related topics.

NORC at the University of Chicago conducted the first round of RANDS during COVID-19 from June 9, 2020 to July 6, 2020 using NORC's AmeriSpeak® Panel and the Dynata Panel as sample sources. AmeriSpeak® is a probability-based panel representative of the US adult, English-speaking, non-institutionalized household population maintained by NORC. This panel is selected using multistage probability sampling involving stratification and clustering.[2] RANDS during COVID-19 was administered to a sample of AmeriSpeak panelists via online web survey (n = 12,610) or phone interviews (n = 410), based on the preferred survey administration mode of each panelist. In addition to the probability-

[1] https://www.cdc.gov/nchs/rands/index.htm

[2] Data users can consult the following website for more information: https://www.amerispeak.org/

based Amerispeak sample, a non-probability online only opt-in sample from the Dynata Panel was included to augment the overall sample size.[3] Both the AmeriSpeak and Dynata data are included in this analysis. As a result, the data are unweighted and the complex design of the AmeriSpeak sample is not taken into account in this analysis. Systematic exploration between the panel types is ongoing but is not included in this report.

The first round of RANDS during COVID-19 focused on health care during the pandemic, impacts on health (mental and physical), financial security, understanding of quarantine and isolation, and behavioral measures.

Because many of the concepts surrounding the COVID-19 pandemic and the public's response to the outbreak were new, it was not possible to conduct extensive and systematic cognitive evaluation of new items before they were fielded. This was due to the urgency of collecting data and producing statistics on coronavirus-related topics in a timely manner. Furthermore, in-person cognitive interviews were impossible and virtual cognitive interviews were not yet available to NCHS before RANDS during COVID-19 was administered. Therefore, in order to understand how these new, coronavirus-related survey items functioned—and to determine what constructs they captured—RANDS included a series of open- and close-ended web probes and experimental designs (e.g., split-ballot experiments comparing open-ended versus closed-end probes), which provided data for *post hoc* evaluations. Web probing is an emerging technique for the evaluation of survey questions, which consists of the use of either closed or open-ended follow-up questions following a target survey question or questions (Geisen & Murphy, 2019). While typically practiced in the context of self-administered web surveys, as noted above, the sample for the current study includes a small number of telephone interviews where respondents also received the probes.

It is important that survey questions be clear about "what time period" is covered by the question (Groves et al. 2011, pp. 244). If the time period is not clearly specified or if it is not consistently understood, responses may vary for that reason, reducing data reliability (Fowler and Cosenza, 2013). This report discusses the analysis and results of two web probes related to the time referenced, to shed light on the use of the phrase "Since the Coronavirus pandemic began" as a time reference in survey questions (e.g., "Since the Coronavirus pandemic began, have you been able, unable, or have not needed…To get medications? To get a doctor's appointment or some other kind of healthcare?").

The wording of the probes, both of which were open-ended, is as follows:

1.) When do you think that the Coronavirus pandemic began? Your best guess is fine.
2.) When did the Coronavirus pandemic first affect your daily life? Your best guess is fine.

These probes were included to examine how consistently people understood the reference period indicator when "the Coronavirus pandemic began", the extent to which their response aligns with when they think it first affected their daily lives, the difference between when people think the pandemic began and when it started affecting their lives, and the extent to which these results may differ by respondent age, education level, or race and ethnicity.

The hypothesis is that respondents will understand when the Coronavirus pandemic began fairly consistently and that there will be substantial overlap between when respondents thought the

[3] For information on NORC's combination methodology for the two panels see the following website: https://amerispeak.norc.org/our-capabilities/Pages/TrueNorth.aspx

pandemic began and when it began affecting their lives. Evidence of a relatively uniform understanding of when the pandemic began would support its use as a time reference point in survey questions. Overlap between when the pandemic began and when it began affecting respondents' lives would lend further support for when the pandemic began as a time reference, as it would suggest that it is also likely to be salient and thus memorable to respondents.

A third probe was also included immediately following the two probes outlined above: "Why do you say that?" to further understand how the Coronavirus pandemic first began to affect people's lives (Probe 2). Data from this probe will be qualitatively analyzed using topic modeling and other computer-assisted approaches, in combination with hand coding, and reported on in a separate follow-up report.

Methods

The data for both probes is open text. While there was some discussion of using date fields, it was decided to keep these open-ended to enable respondents to report a range of responses, from specific dates to broader timeframes – such as "Fall 2019". Nonetheless, because many responses to the first two probes were dates, a rule-based machine learning approach was developed to automate the matching and cleaning of the data as much as possible. A Python script was written building on an existing date-detection program (or "date matcher"). Rules were then added to modify and, in some cases, correct the results of the date-detecting program. For example, the date matcher initially assigned all responses with just a month such as "October" to the current year (2020). However, it could be inferred that months after when the survey was fielded (June-July 2020), such as "October" referred to that month of the prior year (e.g., "October 2019"). Dates were resolved to the closest reported day, month, and year (e.g., "December 1, 2019" was converted to "12/01/2019").

Advanced rules were added to handle responses that included more than one date (e.g., "I think it started in November or December") or referred to a broader time period such as "Fall 2019" or "Winter 2019". All but approximately 500 cases were coded using the date matcher and customized Python script. See Table 2A and Table 2B in the Appendix for more information on Python coding rules.

A full manual review of the data was subsequently carried out to resolve relative dates (e.g., "4months ago") based on the date of the interview as well as responses the date matcher did not handle correctly. For example, in one case, the date matcher identified two dates, but the respondent was only reporting one date: "In **March**. My place of work was forced to close entirely, so I and all of my co-workers lost our jobs in **March**." [emphasis added]. The date matcher also had trouble parsing some responses with dashes between months and/or misspellings, assigning December 2019 to this response: "October - Nobember of 2019". See Table 3A in the Appendix for more information on hand coding rules.

Given that the analysis presented in this report focuses on individuals' responses, and that the data being analyzed is reliant on researcher coding instead of direct responses, the findings presented here are unweighted. The data come from Round 1 of RANDS During COVID-19. Complete cases were used in the analysis. Table 1A in the appendix shows the characteristics of the sample by age, education, and race/ethnicity. There were 13,020 complete cases for the overall survey (6,800 out of 8,663 sampled units Amerispeak and 6,220 opt-in Dynata); of those, 12,705 completed Probe 1 and 12,650 completed Probe 2.

Results

As shown in Table 1, the coding procedures described above resulted in 7,726 in-scope single date responses and 2,273 in-scope multiple date responses to Probe 1 and 9,608 in-scope single date responses and 189 in-scope multiple date responses to Probe 2. Examples of one date responses include "December 2019", "I think it started back in January", "Early February". Responses with multiple dates include those that specifically mention multiple dates (e.g., "I think it started in November or December") or cover more than one month within a broader time frame (e.g., "Fall 2019"). Through the manual review process, several types of out-of-scope responses were identified, also shown in Table 1 and detailed in the table notes.

Table 1. Responses to Probe 1 (n = 12,705) and Probe 2 (n = 12,650), for RANDS during COVID-19, Round 1.

Type of Response	**Probe 1: When do you think that the Coronavirus pandemic began?**	**Probe 2: When did the Coronavirus pandemic first affect your daily life?**
Single date	7,726	9,608
Multiple dates	2,273	189
Dates before 2019†	66	8
Nonsensical‡	46	14
Too relative or no date*	96	134
Refused to say**	0	22
Other Invalid Response	2,498	2,675

Note: Single Date and Multiple Date Reponses are In-Scope, All Other Responses are Out-of-Scope.
Dates before 2019† includes possible mistakes in typing ("4/1/2010"), confusion ("In 2012, because that's when I had it"), and/or those citing conspiracy theories ("Years ago put in place by the government").

Nonsensical‡ includes responses that do not make sense, such as "fmadsjxcj" or "July hi hi".

Too relative or no date* includes responses such as "a lifetime ago", "after I turned 18", "years ago" and "when we were ordered to stay home".

Refused to say** indicates respondents who say the pandemic has not affected their daily lives and, therefore, did not provide a date.

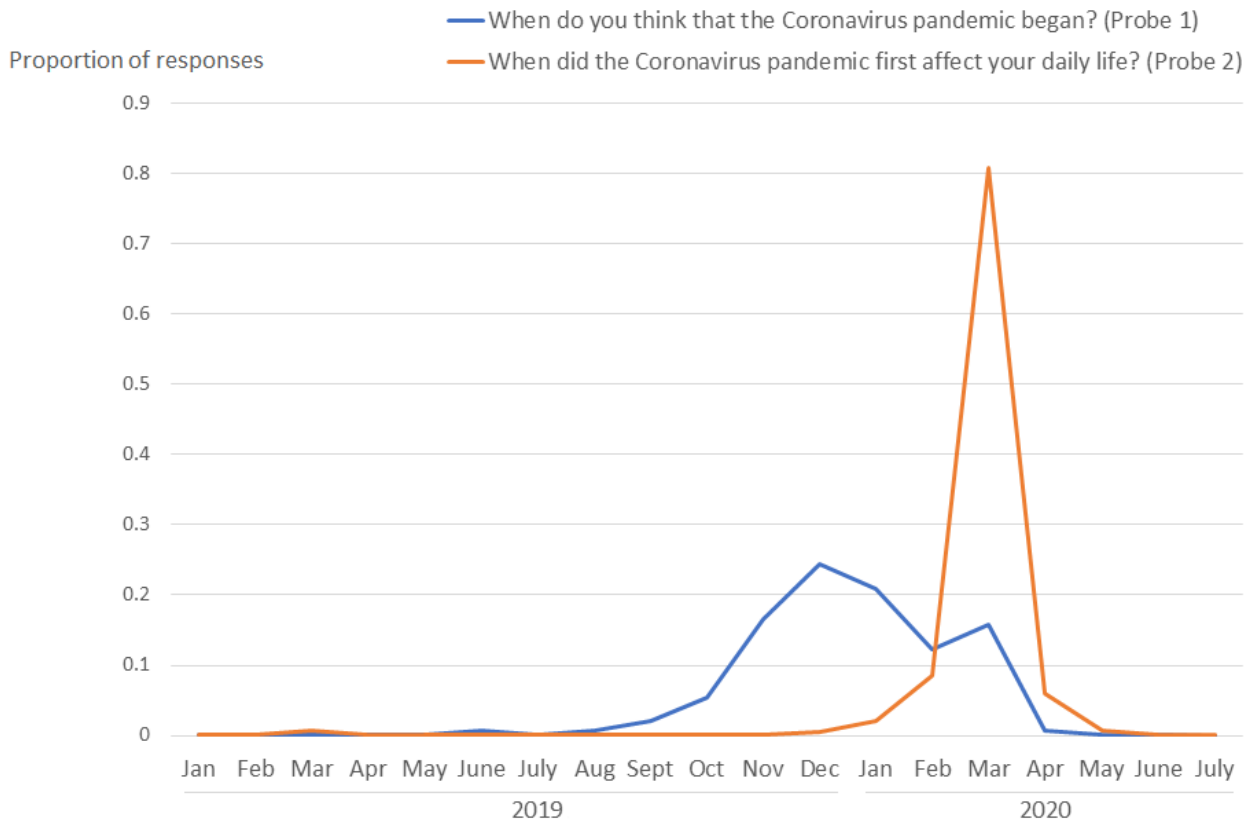
The dates reported in Probes 1 and 2 are plotted in Figure 1. Both single date and multiple date responses are included. We expected a single date from respondents for each probe and thus we chose to weight responses by the number of dates provided. So, someone who answered "Fall" would get 1/3 weight for September, October, and November, respectively, whereas, someone who answered "October 2019" would get one full point for October. In this way, the respondents who provided multiple dates are only counted once, rather than allowing them to have double or triple (or more) representation, compared to those who only gave a single date. The frequencies were then divided by the total number of responses, with the y-axis representing proportion of responses (range is 0 – 1).

Removing respondents who provided multiple dates and plotting Probes 1 and 2 only for those who gave single date responses (e.g., "November, 2019") yields very similar results (not shown).

A wide dispersion of responses is seen for Probe 1 ("When do you think that the Coronavirus pandemic began?") with a gentle peak around the end of 2019 and the beginning of 2020, while for Probe 2 ("When did the Coronavirus pandemic first affect your daily life?"), there is a noticeable spike at March 2020. It is interesting to note that for Probe 1, no single month had more than 25% of the responses (top responses were 24% for December, 21% for January, 17% for November, and 16% for March), whereas, for Probe 2, over 80% of the dates given for that question were in March.

While there is some alignment between the responses to the two probes with a small peak for Probe 1 and the considerable spike for Probe 2 at March 2020, the percentage of respondents reporting March 2020 for Probe 1 (16%) was relatively small.

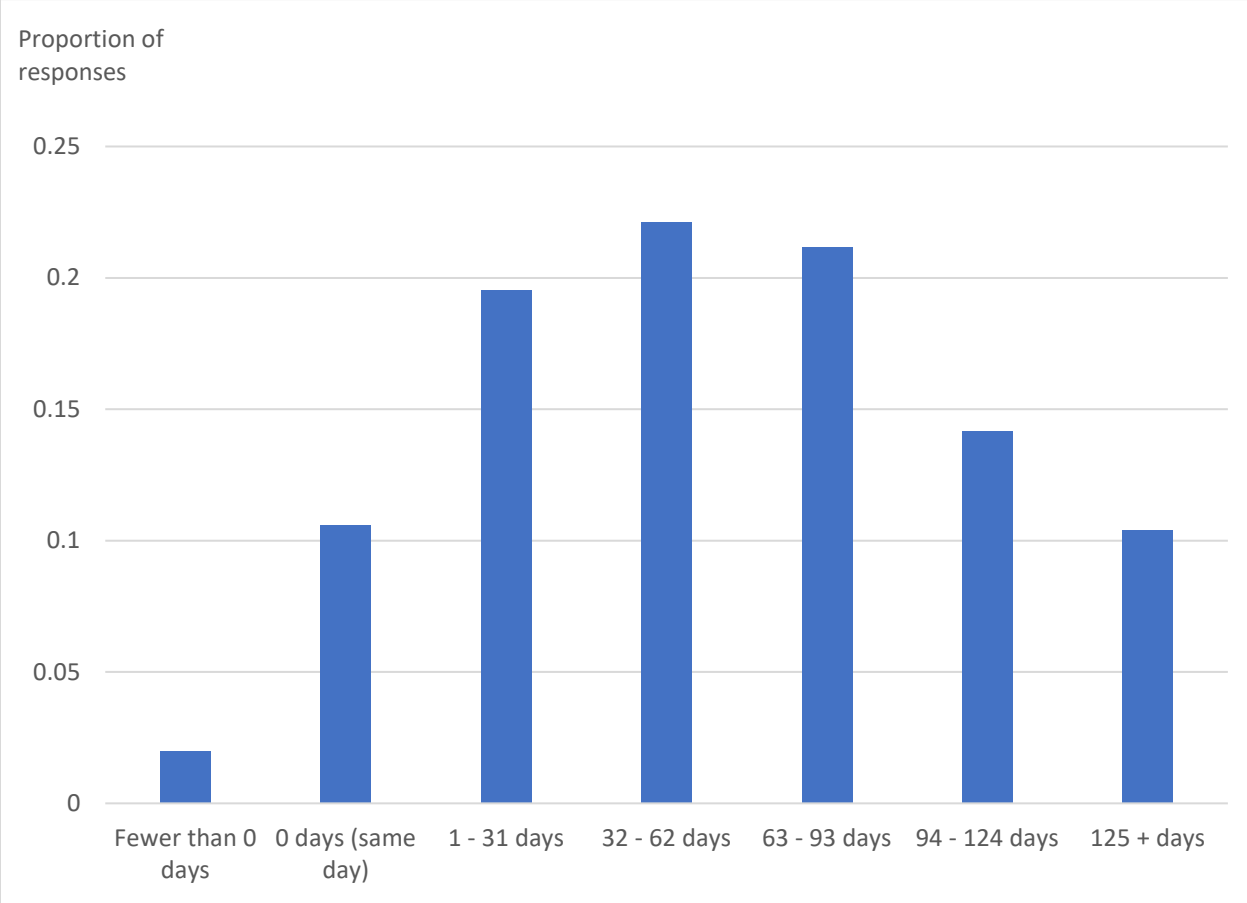
Figure 1. Reported dates to Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?), RANDS during COVID-19, Round 1



We calculated the difference (difference = Probe 2 – Probe 1) between when people think that the coronavirus pandemic began (Probe 1) and when it started affecting their lives (Probe 2) for

respondents who provided a single date for each probe (n = 6,727). The difference ranges from fewer than 0 days (the pandemic started after it began affecting their daily life) to 125 days or more (more than 4 months later), as shown in Figure 2. Most respondents (87%) indicated that the pandemic began affecting their lives sometime after the pandemic began. Examining the differences plotted in Figure 2, there appears to be an approximately normal distribution in the difference between when people think the pandemic began and when it started affecting their lives, with the difference being between 1-31 days (within a month), 32-62 days (within 1-2 months), or 63-93 days (within 2-3 months) for most respondents. Due to the nature of coding responses (dates with only a month specified were assigned a mid-month date, e.g., "March 2020" was coded as 3/15/2020), differences in dates peaked around monthly differences, therefore grouping by month captured the data in a logical way.

Figure 2. Differences between the date reported in Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?), RANDS during COVID-19, Round 1 (Difference = Probe 2 – Probe 1)



Overall, the pattern of responses to the probes by age, education, and race and ethnicity are remarkably similar. Looking at the reported dates by age, shown in Figure 3, we see a similar peak for Probe 1 (when the Coronavirus pandemic began) at December 2019, which is followed by a second, smaller peak around March 2020. However, these peaks are more noticeable for younger respondents (18-29), suggesting that for a slightly larger subset of younger respondents, there is some alignment between the

onset of the pandemic and when they believe that it began to first affect their daily lives, with 21% of younger respondents answering March 2020 for Probe 1. Whereas, reported dates for Probe 1 for the other age groups are more widely dispersed around the peak in December 2019, particularly for those 60+, an equal number of whom (24%) answered December or January for Probe 1. Additionally, these other age groups had a smaller peak at March 2020, compared to the younger respondents. For Probe 2, ("When did the coronavirus pandemic first affect your daily life?"), we see the same noticeable spike at March 2020 for all age groups.

In Figure 4, which shows the distribution of reported dates to Probe 1 and 2 by respondent level of education (Some college or less versus a Bachelor's degree or above), there is a wider dispersion of responses around December 2019 for those with some college or less, compared to a slightly more noticeable peak for those with a Bachelor's degree or above. Otherwise, very little difference can be observed in the distribution of responses by level of education.

A similar difference can be seen in Figure 5 in the response distributions by respondent race and ethnicity. There is a slightly more defined peak around December 2019 for Hispanic; non-Hispanic Black; and non-Hispanic Other[4] respondents, whereas there is a slightly wider dispersion of responses around December 2019 for non-Hispanic White respondents.

[4] The non-Hispanic Other group includes groups which were too small to disaggregate, including American Indian or Alaska Native, Asian, Pacific Islander, Native Hawaiian, and people identify as of more than one race.

Figure 3. Reported dates to Probe 1 and Probe 2 by age (in years), RANDS during COVID-19, Round 1

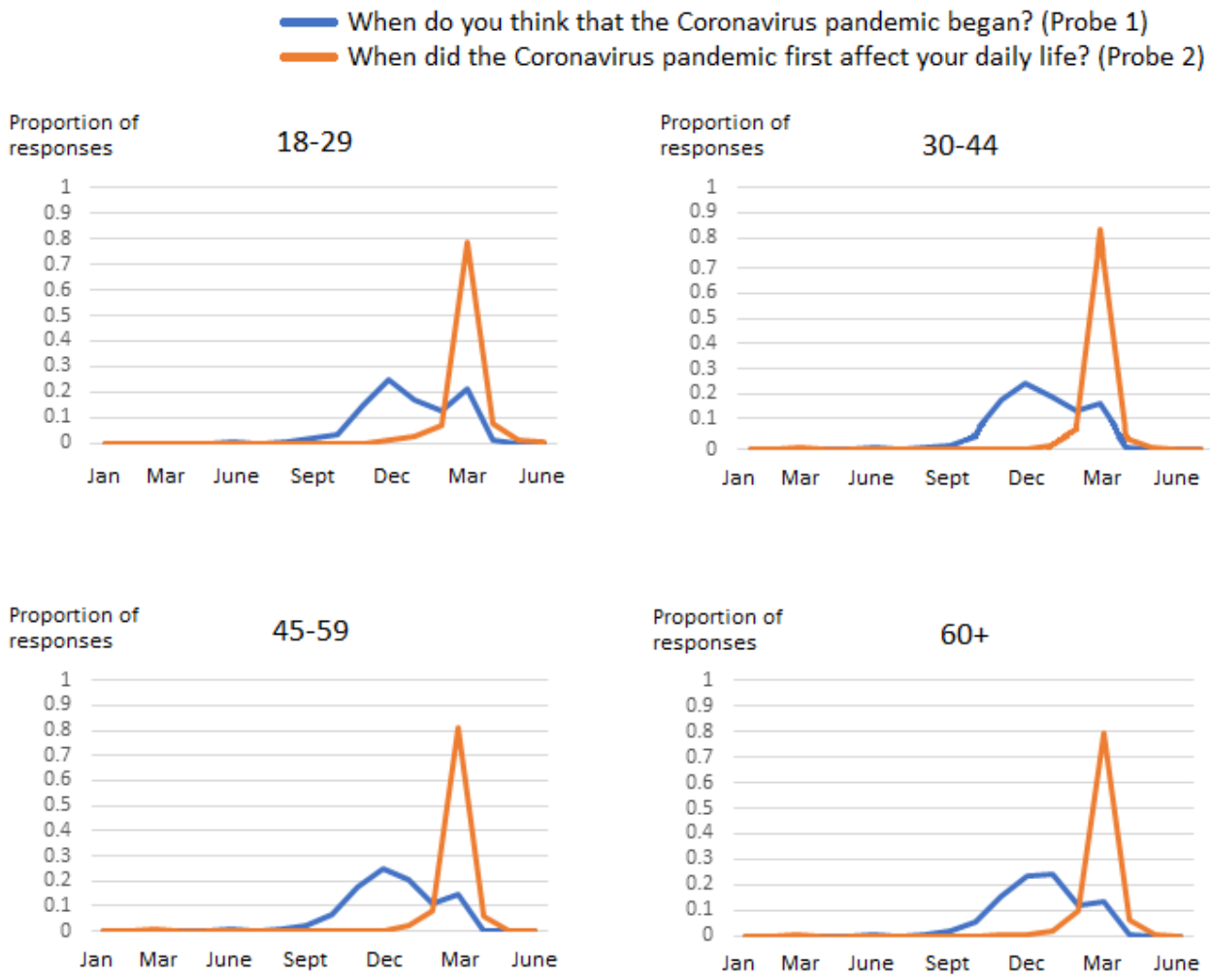


Figure 4. Reported dates to Probe 1 and Probe 2 by education level, RANDS during COVID-19, Round 1

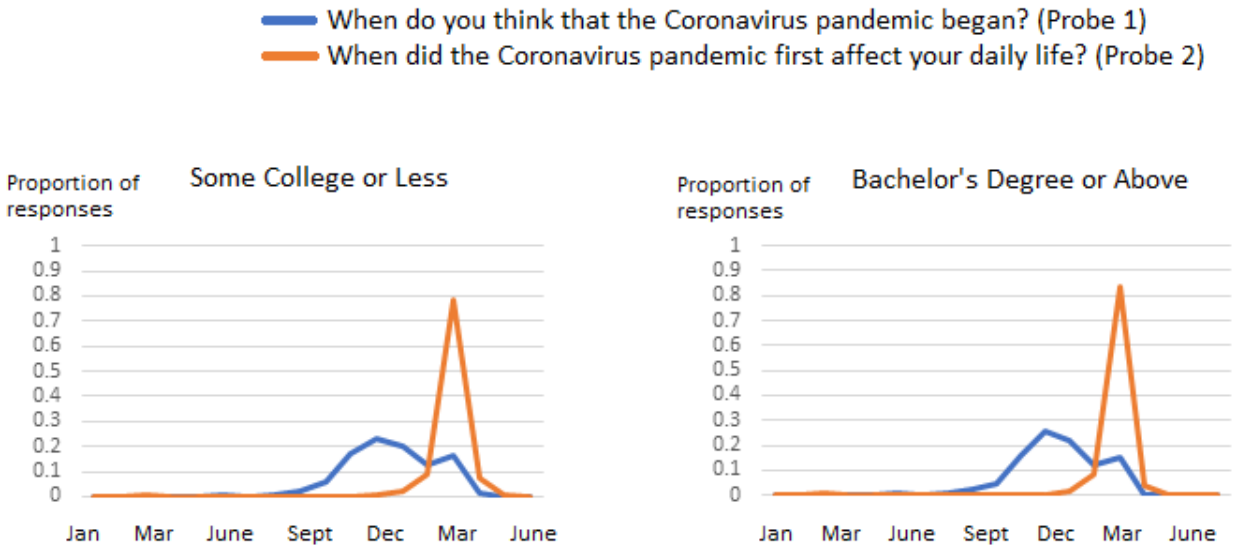
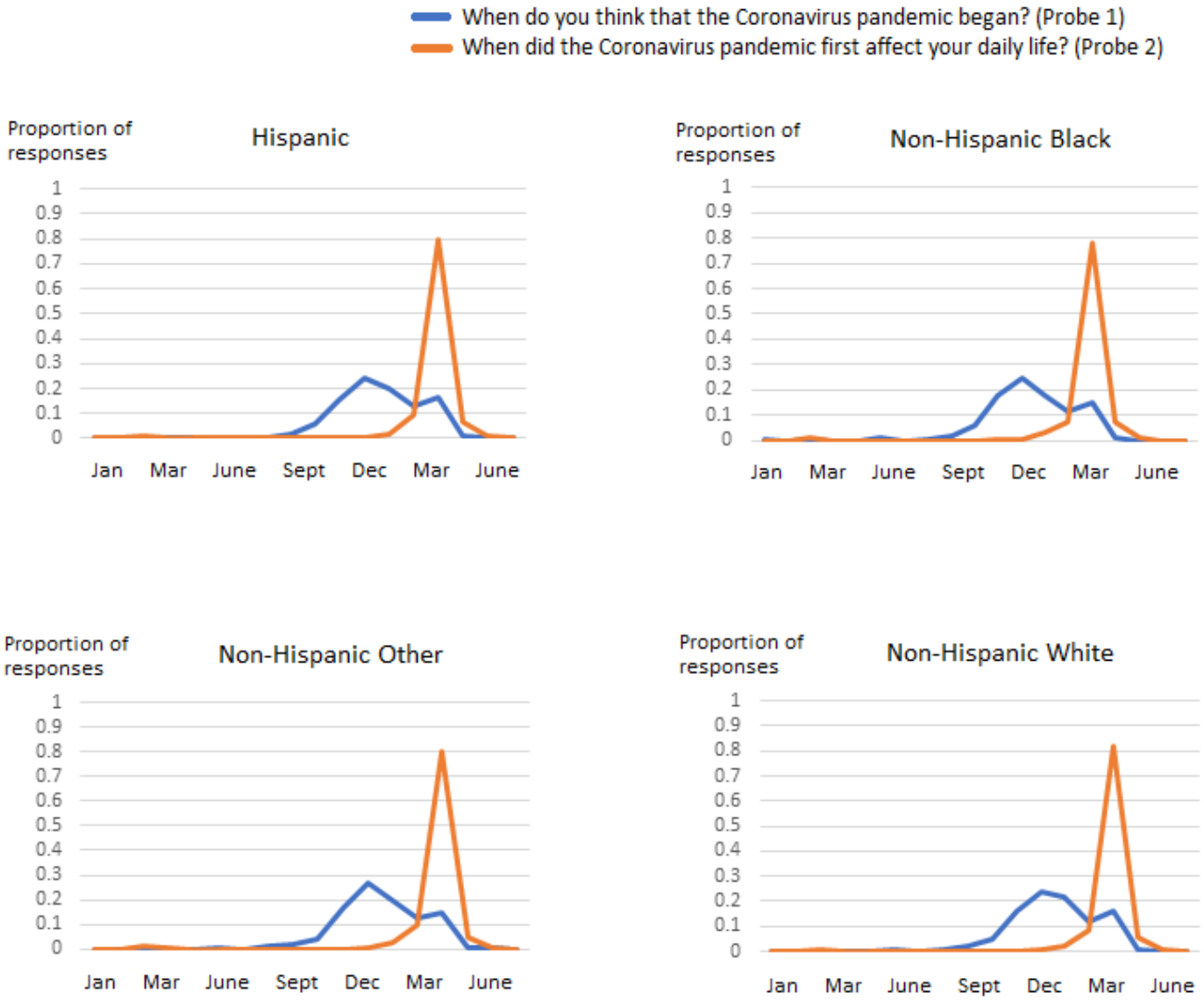


Figure 5. Reported dates to Probe 1 and Probe 2 by race and ethnicity, RANDS during COVID-19, Round 1



Looking at the difference between when people think that the pandemic began (Probe 1) and when it started affecting their lives (Probe 2) by respondent age, education, and race and ethnicity, shown in Table 2, there is very little difference in the median differences between groups. A few groups (people aged 30-44, people aged 60 and older, and non-Hispanic Black persons) have an interquartile range (IQR) that is higher than that for the total sample, indicating a greater spread of the middle 50% of the data for those groups. Also apparent in Figure 3, for people aged 60 and older, their responses were more spread out on when the pandemic began, with 24% of that group choosing January 2020 (which is higher than any other age group) and only 13% choosing March 2020. People between 30 and 44 years of age chose March 2020 as the month the pandemic started affecting them more than any other age group (84%) and had a wide dispersion of responses for when they think the pandemic began, with more people in this group choosing November 2019 and February 2020 than most other age groups. Also apparent in Figure 5, Non-Hispanic Black respondents had more variability in responses, choosing January, February, and March 2020 less often for when the pandemic began and also chose March 2020 less often for when the pandemic started affecting their daily lives, compared to other racial and ethnic groups. However, these differences are less than 15 days (1/2 a month) greater than the sample IQR and as we saw most responses falling in the 1-3 months difference (with a median of 2 months) between Probe 1 and Probe 2 so the difference is not necessarily meaningful.

Table 2. Descriptive statistics of the difference in number of days between Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?) (Probe 2 minus Probe 1), by respondent age in years, education, and race and ethnicity, for RANDS during COVID-19, Round 1.

Group	n	Median	IQR	Min	Max	Range
Total Sample	6,727	60	63	-334	441	775
Age 18-29	946	60	62	-275	433	708
Age 30-44	1,741	60	76	-334	394	728
Age 45-59	1,645	61	62	-289	441	730
Age 60+	2,395	60	75	-304	439	743
Some college or less	3,785	60	66	-303	439	742
Bachelor's degree or above	2,942	60	63	-334	441	775
Hispanic	862	60	63	-289	394	683
Non-Hispanic Black	693	60	76	-303	425	728
Non-Hispanic Other	612	60	63	-334	246	580
Non-Hispanic White	4,560	60	63	-304	441	745

The non-Hispanic other group includes groups which were too small to disaggregate, including American Indian or Alaska Native, Asian, Pacific Islander, Native Hawaiian, and people of more than one race.

Discussion

We analyzed responses to two time-reference web probes ("When do you think that the Coronavirus pandemic began?" (Probe 1) and "When did the Coronavirus pandemic first begin to affect your daily life?" (Probe 2)) to examine how consistently respondents understand when the Coronavirus pandemic began and the extent to which when people think it began aligns with when they think it began affecting their lives. We also examined how the results differ by respondent age, education level, and race and ethnicity. We hypothesized that respondents would understand when the Coronavirus began fairly consistently and that there would be a substantial overlap between when respondents thought the pandemic began and when it began affecting their lives. If there had been a relatively uniform understanding of when the pandemic began, it would make sense to use this as a time reference in survey questions. Overlap between when the pandemic began and when it began affecting respondents' lives would lend further support for when the pandemic began as a time reference as it would likely also be salient to respondents.

However, the wide range of responses to Probe 1 suggests that respondents do not have a uniform or consistent understanding of when the "Coronavirus pandemic began". A closer look at the nature of responses to Probe 1 show that respondents had different interpretations of what was being asked and were, in many cases, answering different questions depending on their interpretation. For example, some are answering the question "when was the Coronavirus declared a pandemic?", e.g., "March 2020 when it was declared a pandemic by the CDC." Others are answering the question "when did the first cases of Coronavirus appear?", e.g., "I think it really began in December when it was first seen in China. This bought the other countries of the world some time, but the US botched that headstart and assumed it wouldn't affect us." For those who report the same or similar dates for Probes 1 and 2, it seems that they are answering the first probe with when the pandemic started to personally affect them, therefore anticipating the second question, such as in this response: "Like factually around December, but like 'for me' March" or this response: "For my area: Early March".

Many respondents reported multiple dates in Probe 1 (n=2,273, in effect answering multiple questions in one response. For example, one respondent said, "Coronavirus started in China around Nov/Dec 2019. It spread throughout several countries beginning 2020. But the pandemic started around March 2020 when all countries shut down and people started staying home to slow the spread." This respondent provided dates for when the coronavirus started in China, when it started spreading to other parts of the world, and when it was declared a pandemic. Others gave similar, multiple date responses, such as "The earliest Covid-19 cases seem to go back to Nov. or Dec. of last fall. WHO declared the pandemic much later - sometime in late February 2020, I think." This respondent contrasted when the first cases likely occurred with when they thought the World Health Organization (WHO) officially declared a pandemic. Another respondent put it this way, "The initial outbreak? August-September 2019. The pandemic? March 2020."

Some respondents noted different dates for different countries, indicating when they heard about the first coronavirus cases in each country. For example, one respondent noted, "I guess it depends on what you set as criteria. It seems the first cases were reported in December in China - maybe first cases then

occur in November? First international cases were confirmed at the end of January... February seems a surge of cases in Europe... USA sees a surge of cases in March...” Another respondent wrote, “In China, September or October. In the USA, December.” These responses suggest that people are answering the question ‘when (and where) did the coronavirus start?’ or they are using where they live (US) as a reference of when either first cases were seen or a ‘surge of cases’ started, compared to other countries, in particular China, as the country of coronavirus ‘origin’; so ‘when did the coronavirus pandemic begin’ is a relative question to these respondents because it began at different times in different places.

In contrast, over 80% of the dates reported for Probe 2 were in March 2020 and very few responses included multiple dates. The peak in responses in March 2020 coincides with when many shutdown orders began, schools closed, people started losing jobs, etc. and point to greater consistency in understanding and/or collective experience of “when the Coronavirus pandemic first” affected daily life for many people (“U.S. state and local government responses to the COVID-19 pandemic”, n.d.). Additionally, 15% of respondents said the pandemic first affected their daily lives in February (9%) or April (6%), possibly reflecting the different times policies were enacted in the United States, with those on the West Coast generally affected earlier, while other areas did not implement stay at home orders or business closures/restrictions until April, if at all (“U.S. state and local government responses to the COVID-19 pandemic”, n.d.) It may also reflect differences in virus transmission as it was not uniform geographically in the beginning of the pandemic. For example, one respondent said, “Mid February when they sent me home to work instead of going in to the office.” Others mentioned policies in April, for example, “Around the time things began to shut down (End of April)”, while another said, “When my state had school closure in April. This changed many things regarding my children's care and my job.” Some respondents who said they were affected in February mentioned preparing in advance of official policies: “February 2020 is when me and mine began preparing for lockdowns and started being more careful. We began about two weeks before the state government did.” Or this response, “Mid-February. I started isolating myself before the government mandated it.”

Comparing results for Probe 1 (“When do you think that the Coronavirus pandemic began?”) and Probe 2 (“When did the coronavirus pandemic first affect your daily life?”) indicates that there is little alignment between when people think the pandemic began and when they think it began affecting their lives. While some alignment was seen with a small peak in the distribution of Probe 1 responses and the considerable spike for Probe 2 at March 2020, this accounts for a relatively small subset of respondents in general.

Slightly more alignment between Probe 1 and Probe 2 was observed for those aged 18-29 suggesting that more respondents in this age group tended to think about the start of the pandemic as coinciding with when it started affecting their daily lives. Additionally, a slightly wider endorsement of dates around December 2019 was reported for Probe 1 by respondents with some college or less education compared to a more noticeable peak at December 2019 for those with a Bachelor’s degree or above and for Black, non-Hispanic; Hispanic; and Other, non-Hispanic respondents compared to White, non-Hispanic respondents. This could potentially indicate that those with less education and White, non-Hispanic respondents were less cohesive in believing that the pandemic began in December, with many choosing November or January as when the pandemic began.

The pattern of reported dates for Probe 2 was remarkably similar across age, education, and racial and ethnic groups, again pointing to a fair amount of consistency in understanding and/or collective experience of "when the Coronavirus pandemic first" affected daily life for most people.

As noted above, data from a third probe will be analyzed using topic modeling and other computer-assisted approaches, in combination with hand review, to explore why people answered the way they did to Probe 2 and the ways in which the pandemic affected their lives based on responses to the third open-ended probe ("Why do you say that?") in a separate follow-up report.

Conclusion

Based on these findings, using the phrase "since the Coronavirus pandemic began" as a time frame in survey questions will yield unreliable results, as there was very little agreement among respondents on when the pandemic began. Providing a specific time frame such as "Since the Coronavirus pandemic began in March 2020…" is therefore advisable. With over 80% of respondents answering that the Coronavirus pandemic first began affecting their lives sometime in March 2020 this time frame is likely to be a salient and memorable reference point for many.

References

Geisen, E. & Murphy, J. (2019). A compendium of web and mobile survey pretesting methods. *Advances in questionnaire design, development, evaluation and testing*, 287-314.

Groves, R.M., Fowler, Jr, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons.

Fowler Jr, F.J., & Cosenza, C. (2009). Design and evaluation of survey questions. *The SAGE handbook of applied social research methods*, 375-412.

U.S. state and local government responses to the COVID-19 pandemic. (n.d.). In Wikipedia. Retrieved October 4, 2021. https://en.wikipedia.org/wiki/U.S._state_and_local_government_responses_to_the_COVID-19_pandemic#Initial_pandemic_responses,_including_full_lockdowns.

Appendix

Table 1A. Sample characteristics for complete cases (n = 13,020), RANDS during COVID-19, Rounds 1, Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?)

Group	n	%
Age 18-29	2,200	17
Age 30-44	3,261	25
Age 45-59	2,946	23
Age 60+	4,613	35
Some college or less	7,722	59
Bachelor's degree or above	5,298	41
Hispanic	1,827	14
Non-Hispanic Black	1,603	12
Non-Hispanic Other	1,207	9
Non-Hispanic White	8,383	65

The non-Hispanic other group includes groups which were too small to disaggregate, including American Indian or Alaska Native, Asian, Pacific Islander, Native Hawaiian, and people of more than one race.

Table 2A. Simple rule fixes for the rule-based date matching approach, RANDS during COVID-19, Round 1: Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?).

What to Change	How it was Changed
October	October 2019
November	November 2019
December	December 2019
January	January 2020
February	February 2020
Febuary [misspelled]	February 2020
March	March 2020
April	April 2020
May	May 2020
June	June 2020
July	July 2020
August	August 2020
September	September 2020
The word "last"	Remove the word
The word "late"	Remove the word

Table 2B. Advanced rule fixes for relative dates implemented in the rule-based date matching approach, RANDS during COVID-19, Round 1: Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?).

What to Change	How it was Changed
Summer	6/1/2020, 7/1/2020, 8/1/2020
Fall	9/1/2019, 10/1/2019, 11/1/2019
Winter	12/1/2019, 1/1/2020, 2/1/2020
Spring	3/1/2020, 4/1/2020, 5/1/2020
2020	1/1/2020, 2/1/2020
2019	11/1/2019, 12/1/2019
Christmas	12/1/2019
Halloween	11/1/2019
Last year	2019
This year	2020

When only given a month and year, the program assigns the day of the month as the day the code was run. This code was run on December 22nd, therefore all dates without a day were assigned the 22nd. In hand coding, these dates were reassigned to the middle of the month (e.g., "December 2019" becomes 12/15/2019). Additionally, many responses indicated a relative time frame, such as "early March" or "the second week of March" and those were recoded by hand as shown in the table below.

Table 3A. Changes made during hand coding, RANDS during COVID-19, Round 1: Probe 1 (When do you think that the Coronavirus pandemic began?) and Probe 2 (When did the Coronavirus pandemic first affect your daily life?)

What to Change	How it was Changed	Example
Month with no date	15th day of the month	"January 2020" becomes 1/15/2020
"early" Month	1st day of the month	"early March" becomes 3/1/2020
"mid" Month	15th day of the month	"mid-April" becomes 4/15/2020
"late" Month	30th day of the month	"late December" becomes 12/30/2019
"first week of" Month	1st day of the month	"first week of November" becomes 11/1/2019
"the [specific] week of" Month	first day of that week in that month	"second week of March" becomes 3/8/2020
[number] weeks ago	date the survey was taken minus number of weeks	"6 weeks ago" becomes 4/21/2020, if the survey was taken on 6/2/2020
[number] months ago	date the survey was taken minus number of months	"2 months ago" becomes 4/12/2020, if the survey was taken on 6/12/2020
"early" season	first month in the season	"early fall" becomes 9/1/2019
"late" season	last month in the season	"late winter" becomes 2/1/2020
winter 2019	the month of winter in 2019 is December	12/1/2019
2nd quarter	use fiscal year segments	4/1/2019, 5/1/2019, 6/1/2019
3rd quarter	use fiscal year segments	7/1/2019, 8/1/2019, 9/1/2019

What to Change	How it was Changed	Example
4th quarter	use fiscal year segments	10/1/2019, 11/1/2019, 12/1/2019
Thanksgiving	date of holiday	11/28/2019
Easter	date of holiday	4/12/2020
spring break	middle of March	3/15/2020
this year; beginning of the year; early 2020	2020	1/1/2020, 2/1/2020
last year; the holidays; sometime before 2020; before the first of the year; end of the year; last couple/few months of 2019; late 2019	2019	11/1/2019, 12/1/2019
last of 2019; end of 2019	December 2019	12/15/2019
before the 2019 holidays	early November 2019	11/1/2019

"Mid-February" or "February" responses were assigned the 14th day of the month. "Late February" responses were assigned the 28th day of the month.

For complex responses, rules are combined, for example "late fall early winter" became two dates: 11/1/2019 and 12/1/2019.

A few relative responses about 2019 and how they were recoded: "beginning or middle of 2019" was recoded as 1/1/2019, 2/1/2019, 3/1/2019, 4/1/2019, 5/1/2019, 6/1/2019; "early 2019" was recoded as 1/1/2019, 2/1/2019; "second half of 2019" was recoded as 7/1/2019, 8/1/2019, 9/1/2019, 10/1/2019, 11/1/2019, 12/1/2019; "mid 2019" was recoded as 6/1/2019.