

## Exploring the Methodological Tradeoffs of Mixed-Mode Surveys with an Experimental Design

David Sterrett, Dan Malato, Jennifer Benz, Trevor Tompson, Ipek Bilgen, Nadarajasundaram Ganesh, J. Michael Dennis, and Vicki Pineau

## Benefits of a mixed-mode design

- Researchers are increasingly using a mixed-mode survey design to obtain a representative sample.
- Allowing people to complete surveys on either the phone or the web can increase coverage and enhance representativeness.
- A large portion of the public does not have access to the internet or prefers not to take surveys on the web (Sterrett et al. 2017).
- Many households are wireless only and are difficult to reach via phone (Blumberg \& Luke 2017).
- Research shows significant attitudinal/behavior differences between those with and without internet access in United States (Dutwin and Buskirk 2017).


## Challenges with a mixed-mode design

- A mixed-mode approach can potentially lead to survey mode effects.
- There could be differences between those who complete the survey on the web and those who complete it on the phone based on the interview mode.
- Phone mode features interviewers while web mode is self-administered.
- Phone mode presents questions verbally while web mode present questions visually.


## Research question

- When there are differences between those who complete the survey on the web and those who complete it on the phone, are those a result of differences in the composition of the two samples or the survey mode?
- Researchers want to maximize representativeness and capture any composition differences between two samples.
- Researchers want to minimize mode effects.
- The challenge is that respondents often select their mode so it is difficult to disentangle differences due to composition and mode.
- A survey experiment is needed to explore whether differences are due to the sample composition or survey mode.


## Survey experiment with AmeriSpeak Panel®

AmeriSpeak is the first U.S. probability-based household panel to combine the speed and cost-effectiveness of panel surveys with enhanced representativeness of the U.S. population.

Enhanced sample
representativeness with FACE-TO-FACE RECRUITMENT

Industry-leading RESPONSE RATE

## AmeriSpeak

Full sample coverage via NORC NATIONAL FRAME

Transparent sample QUALITY MEASUREMENT

## RESPONSE RATE

The AmeriSpeak recruitment protocol includes NORC field interviewers for face-to-face recruitment, resulting in the highest AAPOR response rate among U.S. panels.

## SAMPLE REPRESENTATIVENESS

AmeriSpeak is representative of the entire U.S. population because it leverages NORC's National Frame, the industry leader with $97 \%$ coverage of the U.S. population (better than address-based sampling).

## TRANSPARENCY IN SAMPLING

Through NORC Card, AmeriSpeak delivers a quantitative measurement of sample quality for clients, providing an objective measure of sample representation.

## Experimental design

## Panelist usual mode

## Experimental study mode



## Phone-Phone

Web-Phone

Web-Web

## Sample details

- Drew a nationally representative sample.

| Group | Sample size |
| :--- | :--- |
| Web-Web | 1,801 |
| Web-Phone | 1,017 |
| Phone-Phone | 880 |
| Total | 3,698 |

## Survey details

- A 15-minute survey with wide range of topics including:
- Political attitudes
- Views on social issues and the economy
- Personal finances
- Participation in social groups
- News behavior
- Personal health and medical care


## Evidence of sample composition effects

- Large differences in self-reported health between Phone-Phone and Web-Web groups, and the WebPhone group looks very similar to Web-Web group.
- Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Would you say that in general your health is...? \% Excellent or very good


## Evidence of sample composition effects

- Large differences in likelihood of having a credit card between Phone-Phone and Web-Web groups, and the Web-Phone group looks very similar to Web-Web group.
- Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Thinking about your financial situation, do you have any of the following...a credit card? \% Yes


## Evidence of sample composition effects

- Large differences in likelihood of saying benefits of immigration outweigh the risks between PhonePhone and Web-Web groups, and the Web-Phone group looks very similar to Web-Web group.
- Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

In general, do you think the benefits the United States gets from legal immigration outweigh the risks, or are the risks to the United States great enough so that legal immigration should be further limited?
\% Outweigh


## Evidence of sample composition effects

- Large differences in general trust between Phone-Phone and WebWeb groups, and the Web-Phone group looks very similar to WebWeb group.
- Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Generally speaking, would you say that most people can be trusted or that you can't be too careful in
dealing with people?
\% Most people can be trusted


## Evidence of sample composition effects

- Large differences in view toward legalizing marijuana between Phone-Phone and Web-Web groups, and the Web-Phone group looks very similar to Web-Web group.
- Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Do you think the use of marijuana should be made legal or not? \% Yes, should be made legal


## Evidence of sample mode effects

- There are differences in attitudes about discrimination against blacks between Phone-Phone and WebWeb groups, and the Web-Phone group looks more like Phone-Phone group.
- Questions such as this have potential mode effects related to social desirability bias and differences between interviewer/self administered modes.

Is there a lot of discrimination against each of the following groups, or not? Blacks
\% Yes, a lot of discrimination


## Preliminary takeaways

- There is strong evidence that offering a mixed-mode design improves sample coverage and the representativeness of the survey.
- Those who choose to complete a survey in one mode (phone) are very different on some key attitudinal and behavioral issues than those who choose to complete the survey in another mode (web).
- Many AmeriSpeak phone panelists are recruited by in-person non-response follow-up and these cases help improve sample representativeness (Bilgen 2017).
- Researchers should be aware of potential mode effects and design the questionnaire to mitigate such effects.
- Limit questions prone to social desirability bias.


## Thank You!


at the UNIVERSITY of CHICAGO
insight for informed decisions ${ }^{\text {ma }}$


## Examination of Nonresponse Follow-up (NRFU) Impact on AmeriSpeak Panel Data Quality and Study Estimates

Ipek Bilgen, J. Michael Dennis, and N. Ganesh

73nd Annual AAPOR Conference
Denver, Colorado
May 18, 2018

## Introduction

- Within the last decade, probability-based panels are growing in acceptance by the research community
- Start out with a probability-based sample frame to assure generalizability to the studied population
- Used in variety of fields - Social research, market research, medical research, election studies, etc.
- Nonresponse during the recruitment stage is as a big danger to accurate estimates in probability-based web panels
- In this study, we examine the impact of nonresponse follow-up using face-to-face (F2F) interviewing during recruitment on AmeriSpeak Panel data quality


## Methods: AmeriSpeak Recruitment Methodology



## Research Questions

What is the impact of nonresponse follow-up (NRFU) during panel recruitment on AmeriSpeak Data Quality and Study Estimates?

- What is the impact of F2F NRFU on panel response rates (AAPOR RR3)?
- Does F2F NRFU improve demographic representation of the panel sample?
- Does F2F NRFU have an impact on the study estimates?
- In what ways NRFU panel recruits differ in their opinions, attitudes, and behaviors in comparison to the initial recruited panelists among different Amerispeak studies?


## Response Rate \& Panel Quality Analysis

## AAPOR RR3 (2014-2017 AmeriSpeak Panel Recruits)

| Description | Response Rate <br> (AAPOR RR3, <br> Weighted) |
| :--- | :--- |
| Household Response <br> Rate due to Initial Recruitment | $5.8 \%$ |
| Household Response <br> Rate due to NRFU | $27.9 \%$ |
| Household <br> Response Rate | $33.7 \%$ |

NRFU boosts response rate by 5.8 times

More than half (51.4\%) of AmeriSpeak panelists have been recruited during the NRFU recruitment stage (2014-2017)

## Panel Composition (2014-2017 AmeriSpeak Panel Recruits)

Unweighted Panel Distribution by Panelist Type and Age


## Panel Composition (2014-2017 AmeriSpeak Panel Recruits)

Unweighted Panel Distribution by Panelist Type and Race


## Panel Composition (2014-2017 AmeriSpeak Panel Recruits)

Unweighted Panel Distribution by Panelist Type and Education


## Study Estimates Analysis

## Data \& Methods

## Examined Five Studies from the AmeriSpeak Panel

- Five AmeriSpeak studies which mainly examine different policy issues and political attitudes within the U.S.
- Studies selected based on:
- Types of items/measures (sensitive/non-sensitive; attitudinal/behavioral),
- Study topics, and
- Study target populations
- All studies are offered on both the phone and web
- Bivariate comparisons: One final study weight is used when survey outcomes are compared among recruitment types.
- Final study weights incorporate sampling base weights, adjustment for interview nonresponse, and raking adjustment to person-level population totals.


## Exemplar Studies

$\left.$| Study <br> Name | NASA <br> Scientific <br> Literacy Survey | GSS Web <br> Study | Gun Control <br> Survey | AARP <br> Retirement <br> Study | Justice Gap <br> Survey |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Sponsor | Dr. Jon Miller, <br> ISR, U of <br> Michigan | NORC (internal <br> AmeriSpeak <br> Study) | Johns Hopkins <br> University | AARP | Legal Services <br> Corporation |
| Target <br> Population | General U.S. <br> population | General U.S. <br> population | General U.S. <br> population - <br> oversample of <br> likely gun owner <br> panelists | Working adults |  |
| age 18-64 |  |  |  |  |  |$\quad$| Low Income |
| :--- |
| Households | \right\rvert\,

## NASA Survey: Survey Results Differences (Weighted)

How interested are you in current news events?


## NASA Survey: Survey Results Differences (Weighted)

Science and technology are making our lives healthier, easier, and more comfortable


## GSS Study: Survey Results Differences (Weighted)



## Gun Control Study: Survey Results Differences (Weighted)

Will not reduce gun violence because criminals purchase guns on the underground market


## Gun Control Study: Survey Results Differences (Weighted)

Do you favor or oppose allowing a person convicted of drunk and disorderly conduct to carry a loaded gun in public?


## AARP Study: Survey Results Differences (Weighted)

How would you characterize your political views?


## Justice Gap Study: Survey Results Differences (Weighted)

To what extent do you think people like you are treated fairly in the civil legal system?


[^0]
## Conclusion and Discussion (I)

- F2F NRFU effort during panel recruitment pays off ...

NRFU improves:

- AAPOR panel recruitment response rate by 5.8 times
- Panel sample representation improves for groups traditionally more reluctant to respond to surveys
- Younger individuals (persons age 18 to 34)
- Hispanics
- High school degree or less
- AmeriSpeak Panel members recruited through F2F NRFU have a tendency to report
- More moderate views towards policy attitudes
- Somewhat more conservative strains of political views


## Conclusion and Discussion (II)

- Surveys and panels without a robust non-response follow-up component may overstate political and policy polarization
- Findings provide evidence that individuals with stronger opinions and attitudes towards political issues are more likely to willing to respond to surveys about politics and policy issues in the U.S. than individuals with more moderate views.
- Nonresponse and Measurement Error
- NRFU panelists are more likely to provide DK responses and skip items on the web. BUT ... NRFU recruits provide their opinions for issues they feel strongly about.


## Questions? <br> bilgen-ipek@norc.org

Thank You!


## Panel-based Probability Alternatives for Sampling Racial and Ethnic Minorities

Thursday, May 17, 2018, 1:45 p.m. - 3:15 p.m.

Session Title: "Challenges of Sampling Racial and Ethnic Minorities and Non-English Speakers in the United States"
J. Michael Dennis, SVP, NORC

73nd Annual AAPOR Conference
Denver, Colorado
May 18, 2018

- Statistical surveys using probability-based sampling are operationally challenging and costly for polls of racial and ethnic minorities in the U.S.
- These challenges reflect the simple fact that a large amount of fielded sample is unavoidably wasted (not used) since only a fraction of the in-field screened households qualify for the survey

What is the well-intentioned, well-trained researcher to do?

## Probability Sample Frames Available

In order of descending cost and quality:

- Area probability sample frames
- The USPS address frame (ABS)
- Cellphone sample frames
- Commercial data files of households and/or registered voter lists


## Why not just use Non-Probability Samples?

RE: online opt-in convenience panels, river samples, router-based approaches, use of online sample exchanges, and social media recruitment

But limitations persist:

- Self-selection bias
- Socio-economic upscale bias in the sample
- Higher risk of fraudulent, "fake" respondents, and bots
- "Professional respondent" problem and panel conditioning
- High churn rate, frustrating longitudinal analysis
- "Geo offenders"
- Unintended consequences from respondent remuneration
- Statistical inference and sampling margin of error complications


## If Not Non-Probability, Use Probability Panel or Not a Panel?

- The design decision tree then comes down to whether or not to use a panel of households
- Panels are pre-recruited sample units already consented to participate in research
- Except for studies with very large budgets, probability-based surveys of racial and ethnic minorities are largely conducted using panels of pre-recruited households instead of fresh cross-sectional surveys
- There are disadvantages of probability-based panels for sampling racial/ethnic minorities: Panel conditioning from prior exposure to past panel survey taking, respondent fatigue, and a higher risk of self-selection bias (to the extent that risk is higher for panels than for fresh cross-sectional surveys)


## What is a Probability-Oriented Researcher to do on a Limited Budget?

- The responsible researcher attempts to strike an optimal balance between error reduction and cost-effectiveness

To create the necessary survey infrastructure, there are two solutions that stay within the probability sampling paradigms for panels:

- Alter the panel design itself
- Blend probability samples to obtain the necessary sample sizes

These enhancements to probability-based panels help make it possible to conduct surveys of racial and ethnic minorities (while staying in the probability sampling paradigm)

## 1. Changing the Panel Design Itself with Targeted Oversampling

In order to obtain sufficient sample sizes for surveys of the targeted racial or ethnic groups, the approach here is to change the panel design itself while keeping the sample frame constant

The use case is NORC's construction of AmeriSpeak Latino to support surveys of representative samples of US adult and teen Hispanics

AmeriSpeak Latino is a subset of NORC's AmeriSpeak Panel, supporting 2K interviews of a representative sample of U.S. Latinos

## Methods: AmeriSpeak Recruitment Methodology - Modified for AmeriSpeak Latino

## Oversample Census Tracts with 33\%+ Hispanic share

2017 ABS Frame Supplement:
Oversample HHs identified by commercial auxiliary data as likely Spanish-speaking HHs


## 2. Blending Probability Samples

- The second approach involves a judicious mix of sample frames in building the panel itself to support sufficient sample sizes
- Solution: Obtain sufficient interviewed sample sizes from the gold-standard probability panel while supplementing with race/ethnic targeted recruitment from lower-quality sample frames that still have a probability basis
- The selection of the supplemental sample frame can be determined by the accuracy of the auxiliary data appended to the sample frame and used for sample targeting (e.g., African American or Hispanic households). Other selection criteria include conventional metrics such as sample coverage, accuracy of contact information, and cost.


## Illustration of the University of Chicago GenForward Survey Panel

GenForward is a nationally representative survey panel of adults ages 18-34 recruited and administered by NORC at the University of Chicago and funded by grants to the Black Youth Project at the University of Chicago from the John D. and Catherine T. MacArthur Foundation and the Ford Foundation. The founder and principal investigator is Professor Cathy Cohen, Department of Political Science at the University of Chicago.

## gen FORW>RD

The GenForward Survey is the first of its kind-a nationally representative survey of over 1750 young adults ages $18-34$ conducted bi-monthly that pays special attention to how race and ethnicity shape how respondents experience and think about the world.

How to Build a Probability Panel to Support Polls of Young Adult People of Color?

- The initial sample plan required the collection of approximately 1,250 interviews on a near-monthly basis from a representative sample of African Americans, Hispanics, and Asians age 18 to 30
- AmeriSpeak did not have the scale to collect all the interviews


## BLENDED-SAMPLE SOLUTION

## AmeriSpeak + Targeted Registered Voter Sample Recruitment

 $=$
## gen FORW>RD

## Blending AmeriSpeak and Registered Voter Samples

- The research premise is that the AmeriSpeak Panel would provide 97\% sample coverage rate for the majority of the interviewed sample, while the rest of the interviews would be derived from a lower quality but still probability-based sample frame (about $40 \%$ to $50 \%$ sample coverage)
- Since both samples are probability samples, AmeriSpeak effectively covers the full target population, while the lower quality sample only covers a fraction of the target population
- Registered voter sample is down-weighted to represent their appropriate share of the population


## Conclusions

- Cost-effective survey infrastructure for probability-based surveys of racial and ethnic minorities by ...
- Building into the panel design itself over-sampling routines using commercial auxiliary data and Census tract data in order to obtain larger sample sizes of the targeted racial and ethnic minority segments
- Blend probability samples of varying degrees of sample coverage to construct a panel supporting surveys of racial and ethnic minority groups


## Questions?

## Dennis-Michael@norc.org

Thank You!


## Panel Recruitment for Spanish Speaking Populations: The AmeriSpeak® Case Study

Ilana Ventura<br>Rene Bautista<br>David Gleicher<br>Carolina Milesi<br>Erlina Hendarwan

May 2017, AAPOR

## Background: NORC's AmeriSpeak® Panel

- Nationally Representative Probability Sample of U.S. Households
- AmeriSpeak sample frame: NORC's area probability National Frame
- Sample coverage for over 97\% of U.S. households, enhanced coverage of
- Rural
- Low-income households
- AmeriSpeak Panel Recruitment Design: Two Stages
- Sample units are invited by mail and phone outreach to join AmeriSpeak by visiting the Panel website or by telephone (in-bound/outbound supported)
- Non-response follow-up using face-to-face/ in-person visits
- 34\% AAPOR R3 (weighted) 2014-2016 panel recruitment
- English and Spanish languages supported for online, telephone, and inperson recruitment


## Background: NORC's AmeriSpeak® Latino

- AmeriSpeak® Latino: Representing the full diversity of the Latino population
- Oversample of Hispanic and specifically Spanish-speaking segment
- 5,000 Households
- Demographically balanced sample representation
- Additional demographics: language acculturation, diversity among Latino groups, etc.
- Largest-growing segment of U.S. population
- In demand for marketing, health, policy, and social research
- Recruitment protocol sensitive to cooperation barriers
- Language
- Cultural and behavioral differences
- Privacy concerns


## Introduction

Obtaining representative samples of the Hispanic and Spanish speaking population in the United States for survey research is challenging!

How do we most effectively recruit Hispanic and Spanish Speaking respondents to the AmeriSpeak Panel?


How can we improve printed panel recruitment materials to target Hispanic and Spanish speaking populations?

## Research Questions

- Concepts
- How familiar are Hispanic and Spanish speaking respondents with surveys and social science research?
- What messages for recruitment work well?
- How we can integrate these findings into recruitment material messaging?
- Language
- How do we effectively translate materials from English, paying attention to connotations?
- Design
- How do we make design of materials (and the messages that the designs convey) conducive to Hispanic and Latino respondent recruitment?


## Methods

- Research performed November - December 2016
- Two "Rounds" of qualitative research, in Spanish, with redesign in between rounds.
- Round 1: Three In-Depth Interviews and one Focus Group
-Gender: 2 F, 7 M
-Education: 5 some HS, 3 HS graduates, 1 PhD
- Material Redesign
- Round 2: Three In-Depth Interviews and one Focus Group
-Gender: 9 F, 2 M
-Education: 8 some HS, 3 BA


## Findings

The concept of "surveys"

- Majority of respondents were not familiar with surveys, or how surveys applied to them.
- Those that have heard about surveys didn't necessarily distinguish marketing surveys from social science surveys.
- Respondents were generally skeptical of the materials and wanted to know what we were selling
- Some indicated that they did not know enough about particular topics (i.e., politics) to express an opinion.



## Findings

The concept of "community"

- Community is an important concept for Latinos and Spanish Speakers
- Did not understand the idea of how an online panel or AmeriSpeak could become their community
- They were already part of a community
- Not a place for interaction with other people

Recientemente le hemos enviado una invitación especial para unirse a AmeriSpeak, la comunidad exclusiva para compartir opiniones del NORC.

We recently sent you a special invitation to join AmeriSpeak, NORC's exclusive opinion-sharing community.
 Share your opinion. Represent your community.

- What worked: representation of one's community through surveys


## Findings

Concept \& Language: Being an Influencer vs. Being Heard

- While in English, the concept of "Being an Influencer" may be more positive, it has a negative connotation in Spanish.


SÍ, TODAVÍA HAY TIEMPO
PARA SER UNA PERSONA
INFLUEYENTE.
Gane al menos 20 dólares
directamente al registrarse $y$

YES, THERE'S STILL TIME TO BECOME AN INFLUENCER.
Earn at least $\$ 20$ right away when you register and take your introductory survey. appealing.

- Respondents were receptive to have opinions heard by government, institutions and companies.


## Findings

Language: Corporate vs. Casual

- Overly formal or "corporate" sounding language made respondents feel as though material was not for them.
- Language simplified to be more casual
- Example: "introductory survey," or "encuesta introductoria" in Spanish changed to "first survey," or "primera encuesta" in Spanish.

1) $\frac{\text { LLAMENOS OUSTIE }}{\text { WWWAMERISPEAK.ORG/ES }}$ (800) 860-1178

2 REGíSTRESE USANDO SU CONTRASEÑA

3 OBTENGA \$20 0 MÁS DESPUÉS DE COMPLETAR SU PRIMERA ENCUESTA

צul vuitu imin vuivir klinuv uv ENCUESTA INTRODUCTORIA

## 01

(1) $\frac{\text { CAL OR VISTIUS }}{\text { WWWAMERISPEAK.ORG }}$ (800) 860-1178

2 REGISTER USING YOUR UNIQUE PIN

3 GET \$20 OR MORE FOR
YOUR FIRST SURVEY
ULi mi LLmoi pau mititivu
COMPLETE AN INTRODUCTORY SURVEY

## Findings

Design: Bilingual Example 1

- Respondents reacted positively to the concept of bilingual materials
- Saw as signal that the organization cares about the Latino population.
- Some respondents found text difficult to read when Spanish and English were close to each other on page.
- When possible, placed Spanish on one side and English on the other, with an arrow indicating that the other language is on the other side....



## Findings

Design: Bilingual Example 2

- ...When not possible, visually separated the languages through graphic design elements, such as visual barriers and color.



## Findings

Design: Selection of Images for Trust

- Several respondents indicated that materials looked too corporate.
" Expressed that more "personal" and "warmer" look to the materials would be more trustworthy.
- Changes Made:

1. Graphics more colorful and rounded
2. Replaced "model-looking" photos for more "ordinary" people (include elderly, people with glasses, etc.).
3. Included images of families and cartoonlike elements.


NeRC

## Conclusions

- Concept: Message revisions focused on how AmeriSpeak is an opinion-sharing platform by which respondents could represent themselves and their families and communities to corporate, government, and other entities through taking surveys.
- Language: Reduced the quantity of text and simplified language.
- Design: Materials were redesigned to have a less corporate look-and-feel, through the use of increased color and by highlighting family as a theme.
- Simplified design enough to encompass both languages, readably and comfortably, in the same document


## Conclusions

Next Steps...

- Quantitative assessment of recruitment materials after current recruitment/ data collection wave, with experimental design in heavily Latino census tracts:
- English/ Spanish
- Spanish/ English
- Spanish Dominant
- Does having more Spanish than English on recruitment materials significantly increase Spanish-language recruitment?
- What is the effect of this language imbalance on bilingual Hispanic respondents choosing to respond in Spanish vs. English?

Ilana Ventura<br>ventura-ilana@norc.org

## Thank You!



## N*RC

at the UNIVERSITY of CHICAGO
insight for informed decisions ${ }^{\text {m" }}$


[^0]:    ${ }^{* * *} p<0.0001$

