DSDR | DATA SHARING FOR DEMOGRAPHIC RESEARCH

ICPSR 38417

National Couples' Health and Time Study (NCHAT), United States, 2020-2021

Claire M. Kamp Dush
University of Minnesota

Wendy D. Manning
Bowling Green State University

Methodology Report

Inter-university Consortium for Political and Social Research
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

Terms of Use

The terms of use for this study can be found at:
http://www.icpsr.umich.edu/web/ICPSR/studies/38417/terms

Information about Copyrighted Content

Some instruments administered for studies archived with ICPSR may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

July 8, 2022

National Couples' Health and Time Study (NCHAT)

Methodology Report

Suggested Citation: Marlar, Jenny, Kamp Dush, Claire M., Manning, Wendy D., Berrigan, Miranda N., VanBergen, Alexandra M., Tsabutashvili, Dato, Chattopadhyay, Manas, and Theodorou, Angelina. 2022. *National Couples' Health and Time Study (NCHAT): Methodology Report*. Gallup. doi: 10.18128/MPC2022-03

This research was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD; 1R01HD094081-01A1).

Contents

Overview

Gallup fielded the National Couples' Health and Time Study (NCHAT) from September 1, 2020, to April 25, 2021. The survey covered adults between the ages of 20 and 60 who were cohabiting with or married to same- and different-gender partners. The study included primary respondents and their partners. The sample was drawn from the Gallup Panel and the Gallup Recontact sample. Both sample sources are probability-based and are representative of the U.S. adult population. Respondents were asked to complete a 40-minute survey (which we will refer to as the "main survey") and one 24-hour time-diary (which we will refer to as the "time-diary"). All survey materials were translated into English and Spanish, and 3,642 primary respondents and 1,515 partners completed the survey. Additional details about the survey procedures are provided in the sections below.

Sampling

Respondents were sampled from the Gallup Panel and the Gallup Recontact sample. Both sample sources, and the sampling procedures for each source, are described in greater detail in the following sections. NCHAT aimed to have completed surveys from at least 1,345 different-gender and 1,345 same-gender couples, evenly split between males and females. The sample included 2,507 different-gender, 994 same-gender couples, and 141 couples with at least one partner who identified as non-binary. Further, the study sought to include an oversample of racial and ethnic minorities so that at least 25% of completed main respondent surveys were from these groups. We achieved our target with 1,393 main respondents who identified as racial or ethnic minorities.

Primary respondents were sampled and invited to participate in the study. Upon completion of the main survey, primary respondents were asked to invite their partners to participate in NCHAT.

Gallup Panel Overview

The Gallup Panel was used as the main sample source. The Gallup Panel is a probability-based panel of U.S. adults that is recruited using address-based sampling methods (ABS) and random digit-dial (RDD) phone interviews that cover landline and cell phones. Gallup conducts a nationally representative RDD survey each month with 1,000 completed interviews. At the conclusion of this survey, all respondents are invited to join the Gallup panel. Approximately 65% of respondents who are invited agree to join. Gallup also conducts regular recruiting via ABS. In the ABS recruitment effort, which is conducted approximately four times per year, Gallup sends recruitment materials to randomly selected households via the mail. Respondents are sent a brochure explaining the request and the panel, a small incentive, and are invited to complete a Panel enrollment survey. The ABS recruitment efforts typically oversample households predicted to have demographics needed to replace groups that tend to attrite from the panel at higher rates, including individuals with lower education levels, who are younger, and who are Black or Hispanic. Gallup also regularly conducts nationally representative RDD and ABS surveys. At the conclusion of these surveys, respondents are asked if they are willing to be recontacted for a future Gallup survey. Approximately 80% of respondents agree to be recontacted for a future survey and are eligible to receive a small number of survey invitations from Gallup. Gallup refers to this as the "recontact sample" (described in the next section). Some individuals from the recontact sample are selected for recruitment into the Gallup Panel and are contacted by Gallup and asked to join the Panel.

Approximately 90,000 Panel members can be reached for web, mail, or telephone surveys. Another 20,000 Panel members do not have email access but have provided a mailing address and can be reached for mail or telephone surveys. Members who have consented to receive text messages can also receive survey invitations or related communications via text message.

Panel members receive up to three surveys per month, and the typical survey is 10 to 15 minutes in length. The average Panel member completes approximately one survey per month. Most Gallup Panel surveys are self-administered web surveys, and Gallup typically sends respondents an invitation and two to five

reminders. The average response rate on a Gallup Panel survey is approximately 40% to 50%, depending on the length of the survey, length of the field period, and the survey topic. Incentives are offered to respondents based upon a variety of factors such as survey burden, available sample size, target population, length of field period, and project budget. Many surveys offer no incentive at all, and most surveys do not offer more than an average of $2 to $5 for participation.

Members may remain in the Panel for as long as they would like, given they continue to participate. Gallup frequently reviews participation records and refreshes the Panel sample. Members who continue to be chronic non-responders are removed from the Panel. Gallup conducts regular recruiting efforts to refresh the sample and recruit new members. Adults between the ages of 18 and 34, individuals with lower education levels, and Black and Hispanic participants tend to have lower participation rates than other demographic groups. Gallup's recruiting efforts generally oversample these groups to maintain a demographically balanced sample. Unequal selection probabilities at the selection stage are taken into account in the panel weight assigned to each member.

Gallup maintains a database of demographic attributes on all Panel members, which can be used for efficient and cost-effective sampling of very low incidence populations that would otherwise be extremely costly, if not impossible, to reach. Known panel member information includes variables that were of key interest to NCHAT, including age, gender, race/ethnicity, marital/cohabitation status, and if the respondent identifies as Lesbian, Gay, Bisexual, or Transgender (LGBT).

Gallup Recontact Sample Overview

As described in the Gallup Panel section, Gallup conducts a large number of nationally representative surveys using RDD dual frame (landline and cellphone) and ABS methodologies. The average response rate (AAPOR Response Rate 3) to Gallup's RDD surveys ranges from 8 to 12%, depending on the survey. For ABS surveys, the average response rate is approximately 18 to 20%. From 2008 to 2017, Gallup fielded the Gallup Daily Tracking survey via RDD, and collected 1,000 interviews per night, 350 days per year[1]. From 2018 to early 2019, the Gallup Daily Tracking survey moved from telephone to a mail push-to-web design[2] with monthly data collection with approximately 10,000 completed surveys per month[3]. Surveys are conducted in English and Spanish. In mid-2019, Gallup discontinued Daily Tracking, but does maintain the monthly Gallup Poll Social Series (GPSS), which is a monthly RDD survey with approximately 1,500 completes.

At the end of these national survey efforts, Gallup asks respondents if they would be willing to be recontacted for a future survey. Approximately 80% of respondents agree to be recontacted for a future survey. For respondents who have completed a telephone survey, the respondent's phone number is retained and used for future contacts. For respondents who have completed the ABS survey, an email address is requested for future contact, and the mailing address is also available. Individuals in the recontact pool can be contacted and asked to join the Panel. Individuals who do not join the Panel can be contacted for other studies, however, Gallup limits this contact to a very select number of studies. Most recontact sample has never been recontacted at all, and no recontact sample member has been contacted more than a few times. As with the Gallup Panel, Gallup is able to use respondent demographics and substantive answers to sample specific subpopulations.

Sample Design

During the planning phase of NCHAT, Gallup planned to draw sample from the Gallup Panel and Recontact sample who had provided an email address. However, it was known that these sources alone would not reach the desired racial and ethnic minority oversamples, particularly within the same-gender group. Gallup

[1] https://www.gallup.com/174155/gallup-daily-tracking-methodology.aspx

[2] In a mail push-to-web design, respondents are sent a survey packet through the mail but have the option to go online to complete the survey.

[3] https://news.gallup.com/poll/246200/gallup-national-health-index-work.aspx

planned to add questions to the end of ongoing representative Gallup surveys (such as the GPSS) to recruit respondents into the study over a period of two years. Sample was to slowly be released for all sample groups and data collected over the course of at least two years. However, the COVID-19 pandemic hit in March 2020 as the research team was preparing for fieldwork. The team decided to change course and collect all data during the COVID-19 pandemic, with the understanding that it would condense the field period and limit the ability to recruit harder to reach segments of the study population. Several sampling strategies were used throughout the field period to achieve minimum sample size requirements for analysis, and several samples were pulled and added throughout the study field period. Samples were also added to improve the coverage of groups that had lower than average response rates (see the following sections for more details on these groups and response rates). Table 1 summarizes the samples that were selected during the field period, and each is described in more detail in the section that follows.

Table 1: Overview of Samples Selected to Participate in NCHAT

Sample group description	Date invited	N invited	N completed*	Participation rate**
Initial sample of panel/recontact	September 1, 2020	7,691	1,662	22%
Telephone recruits	November 12 – January 26, 2021	925	431	47%
1st oversample of Black, Hispanic, and low education	November 24, 2020	3,156	516	16%
2nd oversample of Black, Hispanic, low education, and younger	January 13, 2021	5,418	790	15%
Asian oversample	March 19, 2021	724	243	34%

*Number completed only includes final weighted counts. Partially completed survey and cases that did not qualify are not included.
**Participation rate = completed/number invited

Initial Sample Draw

For the initial survey launch on September 1, 2020, Gallup drew a sample of 7,691 English and Spanish speaking respondents who could be reached via a web survey. Sampled cases were selected from the Gallup Panel and the Recontact sample. All individuals sampled from the Recontact sample had originally completed an ABS survey during 2017 or 2018 and had provided an email address.

Gallup Panel members were stratified into two groups based on their response to the question: "Do you, personally, identify as lesbian, gay, bisexual, or transgender?" This question is asked on the Panel's New Member Survey, which is completed shortly after joining the Panel. The question is also asked of Panel members at least one time per year, and the Panel database (from which sample is drawn) is updated based on the most recent response. Although we could sample based on sexual orientation, it was not possible to pre-identify whether a respondent is in a same-gender or different-gender couple. The Panel database has LGBT status and marital/cohabitation status, but there is no known information about the gender of the respondent's partner. The LGBT question was used as a way to identify respondents with the highest likelihood of qualifying for the same-gender group. It is important to note that many (approximately one-third) of the individuals sampled from the LGBT group were in a different-gender couple and qualified for the different-gender group. This resulted in a higher proportion of individuals who are LGBT within the different-gender group than what would be expected in the population. This issue was addressed in weighting (see weighting section).

Within the group of Panel members who said "no" to the LGBT question, respondents were selected if they were between the ages of 20 and 60 and indicated they were currently married or living with a partner/boyfriend/girlfriend. The non-LGBT sample was stratified by age, gender, education level, race, and ethnicity and a subsample of 3,509 respondents were selected from among all eligible respondents.

Martial status had been updated in the Panel database just prior to the start of the NCHAT survey. The Gallup Panel has used two different versions of the marital status question. Both are self-administered – the respondent reads the question and selects the most appropriate response option. Version 1 is the version of the question that was historically asked on panel surveys. All members answer this question when they join the Panel and it is added to surveys throughout the year to keep the Panel database current. The NCHAT team expressed concern that this version of the question may be missing certain types of couples that may be cohabiting. The team provided a revised question (Version 2), and just prior to selecting sample for NCHAT, Gallup asked Version 2 to several large surveys that were asked of the majority of panel members.

> **Version 1:** *What is your current marital status? 1) single/never been married; 2) married; 3) separated; 4) divorced; 5) widowed; 6) domestic partnership/living with a partner (not legally married); 7) don't know*
>
> **Version 2:**
> Q1: *Are you currently married? 1) Yes; 2) No*
> Q2: If not currently married, then ask: *Are you currently not married but living with a partner/boyfriend/girlfriend? By living together, we mean that you are in a relationship and that neither of you have a separate residence. 1) Yes, living with a partner 2) No, not living with a partner*

All Gallup Panel respondents who said "yes" to the LGBT question and who were between the ages of 20 and 60 were selected into the sample (n = 3,146). Marital/couple status was not a criterion for selection within this group. Although marital status demographics had recently been updated in the panel database, marital status can change and not all respondents answered the question. All respondents were selected with the hope of picking up any respondent who may qualify.

The Panel sample alone did not have enough LGBT respondents to meet the desired number of same-gender completes. To supplement the sample, an additional sample of 1,036 respondents who had completed the Gallup Daily tracking survey, were between the ages of 20 and 60, provided an email address, and said they were LGBT were invited to participate. Respondent age was calculated based on the age at the time the original survey was completed and the amount of time that had passed at the time the sample was drawn. As with the Panel sample, marital status was not considered for the recontact sample, as the surveys had been completed over two years ago and marital status may not be current. A small number of respondents from this recontact sample had been selected for the "Generations" or "Transpop" study. Individuals who participated in the Generations or Transpop study were excluded from the NCHAT sample.

Low Education, Black, Hispanic Oversamples

After the start of fieldwork, the NCHAT team determined additional sample would be needed to achieve adequate representation from respondents who are Black, Hispanic, and/or who have lower education levels. On November 24, 2020, Gallup sampled an additional 3,156 respondents who fit these criteria. All respondents sampled were between the ages of 20 and 60, were married or living with a partner, and said "no" or did not respond to the LGBT question. All individuals who identified as LGBT had already been invited to participate in the study, and therefore there was no additional LGBT sample to include.

A second oversample of 5,418 Black, Hispanic, and/or low education respondents was added to the field on January 13, 2021. As with the first oversample, all respondents were between 20 and 60 years of age, married or living with a partner, and not LGBT.

Phone Recruit Sample for Same-Gender Respondents

In October 2020, the NCHAT team determined additional sample would need to be added to achieve a greater number of same-gender completes. However, no additional web-based sample was available.

Gallup leveraged the Gallup recontact sample who could be contacted via phone. Similar to the recontact sample in the initial sample draw (described earlier), sampled participants had to be between the ages of 20 and 60 and had indicated they were LGBT in a previous survey. Individuals who completed the Generations or Transpop survey were excluded from the sample. Marital/couple status was not a criteria for selection into the recruitment sample, as marital status could have changed since they completed the previous survey with Gallup. However, unlike the other sample sources, for *some* respondents in the sample, same-gender or different-gender couple status was known. Individuals in a same-gender couple were given priority in the sample and dialed first. Additionally, respondents with a high school education or less and individuals who were Black or Hispanic were also part of the first samples to be dialed. This was to recruit these groups as early as possible and maximize the amount of time that they had to respond once the invitation to NCHAT had been sent.

Selected respondents were contacted by a Gallup interviewer via telephone, given a short explanation of the study, and asked three screener questions (currently living with a partner, gender, and gender of partner). Individuals who were confirmed as living with a same-gender partner were asked to provide an email address so that the survey invitation could be sent. All respondents who agreed to participate via phone were invited to participate in NCHAT and were sent their survey invitation shortly after the recruitment phone call. A total of 925 respondents completed the screening survey on the phone, provided a valid email, and were sent an email invitation to participate in NCHAT.

The telephone recruit began on November 12, 2020 and concluded on January 26, 2021. Recruiting paused from December 15, 2020 to January 2, 2021 due to the holiday and anticipated low levels of participation during this time period.

Asian American Oversample

On March 19, 2021, Gallup invited respondents to the study who were Asian American. Asian Americans were not previously oversampled. Due to recent events and hate crimes against Asian Americans, the NCHAT team determined it would be important to ensure that the sample included enough Asian American cases to report results. All sampled individuals were Asian, between the ages of 20 and 60, and married or living with a partner. Everyone in the Gallup Panel who met the criteria were selected (n = 724). Because everyone who identified as LGBT had previously been selected into the study, only non-LGBT individuals were a part of this sample.

Survey Design and Administration

Overview

Survey participants were asked to complete a survey that is approximately 40 minutes in length (we will refer to this as the "main survey") and one time-diary which covers a 24-hour time period (we will refer to this as the "time-diary").

The NCHAT team spent more than a year developing survey questions and programming and testing the survey instrument and the web-based time-diary. The survey covers a range of topics related to family functioning, including a household roster, partner history, physical and mental wellbeing, health behaviors, discrimination, COVID-19, and demographics. The time-diary covers daily activities during a 24-hour time period (activities were entered in 10-minute increments and included multitasking). In addition to asking about what you were doing, the time-diary also recollects the household roster and asks who you were with during the activity and if they were directly involved in the activity, where you were, and how you felt (happy, stressed, and engaged) during each activity. It concludes with questions that ask you to summarize some of your experiences throughout the day, such as your life satisfaction, experiences of microaggressions during your time-diary day, and sleep.

Prior to the expected launch of field work, in March 2020, the COVID-19 pandemic forced lockdowns and significant changes to everyday life. In May 2020, the killing of George Floyd sparked social justice movements and a look at racial inequalities in the U.S. Questions were quickly added to the survey to address both of these issues, and the timeline for data collection was accelerated with the intent of collecting all data during the pandemic.

All survey materials were translated into Spanish. Respondents who speak Spanish received Spanish language invitations and reminders. All respondents had the ability to select a Spanish language survey instrument from the main screen of the survey. Gallup works with a professional translation provider for the Spanish language translations. Gallup has partnered with this translation service for more than a decade, and they are familiar with the unique needs of survey instruments and providing translations for survey questions. The translation process begins with the lead translator, who is fluent in English and a native speaker of the target language (in this case, Spanish). A second translator, fluent in both English and the target language, reviews the entire translation for editing (changes to words) and proofing (typos). The two translators resolve any recommended edits.

Pretesting

Prior to the official launch of the survey, Gallup conducted cognitive interviews and two pilot tests of the main survey and time-diary. The pilot tests included qualitative questions at the end of the survey, soliciting feedback about the overall survey experience.

Cognitive Interviews

The primary purpose of cognitive interviewing is to gain a better understanding of how survey questions are being interpreted and to identify questions that may be confusing or misleading. Gallup conducted ten cognitive interviews in May 2020. Six respondents were LBGTQ and four did not identify as LGBTQ. Participants were between the ages of 23 and 54, and five were men, four were women, and one was a gender minority.

Because of the length of the survey instrument, not all items could not be included in the cognitive interviews. The team selected questions that had not been previously validated or were believed to have the greatest potential for confusion or variation in interpretation. Many of the questions selected could have been impacted by COVID – for example, questions about their employment situation.

Overall, the questions were clear and easy for respondents to understand, however, some areas were identified where respondents had difficulty. For example, some respondents struggled to answer questions in the fertility section because they were finished having children because of their age and felt that some questions had response options that did not apply to their situation.

Pilot testing

Two rounds of pilot testing were conducted, and adjustments to the survey and communications were made based on the results of the pilot. The pilot was only conducted in English. During pilot testing, the research team evaluated overall participation rates and the length of the survey. It was also an opportunity to review data from each question to ensure the questions were programmed properly and functioning as intended. The team also ran alphas on indexes by different-gender and same-gender respondents to ensure the indexes performed well for both groups of respondents.

Further, the survey process relies on complex sample management and survey automations that invite respondents to each new activity. The pilot test also ensured that these automations and sample management were working properly before the launch of live data collection.

At the end of the main survey and time-diary, the following open-ended questions were asked so that respondents could provide feedback on the overall survey experience or areas where they may have had difficulty.

- What parts of this survey did you like or not like?

- Were there any questions or sections of the survey that were confusing or that you had a hard time answering?
- Please feel free to share any additional thoughts here

First pilot

The first pilot was conducted using Gallup Panel sample and sample from personal networks. Responses from participants expected to qualify for the different-gender group came from the Gallup Panel. To conserve the same-gender sample for the official survey administration, the research team invited personal contacts who were expected to qualify for the same-gender sample to participate in the survey. An email was sent to potential participants. The email request included a link to a short survey where interested participants could enter their email address, and 208 people provided contact information.

On May 22, Gallup invited 308 people to participate in the first round of pilot testing (208 from the personal network email recruit and 100 people from the Gallup Panel). Of the 308 invited, 148 completed the survey (30 completes from the Gallup Panel sample and 118 from the personal network sample). Of the 148 who completed, 105 referred a partner to complete the survey.

The completion rate for the pilot was 92%, meaning that 92% of people who started the survey completed it. This rate is relatively high for a survey of this length and is consistent with rates observed on the Gallup panel for surveys that are significantly shorter.

Overall, there were no major issues identified with the survey instruments themselves that required corrective action. However, participation in the survey portion of the study and the number of invited partners fell short of expectations. Alternatively, most respondents who participated in the survey went on to complete the time-diary. The team hypothesized that the 45-minute main survey[4] might be a perceived barrier to participating, while the time-diary was generally well-received. To better understand how respondents viewed the main survey, Gallup contacted eight participants who were invited to the pilot, including people who had participated and people who did not participate, to better understand the reasons why they completed some survey activities and not others. The following themes were uncovered in these interviews:

- Some participants felt the survey was too long, and some reported feeling bored by the end. Several respondents did report that the promise of the post-paid survey incentive motivated them to continue to the end of the survey.
- Other participants reported enjoying the survey, including the nature of the topic and the novelty of the time-diary.
- People who did not send the invitation to their partner to participate felt that 1) their partner would not be interested in taking the survey or 2) that the survey was too long, and they were saving their partner from a burdensome activity.
- Most participants reported either not really noticing the pre-paid incentive or felt like it did not motivate their participation.
- The majority (95%) of respondents who completed the main survey completed the time-diary. This, combined with other feedback, indicated that respondents perceived the main survey to be more burdensome than the time-diary.

Based on this feedback, several changes were implemented. The first was to the incentive structure. In the first pilot, respondents received a $5 pre-paid incentive when they received the main survey invitation. They received a $15 post-paid incentive for the main survey and a $15 post-paid incentive if they completed the time-diary. In the second pilot, the incentive structure was changed such that participants received a $5

[4] The survey invitation and consent for the pilot communicated that the survey was expected to be 45 minutes in length. Based on the results of pilot testing, the actual average survey length was 40 minutes and future communications mentioned a 40 minute main survey.

pre-paid incentive at the time they were invited to the main survey, a $20 post-paid incentive for completing the main survey, and a $15 post-paid incentive for completing the time-diary.

Additionally, changes were made to the partner invitations, and the wording was updated to encourage respondents to share the request with their partner, even if they think their partner won't be interested. It also stressed the importance of having partner responses.

Second Pilot

Based on the results of the first pilot and ideas for changes, the team decided to conduct a second pilot, primarily to see if the changes would improve survey participation. In addition to the incentive and communications changes described above, questions about race relations were also added to the survey. Gallup invited 100 Panel members to participate in the study. Of those invited, 50 were White and 50 were Black. All were expected to qualify for the different-gender group, and pilot survey invitations were sent on July 7, 2020. A total of 41 respondents completed the survey, and five respondents started but did not finish.

The changes made to the survey communications and incentive structure did result in improvements to participation rates. Further, no issues were detected with the survey, and the team determined the survey was ready for the official launch. However, some minor modifications were made to the survey initiation. These changes included 1) letting respondents know that they could leave the survey at any time and return later and complete where they left off and 2) moving details about the time-diary request to later in the email (after the link to the main survey).

NCHAT Survey

All sampled primary respondents were invited to participate in the NCHAT survey. Initial survey invitations from the first sample draw were sent on September 1, 2020. The survey invitation included a $5 pre-paid incentive, a promise of an additional incentive for completing (see below for more details on incentives), an explanation of the study and survey request, the survey link, and the consent statement language. The survey link contained an embedded unique respondent I.D. for accessing the survey.

The survey link took respondents to the survey instrument, which first asked for participant consent. Participants were then asked the following screening question: "Are you currently living with a spouse/husband/wife or partner/girlfriend/boyfriend most of the time?"

Respondents who gave consent and answered "yes" to the screening question continued with the survey. For those who did not consent or who answered "no," the survey ended, and they were no longer contacted to participate in NCHAT.

The survey took, on average, 40 minutes to complete. Respondents were able to leave the survey without finishing and come back to it later using their unique link. The survey resumed where they left off (meaning respondents did not have to start over if they left the survey without completing it).

Respondents who did not complete the survey were originally scheduled to receive a series of three email reminders to complete the survey on September 4, September 7, and September 14. Several additional reminder efforts were scheduled throughout the eight-month field period to improve participation. Additional reminders were sent to the first sample group on September 28, October 3, October 8, October 14, October 18, October 23, and October 29.

Respondents who completed the survey initially received a $20 post-paid incentive. However, due to lower than desired participation, the post-paid incentive was increased to $50 on November 16, 2020 (see the incentives section for more details). All communications prior to November 16 mentioned the $20 postpaid incentive, and all communication on or after November 16, 2020 (including reminder emails) mentioned the $50 postpaid incentive.

Sample for the first low education, Black, and/or Hispanic oversamples was added on November 24, 2020, and reminders were scheduled for December 3, December 10, and December 14, 2020.

Sample for the phone recruit was added from November 12, 2020 to January 26, 2021. Invitations to NCHAT were sent on an ongoing basis throughout the recruitment period so that respondents would receive the invite shortly after completing the screening survey. Reminders were scheduled for three, seven, and 14 days after the initial invitation.

Sample for the second low education, Black, and/or Hispanic oversamples was added on January 13, 2021 and reminders were scheduled for January 17, January 22, January 28, and February 3, 2021.

Sample for the Asian American oversample was added on March 19, 2021 and reminders were scheduled for March 24, March 29, and April 5, 2021.

On January 28 and February 3, 2021, all non-responders who had a high school education or less or who were LGBT received reminder emails. Also on January 28, 2021, Gallup's Director of the Gallup Panel sent these respondents an email from her personal work email address. The email encouraged respondents to participate and was signed by Gallup and the University of Minnesota's research leads. The goal of this email was to convey the importance of the activity by sending individual emails from a personal rather than corporate email address. Additionally, the team hoped this email would be less likely to be filtered into promotional or junk mailboxes by the email provider. Respondents who provided consent to receive text messages also received an SMS reminder.

A second personal email was sent on March 5, 2021 to all LGBT non-responders and all non-responders under the age of 30. These individuals also received an email reminder at this time, and for those who consented, an SMS reminder was sent.

Additional reminders were sent to all non-responders on March 10, March 24, March 29, and April 5, 2021. All respondents who had provided consent to be text messaged also received a text message reminder with a link to the survey. The team hoped these additional reminders and text message efforts might increase the number of participants under the age of 30.

Once the respondent completed the survey (defined as reaching the end of the main survey instrument), the survey system sent the respondent the incentive within three hours (see the section on incentives). Completing the survey also triggered two additional activities: an invitation to the time-diary and an email to invite their partner to participate in NCHAT.

Time-diary

Once the main survey was completed, respondents were invited to participate in the time-diary. Respondents were randomly assigned to a time-diary day. The diary covered a 24-hour time period, starting at 4 am on the assigned diary day. The time-diary day was approximately 10 to 14 days after the date the main survey was completed. This allowed time to give respondents advance notice of their day and to increase the chances of being able to assign the primary and partner respondents to the same time-diary day.

The first of several time-diary communications was sent approximately 48 hours after the main survey was completed. It explained to respondents the time-diary component of the study, notified them of their assigned day, and told them to look for more information the evening before their assigned diary day. It also told them about the $15 post-paid incentive for completing the time-diary.

The day before the assigned diary day, a second time-diary communication was sent. It reminded them that their assigned diary day was tomorrow, let them know they would receive a link after midnight, and encouraged them to enter the survey throughout the day to enter activities as they occurred.

Shortly after midnight (central time) on their assigned diary day, respondents received an additional email with the link to the time-diary. It encouraged the respondent to enter the survey as many times as they wanted during the day to record activities. The link could be used as many times as they needed, and respondents picked back up where they left off. During the time-diary day, all respondents received survey reminder notifications throughout the day. The short reminder message was sent every 4 hours between 8 am and 8 pm and encouraged respondents to enter activities. Their unique survey link was also included. Gallup Panel members who had provided the necessary consent also received text message reminders every four hours.

Respondents received a final email reminder the morning after their assigned time-diary day (if the diary had not already been completed in full). For example, if their assigned day was a Wednesday starting at 4am, they received an email reminder on Thursday to complete their day. The time-diary survey was closed 48 hours after the assigned diary day and could no longer be accessed by the respondent. This was to ensure that the time-diary was completed relatively soon after their assigned day and that some diaries would not be subject to greater recall bias than others.

A time-diary was counted as complete if they entered at least five activities and 180 minutes. (Not all of these cases were weighted. Please see the weighting section for more details about cases that received a weight). Respondents who did not start a time-diary or who did not enter the minimum required number of activities were assigned to one new time-diary day. These individuals received a follow-up email that informed them they had been assigned to a new day. All communications after this email followed the same process for sending the time-diary link and reminders.

Respondents who completed the time-diary survey were sent a $15 post-paid incentive, and their participation in NCHAT was complete.

Partner Invitations

The process for inviting partners was carefully considered to protect the confidentiality of primary and partner respondents. At the end of the main survey, respondents were informed they would be sent an email with instructions for inviting their partners. Within a few hours of completing the main survey, the partner invitation was sent to the primary respondent. The partner invitation contained a link to a survey that collected the partner's email address. The primary respondent was asked to share the invitation (or reminder) email with their partner and encourage their partner to participate. The partner used the unique link to go to a short survey that shared a few details about NCHAT and collected their first name and email address (used only for sending NCHAT-related survey invitations and reminders). Reminders to complete the partner email survey were sent 3, 7, and 14 days after the initial request.

Once the partner's email address was collected, the partner was sent the invitation to the NCHAT survey. Further, once this email address was collected, the process for inviting and reminding and all communications mirrored those used for the primary respondent.

As with the primary respondent, the main survey included the consent and screened for being married to or living with a partner. Unlike the primary respondent, partner respondents could be over the age of 60 and were not screened out or removed from the survey data unless they were under the age of 18.

All partner respondents are assigned a unique I.D. (ResponseID) and a partner I.D. (MATCH_ID), which allows their data to be matched to the primary respondent.

Alternative methods were considered for inviting partners, which could have potentially increased the number of partner emails collected. This included asking the primary respondent, at the end of the survey, to share their partner’s email. This option was deemed unacceptable by the Gallup IRB because personally identifiable information (PII) would be shared without their partner’s consent. Another option considered was to share the partner link to the main survey with the primary respondent and ask them to forward it to their partner (avoiding the need to collect partner email). This option was ruled out because if a partner left

the survey unfinished, the primary respondent would have their partner's unique link and could potentially view answered questions.

During the field period, the team discovered that 101 respondents had provided email addresses that were the same as the primary respondent email addresses. As a quality control and confidentiality measure, the Qualtrics survey system can only send one survey invitation per email address. Therefore, the primary and partner respondents cannot share emails. The team learned of this issue when some respondents reported providing an email address but never receiving the survey invitation. Based on information shared by respondents inquiring about the status of the invitation, the team learned that some primary respondents entered their personal email address, with the intent of receiving the survey invitation and forwarding it to their partner. The short survey which collects partner emails was updated so that the email questions included the following statement (text in italics is the new text):

> What is your email? *(Important: To protect respondent confidentiality, your email and your partner's email MUST be different)*

Additionally, all 101 respondents who provided duplicate emails were contacted and asked to provide a new email address for the partner.

Survey Communications

Two sets of survey invitations were originally created – one for respondents who identified as LGBT (as indicated in the sample database) and another set for respondents who did not identify as LGBT. The survey invitations and reminders for the non-LGBT sample had this introductory statement:

> *We would like to invite you to participate in an important research study Gallup is conducting for The Ohio State University and the University of Minnesota. It will help social scientists understand the effects of daily experiences, stress, and family experiences on the health of individuals and families.*

While the LGBT survey invitations and reminders contained this statement, this type of language has successfully motivated participation in other studies of LGBTQI+ individuals conducted by Gallup, such as the Generations study.

> *We would like to invite you to a very important study that aims to expand the research around LGBTQI+ individuals. This survey explores a variety of experiences and makes a special effort to increase the visibility of LGBTQI+ individuals. Gallup, The Ohio State University, and the University of Minnesota are working together on this study. It will help social scientists understand the effects of daily experiences, stress, and family experiences on the health of individuals and families.*

However, on September 23, 2020, a respondent contacted one of the research team members at Gallup. The respondent's partner found the email communication. The participant felt the email had outed them to their partner. The incident was reported to the Internal Review Board (IRB), and all communications were immediately updated so that all respondents received the same set of communications (the version with no mention of LGBTQI+ individuals).

Incentives

All incentives were sent to the respondent via a link to Rybbon. Rybbon is Gallup's incentive partner for all Gallup Panel surveys, and Gallup Panel members are familiar with the redemption process. Survey invitations and reminders let all respondents (Gallup Panel and non-Gallup Panel) know that they would receive an electronic reward which they can use to select a gift card from their choice from stores and restaurants. The link takes respondents to the Rybbon website, where they can redeem their reward for their choice of gift cards, which include major retailers and restaurants. Respondents also have the option

to donate their incentive to a charitable organization. Once the respondent selects their incentive, they receive an immediate email back from Rybbon with a gift card that can be printed or used electronically.

Respondents who completed the survey initially received a $20 post-paid incentive. However, due to lower than desired participation, the post-paid incentive was increased to $50 on November 16, 2020 (see the incentives section for more details). All survey materials and communications were updated to reflect the new amount. Respondents who had not yet completed the survey as of November 16, 2020 were sent an email with the updated incentive amount. Additional reminders were sent on November 27, December 5, and December 10. The $50 incentive amount was paid for all surveys completed on or after November 16, regardless of the date of the communication that a respondent used to access the survey. Prior to the increase, the participation rate was 19%. After the increase, the overall participation rate for the first sample group increased to 22%. Table 2 provides more details about each sample group by incentive amount and number of completes. The incentive for the time-diary was $15 and did not change during the field period.

Table 2: Main survey completes by sample group and incentive type

Sample group description	**Date invited**	**N invited**	**N completed**	**N Completed with $20 incentive**	**N Completed with $50 incentive**	**Participation rate** (useable completes/number invited)
Initial sample of panel/recontact	September 1, 2020	7691	1662	1444	218	22%
Telephone recruits	November 12 – January 26, 2021	925	431	46	385	47%
1st oversample of Black, Hispanic, and/or low education	November 24, 2020	3156	516	NA	516	16%
2nd oversample of Black, Hispanic, low education, and/or younger	January 13, 2021	5,418	790	NA	790	15%
Asian oversample	March 19, 2021	724	243	NA	243	34%

Survey Changes During the Field Period

During fieldwork, two changes were made on November 17, 2020 to address what were believed to be response errors. First, a large number of respondents had missing age data for children in the household roster. After a review of the survey, it was determined that the visual design may be resulting in respondents overlooking this question. Two fixes were made. First, a prompt was added to the survey requesting a response if it was left blank. Second, a follow-up survey was sent to everyone who previously left the question blank (n=464 were invited and n=295 completed). This short survey asked them to provide the number of children in their household and the age of each child.

The second change occurred on the time-diary. Individuals were asked, "Including yourself, how many people live in your household?" A larger than expected number of people gave an answer of 1. Although some people could have uncoupled between the time they completed the main survey and the time-diary, due to the nature of the target population (people in a couple), a household size of one should be very uncommon. The survey was updated so that option "1" in the dropdown would read "1 (I am the only person in the household)."

During the field period, there was some concern that a larger than expected number of different-gender individuals reported not being married or living with a partner and screened out of the survey

(approximately 25% of respondents who started the survey screened out). These individuals were sampled from people who had recently reported that they were married or living with a partner and almost all were expected to qualify. Gallup selected ten respondents who had screened out the survey and called them to ask about the screening question and the respondent's interpretation of the question. Overall, we detected no issues with the interpretation of the question and could categorize the reasons for screening out into the following two groups:

- Screened out: Respondent is married but not living with their spouse
  - "I am currently married but separated from my spouse."
  - "I currently live by myself and therefore answered no."

- Screened out: Respondent living situations have changed
  - "I recently separated from my spouse. I have a partner, but we do not live together."
  - "I found that particular question clear and straightforward. I also appreciated the wording used, which is very inclusive to fit the many scenarios of cohabitation found in our country."
  - "We were living together during the quarantine period but it didn't work out. So we went our separate ways."

The final survey change made during the field period was made at the end of 2020 to update questions to include the year 2021 in response options. Q31, Q32, Q33, Q34, Q35, Q36, Q87, PARENT3, and AGE were all updated to add 2021 to the year option. Additionally, Q12_C was replaced with D12_C2.

The question asked in 2020, D12_C, read:

> *Do you expect your household's income this year (2020) to be more than, less than, or about the same as your household's income last year (2019)?*
> *1 Less than last year*
> *2 About the same as last year*
> *3 More than last year*

The new question asked in 2021, D12_C2, reads:

> *Was your household's income in 2020 more than, less than, or about the same as your household's income in 2019?*
> *1 Less than 2019*
> *2 About the same as 2019*
> *3 More than 2019*

After the conclusion of fieldwork, the research team discovered that there was an error when Q33 (date of marriage) was updated at the end of 2020. The process of updating required creating a new version of the question. The wrong question text was mistakenly pulled in, and the question asked the date the respondent was last tested for HIV. Responses from the incorrect version of the question were removed from the dataset. In early 2022, respondents were recontacted and asked to provide the date they were married to their partner. The question text specified it should be the partner they were married to at the time they completed the NCHAT survey.

Q33 – REASK:

> *Please think about the spouse you were married to when you completed the National Couples' Health and Time Study in <MONTH> <YEAR>.*
>
> *When were you and your spouse legally married?*

Response Rates

Survey Response Rates

The survey achieved an overall response rate of 28% for primary respondents, using AAPOR RR3. The AAPOR RR3 calculation uses the eligibility rate to estimate eligibility (*e*) for unknown cases. This calculation only includes weighted completes in the numerator. The overall eligibility rate was 73%. This calculation is based on the number of people who attempted the survey and were ineligible (either did not pass the screening question (married or living with a partner) or had to be removed from the data because they were not between the ages of 20 and 60). The estimated eligibility rate may be somewhat conservative. It is based on responses given by people who attempted to take the survey and screened out (were not married or cohabiting). In practice, respondents who were not married or cohabiting may have self-screened out of the survey, as all communications emphasized that this was a study for couples. Therefore, the true eligibility rate for all sampled cases may be lower than what has been estimated.

There are several ways to calculate response rates, and some methods include full and partial completes. Partial completes are worth considering in the calculations of this survey, particularly given the length of the survey. If partial completes are included in the numerator of the response rate calculation (AAPOR RR4), the response rate is 31%.

Table 3: Response outcomes – primary respondents

1	Total sampled cases	17,914
2	Completed survey (weighted cases)	3,642
3	Ineligible	1,416
4	Partial complete	397
5	Refused consent	95
6	Attempted survey (2+3+4+5)	5,550
7	Unknown (1-6)	12,364
8	Eligibility rate ((2+4)/6)	73%
9	Estimated eligible unknowns (7*8)	9,026
10	Total eligible sample (2+4+5+9)	13,160
	AAPOR RR3 (2/10)	27.7%
	AAPOR RR4 ((2+4)/10)	30.7%

Response rates did vary by sample group. The sample of Gallup Panel members expected to qualify for the different-gender group had an AAPOR3 response rate of 22%, while LGBT Panel members, who were expected to qualify for the same-gender group, had a response rate of 50%. For the ABS recontact sample, the AAPOR3 response rate was 17%, and the phone recontact sample, which was screened and invited via phone, had an AAPOR3 response rate of 47%.

These response rate differences by sample group are not unexpected. The Gallup Panel sample that was selected for the different-gender group was stratified by age, gender, race/ethnicity, and education level. Individuals with high school education or less and individuals who are Black or Hispanic tend to have significantly lower response rates than individuals who have higher levels of education or who are white. The addition of these oversamples did pull down overall response rates for this group, but they improved the demographic composition and representativeness of the sample. The LGBT sample from the Gallup Panel and the LGBT sample that was recruited via phone had much higher response rates than the other groups. This was in part because oversamples were not possible within this subsample. However, Gallup has found that response rates tend to be strong for this population when the survey is directly related to sexual and gender minority issues, across all demographic groups. Not unexpectedly, the LBGT recontact sample, which was drawn from past ABS surveys and emailed the survey invitation, had the lowest response rates and eligibility rates. It had been at least two years since this sample had any contact with Gallup, and this group would not have been checking their email for survey invitations.

It is worth noting that response rates are just one metric that can be used to evaluate the quality of a survey and non-response bias. Many researchers rely on this simple measure as the primary metric of sample quality and bias, yet research has shown that there is no correlation between response rates and non-response bias (Groves, 2006; Groves & Peytcheva, 2008; Yeager et al., 2011; MacInnis et al., 2018; Davern, 2013; Keeter et al., 2006). Recent work by Yeager et al. (2011) and MacInnis et al. (2018) have explored the accuracy of estimates from RDD, probability, and non-probability samples and found that probability-based panels produce highly accurate estimates and that there was no relationship between response rates and the accuracy of estimates. Further, response rates declined from the 2011 study to the 2018 follow-up study, yet quality did not erode. This is consistent with other research that has found declining response rates have not compromised estimates (Keeter et al., 2006). A paper by Davern (2013) provides an excellent summary and critique of response rates, the focus placed on this sole metric, and the great expense some studies go to in an effort to improve response rates.

Non-response bias occurs when respondents are not missing at random and the missingness is correlated with the survey measures (Groves, 2006). Further, non-response bias occurs at the individual estimate level, not the survey level. Response rates alone cannot evaluate any of these components, and a non-response bias analysis must be carried out to get a true understanding of potential bias. Gallup conducted a non-response bias analysis, and the results can be found in the non-response bias analysis section.

Partner Response Rates

Primary respondents were asked to invite their partners to complete the survey. Partners provided an email to Gallup and were then invited to all survey activities. A total of 1,968 partners provided a valid email address. Of those,1,515 eligible partners completed the survey, for an AAPOR RR1 partner response rate of 42%. This rate is calculated based on the number of completed partner surveys (n=1,515) and the number of primary respondent surveys that were completed, from which the partners were recruited (n=3,642). Eligibility has not been estimated for partners who did not complete the survey, as very few partners screened out of the survey, and in theory, all eligible primary respondents also had an eligible partner.

Time-Diary Response Rates

The time-diary was sent to all respondents who completed the main survey and had a participation rate of 75%. Interestingly, the rate was 75% for both primary and partner respondents. The participation rate is the number of completed time-diaries divided by the number of completed main surveys. Some time-diaries did not provide enough data to be retained, and 55% of main respondents had a time-diary that was included in the final weighted time-diary dataset. To receive a weight, respondents had to report at least five activities, report at least 2 of 3 basic activities, including eating, sleeping, and personal care (as primary or secondary minutes), report fewer than 3 hours of don't know/cannot remember primary minutes, and complete the time-diary through at least 8pm on their time-diary day which is when the last reminder that day was sent out.

Data Cleaning

Prior to weighting the data, several data cleans were implemented.

Age was calculated for all respondents. Age was asked on the survey as month and year of birth. Age was calculated using the date the survey was completed and the provided birth date. The 15th of the month was used as the exact date of birth for purposes of the age calculation. AGE_1 (month) and AGE_2 (year) are the original date-of-birth questions asked on the survey, and AGE_SD is the computed age variable. Respondents who were under the age of 20 or over the age of 60 were flagged as ineligible for the survey. Most respondents who were excluded from the weighted data were aged 61 at the time the survey was completed. These respondents were aged 60 at the time they were sampled but had aged out of the study by the time they completed the survey.

Respondents needed to answer age, sex, gender, partner gender, partner sex, sexual identity, race, ethnicity, and education level to determine study eligibility and to be weighted. Age, gender, and partner gender are necessary to determine age eligibility for the study and if the respondent is part of the same-gender or different-gender group. Sexual identity, race, ethnicity, and education level are necessary for post-stratification weighting. If a respondent was missing all of these demographic variables after data cleaning and imputation (described in the next paragraph), they could not be retained as a completed survey. Most of these variables were asked at the end of the survey. If a respondent made it to this point in the survey and had answered the substantive questions without breaking off, they also were very likely to answer the demographic questions.

Some respondents did have item-level missing data on one or more key demographic questions, despite answering the majority of other survey questions. Age had the most missing data, which is likely due to the way the question was asked (asking date of birth versus asking age). To have as much data as possible, the team decided to impute missing demographic questions using known demographic information in the sample records. This was carried out for age, gender, race, ethnicity, and education level. The original survey data has been retained in the file. A second imputed variable also appears in the file. If a case was missing a response in the original variable, the imputed variable contains the response from the sample frame database, if a response was available. In some cases, a response was also missing on the sample frame. All imputed variables have _SUPP at the end of the original variable name.

Weights

Primary Respondent Weights

Sample data were weighted to minimize bias in survey-based estimates. Eligible respondents who completed the survey were assigned final weights.

Base-weights

Gallup Panel members have base-weights that reflect member's selection probabilities into the Panel. In addition to Panel members, an additional sample of LGBT people were selected from Gallup's recontact dataset. The recontact dataset is a list of previously surveyed people in various surveys who have agreed to be recontacted but are not members of the Gallup Panel (see sampling section for more details). Base-weights were calculated for the recontact sample.

Sampling was carried out separately for LGBT adults and non-LGBT adults. All LGBT Panel members who met the age criteria were sampled, while only a sub-sample of non-LGBT cohabitating or married adults was sampled. This was done with the purpose of ending up with large enough sample sizes for the two categories of adults. This resulted in an oversample of LGBT adults. Base-weights were adjusted to account for this oversampling.

Targets for Weighting

One complexity of this project is there is no single data source that can be used to generate distributions for weighting. The 2019 American Community Survey (ACS)[5] is used to form the basis of the same-gender and different-gender targets, and the National Health Interview Survey (NHIS)[6] is the basis of the sexual minority targets.

[5] American Community Survey. 2019. https://www.census.gov/programs-surveys/acs
[6] National Health Interview Survey, 2019. https://www.cdc.gov/nchs/nhis/index.htm. More information about the 2019 redesign can be found here: https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm

One of the primary purposes of the study is to perform analyses of adults living in same-gender versus different-gender couples. Therefore, the weighting aimed to adjust demographic distributions within these two main groups independently and then merge and put them in the correct proportions.

The 2019 American Community Survey (ACS) provides variables to subset the data to the NCHAT target population. All targets were pulled based on individuals who are 20 to 60 years old and cohabiting or married to a same-sex or different-sex partner. This sample unweighted consists of about 10,000 individuals. The following distributions were calculated from the 2019 ACS according to age, sex, race, ethnicity, education, and marital status. All targets were calculated for adults 20 to 60 who are cohabiting (either married or not married) and part of a same-sex or different-sex couple. (The term same-sex and opposite-sex were used in the ACS questions about household composition, so we reference sex and not gender).

Age: (1) 20-30; (2) 31-40; (3) 41-50; (4) 51-60;
Sex: (1) Male; (2) Female; *(male and female were used to match ACS targets and avoid very small groups for weighting. All other categories were set to missing and then imputed in a multiple imputation process)*
Race: (1) White only; (2) Black only (3) Asian only; (4) Others
Ethnicity: (1) Hispanic; (2) Non-Hispanic;
Education: (1) High school or less; (2) Above high school but less than Bachelor's; (3) Bachelor's or more
Marital status: (1) Married; (2) not married (cohabiting)

Within the NCHAT survey data, multiple variables were used to identify *couples* in the survey data as either same-sex or different-sex. (D1 – What sex appears on your original birth certificate", If D1 was missing, then D2 "Which of the following best describes your gender," and then D2_1 series – "do any of the following describe your gender." Similar variables were available about the main respondent's partners. Respondents were weighted if, based on these questions, they could be classified either as an adult living in a same-sex couple or an adult living in a different-sex couple. There were 33 cases in the dataset where the respondent indicated they were cohabiting with more than one partner and provided demographic information for the additional partner(s). In this situation, they were asked to indicate which of these partners they considered to be the main partner and were told to think about this partner when answering the survey questions. For these cases where the respondent had more than one survey partner, the partner indicated as the "main" partner was used in weighting.

Although the ACS includes a large number of same-sex couples, the ACS does not include a direct question asking about sexual orientation. Therefore, for weighting purposes, the distribution of sexual orientation for cohabiting and married couples aged 20 to 60 was drawn from the 2019 NHIS.

The LGB category in the NHIS data was based on a question asking whether the respondent identified as "Lesbian, Gay, Bisexual, and something else." This direct question is of particular interest because the gender composition of the couple cannot be used to identify most bisexuals. In the NCHAT sample, the majority of bisexuals live with a different-gender partner/spouse rather than a same-gender partner/spouse. Due to the sample design, individuals who are bisexual are overrepresented in the survey results and are largely observed within the different-gender couple group. Therefore, it is also important that the weighting also adjusts for the proportion of bisexual individuals. The NHIS is used to calculate the sexual orientation distribution targets and demographic distributions. Non-LGB cohabiting/married adults were also calculated from NHIS 2019 data.

The following distributions were calculated from the 2019 NHIS: overall sexual orientation distribution and the distributions within non-LGB and LGB adults according to age, sex, race, ethnicity, education, and marital status. The NHIS sexual orientation question was collapsed into two categories: Straight/non-LGB and LGB. The LGB category includes people who answered "something else." In the 2019 NHIS, the sample size of non-LGB people was small, and age, sex, education, race, ethnicity, and marital status distributions for this group are based on less than 400 people (unweighted N).

Within the NCHAT survey data, the variable IDENTITY was used to code sexual identity for weighting. The identity question asks people which of the following they consider themselves to be and has 11 "select all that apply" options. Individuals who selected the "heterosexual or straight" response option and no other option were coded as non-LGB for purposes of weighting. Individuals who selected an option other than "heterosexual or straight" were coded as LGB for purposes of weighting.

Both ACS and NHIS targets are weighted distributions. The weights the two datasets use are projectable to the total number of target adults living in the United States.

Raking

Base-weights were then adjusted using a multiple iterative raking procedure to post-stratify by age, sex, education, race, ethnicity, and marital status.

While not a large issue, some cases have missing data in one or more raking variables. Therefore, a multiple imputation procedure is applied to impute missing values. This procedure uses sequential regression models in multiple iterations to predict missing values in multiple variables. Each of the raking variables are used both as predictor and dependent variables in the imputation models. Since predicted values are not as accurate as actual reported values, in order to reduce their influence on weights, ten different values are predicted for each missing, generating ten different datasets. Each of these ten datasets is raked and weights across ten datasets are averaged to produce a single weight value for each case.

The poststratification weighting was conducted in two stages[7]. In the first stage, two groups (LGB and non-LGB) were weighted to the NHIS targets. In the second stage, the weights from the first stage were further adjusted by couple type (adults in same-sex couples vs. adults in different-sex couples) from the ACS data targets. The NHIS was used for the first stage and the ACS for the second stage due to the small NHIS sample sizes and the importance of generating final weights based on targets from the larger sample sizes of the ACS.

Different trimming options were considered for the final weights and the weights were trimmed to arrive at an optimal balance between variance in weights and fit to the target distributions. The two-stage weighting produced weighted distributions that were within a few percentage points of the ACS targets for all weighting variables for both same-sex and different-sex couples. The weights also produced demographic distributions by sexual orientation that are comparable to the NHIS targets. A comparison of the weighted NCHAT demographic distributions to the ACS demographics are provided in the next two tables.

Table 4: Demographic Comparison of Weighted American Community Survey (ACS) and Weighted National Couples' Health and Time Study (NCHAT) (Primary Respondents)

	Weighted ACS		Weighted NCHAT	
	Same-sex	Different-sex	Same-sex	Different-sex
Marital Status				
Married	56%	85%	56%	85%
Cohabiting	44%	15%	44%	15%
Sex				

[7] Prior to this two-stage raking procedure, an alternative strategy was considered. The ACS targets were used in the first stage to match population targets specific to same-sex and different-sex couples, regardless of sexual orientation of the individual. We hoped that weighting by same-sex and different-sex couples would produce sexual orientation distributions that were comparable to NHIS data. While the weights produced demographics that matched most ACS targets within a few percentage points, a comparison of demographic distributions by sexual orientation to the NHIS targets showed some large differences. Unfortunately, there is no source that gives researchers the ability to look at demographic distributions by sexual orientation and same-sex or different-sex couple.

Male	48%	49%	50%	49%
Female	52%	51%	50%	51%
Age				
20-30	24%	15%	23%	14%
31-40	28%	28%	29%	28%
41-50	23%	28%	24%	29%
51-60	24%	29%	24%	29%
Education				
High school or less	22%	32%	20%	31%
Above High school below Bachelor's	31%	29%	32%	29%
Bachelor's or above	47%	40%	48%	40%
Ethnicity				
Hispanic	17%	17%	16%	18%
Non-Hispanic	83%	83%	84%	82%
Race				
White only	77%	77%	77%	76%
Black only	10%	8%	10%	8%
Asian only	4%	7%	4%	7%
Other	9%	8%	9%	8%

Table 5: Comparison of Weighted National Health Interview Survey (NHIS) and Weighted National Couples' Health and Time Study Primary Respondents (NCHAT)

	Weighted NHIS						**Weighted NCHAT**	
	LGB	Lower 95% INT	Upper 95% INT	**Non-LGB**	Lower 95% INT	Upper 95% INT	**LGB**	**Non-LGB**
Marital Status								
Married	54.4%	47.8%	60.8%	82.8%	81.8%	83.7%	58%	86%
Cohabiting	45.6%	39.2%	52.2%	17.2%	16.3%	18.2%	42%	14%
Sex								
Male	30.3%	25.0%	36.1%	49.6%	48.5%	50.7%	34%	49%
Female	69.5%	63.6%	74.7%	50.4%	49.3%	51.5%	66%	51%
Age								
20-30	37.3%	31.3%	43.8%	17.8%	16.8%	18.7%	33%	14%
31-40	27.8%	22.6%	33.6%	27.6%	26.6%	28.5%	31%	28%
41-50	18.8%	14.5%	24.1%	27.6%	26.6%	28.6%	20%	29%
51-60	16.1%	12.1%	21.0%	27.1%	26.1%	28.1%	16%	29%
Education								
High school or less	29.4%	23.5%	36.1%	35.4%	34.0%	36.7%	25%	31%
Above High school below Bachelor's	33.7%	28.1%	39.9%	29.7%	28.6%	30.9%	33%	29%
Bachelor's or above	36.9%	31.3%	42.9%	34.9%	33.6%	36.3%	42%	40%
Ethnicity								
Hispanic	14.1%	10.1%	19.4%	19.1%	17.6%	20.7%	15%	18%
Non-Hispanic	85.9%	80.6%	89.9%	80.9%	79.3%	82.4%	85%	82%
Race								
White only	77.7%	71.3%	82.9%	79.2%	77.8%	80.5%	75%	76%
Black only	12.8%	8.6%	18.7%	9.8%	9.0%	10.8%	10%	8%
Asian only	4.0%	2.1%	7.7%	7.7%	7.0%	8.6%	4%	7%
Other	5.5%	3.2%	9.4%	3.3%	2.6%	4.0%	11%	8%

Note. INT = Confidence Interval; LGB = Lesbian, Gay, or Bisexual

Other External sources

Other sources of weighting targets were considered. However, data sources that could be used had either a small sample size for LGBT people for the purposes of creating subgroups for weighting (Gallup Poll Social Series), were dated (older years of the National Health Interview Survey), or were not based on the exact same target population (e.g., the National Survey of Family and Growth had a younger target population (upper age limit of 50) and only identified different-sex couples).

We conducted a comparison of the final weighted NCHAT data to these sources to determine whether there were any systematic indications that certain subgroups were misrepresented with the final weights. Even though we did not observe such systematic indications, there is one key distinction. The final weights produced an NCHAT sexual orientation distribution within different-sex couples of 1.9% LGB vs. 98.1% non-LGB. This 1.9% is lower than NSFG's 4.7%. Even when sub-setting the NCHAT data to match the younger age restriction of the NSFG (20-50), the NSFG analysis resulted in a weighted estimate with 3.3% of different-sex couples consisting of an LGB identified respondent.

Couple Weights

Weights were generated for interviews where both the primary respondent as well as their partner completed the main survey. The application of couple weights will produce representative estimates of the couples where at least one partner is 20-60 years old and the partners are married or cohabiting in the United States. The couple weights can also be used to generate estimates for same and different gender couples.

All stages of weighting used a similar strategy to that used for primary respondents' weights. The base weights constructed for primary respondents were used for couples (see the previous section).

The 2019 ACS and NHIS were also used for weighting target construction. In addition to personal weights, both of these sources have household-level weights that can be used to generate couple-level estimates. The 2019 ACS was used to construct targets for same-sex and different-sex couples.

The following distributions were calculated from the 2019 ACS within same-sex and different-sex couples:
Age: (1) both under 34; ()2 One under 34, the other 34-47; (3) One under 34, the other 48+; (4) both 34-47; (5) One 34-47 the other 48+; (6) both 48+;
Sex: (1) both male; (2) both female; (3) different sex *(In order to match categories with ACS data and also to avoid very small groups for weighting, the small number of responses that were "Trans man," "Trans woman," "Do not identify with any of the above, "as well as non-response from either the main respondent or their partner, were not included in the sex categories of couples. Missing couples' sex categories, due to this, were later filled in the multiple imputation procedure).*
Race: (1) both white; (2) both Black; (3) both Asian; (4) All Others
Ethnicity: (1) both Hispanic; (2) one Hispanic; (3) neither Hispanic
Education: (1) both with high school or less; (2) one with high school or less, the other with between high school and Bachelor; (3) one with High school or less, the other with Bachelor's +; (4) both with between high school and Bachelor; (5) one with between high school and Bachelor, the other with Bachelor's+; (6) both with Bachelor's+
Marital status: (1) Married (2) not married (cohabiting)

The NHIS was used to construct targets for couples with an LGB (gay/Lesbian, bisexual, or something else other than straight) primary respondent and couples with a non-LGB primary respondent. NHIS does not provide data for primary respondents' partners' sexual orientation. For this reason, NHIS data could not be grouped with more precise categories of couples based on both members' sexual orientation.

The same demographic distributions from NHIS 2019 were calculated within couples with an LGB main respondent and couples with a non-LGB main respondent. Responses on these demographics for main respondents were used from the main respondent survey responses. For partners, the partner responses on their survey were used for the partner demographics. Some missing values were imputed based on

information in the panel database (see cleaning section above). Values that were still missing after cleaning were imputed in a multiple imputation process.

Like the main respondents' weights, couples' weights were adjusted in two stages. In the first stage, the two groups (couples with an LGB main respondent and couples with a non-LGB main respondent) were weighted to the NHIS targets. In the second stage, the weights from the first stage were further adjusted by couple type (same-sex couples vs. different-sex couples) from the ACS data targets.

Different trimming options were considered for the final weights and the weights were trimmed to arrive at an optimal balance between variance in weights and fit to the target distributions. Like the primary respondents' weights, the two-stage weighting produced weighted distributions that were within a few percentage points of the ACS targets for all weighting variables for both same-sex and different-sex couples. The weights also produced demographic distributions by sexual orientation that are comparable to the NHIS targets.

Time-diary Weights

The procedures for time-diary weighting were the same as the main survey weights. The main survey respondents were subset to those who had provided a valid response to the time diary. The weights from the main survey were used as the base-weight. Next, the sample was post-stratified to the targets, using the same procedures as outlined in the section above. These procedures were carried out for the sample of main respondents and partner respondents.

Weighting variables

Weight_mainrespondent is the constructed weight variable to be used for analysis of the main respondent file. This weight should be used when analyzing data from primary respondents and will generate nationally representative estimates of individuals aged 20 to 60 who are living or cohabiting with a partner. This weight can also be used to generate representative estimates when analyzing sub-groups, such as individuals who are in a same-gender or different-gender couple, or individuals who are gay, lesbian, or bisexual.

Weight_couple is the constructed weight variable to be used for analysis of couple level data from primary respondents and their partners. This weight should be used when analyzing the dyadic data (primary and partner responses). Cases from the primary respondents and partner respondents should not be treated as individual cases and combined to create a larger sample of individual adults. Weights are not valid for this purpose. Further, this type of analysis would require accounting for clustering within couples and potentially highly correlated responses (intraclass correlation). It is recommended that researchers do not take this approach.

To appropriately account for the design effect from weighting, and to obtain accurate standard errors, an analysis package that can take into account the complex study design should be used. In Stata, this can be done using the svyset commands.

Non-Response Bias Analysis

Response rates are a measure of non-response, and high response rates may indicate lower *risk* of non-response bias than a study with low response rates. However, response rates are very poorly correlated with non-response bias and are not a measure of bias (Groves, 2006; Groves & Peytcheva, 2008; Yeager et al., 2011; MacInnis et al., 2018; Davern, 2013; Keeter et al., 2006). For bias to occur, non-respondents must have systematically different responses from the respondents on the substantive variables of interest.

To assess the potential for bias, the National Academy of Sciences and the Office of Management and Budget (OMB) have outlined several methods that can be used when response rates indicate a potential risk of non-response bias[8]. These methods include:

- comparing the weighted and unweighted samples and results
- benchmarking survey estimates and respondent characteristics to gold standard estimates
- conducting a non-response follow-up survey or executing intense non-response follow-up
- time-of-return analysis (comparing late and early responders)
- comparing respondents and non-respondents
- comparing response rates across sub-groups

This non-response bias analysis utilized several of these methods, which are described in the following sub-sections. This study had the advantage of utilizing the Panel and recontact samples, which provided researchers with a demographic profile of all respondents who were sampled, including non-respondents. Further, these individuals have completed previous surveys, which could be used to look at the psychographics of respondents and non-respondents. Finally, the study went to considerable efforts to convert non-respondents, which were outlined in previous sections. The survey was in the field for more than six months, which allowed us to explore early and late responders. Taken as a whole, the findings outlined in each of the sections below do not seem to indicate the potential for significant non-response bias.

Comparison of Demographic Characteristics of Invited Participants and Responding Participants

Table 6 presents the number of invited participants by demographic groups. Although these numbers are limited in their ability to help assess non-response bias, they do provide helpful context for understanding the sample composition and the rest of the analyses in this section.

All numbers in the “invited” column are based on the known demographic attributes of the Panel and recontact sample members at the time they were sampled. All numbers in the “completed” column are based on responses provided in the NCHAT survey.

It is important to note that the sample of invited participants was based on oversamples of several demographic groups to meet minimum sample size requirements for certain subgroups and to account for differential non-response (see sampling section). Therefore, on an unweighted basis the invited sample and completed sample deviate from the population targets. Weighting corrected for deviations from the population targets. Additionally, not everyone who was invited was a part of the target population. The difference between invited and completed includes both people who did not respond and people who were ineligible to participate. The next section shows response rates by demographic group, and accounts for the ineligible sample.

Table 6: Comparison of demographics of invited participants and responding participants

	N Invited (N = 17914)	**N Completed*** (N=3642**) (unweighted)
Gender		
Male/Man***	10062	1774
Female/Woman	7754	1737
Something other than man/woman	NA	127
Age		
20-30	3637	435
31-40	4779	959
41-50	3912	934
51-60	5499	1257

[8] National Research Council 2013. *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18293

Education		
HS or less	6636	640
Above HS below BA	4973	947
BA or above	6185	2050
Ethnicity		
Hispanic	3676	576
Non-Hispanic	14115	2926
White vs. non-white		
non-Hispanic white	9544	2329
non-White	8374	1313
Black vs. non-Black		
non-Hispanic Black	3436	380
non-Black	14393	3262
Census Region		
Northeast	3289	667
Midwest	3315	681
South	6800	1229
West	4422	1065

*Includes only respondents who met eligibility requirements and who were assigned a weight.
**Numbers do not sum to the total invited or completed because of missing data. The unsupplemented demographic variables were used to create this table.
***Gallup respondent database (which is used to evaluate who was invited) has gender with the categories male and female. The NCHAT survey (which is used to evaluate who completed) asked gender with the categories man, woman, trans man, trans woman, and do not identify as any of the above.

Comparison of Response Rates

This analysis examines the relationship between demographics and response rate. For example, we calculate the response rate (R.R.) for different age groups and examine whether R.R. is different across age groups. A similar analysis was done for several other variables, including gender, race, ethnicity, marital status, and employment status, to see if certain subgroups had a lower propensity to respond, which could lead to potential bias in estimates if not corrected for in weighting.

All R.R. calculations were carried out following the AAPOR3 response rate calculation. Each person in the sample was classified into one of three eligibility categories: Known Eligible, Known Ineligible, and Unknown Eligibility. For this study, a person was considered eligible if they were between the ages of 20 and 60 and cohabiting with a partner. A person was identified as 'Known Eligible' if they confirmed meeting these criteria. A person was treated as 'Known Ineligible' if they did not meet the eligibility criteria. Finally, a person was included in the 'Eligibility Unknown' category if they did not answer the necessary question(s) related to living with a partner. This unknown eligibility group primarily includes people who never started the survey but also includes people who started the survey but did not answer the initial screening questions.

The overall R.R. was 28%. The R.R. was also calculated by demographic categories for age, gender, race, ethnicity, marital status, educational attainment, census region, population density, employment status, and sexual identity. These demographic indicators were available on the sample frame and based on earlier self-reports. Response rates were expected to vary by group for several of the demographic variables. For example, younger respondents, racial and ethnic minorities, and those with low levels of education tend to respond to surveys at lower rates. The demographic groups selected for this analysis are also expected to be related to the key variables of interest. For example, racial and ethnic minorities may have very different daily experiences and stressors as compared to white respondents. As another example, we hypothesized that employed and not employed individuals may have very different experiences during the COVID-19 pandemic, which could also be correlated with participation in the survey and substantive responses to the survey questions. Response rates by group are presented in the table below.

Table 7: Response rates by demographic groups

Demographic Group	AAPOR3 RR
Age	
20-30	20%
31-40	27%
41-50	31%
51-60	33%
Gender	
Male	26%
Female	31%
White vs. non-white	
Non-Hispanic white	33%
Non-white	23%
Black vs. non-Black	
Non-Hispanic Black	20%
Non-Black	30%
Ethnicity	
Hispanic	21%
Non-Hispanic	30%
Education	
High School or Less	18%
Some College	29%
College Grad	38%
Marital Status	
Married	37%
Cohabiting	28%
Sexual Identity	
Panel non-LGBT	22%
Panel LGBT	50%
ABS recontact LGBT	17%
Phone recontact LGBT	47%
Census Region	
Northeast	30%
Midwest	29%
South	25%
West	32%
Population Density	
1 – Top 20%	32%
2	30%
3	31%
4	30%
5 – Bottom 20%	28%
Employment Status	
Employed	34%
Not Employed	32%

As expected, the R.R. was lower for some demographic groups than others. Individuals who are between the ages of 18 and 30, male, non-white, non-Hispanic Black, Hispanic, high school education or less, unmarried, and not-LGBT had lower response rates than their counterparts. These factors were all taken into account in the weighting, thereby minimizing the potential for bias due to underrepresentation in the

sample. The response rate was nearly identical across subgroups, within region, and population density, with the exception of the R.R. for the South, which was slightly lower than other groups.

Weighted and Unweighted Results – Key Outcomes

This section presents the weighted and unweighted distributions of survey variables that are of key interest to the research team. A comparison of weighted and unweighted distributions gives information about the impact the weighting adjustments had on the estimates. For most questions and response options shown in this analysis, the weighting adjustments made relatively minor differences in the estimates.

Table 8: Relationship Quality Measures – Unweighted and Weighted

I have a warm and comfortable relationship with my spouse/partner (Q19)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Not at all true	59	1.6	1.7
A little true	141	3.9	4.3
Somewhat true	288	7.9	7.7
Mostly true	685	18.8	17.7
Almost completely true	840	23.1	22.1
Completely true	1620	44.5	46.4
How rewarding is your relationship with your spouse/partner (Q20)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Not at all rewarding	73	2.0	1.7
A little rewarding	198	5.4	5.8
Somewhat rewarding	348	9.6	8.8
Mostly rewarding	859	23.6	23.2
Almost completely rewarding	1044	28.7	26.1
Completely rewarding	1119	30.7	34.4
In general, how satisfied are you with your relationship? (Q21)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Not at all satisfied	98	2.7	2.4
A little satisfied	152	4.2	3.8
Somewhat satisfied	289	7.9	7.4
Mostly satisfied	774	21.3	19.8
Almost completely satisfied	1257	34.5	32.1
Completely satisfied	1070	29.4	34.5
In general, how committed are you to your current spouse/partner? (Q22)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Not at all committed	18	.5	0.4
A little committed	41	1.1	0.8
Somewhat committed	102	2.8	1.8
Mostly committed	248	6.8	5.0
Almost completely committed	408	11.2	8.5
Completely committed	2819	77.4	83.4

Table 9: Depression Index Measures – Unweighted and Weighted

I was bothered by things that don't usually bother me (Q55A)

	Unweighted N	Unweighted %	Weighted %
Missing	21	.6	0.8
Rarely or none of the time (Less than 1 day)	2057	56.5	60.9
Some or a little of the time (1-2 days)	1080	29.7	27.5
Occasionally or a moderate amount of time (3-4 days)	398	10.9	9.2
Most or all of the time (5-7 days)	86	2.4	1.7
I had trouble keeping my mind on what I was doing (Q55B)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Rarely or none of the time (Less than 1 day)	1438	39.5	44.2
Some or a little of the time (1-2 days)	1167	32.0	32.2
Occasionally or a moderate amount of time (3-4 days)	714	19.6	17.2
Most or all of the time (5-7 days)	318	8.7	6.2
I felt lonely (Q55C)			
	Unweighted N	Unweighted %	Weighted %
Missing	16	.4	0.5
Rarely or none of the time (Less than 1 day)	2351	64.6	70.1
Some or a little of the time (1-2 days)	747	20.5	17.5
Occasionally or a moderate amount of time (3-4 days)	348	9.6	8.0
Most or all of the time (5-7 days)	180	4.9	3.9
My sleep was restless (Q55D)			
	Unweighted N	Unweighted %	Weighted %
Missing	13	.4	0.2
Rarely or none of the time (Less than 1 day)	1235	33.9	37.4
Some or a little of the time (1-2 days)	1204	33.1	33.0
Occasionally or a moderate amount of time (3-4 days)	716	19.7	18.9
Most or all of the time (5-7 days)	474	13.0	10.4
I felt depressed (Q55E)			
	Unweighted N	Unweighted %	Weighted %
Missing	14	.4	0.3
Rarely or none of the time (Less than 1 day)	2033	55.8	62.8
Some or a little of the time (1-2 days)	938	25.8	22.7
Occasionally or a moderate amount of time (3-4 days)	422	11.6	9.9
Most or all of the time (5-7 days)	235	6.5	4.3
I felt like everything I did was an effort (Q55F)			
	Unweighted N	Unweighted %	Weighted %
Missing	14	.4	0.2
Rarely or none of the time (Less than 1 day)	1689	46.4	50.3
Some or a little of the time (1-2 days)	1101	30.2	30.3
Occasionally or a moderate amount of time (3-4 days)	517	14.2	12.8
Most or all of the time (5-7 days)	321	8.8	6.4
I felt hopeful about the future (Q55G)			
	Unweighted N	Unweighted %	Weighted %
Missing	10	.3	0.2
Rarely or none of the time (Less than 1 day)	574	15.8	15.3
Some or a little of the time (1-2 days)	910	25.0	23.3
Occasionally or a moderate amount of time (3-4 days)	1164	32.0	31.6
Most or all of the time (5-7 days)	984	27.0	29.7
I felt fearful (Q55H)			
	Unweighted N	Unweighted %	Weighted %
Missing	20	.5	0.5

Rarely or none of the time (Less than 1 day)	2032	55.8	60.0
Some or a little of the time (1-2 days)	953	26.2	25.3
Occasionally or a moderate amount of time (3-4 days)	453	12.4	10.3
Most or all of the time (5-7 days)	184	5.1	3.8
I was happy (Q55l)			

	Unweighted N	Unweighted %	Weighted %
Missing	14	.4	0.4
Rarely or none of the time (Less than 1 day)	201	5.5	5.4
Some or a little of the time (1-2 days)	662	18.2	16.3
Occasionally or a moderate amount of time (3-4 days)	1264	34.7	32.7
Most or all of the time (5-7 days)	1501	41.2	45.1
I could not get "going" (Q55l)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Rarely or none of the time (Less than 1 day)	1656	45.5	48.4
Some or a little of the time (1-2 days)	1148	31.5	30.8
Occasionally or a moderate amount of time (3-4 days)	559	15.3	14.0
Most or all of the time (5-7 days)	274	7.5	6.7

Table 10: Health and Stress Measures – Unweighted and Weighted

In general, would you say your health is: (Q46A)			
	Unweighted N	Unweighted %	Weighted %
Missing	<10	-	-
Poor	83	2.3	1.9
Fair	523	14.4	14.0
Good	1396	38.3	38.1
Very good	1264	34.7	35.6
Excellent	367	10.1	10.2
In the past week, how stressed have you been? (STRESS1)			
	Unweighted N	Unweighted %	Weighted %
Missing	27	.7	1.0
1 – Not at all stressed	521	14.3	17.9
2	1009	27.7	29.0
3	1080	29.7	28.2
4	669	18.4	16.3
5 – Very stressed	336	9.2	7.6
How stressed have you been about the following? Getting coronavirus (STRESS3_A)			
	Unweighted N	Unweighted %	Weighted %
Missing	27	.7	0.7
1 – Not at all stressed	842	23.1	31.6
2	877	24.1	23.6
3	911	25.0	22.9
4	636	17.5	14.1
5 – Very stressed	349	9.6	7.2

Time-of-Return Comparison

The time-of-return analysis compares the "early" (easy-to-obtain) respondents to the "late" (hard-to-obtain) respondents on selected key variables of interest. This analysis is conducted based on the assumption that the latter group may, in some ways, more closely resemble the population of non-respondents than the early responders. The total pool of 3,642 respondents to the NCHAT survey was split into two groups (early and late) based on the time they took to complete the survey. Not all respondents received their survey invitation at the same time; in fact, some respondents received their initial survey invitation several months after other respondents. To take this into account, time of response was calculated based on each respondent's unique invitation date. Respondents who completed the survey within 14 days of receiving their initial invitation were treated as 'early' respondents. Those who took more than 14 days to respond after receiving their initial invitation were identified as 'late' respondents. Based on this criterion, the 'early' group included 2,448 respondents (67.2 percent), while the remaining 1,194 respondents (32.8 percent) formed the 'late' group for comparison.

For the purpose of comparing the 'early' and 'late' respondents, several NCHAT survey questions were selected, and statistical tests (chi-square tests) were applied to see if the two groups were significantly different on those questions. The distribution of responses across the different response categories and related statistical tests are presented below (both unweighted and weighted) for 'Early' and 'Late' respondents. These tables also show the weighted and unweighted results for both early and late responders and demonstrate overall very similar findings between early and late responders and between weighted and unweighted data.

The first set of measures considered were a battery of three relationship satisfaction measures, which are presented in Tables 11, 12, and 13. No significant differences between early and late responders were detected for any of the three measures.

Next, we examined life evaluation both now and in the future. The results are in Tables 14 and 15. There was no difference in early and late responders in the measure of current life evaluation, but the Chi-square test of life evaluation five years from now was significant. However, when groups were collapsed into those with high future life evaluation (code 7-10) and those with lower future life evaluation (0-6), there we no significant differences.

Self-reported health was also considered, and there were significant differences between early and late responders (Table 16). However, an examination of the results indicates minimal meaningful differences between early and late responders. Overall, both groups report being relatively healthy, which is not unexpected given the age of the study participants.

We also examined the ten measures from the depression index (Q55A through Q55J). For eight of the ten items, there was no statistically significant difference between early and late responders (Q55A – bothered by things; Q55C – lonely; Q55D – restless sleep; Q55E – depressed; Q55F – everything an effort; Q55G – hopeful; Q55H – fearful; Q55I – happy).

Two of the ten questions on the depression index did have a significant difference between early and late responders (Q55B – trouble keeping mind on what I was doing; Q55J – could not get going). Results of the tests for Q55B and Q55J are shown in Tables 9 and 10. Similar to what was found for self-reported health, although the findings were statistically significant, they do not show that early and late responders have substantively different profiles.

Finally, we looked at eight measures of experiential wellbeing. Seven of these measures had no significant difference between early and late responders (W3A – enjoyment; W3B – worry; W3C – sadness; W3D – stress; W3E – anger; W3F – happiness; W3H - loneliness. Early and late responders were significantly different on “boredom” (W3G) (see Table 17).

Overall, the results of the time-of-return analyses do not indicate meaningful differences between early and late responders on key outcome variables and do not suggest that late responders were systematically different in ways that would introduce bias.

Table 11: Early versus late responders - *I have a warm and comfortable relationship with my spouse/partner (Q19)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Not at all true	1.6%	1.7%	60	1.9%	1.2%
A little true	3.7%	4.3%	150	4.4%	3.5%
Somewhat true	7.4%	8.9%	281	6.6%	9.7%
Mostly true	18.2%	20.0%	652	17.3%	19.0%
Almost completely true	23.7%	21.7%	797	22.2%	21.3%
Completely true	45.1%	43.3%	1695	47.3%	45.0%

The chi-square test was not significant (p = 0.19) at the 5% level of significance

Table 12: Early versus late responders - *How rewarding is your relationship with your spouse/partner?* (Q20)

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Not at all rewarding	1.9%	2.2%	73	2.0%	1.3%
A little rewarding	5.1%	6.1%	198	5.4%	6.2%
Somewhat rewarding	8.9%	10.8%	348	8.5%	9.8%
Mostly rewarding	23.7%	23.4%	859	23.1%	23.9%
Almost completely rewarding	29.6%	26.8%	1044	26.1%	25.5%
Completely rewarding	30.7%	30.7%	1119	34.9%	33.2%

The chi-square test was not significant (p = 0.69) at the 5% level of significance

Table 13: Early versus late responders - *In general, how satisfied are you with your relationship? (Q21)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Not at all satisfied	2.5%	3.0%	98	2.6%	2.0%
A little satisfied	4.2%	4.0%	152	4.3%	2.7%
Somewhat satisfied	7.6%	8.6%	289	7.0%	8.4%
Mostly satisfied	20.3%	23.2%	774	18.6%	22.5%
Almost completely satisfied	35.7%	32.1%	1257	32.6%	31.0%
Completely satisfied	29.6%	29.0%	1070	34.8%	33.3%

The chi-square test was not significant (p = 0.13) at the 5% level of significance

Table 14: Early versus late responders - *On which step of the ladder would you say you personally feel you stand at this time? (Q4)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
0 – Worst possible	-	-	<10	-	-
1	-	-	<10	-	-
2	0.7%	0.6%	24	0.6%	0.2%
3	1.8%	2.8%	79	1.9%	2.6%
4	3.8%	4.9%	152	3.4%	4.9%
5	8.1%	7.5%	287	8.0%	7.9%

6	14.0%	17.0%	545	13.1%	16.0%
7	28.1%	26.1%	1000	27.7%	26.6%
8	28.7%	26.7%	1021	29.3%	24.3%
9	11.0%	9.9%	387	11.4%	12.2%
10 – Best possible	3.5%	3.9%	132	4.2%	4.5%

The chi-square test was not significant (p = 0.15) at the 5% level of significance

Table 15: Early versus late responders - *On which step do you think you will stand about five years from now? (Q5)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
0 – Worst possible	-	-	<10	-	-
1	-	-	<10	-	-
2	0.4%	0.5%	16	0.5%	0.4%
3	0.7%	0.5%	24	0.8%	0.4%
4	1.3%	1.3%	49	1.8%	1.7%
5	3.6%	3.0%	124	3.9%	2.7%
6	4.5%	4.9%	168	4.4%	6.3%
7	2.0%	12.0%	437	11.3%	10.0%
8	29.0%	27.9%	1042	28.3%	25.6%
9	34.8%	32.5%	1241	35.0%	33.0%
10 – Best possible	13.2%	16.6%	522	13.5%	18.7%

The chi-square test *was* significant (p = 0.013) at the 5% level of significance. Groups 0-6 were collapsed into one group and 7-10 into another. The means of these two groups were not significantly different (p = 0.61) at the 5% significant level.

Table 16: Early versus late responders - *In general, would you say your health is? (Q46a)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Poor	2.0%	2.9%	83	1.5%	2.7%
Fair	13.9%	15.2%	523	13.0%	15.8%
Good	37.1%	41.0%	1396	37.1%	40.3%
Very good	36.1%	31.9%	1264	36.7%	33.0%
Excellent	10.7%	8.9%	367	11.4%	8.3%

The chi-square test was significant (p = 0.01) at the 5% level of significance

Table 17: Early versus late responders - *Below is a list of the ways you might have felt or behaved. How often have you felt this way in the past 7 days? I had trouble keeping my mind on what I was doing (Q55B)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Rarely or none of the time (Less than 1 day)	40.5%	37.4 %	1148	46.2%	39.6%
Some of a little of the time (1 – 2 days)	31.1%	33.9 %	1167	31.2%	34.6%
Occasionally or a moderate amount of the time (3 – 4 days)	19.3%	20.2 %	714	16.1%	19.8%
Most or all of the time (5 – 7 days)	9.0%	8.2%	318	6.5%	5.7%

The chi-square test was significant (p = 0.014) at the 5% level of significance

Table 18: Early versus late responders - *Below is a list of the ways you might have felt or behaved. How often have you felt this way in the past 7 days? I could not get going (Q55J)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	-	-	<10	-	-
Rarely or none of the time (Less than 1 day)	47.3%	41.7 %	1656	50.5%	43.4%
Some of a little of the time (1 – 2 days)	30.7%	33.2 %	1148	29.6%	33.4%
Occasionally or a moderate amount of the time (3 – 4 days)	14.8%	16.4 %	559	13.8%	15.1%
Most or all of the time (5 – 7 days)	7.0%	8.5%	274	6.0%	8.0%

The chi-square test was significant (p = 0.017) at the 5% level of significance

Table 19: Early versus late responders - *In the past day, did you experience any of the following feelings during a lot of the day: Boredom (W3G)*

	Unweighted			Weighted	
	Early	Late	Unweighted n	Early	Late
Missing	0.4%	0.7%	17	0.4%	0.7%
Yes	33.0%	35.7%	1234	31.8%	38.4%
No	66.6%	63.7%	2391	67.7%	60.9%

Effect of Pandemic Experiences on Response Rate

One data collection concern during the COVID-19 pandemic has been whether or not the negative impacts of COVID-19 could influence participation in surveys, thereby biasing the results. Overall, during the pandemic, including the time that the NCHAT data were collected, Gallup and other organizations observed response rates to web and telephone surveys that were either similar to or higher than pre-pandemic response rates[9;10]. Beginning March 13, 2020, Gallup fielded a daily COVID tracking survey via web, using the Gallup Panel sample. The survey had an average response rate of 45 – 50% per field period and was one of the higher response rates seen on Gallup Panel surveys in recent years. Further, this survey detected significant declines in well-being and an increase in stress and worry, to levels not seen since the start of the Great Recession[11]. These findings have strong face validity and indicate that surveys were still being completed during the pandemic, including by individuals who were negatively impacted by the pandemic.

A large number of the Gallup Panel members who were invited to participate in NCHAT also completed the COVID tracking survey. The COVID tracking survey contained a large number of questions related to health, well-being, and the impact of COVID on their lives. Of the participants sampled for NCHAT, 6,608 had data from the COVID tracking survey. Of these cases, 2,792 completed the NCHAT survey. The COVID tracking survey has a response rate that is nearly twice that of the NCHAT survey. Further, the COVID survey takes less than 15 minutes to complete and covers a highly salient topic. It is important to note that more than half of individuals sampled for NCHAT do not have COVID tracking data and are not included in this analysis.

[9] Marlar, J., Jones. J. 2021. “Polling in the Time of COVID.” Presented at the 2021 Annual Conference of the American Association for Public Opinion Research. Virtual. May, 2021.
[10] https://www.nytimes.com/2020/04/17/us/politics/polling-coronavirus.html
[11] https://news.gallup.com/poll/308276/life-ratings-plummet-year-low.aspx

Similar to the demographic analysis, response rates were calculated for responses to COVID questions. We examined questions from the COVID survey that were believed to have the potential to be correlated with participation in future surveys or to be correlated with key outcomes from the NCHAT survey. The questions, response options, and response rates for each COVID response are summarized in Table 20.

Overall, there are very few differences in response rates by attitude that would indicate NCHAT respondents are systematically different from respondents in the COVID survey (which had a significantly higher response rate). For example, being laid off or permanently let go from a job during COVID could be a significant stressor on an individual, yet individuals who reported having experienced a job loss had NCHAT response rates that were nearly identical to those who did not experience a job disruption. When looking at daily experiences, the COVID survey found significant increases in negative emotions, such as stress, worry, depression, and loneliness. Yet individuals who reported these emotions had nearly identical responses to those who did not report these emotions.

Table 20: NCHAT Response rates by COVID tracking survey responses

Question	AAPOR3 Response Rate
How confident are you that you can protect yourself when out in public from being infected by the coronavirus?	
Not confident at all	67%
Not too confident	58%
Somewhat confident	68%
Very confident	71%
To what extent is your own life being affected or disrupted by the coronavirus situation?	
Not at all	74%
Not much	83%
A fair amount	79%
A great deal	81%
Have there been times in the past 30 days when you did not have enough money to pay for healthcare or prescriptions that you or your family needed?	
Yes	65%
No	68%
Have you been temporarily laid off?	
Yes	56%
No	62%
Does not apply	59%
Have you been permanently let go?	
Yes	63%
No	61%
Does not apply	59%
Did you experience the following feelings during a lot of the day yesterday? Enjoyment	
Yes	61%
No	59%
Did you experience the following feelings during a lot of the day yesterday? Worry	
Yes	59%
No	62%
Did you experience the following feelings during a lot of the day yesterday? Sadness	
Yes	58%
No	62%

Did you experience the following feelings during a lot of the day yesterday? Stress	
Yes	60%
No	61%
Did you experience the following feelings during a lot of the day yesterday? Anger	
Yes	57%
No	62%
Did you experience the following feelings during a lot of the day yesterday? Happiness	
Yes	60%
No	61%
Did you experience the following feelings during a lot of the day yesterday? Boredom	
Yes	56%
No	63%
Did you experience the following feelings during a lot of the day yesterday? Loneliness	
Yes	59%
No	61%
Did you experience the following feelings during a lot of the day yesterday? Depression	
Yes	66%
No	65%
Did you experience the following feelings during a lot of the day yesterday? Anxiety	
Yes	66%
No	64%
Did you experience the following feelings during a lot of the day yesterday? Isolation	
Yes	74%
No	75%

Analysis of Correlates of Key Outcome Variables

This analysis uses logistic regression models to model predictors of key outcome variables (relationship satisfaction and depression batteries). All models control for the demographics used in weighting - age, marital/cohabitation status, education level, sex, couple type, and orientation (same-gender or different-gender), race, and ethnicity. Additional variables were included in the model that were not included in weighting but are hypothesized to potentially be correlated with survey participation and the variable of interest. These variables include children in the household, self-reported health, self-reported stress during COVID, employment status, and income. For there to be evidence of potential non-response bias related to these factors, the following must occur: 1) the additional variables must be correlated with the dependent variable, controlling for the variables included in weighting, 2) there must be differential non-response rates related to these factors, resulting in an underrepresentation of certain groups, and 3) responders and non-responders need to give significantly different responses.

For example, one might speculate that some individuals felt stress related to COVID at the time the survey was fielded and were less likely to respond to the survey. However, if stress is not correlated with responses to the dependent variables, no non-response bias is introduced. However, if stress is correlated with the dependent variable, controlling for demographics that were included in weighting, people who experienced more stress were underrepresented in the sample, and the responses from those who are

stressed are meaningfully different from those who are not stressed, there is a *potential* for non-response bias.

Table 21 shows the outcome of the logistic regression model for questions Q19, Q20, Q21, and Q22, which all ask about marital satisfaction (Table 23 has Q55A-Q55E and Table 25 has Q55F-Q55J). Q19 through Q22 were collapsed into two categories – codes 4, 5, and 6 (representing more positive responses) were assigned to code 1 and code 1, 2, and 3 (less positive responses) were assigned to code 0. In the depression series, codes 1 and 2 (fewer reports of depressed feelings) were assigned to code 0 and codes 3 and 4 (more reports of depressed feelings) were assigned to code 1. All analyses presented in this section also controlled for demographics used in weighting - age, marital/cohabitation status, education level, sex, couple type, and orientation (same-gender or different-gender), race, and ethnicity

The presence of a child in the household, stress in the past week, stress related to COVID, self-reported health, employment status, and income were all included in the model. These variables are hypothesized to potentially be related to survey response and the outcome variables of interest, but targets are not available for weighting for the target population. The variables used in weighting were also included in the model. Table 22, Table 24, and Table 26 provide predicted probabilities of giving a certain response, by demographic group, which is valuable for interpreting the potential for non-response bias.

As an example of interpretation of these tables, consider Q19 (warm and comfortable relationship) in Table 21. Children in the household, stress in the past week, and income are all significant predictors of Q19, controlling for the demographics included in weighting.

As an example of the interpretation of predicted probabilities, consider Q19 in Table 22. The presence of a child in the household is a significant predictor of response in Q19 (as shown in Table 21). Respondents who have a child are less likely to give Q19 a 4, 5, or 6. For respondents with a child in the household, the predicted probability of answering 4, 5, or 6 in Q19 is .84 compared to .88 for those without a child. Respondents with children accounted for 50% of the total sample on an unweighted basis. We will make an assumption, strictly for the purposes of this example, that parents were less likely to respond, and are therefore underrepresented in the sample, and should comprise 60% of the weighted sample. The predicted probability of the sample as a whole answering 4, 5, or 6 in Q19 would drop from 86.4 to 85.9 because of the 10% increase in parents in the sample. If the percent of parents in the sample were to increase to 70% of the sample, the predicted probability would be 85.4. Although parental status is predictive of dependent variables, even if they were under-represented in the sample, answers of parents with a child in the household are not different enough from people without a child in the household to meaningfully change results.

Self-reported stress over the past 7 days is a significant predictor of the dependent variable in most of the analyses, after controlling for the demographics included in weighting. Additionally, the predicted probabilities of being a code 1 in the relationship satisfaction battery and depression battery differ by the level of stress experienced. People who report low levels of stress have a higher predicted probability of reporting satisfaction with their relationship and have a lower predicted probability of reporting feelings of depression.

Importantly, the analysis of response rates to the COVID survey (presented in the previous section) found very few differences in response rates to the NCHAT survey by reported emotions on the COVID survey. In other words, those who reported stress, worry, or depression on the COVID survey had similar responses to those who did not report stress, worry, or depression. Although reported stress is related to the variables on the survey, there is no evidence that individuals who experienced stress or other negative emotions were less likely to respond to the survey, thereby introducing bias.

Table 21: Results of logistic regression – relationship satisfaction questions

	Q19		Q20		Q21		Q22	
	Wald F	Sig	Wald F	Sig	Wald F	Sig	Wald F	Sig
Corrected model	3.437	0.000	3.574	0.000	3.020	0.000	3.570	0.000

Intercept	32.307	0.000	37.900	0.000	17.410	0.000	79.645	0.000
Child in HH	9.391	0.002	3.281	0.070	5.180	0.023	0.192	0.661
Stress past week	12.219	0.000	11.730	0.000	14.317	0.000	1.227	0.297
COVID stress	1.409	0.228	1.008	0.402	1.377	0.239	0.149	0.963
Self Reported Health	2.293	0.057	5.194	0.000	2.113	0.077	3.551	0.007
Employment	0.787	0.501	0.516	0.671	0.378	0.769	3.429	0.016
Income	2.718	0.028	1.731	0.140	1.994	0.093	5.239	0.000

Table 22: Predicted probabilities of relationship satisfaction by response

		Q19	Q20	Q21	Q22
In the past week, how stressed have you been? (STRESS1)	1 Not stressed at all	0.943	0.934	0.959	0.979
	2	0.920	0.894	0.913	0.981
	3	0.827	0.786	0.826	0.956
	4	0.812	0.782	0.826	0.973
	5 Very stressed	0.721	0.706	0.690	0.970
How stressed have you been about the following? Getting coronavirus (STRESS3_A)	1 Not stressed at all	0.876	0.859	0.885	0.977
	2	0.897	0.854	0.892	0.971
	3	0.852	0.831	0.854	0.972
	4	0.826	0.792	0.803	0.969
	5 Very stressed	0.822	0.798	0.839	0.949
Self-reported health (Q46A)	Poor	0.678	0.484	0.718	0.802
	Fair	0.775	0.739	0.781	0.980
	Good	0.853	0.825	0.861	0.967
	Very good	0.905	0.889	0.905	0.979
	Excellent	0.918	0.899	0.877	0.980
Child in household	No child	0.888	0.849	0.885	0.971
	Child in HH	0.841	0.826	0.844	0.972
Employment status	Full time employed	0.878	0.851	0.872	0.971
	Part time employed	0.843	0.812	0.880	0.952
	Employed but not working	0.768	0.742	0.812	0.949
	Unemployed	0.847	0.821	0.841	0.983
Household Income	0 thru 47,999	0.788	0.776	0.812	0.939
	48,000 thru 89,999	0.874	0.842	0.870	0.974
	90,000 thru 119,999	0.870	0.863	0.877	0.986
	120,000 thru 179,999	0.872	0.825	0.870	0.974
	180,000 and over	0.903	0.874	0.886	0.978

Table 23: Results of logistic regression – Part 1 of depression battery

	Q55A		Q55B		Q55C		Q55D		Q55E	
	Wald F	Sig	Wald F	Sig	Wald F	Sig	Wald F	Sig	Wald F	Sig
Corrected model	4.012	0.000	6.161	0.000	3.902	0.000	5.453	0.000	5.694	0.000
Intercept	42.536	0.000	18.230	0.000	12.153	0.000	7.108	0.008	14.966	0.000
Child in HH	0.123	0.725	1.627	0.202	3.342	0.068	3.820	0.051	0.299	0.584
Stress past week	28.740	0.000	37.440	0.000	22.266	0.000	35.114	0.000	49.524	0.000
COVID stress	0.366	0.833	1.529	0.191	0.415	0.798	1.980	0.095	3.189	0.013
Self Reported Health	3.346	0.010	3.621	0.006	5.786	0.000	7.477	0.000	7.670	0.000
Employment	1.546	0.201	0.609	0.609	3.424	0.016	2.912	0.033	3.673	0.012
Income	2.088	0.080	0.972	0.421	1.034	0.388	0.178	0.950	0.827	0.508

Table 25: Predicted probabilities of depression (Part 1 of battery) by response

		Q55A	Q55B	Q55C	Q55D	Q55E
In the past week, how stressed have you been? (STRESS1)	1 Not stressed at all	0.018	0.048	0.024	0.078	0.007
	2	0.027	0.080	0.041	0.154	0.035
	3	0.113	0.279	0.124	0.352	0.122
	4	0.216	0.444	0.230	0.441	0.309
	5 Very stressed	0.372	0.633	0.374	0.777	0.565
How stressed have you been about the following? Getting coronavirus (STRESS3_A)	1 Not stressed at all	0.082	0.141	0.083	0.195	0.072
	2	0.086	0.210	0.101	0.236	0.148
	3	0.097	0.266	0.134	0.327	0.135
	4	0.157	0.317	0.150	0.418	0.206
	5 Very stressed	0.221	0.449	0.215	0.543	0.308
Self-reported health (Q46A)	Poor	0.205	0.524	0.499	0.615	0.526
	Fair	0.212	0.353	0.213	0.513	0.294
	Good	0.107	0.260	0.115	0.310	0.141
	Very good	0.064	0.175	0.068	0.213	0.082
	Excellent	0.100	0.124	0.107	0.145	0.067
Child in household	No child	0.103	0.229	0.103	0.267	0.136
	Child in HH	0.112	0.237	0.133	0.316	0.145
Employment status	Full time employed	0.089	0.217	0.091	0.251	0.111
	Part time employed	0.133	0.229	0.135	0.318	0.144
	Employed but not working	0.182	0.275	0.238	0.449	0.244
	Unemployed	0.138	0.277	0.176	0.380	0.212
Household Income	0 thru 47,999	0.144	0.272	0.157	0.386	0.215
	48,000 thru 89,999	0.139	0.238	0.143	0.295	0.166
	90,000 thru 119,999	0.087	0.215	0.104	0.301	0.106
	120,000 thru 179,999	0.065	0.197	0.094	0.246	0.102
	180,000 and over	0.099	0.250	0.092	0.252	0.117

Table 26: Results of logistic regression – Part 2 of depression battery

	Q55F		Q55G		Q55H		Q55I		Q55J	
	Wald F	Sig	Wald F	Sig	Wald F	Sig	Wald F	Sig	Wald F	Sig

Corrected model	6.412	0.000	4.442	0.000	5.216	0.000	4.948	0.000	5.444	0.000
Intercept	5.903	0.015	0.235	0.628	22.393	0.000	7.282	0.007	25.045	0.000
Child in HH	0.384	0.536	0.491	0.483	0.026	0.871	4.376	0.037	0.001	0.980
Stress past week	32.504	0.000	30.080	0.000	40.089	0.000	36.398	0.000	21.901	0.000
COVID stress	1.632	0.163	0.661	0.619	1.683	0.151	1.564	0.181	2.781	0.025
Self Reported Health	11.264	0.000	7.273	0.000	2.169	0.070	9.033	0.000	10.693	0.000
Employment	2.861	0.036	1.980	0.115	3.799	0.010	0.989	0.397	5.951	0.000
Income	0.794	0.529	2.476	0.042	3.451	0.008	1.119	0.346	0.333	0.856

Table 27: Predicted probabilities of depression (Part 2 of battery) by response

		Q55F	**Q55G**	**Q55H**	**Q55I**	**Q55J**
In the past week, how stressed have you been? (STRESS1)	1 Not stressed at all	0.035	0.759	0.018	0.930	0.074
	2	0.076	0.769	0.028	0.891	0.099
	3	0.203	0.561	0.139	0.758	0.249
	4	0.340	0.452	0.278	0.628	0.289
	5 Very stressed	0.621	0.257	0.572	0.421	0.594
How stressed have you been about the following? Getting coronavirus (STRESS3_A)	1 Not stressed at all	0.121	0.656	0.100	0.812	0.123
	2	0.169	0.671	0.084	0.808	0.188
	3	0.193	0.621	0.137	0.783	0.224
	4	0.249	0.519	0.247	0.716	0.288
	5 Very stressed	0.443	0.436	0.309	0.664	0.415
Self-reported health (Q46A)	Poor	0.640	0.194	0.443	0.440	0.734
	Fair	0.397	0.450	0.237	0.585	0.381
	Good	0.206	0.570	0.143	0.779	0.231
	Very good	0.102	0.717	0.104	0.839	0.121
	Excellent	0.080	0.737	0.075	0.905	0.085
Child in household	No child	0.183	0.608	0.140	0.761	0.204
	Child in HH	0.197	0.625	0.142	0.800	0.209
Employment status	Full time employed	0.151	0.628	0.114	0.798	0.170
	Part time employed	0.220	0.553	0.176	0.757	0.181
	Employed but not working	0.319	0.479	0.307	0.676	0.234
	Unemployed	0.276	0.628	0.183	0.751	0.320
Household Income	0 thru 47,999	0.296	0.499	0.166	0.701	0.310
	48,000 thru 89,999	0.220	0.588	0.169	0.754	0.220
	90,000 thru 119,999	0.177	0.613	0.133	0.814	0.177
	120,000 thru 179,999	0.143	0.665	0.136	0.821	0.170
	180,000 and over	0.127	0.699	0.095	0.809	0.168

Comparison of Demographic Characteristics of Primary Respondents who Completed the Main Survey and had a Complete Time Diary

Table 28 presents the number of primary respondents who completed a time diary by demographic groups.

Table 28: Comparison of demographics of primary respondents with primary respondents who completed time diaries - Unweighted

	N Overall Sample of Primary Respondents	Percent	N Completed Time Diary Primary Respondents	Percent
	(N = 3,642)		(N = 1,975)	
*Gender**				
Man	1776	48.8	897	45.4
Woman	1739	47.8	1008	51.0
Do not identify as man or woman	127	3.5	70	3.5
Age				
20-29	359	9.9	209	10.6
30-39	975	26.8	559	28.3
40-49	923	25.3	506	25.6
50-60	1385	38.0	701	35.5
Education				
High school or less	641	17.6	317	16.1
Above high school below Bachelor's	949	26.1	475	24.1
Bachelor's or above	2051	56.3	1183	59.9
missing	<10	-	-	-
Ethnicity				
Hispanic	585	16.1	323	16.4
Non-Hispanic	3056	83.9	1651	83.6
missing	<10	-	<10	-
White vs. non-white				
non-Hispanic white	2247	61.7	1211	61.3
non-white	1389	38.1	759	38.4
missing	<10	-	<10	-
Black vs. non-Black				
non-Hispanic Black	336	9.2	164	8.3
non-Black	3300	90.6	1806	91.4
missing	<10	-	<10	-
Martial Status				
Married	2682	73.6	1476	74.7
Cohabiting	956	26.3	498	25.2
missing	<10	-	<10	-
Sexual Identity				

non-LGBT	2021	55.5	1096	55.5
LGBT	1621	44.5	879	44.5
Employment Status				
Employed	2950	81.0	1596	80.8
Not Employed	688	18.9	377	19.1
missing	<10	-	<10	-

*D2_SUPP was used for Gender. Trans man, trans woman, and do not identify as any of the above were categorized into "Do not identify as a man or woman"; LGBT = Lesbian, Gay, Bisexual, Trans (non-heterosexual)

Comparison of Demographic Characteristics of Primary Respondents who had a Partner Participate

Table 29: Comparison of demographics of primary respondents with who had a partner participate - Unweighted

	N Overall Sample of Primary Respondents	Percent	N Completed Time Diary Primary Respondents	Percent
	(N = 3,642)		(N = 1,515)	
*Gender**				
Man	1776	48.8	712	47.0
Woman	1739	47.8	736	48.6
Do not identify as man or woman	127	3.5	67	4.4
Age				
20-29	359	9.9	195	12.9
30-39	975	26.8	482	31.8
40-49	923	25.3	391	25.8
50-60	1385	38.0	447	29.5
Education				
High school or less	641	17.6	240	15.8
Above high school below Bachelor's	949	26.1	380	25.1
Bachelor's or above	2051	56.3	895	59.1
missing	<10	-	-	-
Ethnicity				
Hispanic	585	16.1	264	17.4
Non-Hispanic	3056	83.9	1250	82.5
Missing	<10	-	<10	-
White vs. non-white				
non-Hispanic white	2247	61.7	919	60.7
non-white	1389	38.1	596	39.3
missing	<10	-	-	-
Black vs. non-Black				
non-Hispanic Black	336	9.2	130	8.6

non-Black	3300	90.6	1385	91.4
missing	<10	-	-	-
Martial Status				
Married	2682	73.6	1100	72.6
Cohabiting	956	26.3	414	27.3
missing	<10	-	<10	-
Sexual Identity				
non-LGBT	2021	55.5	787	52.0
LGBT	1621	44.5	728	48.1
Employment Status				
Employed	2950	81.0	1229	81.1
Not Employed	688	18.9	285	18.8
missing	<10	-	<10	-

*D2_SUPP was used for Gender. Trans man, trans woman, and do not identify as any of the above were categorized into "Do not identify as a man or woman"; LGBT = Lesbian, Gay, Bisexual, Trans (non-heterosexual)

Demographic characteristics of partners

Table 30: Demographic characteristics of partners - Unweighted

	N Overall Sample of Partners	Percent
	(N = 1,515)	
*Gender**		
Man	745	49.2
Woman	667	44.0
Do not identify as man or woman	60	4.0
missing	43	2.8
Age		
20-29	175	11.6
30-39	470	31.0
40-49	366	24.2
50-60	350	23.1
missing	154	10.2
Education		
High school or less	199	13.1
Above high school below Bachelor's	448	29.6
Bachelor's or above	820	54.1
missing	48	3.2
Ethnicity		
Hispanic	203	13.4
Non-Hispanic	1223	80.7
Missing	89	5.9
White vs. non-white		
non-Hispanic white	958	63.2
non-white	468	30.9
missing	89	5.9
Black vs. non-Black		
non-Hispanic Black	96	6.3
non-Black	1330	87.8
missing	89	5.9
Martial Status		
Married	1074	70.9
Cohabiting	411	27.1
missing	30	2.0
Sexual Identity		
non-LGBT	896	59.1
LGBT	577	38.1
missing	42	2.8
Employment Status		

Employed	1177	77.7
Not Employed	318	21.0
missing	20	1.3

*D2_SUPP was used for Gender. Trans man, trans woman, and do not identify as any of the above were categorized into "Do not identify as a man or woman"; LGBT = Lesbian, Gay, Bisexual, Trans (non-heterosexual)
Note. These demographics are the partners own reports of their characteristics.

Comparison of Demographic Characteristics of Partners who Completed a Time Diary

Table 31: Comparison of demographics of partner respondents with partner respondents who completed time diaries - Unweighted

	N Overall Sample of Partners	Percent	N Completed Time Diary Partners	Percent
	(N = 1,515)		(N = 839)	
*Gender**				
Man	745	49.2	417	49.7
Woman	667	44.0	389	46.4
Do not identify as man or woman	60	4.0	33	3.9
missing	43	2.8	-	-
Age				
20-29	175	11.6	110	13.1
30-39	470	31.0	276	32.9
40-49	366	24.2	211	25.2
50-60	350	23.1	180	21.5
missing	154	10.2	62	7.4
Education				
High school or less	199	13.1	109	13.0
Above high school below Bachelor's	448	29.6	218	26.0
Bachelor's or above	820	54.1	509	60.7
missing	48	3.2	<10	-
Ethnicity				
Hispanic	203	13.4	105	12.5
Non-Hispanic	1223	80.7	708	84.4
Missing	89	5.9	26	3.1
White vs. non-white				
non-Hispanic white	958	63.2	569	67.8
non-white	468	30.9	244	29.1
missing	89	5.9	26	3.1
Black vs. non-Black				
non-Hispanic Black	96	6.3	50	6.0

non-Black	1330	87.8	763	90.9
missing	89	5.9	26	3.1
Martial Status				
Married	1074	70.9	618	73.7
Cohabiting	411	27.1	219	26.1
missing	30	2.0	<10	-
Sexual Identity				
non-LGBT	896	59.1	500	59.6
LGBT	577	38.1	337	40.2
missing	42	2.8	<10	-
Employment Status				
Employed	1177	77.7	663	79.0
Not Employed	318	21.0	175	20.9
missing	20	1.3	<10	-

*D2_SUPP was used for Gender. Trans man, trans woman, and do not identify as any of the above were categorized into "Do not identify as a man or woman"; LGBT = Lesbian, Gay, Bisexual, Trans (non-heterosexual)
Note. These demographics are the partners own reports of their characteristics.

Demographic Characteristics of Primary Respondents who completed a Time Diary and their Partner completed a Time Diary

Table 32: Demographics of main respondents who have a time diary and their partner also has a time diary - Unweighted

	N Overall Sample of Partners	Percent
	(N = 685)	
*Gender**		
Man	304	44.4
Woman	350	51.1
Do not identify as man or woman	31	4.5
missing	-	-
Age		
20-29	92	13.4
30-39	221	32.3
40-49	192	28.0
50-60	180	26.3
missing	-	-
Education		
High school or less	94	13.7
Above high school below Bachelor's	156	22.8
Bachelor's or above	435	63.5

missing	-	
Ethnicity		
Hispanic	120	17.5
Non-Hispanic	564	82.3
Missing	<10	-
White vs. non-white		
non-Hispanic white	424	61.9
non-white	261	38.1
missing	-	-
Black vs. non-Black		
non-Hispanic Black	45	6.6
non-Black	640	93.4
missing	-	-
Martial Status		
Married	518	75.6
Cohabiting	167	24.4
missing	-	-
Sexual Identity		
non-LGBT	340	49.6
LGBT	345	50.4
missing	-	-
Employment Status		
Employed	563	82.2
Not Employed	121	17.7
missing	<10	-

*D2_SUPP was used for Gender. Trans man, trans woman, and do not identify as any of the above were categorized into "Do not identify as a man or woman"; LGBT = Lesbian, Gay, Bisexual, Trans (non-heterosexual)
Note. These demographics are the partners own reports of their characteristics.

References

Davern, M. (2013). Non-response Rates are a Problematic Indicator of Nonresponse Bias in Survey Research. *Health Services Research, 48*(3), 905-912.

Groves, R. (2006). Non-response Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly, 70*(5), 646-675.

Groves, R., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly, 72*(2), 167-189.

Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. *Public Opinion Quarterly, 70*(5), 759-779.

MacInnis, B., Krosnick, J., Ho, A., and Cho, M. (2018). The Accuracy of Measurements with Probability and NonProbability Survey Samples. *Public Opinion Quarterly*, 82(4), 707-744.

Yeager, D., Krosnick, J., Chang, L., Javitz, H., Levendusky, M., Simpser, A., and Wang, R. (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly*, 75(4), 709-747.