

Assessing the Impact of Modifying the Introduction on the National Immunization Survey

Megha S. Ravanam¹, Benjamin Skalland¹, Christopher Scott¹, Qiao Ma¹, Vincent Welch¹, Tiffani Balok¹, Laurie D. Elam-Evans², Chalanda Smith²

¹NORC at the University of Chicago

²National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention

Abstract

The National Immunization Survey (NIS) Family of Surveys are random-digit-dial telephone surveys used to monitor vaccination coverage in the United States among children age 19-35 months (NIS-Child), adolescents 13-17 years (NIS-Teen), and influenza vaccination for children 6 months-17 years (NIS-Flu). The surveys collect household-reported demographic and access-to-care data from a parent or guardian during telephone interviews. In the NIS-Child and NIS-Teen, the parent or guardian is asked for consent to contact the child's vaccination provider(s) to obtain the child's immunization history using a mailed questionnaire. In 2018, the NIS transitioned from a dual-frame landline and cell-phone sample design to a single-frame cell phone sample design. As response rates continue to be lower for cell phones than for landlines, it is important to identify methods for increasing response rates for the NIS.

Evaluations were conducted on the NIS Surveys in Quarters 2 and 3 of 2018 to determine the impact of modified introduction scripts on survey completion rates. In Quarter 2, approximately half of the telephone interviewers were trained to use a progressive engagement methodology on the current introduction text. The remaining interviewers were trained to continue to read the text in the traditional way. Progressive engagement methodology involves interviewers pausing at appropriate times so that the respondent can participate in a conversational manner. It was hypothesized that using a progressive engagement methodology would increase the screener completion rate. In Quarter 3, interviewers were randomly selected to be trained on one of four conditions: the current version, a shortened version, a version emphasizing the purpose and importance of the study, or an informal version of the introduction text. It was hypothesized that the three new versions of the introduction would increase the screener completion rate relative to the current version. This paper presents the study design, results, conclusions, limitations, and recommendations for future research.

Key Words: National Immunization Survey, Introduction Text, Conversational, Progressive Engagement

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

1. Introduction

In telephone surveys, the introduction is the first thing a respondent hears on the phone upon answering. This introductory contact is especially important on cell phones given that an advance letter is not usually possible (National Research Council, 2013) and that not all cell phones display the caller ID name of the caller, which can arguably serve as a “business card” for the call (Callegaro et al., 2005). Therefore, it is critical to have the best introduction text scripted in a computer-aided telephone survey (CATI) to let the respondent know who the caller is, the reason for the call, and to keep them on the phone.

Previous research regarding the topics of progressive engagement methodology and introduction text content has been widely published. Progressive involvement/engagement is the idea that inserting pauses after every sentence or two allows the respondent an opportunity to react or respond to what the interviewer is saying to them, resulting in an interaction where the interviewer is engaging the respondent in conversation. This is different from traditional reading of introduction scripts where the full introduction is read to the respondent before allowing the respondent an opportunity to speak. Published research has suggested that a “progressive involvement” linguistic approach for gaining cooperation had significantly higher cooperation rates than those that employed a non-two-way conversation approach to introductions (Burks et al., 2007). Additionally, studies have shown that certain key elements such as sponsorship, importance, statement of non-solicitation, and purpose can have a significant impact on outcomes (Vaden-Kiernan et al., 1997; Cowling et al., 2003), and that scripted introductions generated lower response rates than unscripted (Morton-Williams, 1993; Houtkoop-Steenstra and van den Bergh, 2000).

Although these findings were significant in the telephone survey industry, this research had all been done with landline samples and these studies are now several years old. The National Health Interview Survey estimates that 55.3% of adults and 64.9% of children are living in households with only wireless telephone service (Blumberg and Luke, 2018). The continued increase in wireless-only households makes it imperative that methods tested on landlines are reassessed on cell phones, as results may differ. NORC at the University of Chicago conducted a series of evaluations on a large scale, single-frame random-digit-dial (RDD) telephone survey of cell phones in 2018 to assess the impact of using modified introduction scripts on respondent participation.

2. Design of the Experiment

The National Immunization Surveys (NIS), sponsored by the Centers for Disease Control and Prevention (CDC), are RDD telephone surveys that are conducted annually. The surveys provide U.S. national, state, and select local area estimates of vaccination coverage among young children age 19-35 months (NIS-Child) and adolescents age 13-17 years (NIS-Teen), and assess influenza vaccination coverage estimates among children age 6 months-17 years (NIS-Flu). Sampled telephone numbers are dialed and screened to identify households with children in the target age ranges. If the household has an eligible child, information on socio-demographic characteristics and access to care are collected; for NIS-Child and NIS-Teen, consent to contact the child’s vaccination provider(s) is also collected. The NIS-Child and NIS-Teen RDD surveys are followed by mail surveys to nominated vaccination providers from the telephone interview to obtain the vaccination histories of the selected children or adolescents. From 2011 through 2017, the NIS was a dual-frame RDD survey, sampling from both the landline and cell-phone frames. In 2018, the NIS moved to a single-frame RDD survey design, sampling only from the cell-phone frame.

Prior to the start of the experiments in Quarter 2, 2018, the general introduction text for outbound calls on cell phones read: “Hello, my name is _____. I’m calling on behalf of the Centers for Disease Control and Prevention. We’re conducting a survey with cell phone users regarding childhood immunizations. Your cell phone number has been selected at random. This call will be recorded or monitored.” In order to assess the impact of modifying the introduction text, a series of experiments were designed to answer two questions: (1) does utilizing a progressive engagement methodology have a positive impact on respondent participation rates on cell phones, as had been found previously for landlines?; (2) does varying elements of the content of the introduction script text result in higher participation rates?

The first experiment focused on testing the progressive engagement methodology. Interviewers were randomized into groups of “traditional reading” where interviewers read the current script, and “scripted progressive engagement reading” with two-second pauses written into the script to remind the interviewers to pause, allowing the respondent an opportunity to respond or interrupt. Interviewers were provided their assigned scripted introductions on-screen in CATI:

Traditional Reading (No Pausing):

Hello, my name is _____. I’m calling on behalf of the Centers for Disease Control and Prevention. We’re conducting a survey with cell phone users regarding childhood immunizations. Your cell phone number has been selected at random. This call will be recorded or monitored.

Scripted Progressive Engagement (Pausing):

Hello, my name is _____. I’m calling on behalf of the Centers for Disease Control and Prevention.

[PAUSE]

We’re conducting a survey with cell phone users regarding childhood immunizations.

[PAUSE]

Your cell phone number has been selected at random. This call will be recorded or monitored.

The second experiment looked at the content of the script. Interviewers were randomly assigned to one of four groups. Depending on the group assignment, the interviewer’s screen would display the current scripted version, a condition with a shortened script, a condition emphasizing the purpose/importance of the survey, or a scripted condition that had an informal tone to it:

Control Version, Current Introduction Script:

Hello, my name is _____. I’m calling on behalf of the Centers for Disease Control and Prevention. We’re conducting a survey with cell phone users regarding childhood immunizations. Your cell phone number has been selected at random. This call will be recorded or monitored.

Condition 1, Shortened:

Hi, my name is _____. I'm calling on behalf of the Centers for Disease Control and Prevention. The CDC is conducting a survey about the vaccinations and health of children and teens. This call will be recorded or monitored.

Condition 2, Emphasize Purpose and Importance:

Hi, my name is _____, and I'm calling on behalf of the U.S. Centers for Disease Control and Prevention. The CDC is conducting an important study about the health and vaccinations of children and teens, which will provide crucial information about the risk of diseases in our communities. This call will be recorded or monitored.

Condition 3, Informal:

Hi, my name is _____, and I'm calling on behalf of the Centers for Disease Control and Prevention. How are you today? [PAUSE FOR RESPONSE, REPLY APPROPRIATELY]

The CDC is conducting a really important study about kids' health and vaccinations to find out about the risk for certain diseases in our communities, and we're asking for your help. I'd like to ask you a few questions to see if your household is eligible for the study; and just to let you know: my call will be recorded or monitored for quality purposes.

The first experiment took place in Quarter 2, 2018 and the second experiment started with Quarter 3, 2018. Data analysis was limited to calls on which contact was first made with the respondent; that is, analysis was limited to calls on which the introduction was first read. Several key first-contact call outcomes were examined:

- Age-screener completion rate among contacts
- Age-eligibility rate among screener completes
- Interview completion rate among identified age-eligible households
- Interview yield rate among contacts ((age-screener completion rate among contacts) x (age-eligibility rate among screener completes) x (interview completion rate among identified age-eligible households))

3. Results

For the first experiment assessing progressive engagement, a total of 898 telephone interviewers were randomly assigned to either the treatment condition (reading the script utilizing the progressive engagement methodology) or the control condition (reading the traditional script). Figures 1 and 2 present the key component rates for the control and treatment groups, 95% confidence intervals for these rates, and p-values for tests of no difference between treatment and control.¹

¹ In the experiments that are the subject of this paper, the assignment to treatment group was made at the interviewer level, not at the call level. The 95% confidence intervals and chi-square tests reported in this paper account for this design by treating the first-contact calls as nested within interviewers (clusters).

Figure 1: Key Component Rates for Experiment 1, Progressive Engagement (Treatment) vs. Traditional Methodology (Control); Quarter 2, 2018

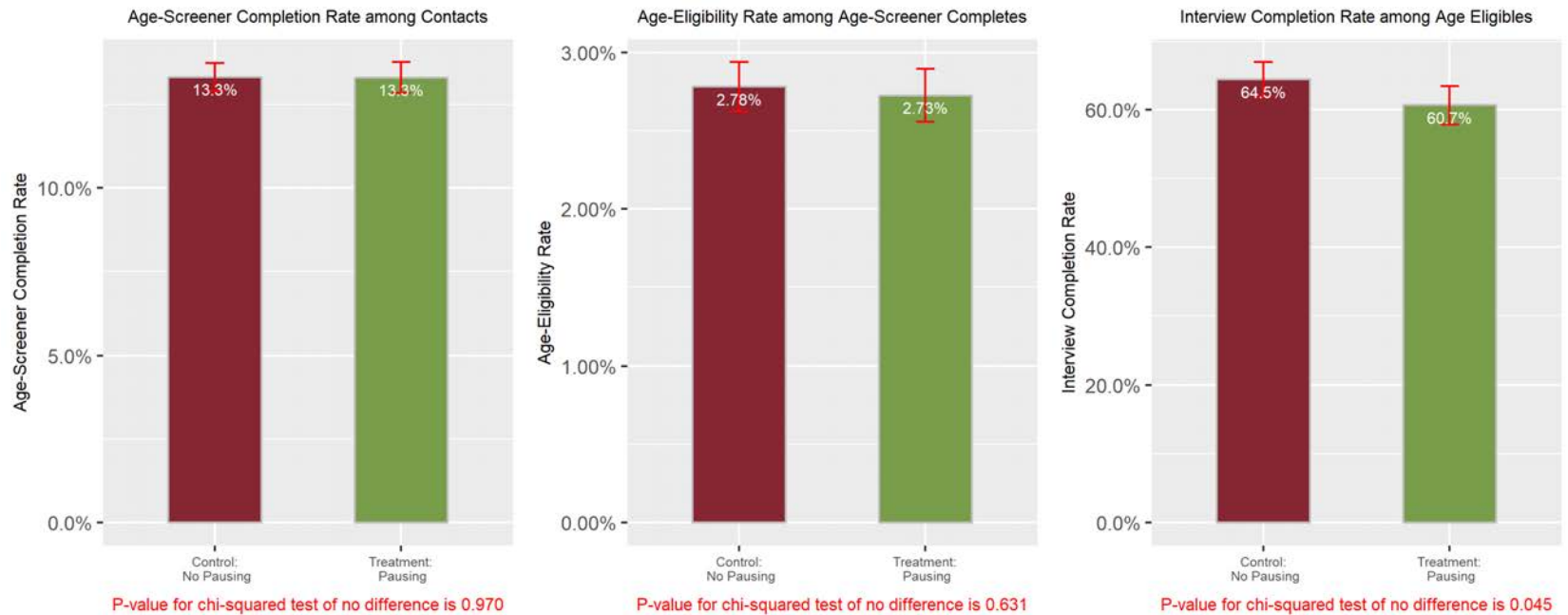
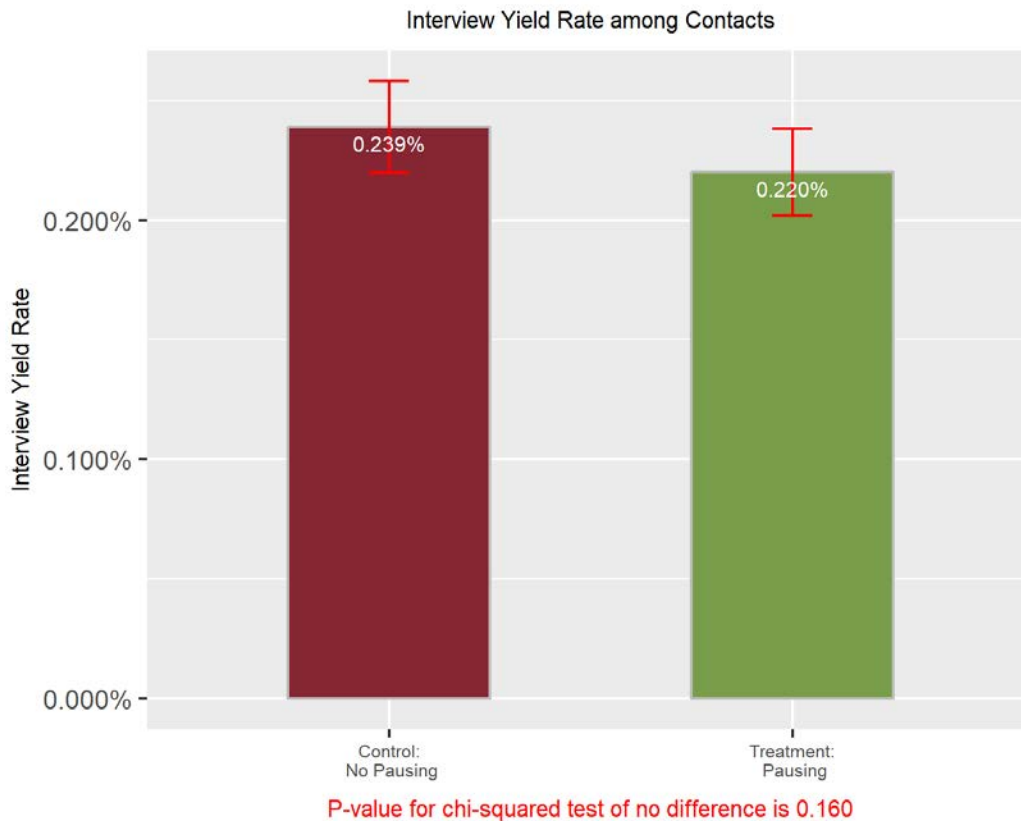


Figure 2: Interview Yield Rate among Contacts for Experiment 1, Progressive Engagement (Treatment) vs. Traditional Methodology (Control); Quarter 2, 2018



The progressive engagement pausing treatment was not found to have a positive impact on respondent participation rates. Differences were not observed between treatment and control for the age-screener completion rate among contacts or the age-eligibility rate among the age-screener completes. The interview completion rate among age-eligibles was lower for the treatment group than the control group ($p=0.045$). This led to a lower overall interview yield rate among contacts for the treatment group than for the control group (0.220% vs. 0.239%, an 8% decrease) though this was not a statistically significant difference ($p=0.160$) (Figure 2).

The second experiment evaluated the effect of modifying the content of the scripted introduction text. The 937 interviewers in Quarter 3 were randomly assigned to administer the traditional scripted text (Control), a condition utilizing a shortened version of the traditional scripted text (V1: Shortened), a condition using text that emphasized the purpose/importance (V2: Purpose/Importance), or a condition with informal sounding text that was fully scripted (V3: Informal). Figures 3 and 4 present the key component rates and overall yield rate of completed interviews for first-contact calls in Quarter 3, along with 95% confidence intervals for these rates and p-values for tests of no difference between groups.

Figure 3: Key Component Rates for Experiment 2, Modified Introduction Scripts; Quarter 3, 2018

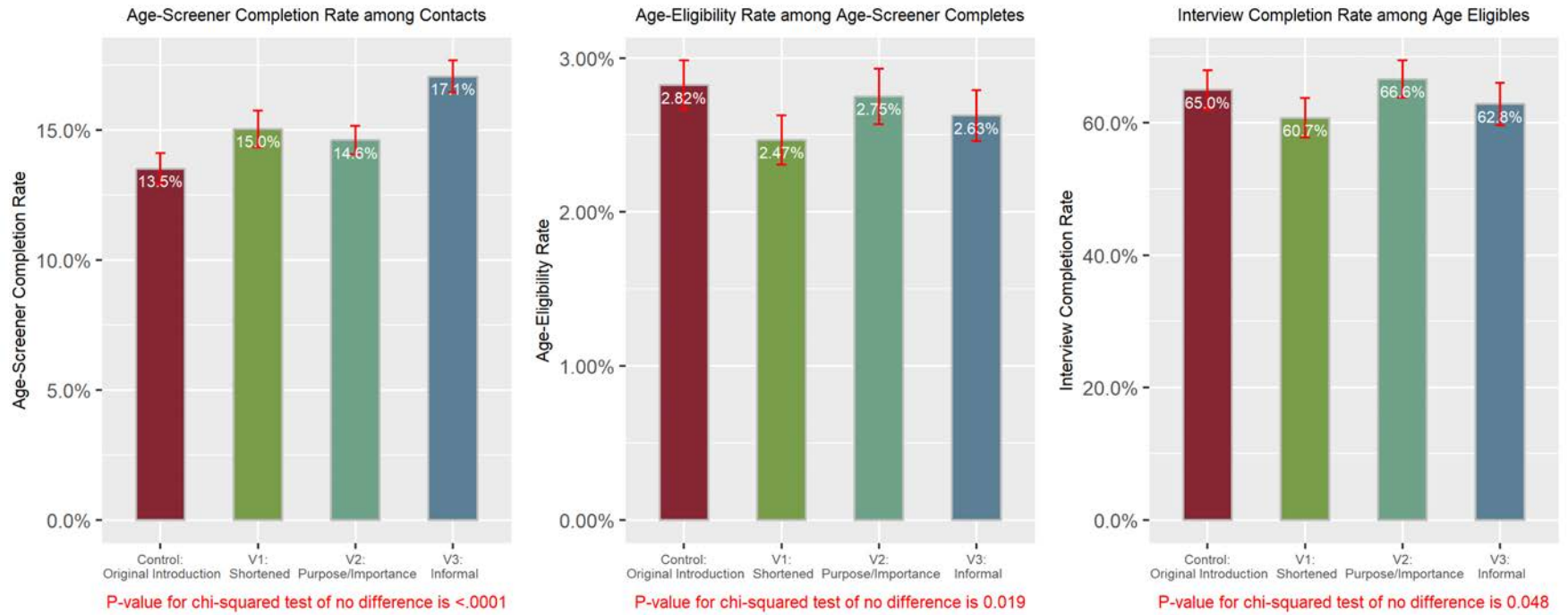
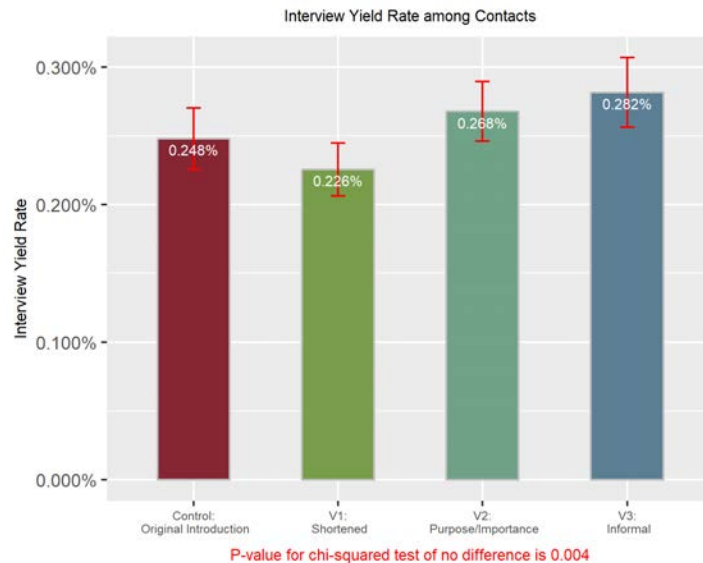


Figure 4: Interview Yield Rate among Contacts for Experiment 2, Modified Introduction Scripts; Quarter 3, 2018



The age-screener completion rate among contacts differed by treatment group ($p < 0.0001$), with the informal condition having the highest completion rate of the four and the control version having the lowest (17.1 vs. 13.5%, respectively). Differences between groups were also observed for the age-eligibility rate among age-screener completers ($p = 0.019$) and the interview completion rate among age-eligibles ($p = 0.048$), with the shortened version performing the worst. The overall yield rate of completed interviews among contacts (Figure 4) also differed by condition ($p = 0.004$), with the purpose/importance and informal versions having the highest rates of completed interviews. Based on these results, only the purpose/importance condition and the informal condition were retained in Quarter 4, and the control and shortened versions were dropped.

The first-contact call key component rates and overall yield rate of interviews among contacts for Quarter 4 are presented in Figures 5 and 6. A significantly higher age-screener completion rate among contacts was observed for the informal condition when compared to the purpose/importance condition (16.9% vs. 14.4%, $p < 0.0001$). The higher age-screener completion rate led to a higher interview yield rate among contacts for the informal condition compared to the purpose/importance condition (0.284% vs. 0.236%, $p < 0.0001$).

Figure 5: Key Component Rates for Experiment 2, Modified Introduction Scripts, Refined; Quarter 4, 2018

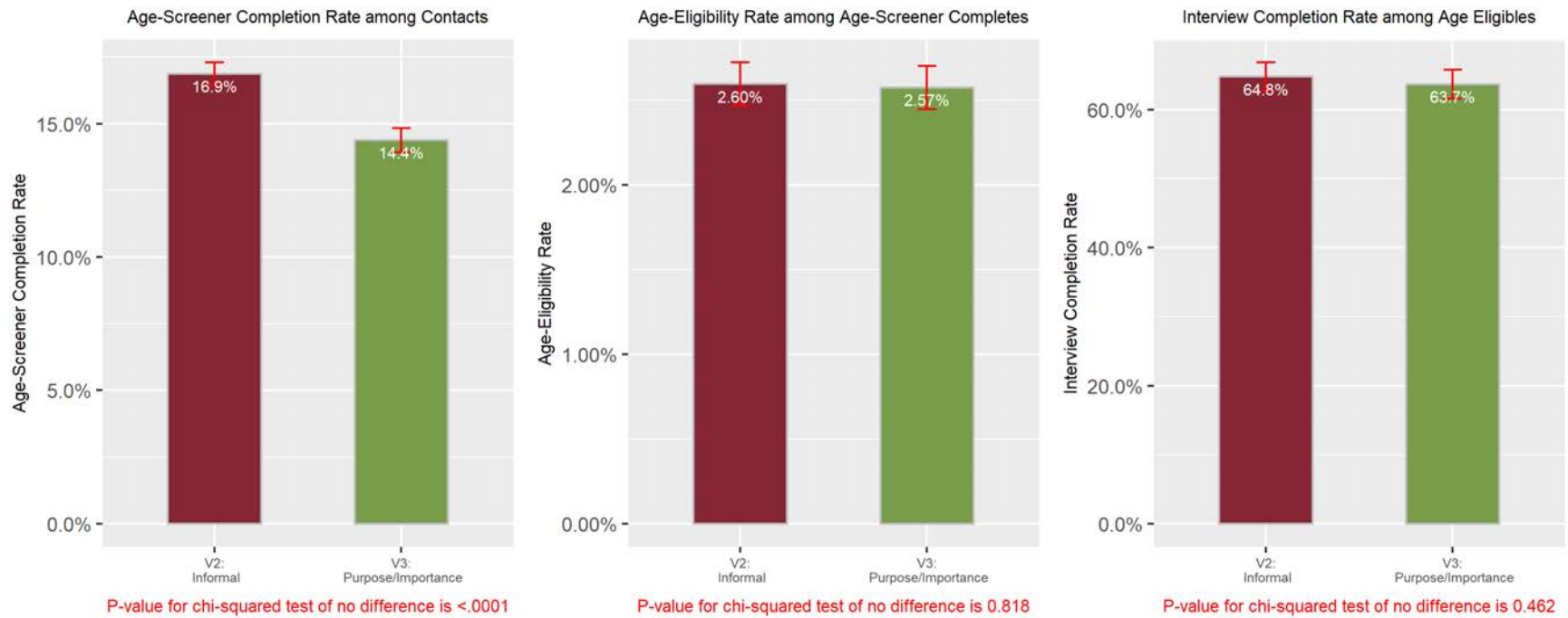
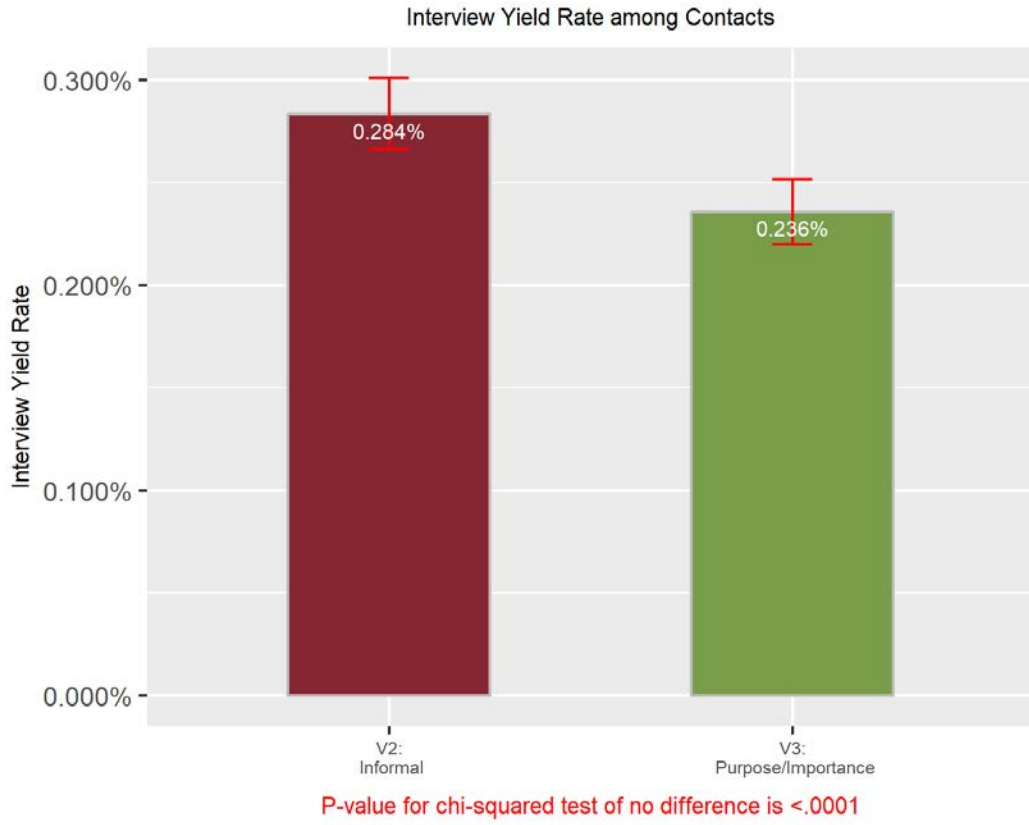


Figure 6: Interview Yield Rate among Contacts for Experiment 2, Modified Introduction Scripts, Refined; Quarter 4, 2018



4. Discussion

Contrary to results from previous studies, a positive impact on respondent participation of progressive engagement pausing methodology compared to the traditional reading of the introduction was not found in this evaluation. The results in the current evaluation may have differed from previous studies because this experiment was conducted using a cell phone sample, whereas previous studies were conducted on a landline sample. Respondent behavior may differ on cell phones than on landlines because they may view the cell phone as more of a personal than shared device compared to a landline phone.

The lack of a positive impact of progressive engagement in the current study could also be due to failure of interviewers to completely adhere to the protocol. Although interviewers were monitored to confirm that they were using their assigned methodology for reading the introduction, it is possible that some interviewers did not pause between statements as they should have, and were subsequently provided feedback and retrained if identified in monitoring. Reading statements with a full 2 second pause after every statement can be uncomfortable for interviewers; it not only might feel unnatural to speak to someone that way, but many interviewers have had years of experience in telephone interviewing during which they were trained to read through introductions quickly and get to the first question to a respondent as soon as they can.

When assessing the impact of using different styles of introduction text, it was found that using a more informal, casual-sounding scripted introduction text outperformed the scripted versions that had a more formal-sounding tone. In both Quarter 3 and Quarter 4, the informal condition had a higher age-screener completion rate among contacts, which ultimately resulted in a higher interview yield rate among contacts – particularly, in the Quarter 4 experiment where the informal condition was tested against the purpose/importance, the interview yield rate among contacts was 20% higher for the informal condition. This difference is quite large and can result in a significant cost impact on a survey as large as the NIS.

Vaden-Kiernan et al (1997) and Cowling et al (2003) emphasized the importance of the content of key elements of scripts, and our results were consistent with those findings. Seemingly innocuous stylistic changes to the introduction script, such as a scripted pause to ask the respondent “How are you today?” or mentioning that we want to ask a few questions “to see if your household is eligible for the study,” resulted in more respondents continuing the interview long enough to complete the screener. These results are suggestive of an opportunity to achieve additional rate increases (and eventual cost savings) through other changes to the introduction scripts.

5. Limitations

Although interviews were routinely monitored, it is possible interviewers were not always using their assigned introduction and instead using another one they prefer, either in the progressive engagement methodology or in scripted text. This would have resulted in the inability to measure a difference where one may have truly existed, potentially biasing the results.

Additionally, these results are specific to the use of CDC sponsorship for an immunization study targeting a population of households with young children. These results may not be

generalizable to other surveys, and these conclusions may not be applicable to other surveys with a different subject matter, target population, sponsor, or data collection contractor.

6. Future Research

As a result of the study findings, further-modified versions of the informal script were tested beginning in Quarter 1, 2019. Given the success of the scripted informal text, another condition was added to Quarter 1, 2019 to test a fully unscripted, conversational introduction. Interviewers placed in this group were trained to craft their own introductions while maintaining a few key elements of the introduction. On-screen instructions were provided, and read:

YOU MUST INCLUDE:

1. Greeting: Greet and introduce yourself to the respondent (first and last name)
2. Survey Sponsor: Calling on behalf of the Centers for Disease Control and Prevention
3. Purpose of Call: Explain why we are calling
4. The call will be recorded or monitored: Explain that the call will be recorded or monitored for quality purposes. If they decline to be recorded, proceed as usual by selecting “continue without recording.”

This experiment is still ongoing as of Quarter 3, 2019.

References

- Blumberg, S.J., and Luke, J.V. (2018). “Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, January–June 2018.” Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention. Available at: <https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201812.pdf>
- Burks, A.T., Camayd, E., Lavrakas, P.J., and Bennett, M.A. (2007). “The use of progressive involvement techniques in a telephone survey introduction.” 62nd annual conference of the American Association for Public Opinion Research, Anaheim, CA.
- Callegaro, M., McCutcheon, A. and Ludwig, J. (2005). “Who’s Calling?: The Impact of Caller ID on Telephone Survey Response.” Paper presented at the Second International Conference on Telephone Survey Methodology, Miami.
- Cowling, David, W., T.P. Johnson, B.C. Holbrook, R. B. Warnecke, and H. Tang. (2003). “Improving the Self-reporting of Tobacco Use: Results of a Factorial Experiment.” *Tobacco Control*, 12:178-183.
- Houtkoop-Steenstra, H., and van den Bergh, H. (2000). “Effects of introductions in large-scale telephone survey interviews.” *Sociological Methods and Research* 28:281–300.
- Morton-Williams, J. (1993). *Interviewer Approaches*. Aldershot, England: Dartmouth.
- National Research Council. (2013). *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/18293>.
- Vaden-Kiernan, Nancy, David Cantor, Pat Cunningham, Sarah Dipko, Karen Malloy, and Patricia Warren. (1997). “1997 NSAF Telephone Survey Methods (report no. 9).”

The Impact of Voicemail Messages on RDD Cell-Phone Response Rates in the National Immunization Surveys

Benjamin Skalland¹, Vince Welch¹, Holly A. Hill², Benjamin Fredua^{2,3},
Laurie D. Elam-Evans², Chalanda Smith²

¹NORC at the University of Chicago

²National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention

³Leidos Health, Inc.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Abstract

The National Immunization Surveys (NIS) are cell-phone random-digit-dial (RDD) telephone surveys used to assess vaccination coverage in the United States among children age 19-35 months (NIS-Child) and adolescents 13-17 years (NIS-Teen), and influenza vaccination coverage among children 6 months-17 years (NIS-Flu). Among working telephone numbers in the cell-phone RDD samples, the most common NIS call outcome is reaching a potential respondent's voicemail. Standard NIS protocol is to leave a message the fourth time a respondent's voicemail is encountered. Under current procedures, the message that is left is a pre-recorded message rather than a message left live by the interviewer making the call.

In Q3/2018, an evaluation was conducted in the NIS that varied both the timing of the message – i.e., whether the message was left on the first, second, third, or fourth voicemail event – and the gender and age (young versus mature) of the person associated with the pre-recorded voice. In Q4/2018, a similar evaluation was conducted that varied the timing, the gender of the person associated with the pre-recorded voice, and the pacing of the pre-recorded message being left.

In this paper, a comparison of respondent contact and cooperation rates is made among the various treatment groups to estimate the effect of (1) leaving vs. not leaving a voicemail message when a voicemail is encountered, (2) leaving a message on the first vs. the second voicemail event, (3) the voice used on the message, and (4) the pacing of the message. The study design, study results, conclusions, implications, and plans for future research are presented. We find that voicemail messages for a cell-phone RDD survey decreased contact rates but increased cooperation among the contacted, resulting in a much higher yield rate of completed interviews among the dialed numbers; we did not observe a difference in yield rates by whether the message was left on the first vs. second voicemail event; and we find that the voice used to record the message had an impact on survey participation.

Key Words: National Immunization Surveys, Answering Machine Message, Voicemail Message

1. Introduction

In surveys with telephone data collection, voicemail messages are a way to inform potential respondents about the survey even if they do not answer the phone (Holbrook et al. 2007). The literature on the effectiveness of such messages is mixed, with some studies finding no effect of messages on survey participation (e.g., Baumgartner, 1990; Daves 1990; Tuckel and Schulman, 2000) and others finding a positive effect (e.g., Harlow et al., 1993; Xu et al., 1993; Koepsell et al., 1996; Roth et al., 2001; Benford et al., 2010). All but one of these studies were based on surveys using landline samples; only Benford et al. (2010) utilized cell-phone samples, finding lower contact rates but higher cooperation rates when voicemail messages were left compared to when voicemail messages were not left.

In this paper, we present results of a study to gauge the impact of leaving voicemail messages in the National Immunization Surveys (NIS), which are large, national cell-phone random-digit-dial (RDD) telephone surveys. Our research questions are:

1. What is the impact of leaving vs. not leaving voicemail messages on survey participation in a cell-phone RDD survey?
 - a. What is the impact on the outcome of the next call?
 - b. What is the impact on the survey campaign as a whole?
 - c. Does the impact differ if the messages are left on the first versus the second voicemail event?
2. Do the voice characteristics of a pre-recorded voicemail message impact survey participation in a cell-phone RDD survey?
 - a. What is the impact of the voice characteristics of the person leaving the message?
 - b. What is the impact of the pacing of the message?

Further details of our evaluations are provided below. In Section 2, we describe the NIS and its voicemail message protocol; in Section 3, we present the design of our study; in Section 4, we present the methods and results of our evaluations; in Section 5, we discuss the results and their implications; in Section 6, we state the limitations of our study; and in Section 7, we discuss future research.

2. Description of the NIS and the NIS Voicemail Protocol

The NIS family of surveys are annual RDD telephone surveys to provide U.S. national, state, and selected local area estimates of vaccination coverage among children age 19-35 months (NIS-Child) and adolescents age 13-17 years (NIS-Teen), and influenza vaccination coverage estimates among children age 6 months-17 years (NIS-Flu). These surveys are sponsored by the Centers for Disease Control and Prevention (CDC), and, since 2005, have been conducted by NORC at the University of Chicago. NIS data are collected via a RDD telephone survey of parents and guardians of children in the target age ranges. From 2011 through 2017, the NIS was a dual-frame RDD survey, sampling both landline and cell-phone numbers; as of 2018, the NIS is a single-frame RDD survey, sampling only cell-phone numbers. For NIS-Child and NIS-Teen, the RDD telephone interview is followed by a mail survey sent to the vaccination providers of the children and adolescents identified in the RDD phase to obtain their vaccination histories with consent from a parent or guardian.

Because the telephone survey targets a rare population (children in specific age ranges) and has sample size requirements at the state- and local-area levels, the NIS places a very large number of dials each year – over 60 million in 2018. While a large portion of those dials are to numbers that turn out to be not in service, among working numbers the most common call outcome is reaching a potential respondent’s voicemail; over 75 percent of NIS calls to working numbers reached a voicemail in 2018. Leaving live messages on all of those calls would be quite expensive. Therefore, the long-standing NIS protocol has been to leave a message, read live by the interviewer, on the fourth voicemail event and on every third voicemail event thereafter.¹

Because leaving live messages before the fourth event has been cost-prohibitive, in recent years the NIS has been exploring alternatives to live messages. Beginning in Quarter 4 of 2017 and continuing through Quarter 2 of 2018, the NIS conducted an evaluation of the use of pre-recorded messages. Pre-recorded messages were found to be nearly as effective as live messages in inducing post-message survey participation with respondents and were found to be more cost effective than live messages because the interviewer no longer had to spend time reading the message (Skalland et al., 2018).

When leaving pre-recorded messages from Quarter 4 of 2017 to Quarter 2 of 2018, interviewers would wait for the beep before initiating the pre-recorded message. Since Quarter 3 of 2018, to further reduce data collection costs, the interviewer simply identifies a call outcome as a voicemail event, and an automated telephony system waits for silence after the beep and initiates the pre-recorded message.

As interviewers are no longer waiting for the beep and leaving live messages but are instead simply coding call outcomes as voicemails and letting an automated system wait for the beep and leave the message, the amount of interviewer time spent on a call on which a message is left is now very similar to the amount of interviewer time spent on a call on which a message is not left. The reduction in interviewer time on calls on which messages are left means that instead of waiting for the fourth voicemail event before leaving a message, as has been the long-standing NIS protocol, it is now cost effective to leave messages earlier in a potential respondent’s call history. Furthermore, the use of pre-recorded messages allows for experimentation with the voice and pacing used on the pre-recorded message. In Quarters 3 and 4 of 2018, evaluations were conducted that varied the timing of when messages were left and the voice and pacing of the pre-recorded message, as described in the next section. Throughout Quarters 3 and 4 of 2018 and across the conditions in these studies, the content of the voicemail message remained constant:

“Hello. I am calling on behalf of the Centers for Disease Control and Prevention. We are conducting a survey about childhood immunization. Would you please call us at 1-877-XXX-XXXX to let us know whether or not there are any children between 12 months and 4 years old living or staying in this household? The number again is 1-877-XXX-XXXX. Thank you.”

¹ There are a few rare exceptions to this protocol. For example, if the respondent had scheduled an appointment but a voicemail is reached when the respondent is called for that appointment, a message is always left on that call.

3. Design of the NIS Voicemail Studies

In Quarters 3 and 4 of 2018, the NIS-Child sample was randomly divided into four voicemail timing treatments, as depicted in Figure 1. For those in the Treatment 1 group, voicemail messages were left on the first voicemail event, i.e., the first time a call reached a respondent's voicemail; in Treatment 2, voicemail messages were left on the second voicemail event; in Treatment 3, voicemail messages were left on the third voicemail event; and in Treatment 4, voicemail messages were left on the fourth voicemail event.

Figure 1: Voicemail Timing Treatments, Quarters 3 and 4, 2018

Voicemail Event Number	Treatment 1	Treatment 2	Treatment 3	Treatment 4
1	Leave Message	X	X	X
2	X	Leave Message	X	X
3	X	X	Leave Message	X
4	X	X	X	Leave Message

Independent of the assignment of cases to a voicemail timing treatment group, in Quarter 3 of 2018 cases were also randomly assigned to one of four pre-recorded message groups that differed in the age and gender of the person recording the voicemail message, as shown in Figure 2. One recording used a mature (i.e., older) female voice, one used a young female voice, one used a mature male voice, and one used a young male voice. As noted earlier, the content of the message was the same across these treatment groups; only the voice on the recorded message differed.

Figure 2: Voicemail Pre-Recorded Voice Treatments, Quarter 3, 2018

Treatment 1	Treatment 2	Treatment 3	Treatment 4
Mature Female Voice	Young Female Voice	Mature Male Voice	Young Male Voice

In Quarter 4 of 2018, independent of the assignment of cases to a voicemail timing treatment group, cases were also randomly assigned to one of four pre-recorded message groups that differed in the pacing of the pre-recorded message and gender of the person recording the voicemail message, as shown in Figure 3. One recording used a female voice with a slower pacing (the duration of the message was ~30 seconds), one recording used the same female voice but with a faster pacing (~24 seconds), one recording used a male voice with a slower pacing (~30 seconds), and one recording used the same male voice but with a faster pacing (~24 seconds).

Figure 3: Voicemail Pre-Recorded Pacing and Voice Treatments, Quarter 4, 2018

Treatment 1	Treatment 2	Treatment 3	Treatment 4
Slow Pacing, Female Voice	Fast Pacing, Female Voice	Slow Pacing, Male Voice	Fast Pacing, Male Voice

4. Methods and Results

Data from the evaluations in Quarters 3 and 4 were used to estimate the impact of leaving versus not leaving a voicemail message on the outcome of the next call, the impact of leaving vs. not leaving voicemail messages on the survey campaign as a whole, the impact of leaving messages on the first versus the second voicemail event, the impact of the age and gender of the person that recorded the voicemail message, and the impact of the pacing of the message and the gender of the person that recorded the message.

4.1 Impact on the Outcome of the Next Call

To estimate the impact of leaving a message on the outcome of the next call, the analysis was limited to cases that had a voicemail call outcome on their first dial, and comparisons of key outcome rates on the next call were made between cases that had a voicemail message left versus not left on the first dial. By limiting to first-dial voicemail cases and examining the outcome of the next dial, the impact of leaving versus not leaving a message on the first-dial voicemail event can be directly estimated by comparing cases in Treatment 1 to cases in Treatments 2-4, as shown in Figure 4.

Figure 4: Estimating the Impact of Leaving a Message versus Not Leaving Message on the Outcome of the Next Dial

Voicemail Event Number	Treatment 1	Treatment 2	Treatment 3	Treatment 4
1	Leave Message	X	X	X
2	X	Leave Message	X	X
3	X	X	Leave Message	X
4	X	X	X	Leave Message

Note: By limiting to first-dial voicemail event outcome cases and comparing the outcome of the second dial between cases in Treatment 1 and cases in Treatment 2-4, the impact of leaving versus not leaving the message on the first-dial voicemail call can be directly estimated.

When looking at the next call after the first-dial voicemail event, the key outcome rates examined were:

- **Contact rate among working numbers:** Of calls to working numbers (i.e., numbers that are in service), the proportion for which a human being answered the phone.
- **Age-screener completion rate among contacts:** Of calls on which a human being answered the phone, the proportion that completed the screener to determine age-eligibility for the NIS-Child.
- **Age-eligibility rate among age-screener completes:** Of calls on which a human being answered the phone and completed the age-screener, the proportion that were age-eligible for the NIS-Child.
- **Interview completion rate among age-eligibles:** Of calls on which a human being answered the phone and completed the age-screener as age-eligible for NIS-Child, the proportion that completed NIS-Child interview.
- **Consent-to-contact-providers rate among interview completes:** Of calls on which a human being answered the phone, completed the age-screener as age-eligible for NIS-Child, and completed the NIS-Child interview, the proportion that gave consent to contact the child's vaccination providers to obtain vaccination records.

We refer to the outcome rates above as *component rates*, because each measures completion of a key component of the NIS data collection process: gaining human contact, completing the age-screener, identifying eligible households, completing the interview, and obtaining consent to contact vaccination providers. We also examined a key *summary rate*:

- **Yield rate of completed interviews with consent to contact providers among working numbers:** Of calls to working numbers (i.e., numbers that are in service), the proportion on which a human being answered the phone, completed the age-screener as age-eligible for NIS-Child, completed the NIS-Child interview, and gave consent to contact the child's vaccination providers to obtain vaccination records.

This summary rate is an overall productivity rate and is equal to the product of the component rates listed above.

Figures 5-9 present the component rates for the call after the first-dial voicemail event, first for cases that did not have a voicemail message left and then for cases that had a voicemail message left. Also shown are 95 percent confidence intervals for the estimated rates and the p-value of a chi-square test of no difference in outcome rate between those with and without a voicemail message left.

Contrary to expectation, leaving a message on the first-dial voicemail event actually *decreased* the contact rate among working numbers on the next dial, as can be seen in Figure 5. Figures 5a and 5b offer an explanation. Figure 5a presents the rate of inbound calls (i.e., the respondent calling us) immediately following a first-dial voicemail event, and Figure 5b presents the next-dial contact rate following a first-dial voicemail event, excluding cases that called inbound from both the numerator and denominator. The rate of inbound dials was significantly lower when a voicemail message was left than when a

message was not left (Figure 5a); after excluding those inbound dials, we do not observe a difference in the contact rate between cases where a voicemail message was left versus not left. This suggests that the lower overall contact rate when a message was left was entirely driven by a lower inbound call rate when a message was left.

Among cases where contact was made, cooperation rates were higher when a message was left on the first-dial voicemail event than when a message was not left (Figures 6 through 9). The age-screener completion rate ($p=0.000$), the age-eligibility rate ($p=0.011$), the interview completion rate ($p=0.010$), and the consent rate to contact providers ($p=0.143$) were all higher when a message was left.

Figure 5: Second Dial Outcome among First-Dial Voicemail Cases: Contact Rate among Working Numbers

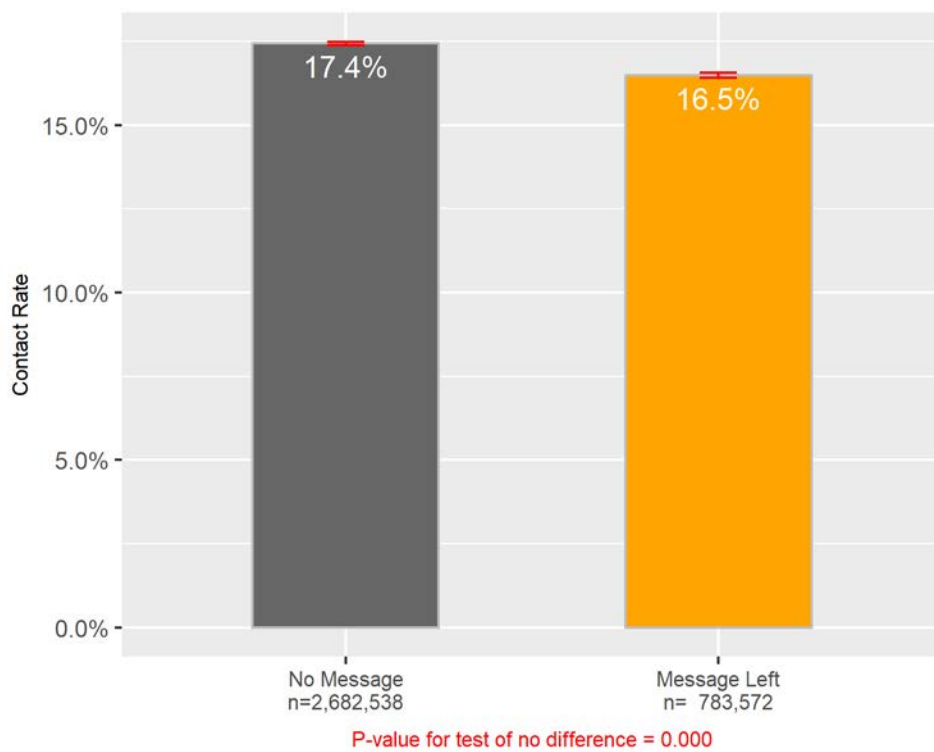


Figure 5a: Second Dial Outcome among First-Dial Voicemail Cases: Inbound Call Rate

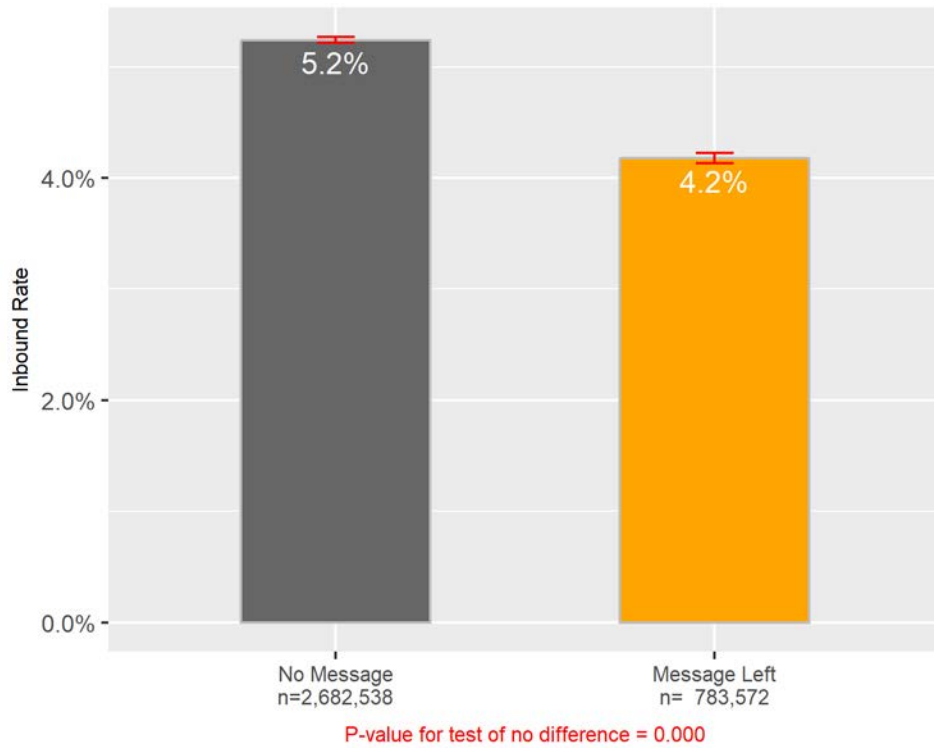


Figure 5b: Second Dial Outcome among First-Dial Voicemail Cases: Contact Rate among Non-Inbound Call Working Numbers

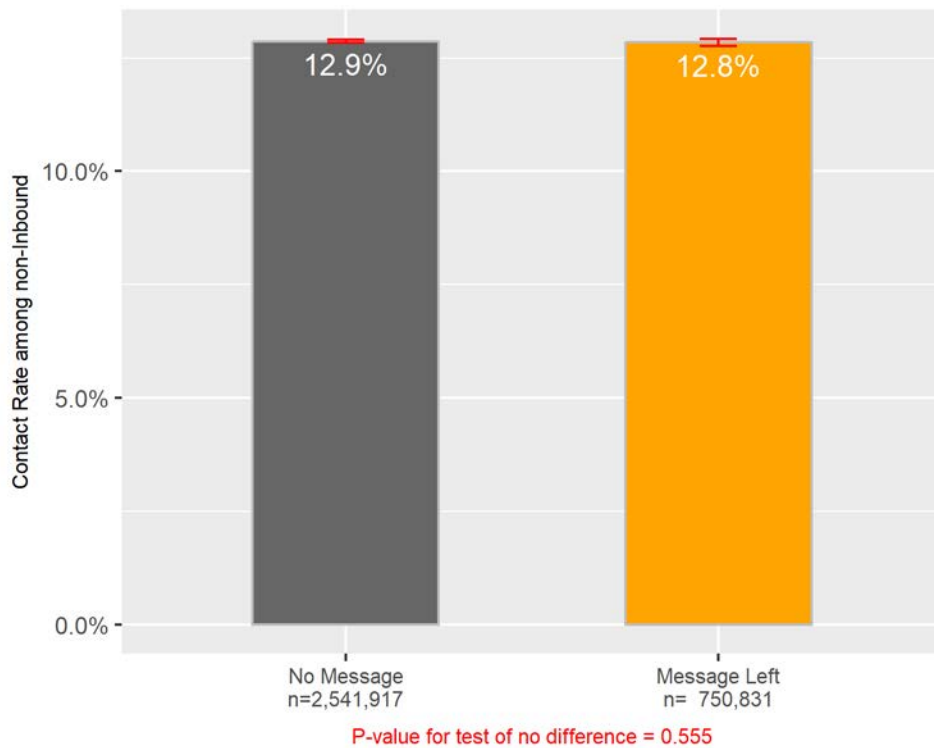


Figure 6: Second Dial Outcome among First-Dial Voicemail Cases: Age-Screener Completion Rate among Contacts

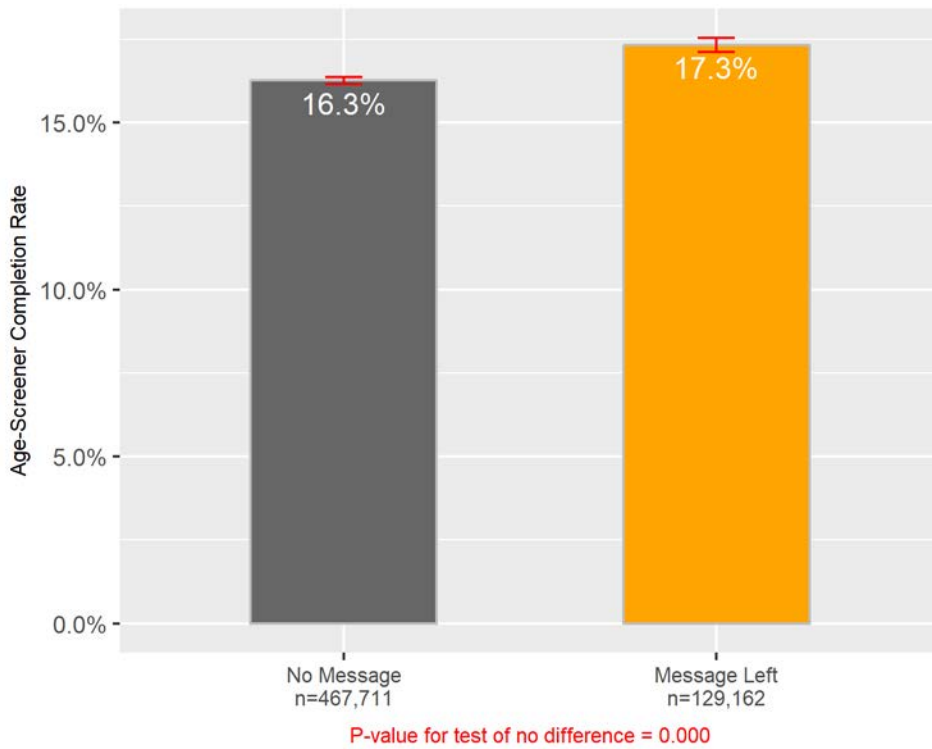


Figure 7: Second Dial Outcome among First-Dial Voicemail Cases: Age-Eligibility Rate among Age-Screener Completes

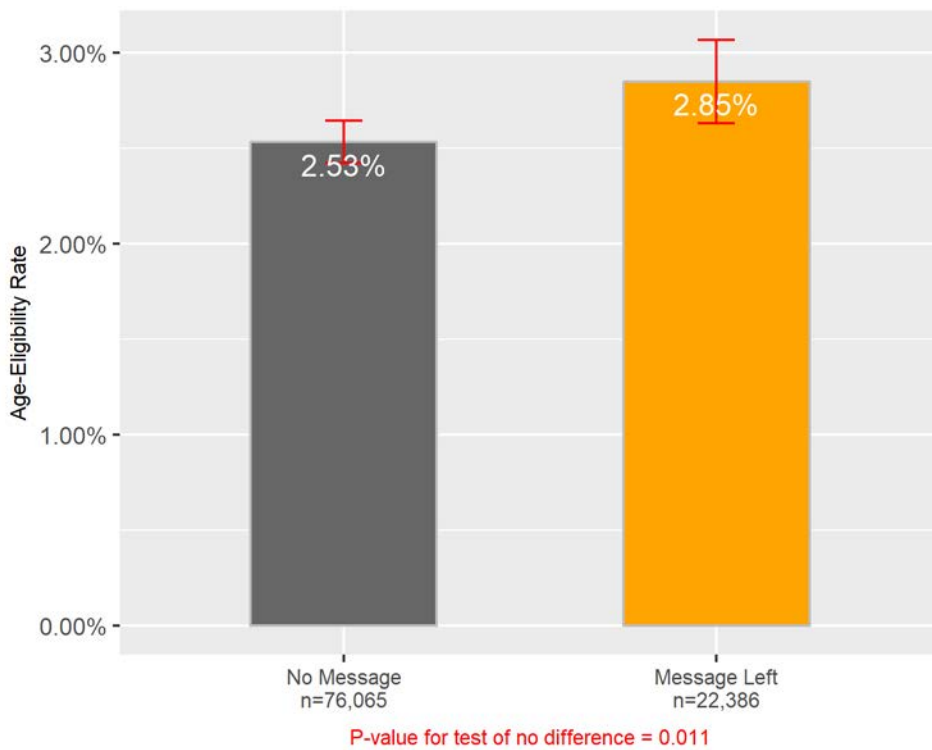


Figure 8: Second Dial Outcome among First-Dial Voicemail Cases: Interview Completion Rate among Age-Eligibles

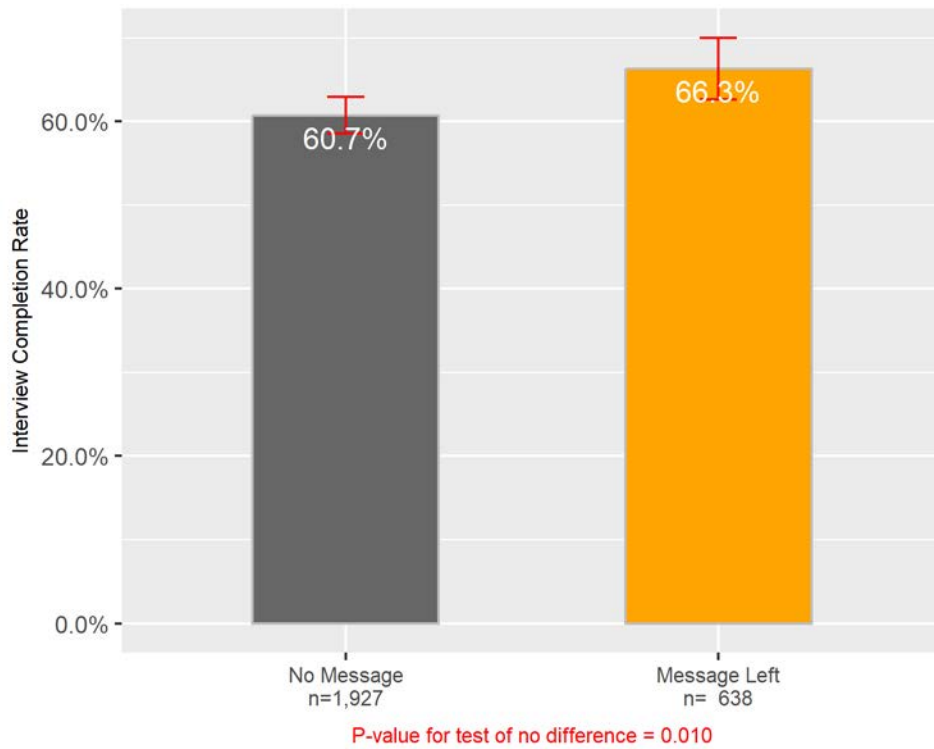


Figure 9: Second Dial Outcome among First-Dial Voicemail Cases: Consent Rate to Contact Providers among Interview Completes

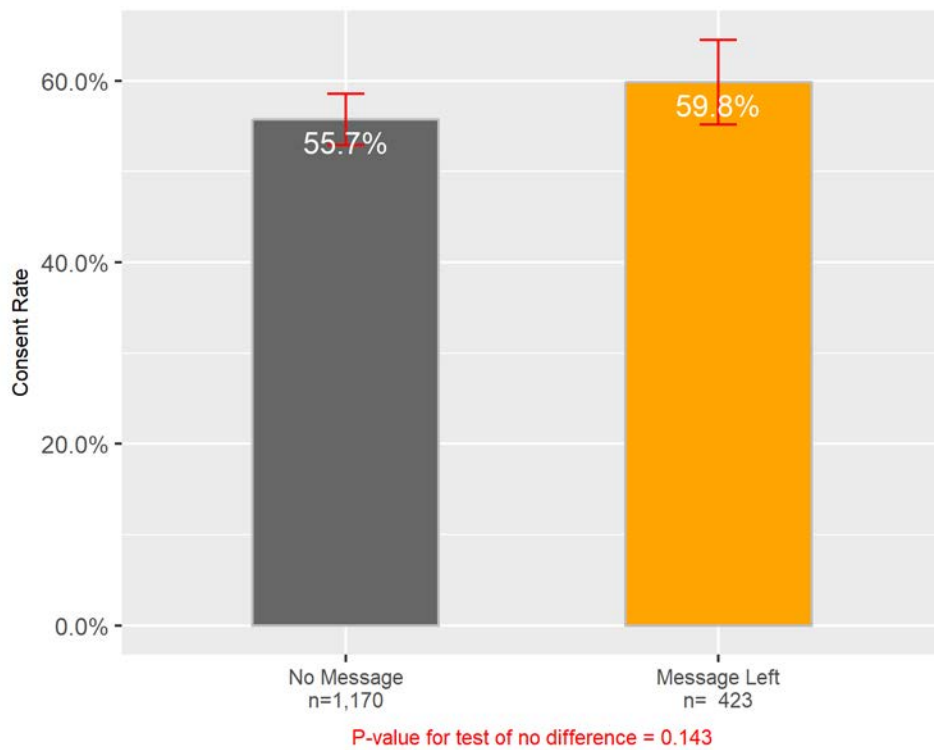
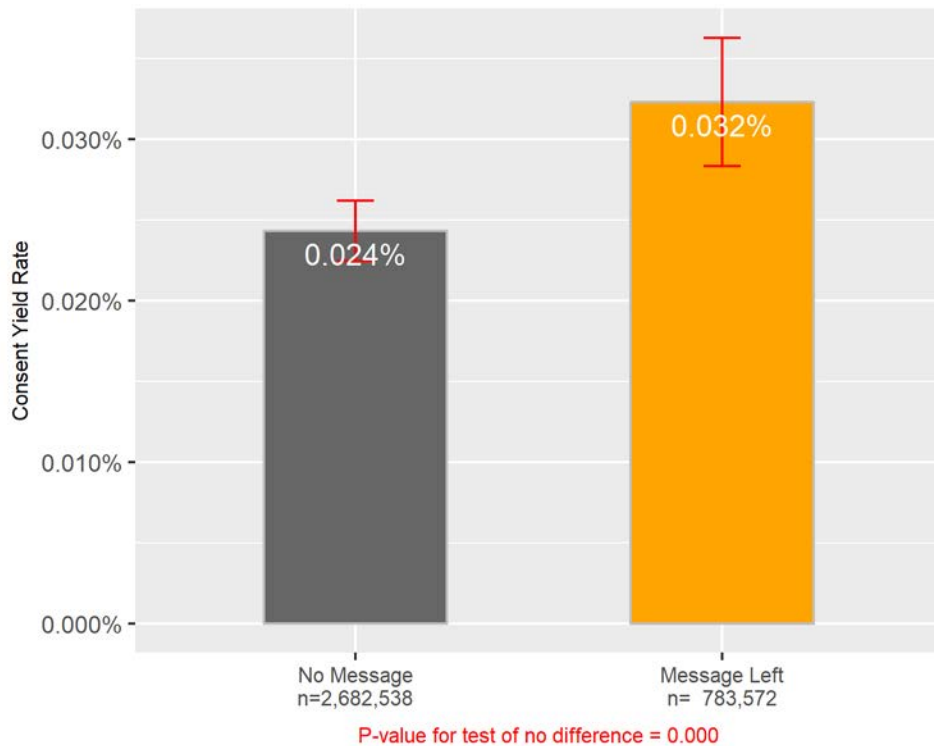


Figure 10: Second Dial Outcome among First-Dial Voicemail Cases: Yield Rate of Completed Interviews with Consent to Contact Providers among Working Numbers



The overall yield rate of completed interviews with consent to contact providers among the working numbers was significantly higher ($p < 0.001$) when a message was left (Figure 10). That is, despite the lower contact rate when a message was left, the other component rates were so much higher when a message was left that the resulting overall yield rate was higher when a message was left (0.032 percent) than when a message was not left (0.024 percent), for a very large effect size of 33 percent ($0.032/0.024 - 1$).

4.2 Impact on the Survey Campaign as a Whole

The previous section examined the impact on the outcome of the second dial of leaving versus not leaving a voicemail message on the first dial, among cases with a voicemail event on the first dial. In this section, we examine the impact of leaving versus not leaving a voicemail message on outcomes for a survey campaign as a whole. To estimate the impact on a survey campaign as a whole, all cases were included in the analysis, not just those with a voicemail event, and cumulative case outcomes as of the end of the third dial were examined. By examining the outcomes as of the end of the third dial, we can directly estimate the impact of leaving versus not leaving a message by comparing cases in Treatments 1 and 2 to cases in Treatments 3 and 4, as shown in Figure 11. By examining outcomes as of the end of the third dial, this analysis simulates the impact for a survey with a three-dial protocol.

Figure 11: Estimating the Impact of Leaving a Message vs. Not Leaving Message on a Survey with a Three-Dial Protocol

Voicemail Event Number	Treatment 1	Treatment 2	Treatment 3	Treatment 4
1	Leave Message	X	X	X
2	X	Leave Message	X	X
3	X	X	Leave Message	X
4	X	X	X	Leave Message

Note: By comparing the cumulative outcomes as of the end of the third dial between cases in Treatments 1-2 and cases in Treatment 3-4, the impact of leaving versus not leaving a voicemail message on the first or second dial can be directly estimated for a survey with a three-dial protocol.

Figure 12: Survey Campaign as a Whole for a Three-Dial Protocol: Yield Rate of Completed Interviews with Consent to Contact Providers among Working Numbers

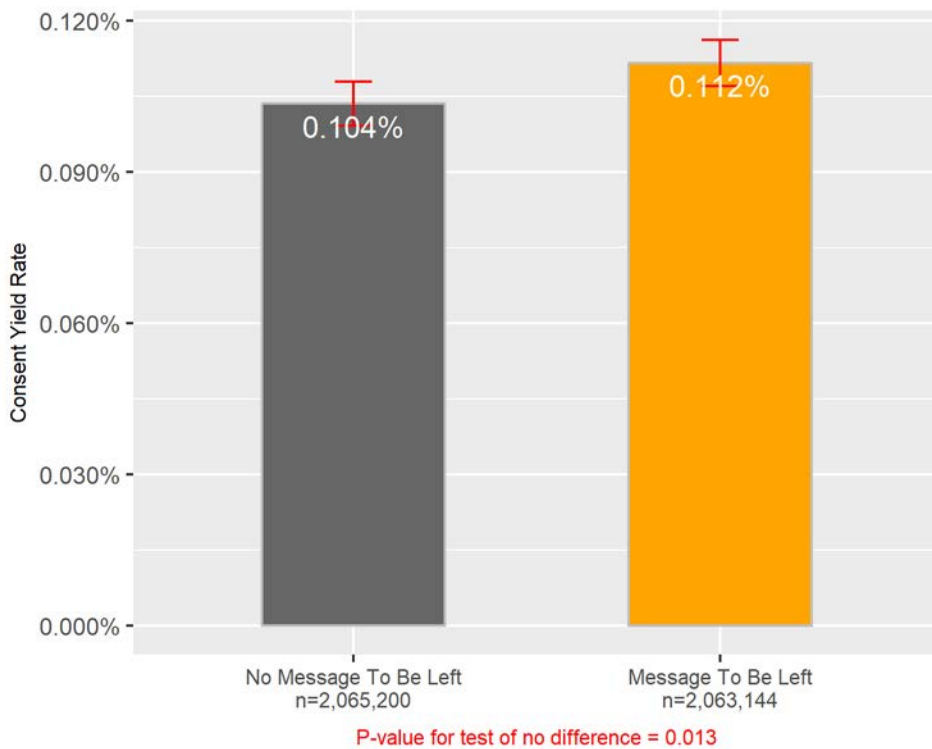


Figure 12 presents the overall yield rate of completed interviews with consent to contact providers among the working numbers, as of the end of the third dial, for cases with no voicemail message to be left if a respondent's voicemail is encountered before the third dial (Treatments 3 and 4) versus cases with a message to be left if a respondent's voicemail is encountered before the third dial (Treatments 1 and 2). The overall yield rate of completed interviews with consent to contact providers among the working numbers was significantly higher ($p=0.013$) when messages were to be left (0.112 percent) than when messages were not to be left (0.104 percent), for an effect size of 8 percent (0.112/0.104 – 1).

4.3 Impact of First- versus Second-Event Messages on the Survey Campaign as a Whole

The study design can also be used to examine the impact on the survey campaign as a whole of leaving a voicemail on the first voicemail event versus the second voicemail event. By examining case outcomes as of the end of the third dial separately for cases in Treatment 1 and cases in Treatment 2 (see Figure 13), we can simulate and compare case outcomes for a three-dial protocol for a survey in which voicemail messages are left on the first versus second voicemail event.

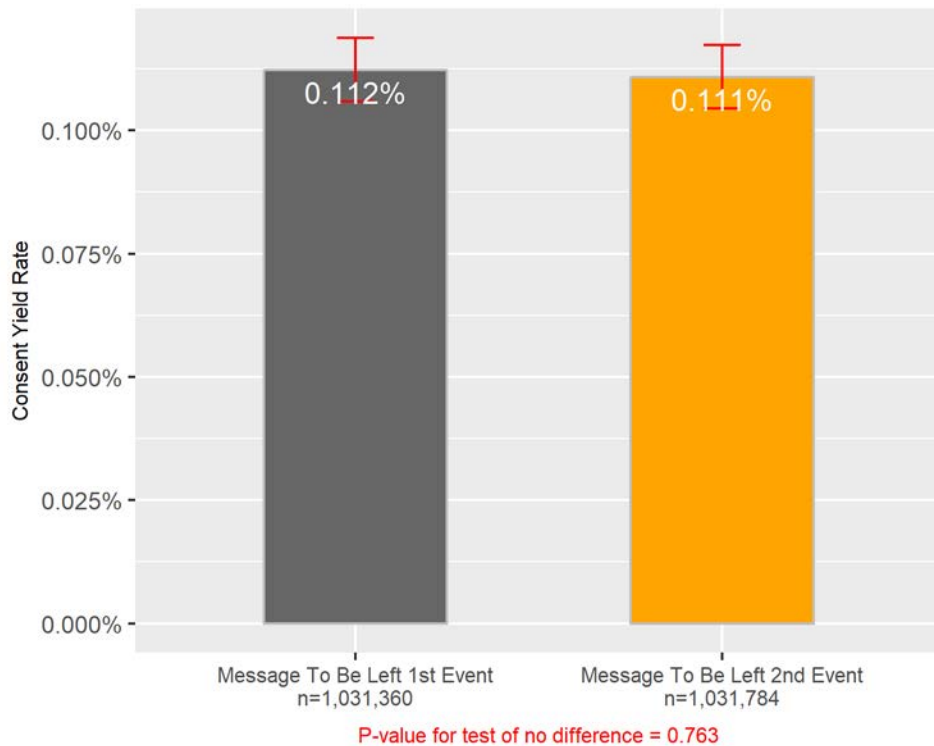
Figure 14 presents the overall yield rate of completed interviews with consent to contact providers among the working numbers, as of the end of the third dial, for cases with a voicemail message to be left on the first versus second voicemail event if a respondent's voicemail is encountered. The overall yield rate of completed interviews with consent to contact providers among the working numbers was not different ($p=0.763$) when messages were to be left on the first voicemail event (0.112 percent) than when messages were to be left on the second voicemail event (0.111 percent).

Figure 13: Estimating the Impact of First-Event vs. Second-Event Voicemail Messages on a Survey with a Three-Dial Protocol

Voicemail Event Number	Treatment 1	Treatment 2	Treatment 3	Treatment 4
1	Leave Message	X	X	X
2	X	Leave Message	X	X
3	X	X	Leave Message	X
4	X	X	X	Leave Message

Note: By comparing the cumulative outcomes as of the end of the third dial between cases in Treatment 1 and cases in Treatment 2, the impact of leaving a message on the first voicemail event versus the second voicemail even can be directly estimated for a survey with a three-dial protocol.

Figure 14: Survey Campaign as a Whole for a Three-Dial Protocol: Yield Rate of Completed Interviews with Consent to Contact Providers among Working Numbers



4.4 Impact of the Age and Gender of the Person Recording the Voicemail Message

In Quarter 3 of 2018, cases were randomly assigned to one of four treatment groups that differed in the age and gender of the person who recorded the voicemail message that the case would receive if a voicemail was encountered: a mature female voice, a mature male voice, a young female voice, or a young male voice. Figure 15 presents the overall yield rate of completed interviews with consent to contact providers among the working numbers for the simulated three-dial protocol by treatment group. The consent yield rate differed significantly by treatment group ($p=0.020$), with cases assigned to receive the mature male voice recorded message performing significantly worse than cases assigned to receive the mature female or young male voice.

Figure 15: Impact of Pre-Recorded Message Voice on Survey Campaign as a Whole for a Three-Dial Protocol: Yield Rate of Completed Interviews with Consent to Contact Providers among Working Numbers

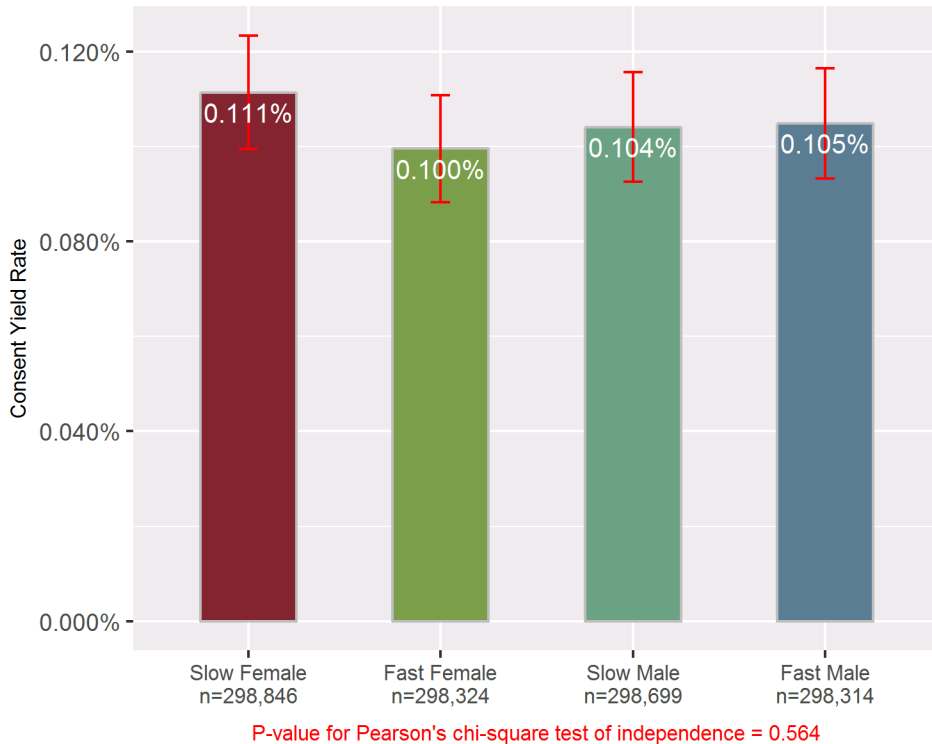


4.5 Impact of the Pacing of the Pre-Recorded the Voicemail Message

In Quarter 4 of 2018, cases were randomly assigned to one of four treatment groups that differed in the pacing of the pre-recorded message that the case would receive if a voicemail was encountered and the gender of the person who recorded the message: a slower-paced female voice, a faster-paced female voice, a slower-paced male voice, or a faster-paced male voice.² Figure 16 presents the overall yield rate of completed interviews with consent to contact providers among the working numbers for the simulated three-dial protocol by treatment group. The consent yield rate did not differ significantly across the treatment groups ($p=0.564$).

² The female voice used in the slower- and faster-paced female voice treatments was the same as the mature female voice used in the Quarter 3 experiment; the male voice used in the slower- and faster-paced male voice treatments was a new, middle-aged male voice that had not been used in Quarter 3.

Figure 16: Impact of Pre-Recorded Message Pacing and Voice on Survey Campaign as a Whole for a Three-Dial Protocol: Yield Rate of Completed Interviews with Consent to Contact Providers among Working Numbers



5. Discussion

Reaching a respondent's voicemail is the most common call outcome among working numbers in the NIS, and leaving voicemail messages can inform potential respondents about the survey. Consistent with Benford et al. (2010), we found that voicemail messages for a cell-phone sample decreased contact rates but increased cooperation among the contacted, and we uncovered a possible explanation for the effect: leaving messages decreased the rate of respondents calling back following the voicemail event. Perhaps some cell phone users redial following a missed call from an unrecognized number out of curiosity as to the identity of the caller and the purpose of the call; if a message is left indicating the caller's identity and purpose, the curiosity factor is removed and these potential respondents may therefore be less likely to redial following the missed call, lowering the inbound call rate and ultimately lowering the contact rate.

Despite the lower contact rate, the increased cooperation rate resulted in a much higher overall yield rate of completed interviews with consent to contact vaccination providers when a message was left. When looking just at the outcome of the second dial for cases with a first-dial voicemail event, the overall yield rate was 33 percent higher when a message was left; when looking at the outcomes for the survey campaign as a whole under a simulated three-dial protocol, the effect remained fairly large, with the overall yield rate being 8 percent higher when a message was left. These results suggest that voicemail messages are an effective way to increase survey participation in the NIS.

We did not observe a difference in the overall yield rate of completed interviews with consent to contact vaccination providers by whether the message was left at the first voicemail event or the second voicemail event, suggesting that it is the act of leaving an early message rather than the timing of the message (first event or second event) that is important.

The NIS utilizes pre-recorded voicemail messages, and we found that the voice used to record the message can have an impact on survey participation. Our study used four particular voices, so we cannot draw general conclusions about how male versus female voices or young versus mature voices perform, but our results demonstrate that the voice can make a difference. No effect was observed for the pacing of the pre-recorded message.

6. Limitations

Our evaluations and conclusions are subject to several limitations. In this study, pre-recorded messages initiated by an automated telephony system were utilized, as opposed to live voicemail messages left by an interviewer; our results may not apply to live messages. The NIS is a survey about the vaccination of children and is sponsored by the CDC; the effect of leaving messages may differ for surveys with a different target population, subject matter, or sponsorship. A particular voicemail message script was used; our conclusions may not be applicable to messages with different content. We used particular male and female, mature and young voices on the pre-recorded messages in this study; our conclusions may not hold for other voices or vocal characteristics (e.g., accent, pitch, etc.). All of the messages in this study were in English; our conclusions may not apply to other languages.

7. Future Research

The use of pre-recorded messages facilitates not only the testing of different voices on the message, but also the testing of different message content. Beginning in Quarter 1 of 2019, the NIS is testing four different versions of the content of the voicemail message:

- The current NIS message;
- A more informal message;
- An informal message, without including the eligibility criteria of the survey;
- A message emphasizing the purpose and importance of the survey, without including the eligibility criteria.

References

- Baumgartner, Robert M. (1990). Telephone answering machine messages and completion rates for telephone surveys. Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.
- Benford R, Lavrakas P, Tompson T, and Fleury C (2010). An experiment testing the impact of leaving voice messages in cell phone surveying. Paper presented to the American Association for Public Opinion Research, Chicago, Illinois.
- Daves, RP (1990). You know what to do at the beep, but do survey researchers? Paper presented to the Midwest Association for Public Opinion Research, Chicago, Illinois.
- Harlow BL, Crea EC, East MA, Oleson B, Fraer CJ, and Cramer DW (1993). Telephone answering machines: the influence of leaving messages on telephone interviewing response rates. *Epidemiology*. 4(4):380-383, July 1993.
- Holbrook A, Krosnick J, and Pfent A (2007). The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In *Advances in telephone survey methodology*, ed. James M. Lepkowski, N. Clyde Tucker, J. Michael Brick, Edith D. De Leeuw, Lilli Japiec, Paul J. Lavrakas, Michael W. Link, and Roberta L. Sangster. New York: Wiley.
- Koepsell TD, McGuire V, Longstreth WT, Nelson LM, and van Belle G (1996). Randomized trial of leaving messages on telephone answering machines for control recruitment in an epidemiologic study. *Am J Epidemiol* 1996; 144:704–6.
- Roth SB, Montaquila J, and Brick JM (2001). Effects of telephone technologies and call screening devices on sampling, weighting and cooperation in a random digit dialing (RDD) survey. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Montreal, Quebec, Canada.
- Skalland B, Reimer B, Ma Q, Welch V, Kornyllo S, Hobson K, Hill HA, and Fredua B. (2018). Evaluating the impact of using pre-recorded voicemail messages in the National Immunization Survey." *JSM Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association.
- Tuckel P and Schulman M (2000). The impact of leaving different answering machine messages on nonresponse rates in a nationwide RDD survey. *Proceeding of the Section on Survey Research Methods, American Statistical Association*, 901–906.
- Xu M, Bates BJ, and Schweitzer JC. The impact of messages on survey participation in answering machine households (1993). *Public Opinion Quarterly* 57(2): 232-237.