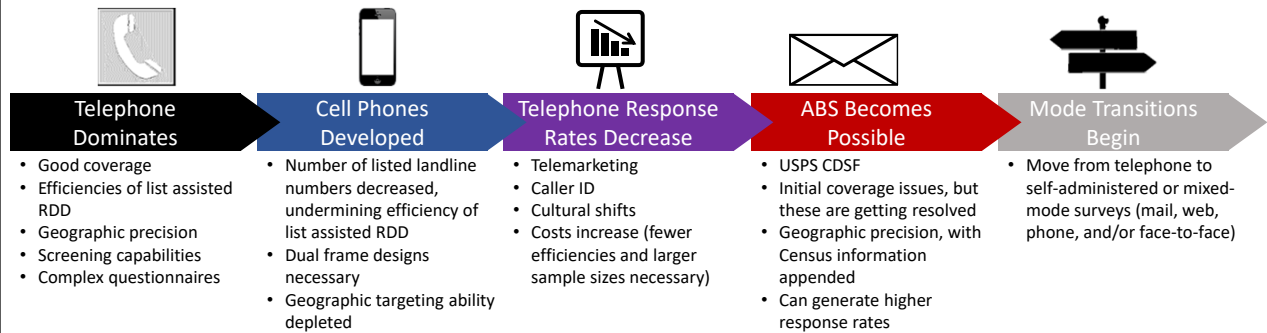


AAPOR Task Force Report on Transitions from Telephone Surveys to Mixed Mode Surveys

Kristen Olson
Jolene D. Smyth

AAPOR Webinar
September 10, 2019

Context for this Task Force



The Charge of the Task Force

- Bringing together a diverse set of experts to provide a “lay of the land”
 - What surveys have done in this transition
 - What is known
 - What areas need additional insights and research

- Review existing methods reports, technical advisory panel reports, peer-reviewed literature and survey practices

- Avoid simply re-summarizing the general mixed-mode literature; instead, focus on issues that emerge when transitioning existing telephone surveys to self-administered or mixed mode surveys.

3

Task Force Members

- Kristen Olson, University of Nebraska-Lincoln (Chair)
- Jolene D. Smyth, University of Nebraska-Lincoln (Co-Chair)
- Rachel Horwitz, US Census Bureau
- Scott Keeter, Pew Research Center
- Virginia Lesser, Oregon State University
- Stephanie Marken, Gallup
- Nancy Mathiowetz, University of Wisconsin-Milwaukee
- Jaki McCarthy, National Agricultural Statistics Service
- Eileen O'Brien, US Energy Information Administration
- Jean Opsomer, Westat
- Darby Steiger, Westat
- David Sterrett, NORC at the University of Chicago
- Jennifer Su, SSRS
- Z. Tuba Suzer-Gurtekin, University of Michigan
- Chintan Turakhia, SSRS
- James Wagner, University of Michigan

4

How did we do the work?

- Extensive review of published articles, technical reports, conference presentations, and internal reports supplied by task force members.
- AAPORnet request for any description, papers, or documentation about surveys that transitioned or were thinking about doing so.
- Convenience sample survey of the AAPOR community to get general insights about transitions.
 - Participation solicited by AAPORnet and by personal contacts
 - 21 organizations responded about 25 different data collection efforts

5

Contents of the full final report

- | | |
|---|---|
| 1. Introduction | 6. Recruitment, Nonresponse, and Operational Issues |
| 2. Coverage and Sample Design | 7. Data Preparation, Processing, and Management |
| 3. Within-Household Selection and Screening of Respondents | 8. Survey Estimation |
| 4. Questionnaire Design | 9. Costs |
| 5. Testing Strategies for Getting Questionnaires and Other Materials from One Mode to Another | 10. Human Subjects Issues |
| | 11. Communicating the Impact of the Change of Modes |

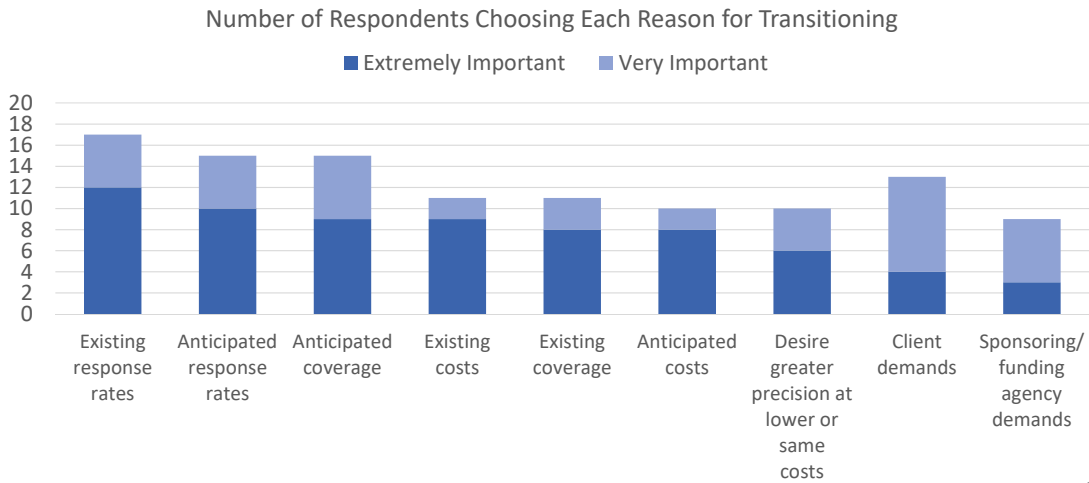
6

Today's Objective

- Give a high level overview of major themes from the report.
 - Why do surveys transition
 - What kind of surveys have transitioned
 - To what modes are they transitioning
 - Is telephone still used
 - What are the key benefits from transitioning
 - What are the major challenges from transitioning
 - What design decisions have received little empirical attention
 - What new reporting requirements may be needed
 - What organizational level shifts are occurring with transitions

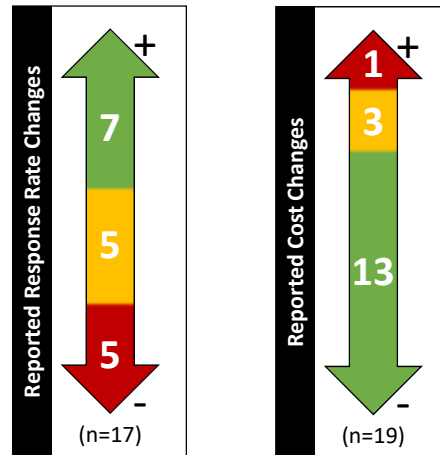
7

Why do surveys transition?



8

What were reported outcomes of transitions and lessons learned?



9

One fundamental decision that has to get made is what the goal of the transition is for data quality.

- Consistency in the survey estimates over time
 - Likely for organizations with frequent or recent telephone administrations or those retaining telephone as a mode.
 - e.g., University of Michigan's Survey of Consumers (Elkasabi, et al. 2014)
- Minimize error in the new modes, independent of the telephone mode
 - Likely for organizations with more distant telephone administrations or those moving away from telephone entirely.
 - e.g., National Household Travel Survey (Transportation Research Board, 2016)

10

What kind of surveys are transitioning?

- We saw a range of different types of surveys in both our literature review and our survey of organizations that had transitioned a survey.
 - National samples AND geographically focused samples
 - General population AND specialized population
 - Government, academic, nonprofit, AND commercial sponsored surveys
 - Largely, although not exclusively
 - household surveys and
 - cross sectional surveys
- In the full report, we illustrate points by providing examples of specific practices and methods from surveys that have transitioned.

11

What modes are being used in self-administered and mixed-mode surveys?

- Mail is the most common recruitment mode. It is used with...
 - Mail surveys
 - Web surveys
 - Concurrent mail and web surveys
 - Sequential mail+web surveys
 - Sequential web+mail surveys
 - Sequential mail+telephone surveys
 - Sequential web+telephone surveys
 - Sequential web+mail+telephone surveys
- Mail is also used to support screening for special populations
 - Mail screener with telephone, mail, or web topical surveys
 - Mail letter recruiting for mixed-mode screener

Examples and references are provided in the report.

12

- Email
 - Less commonly used for recruitment.
 - Used with
 - special populations with email lists (e.g., students, employees, etc.)
 - probability and nonprobability panels (e.g., Penn State Harrisburg Lion Poll)
 - when emails have been collected in a prior survey (e.g., screener, longitudinal survey, etc.)
 - Used to recruit for
 - Web surveys
 - Sequential web+telephone+face-to-face

- Text
 - Must have consent to send a text prior to doing so
 - Can be used in screening, web panels, and longitudinal surveys where consent can be obtained
 - e.g., 2017 National Household Travel Survey

13

Is telephone still part of the mix in mixed mode surveys?

- Many early studies that transitioned tried to get phone numbers for sampled cases to facilitate phone attempts
 - Phone numbers matched to addresses from the DSF
 - Mail used to request phone numbers (e.g., Wisconsin Family Health Survey, California Health Interview Survey)

- Eventually many researchers began to move away from telephone as a primary mode and instead used it as a nonresponse-follow-up mode.
 - e.g., Racial and Ethnic Approaches to Community Health Across the US Risk Factor Survey (REACH US)

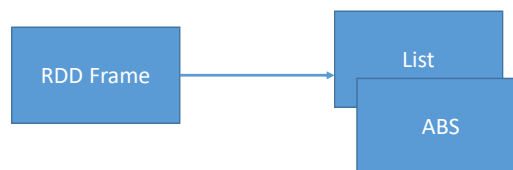
- Mail and web surveys are certainly more common now, but transitioning does not necessarily mean that interviewer-administered modes are abandoned.

14

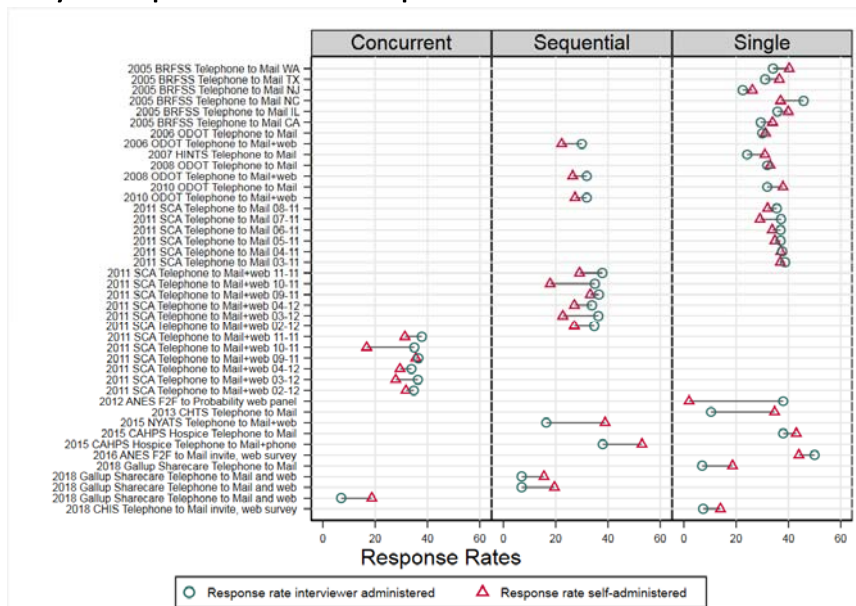
Benefits from transitioning

Improved frame coverage and geographic targeting

- Geographic linkage is easy with addresses
 - States, cities, neighborhoods
- Multiple frames may be used, or information appended to DSF for stratification purposes and/or targeting rare populations
 - Example: Coastal Household Telephone Survey (CHTS, Brick, et al., 2016)
 - Transitioned from landline RDD to a dual frame design:
 1. List of angler licensees
 2. ABS frame
 - Example: Surname listing to target Spanish, Korean, and Vietnamese-speaking households in the California Health Interview Survey pilot test (Wells, et al. 2018)



Potentially improved response rates



Need for a relook at concurrent, sequential and single mode designs

- Response rates for
 - **Single mode** web-only or mail-only designs and **concurrent** web and mail designs = No notable difference
 - **Single mode** web-only or mail-only designs and **sequential** web and mail designs = No notable difference or higher for sequential modes
 - **Concurrent** web and mail designs and **sequential** web and mail designs = No notable difference or higher for concurrent designs

Increased flexibility in the use of incentives

- Prepaid incentives of \$2 and \$5 are the most common incentive levels used in surveys that have transitioned or examined transitioning
 - Others have used prepaid incentive levels of \$1, \$10, \$20, and \$30
- Promised incentives have higher values than prepaid incentives, with a promised value of \$20 being the most common
 - Others have used promised incentive levels of \$5, \$10, \$15, or more than \$20
- A combination of prepaid and promised incentives can be effective in pushing respondents to a new mode (e.g., Biemer, et al. 2018).
- Nonmonetary incentives have seen less success.

Innovative measurement possibilities

- Surveys that transitioned have taken advantage of innovative opportunities for measurement, adding visual cues and automated calculations to web surveys
 - Mapping functions
 - Automated calculations
 - Images to help with recall

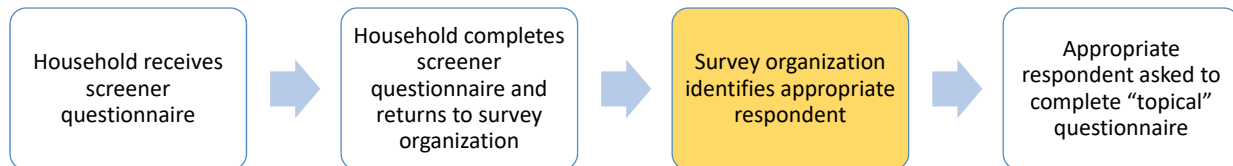
Potentially lower costs

- Few studies directly compare costs from a previously administered telephone survey to a self-administered or mixed mode surveys
 - 2012 ANES: >\$2000/complete for F2F; \$240/complete for GfK panel web;
 - 2016 ANES general pop'n web about \$150-\$200/complete (only mailing and incentive costs)
 - Oklahoma tobacco survey: Total phone costs (\$39K) lower than web+phone (\$44K), but costs per complete lower for web+phone (\$5.66) than phone (\$6.25)
 - Link, et al. (2008): Phone: \$79K/1000 completes; Mail: \$70K/1000 completes
- In our survey
 - 56% reported redesign was to reduce total survey costs; 68% reported total costs were reduced; 81% reported costs per complete were reduced

Challenges from transitioning

Within-household selection may be problematic

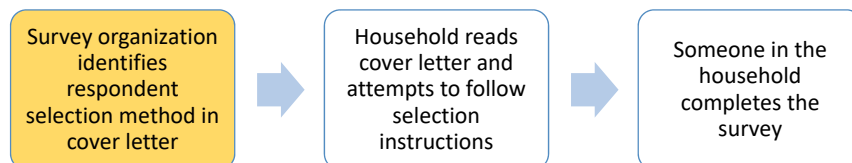
- Approach #1: Separate the rostering step (by the household) from the selection step (by the survey organization)



- Example: National Household Education Surveys (Brick, Williams, and Montaquila 2011; Montaquila, et al. 2013)
- Mode matters – mail screeners more likely to be completed than telephone or web modes; web screeners more likely to yield a completed “topical” questionnaire (Han, et al. 2010; Montaquila, et al. 2013; Amaya, et al. 2015; DeBell, et al. 2017; McPhee, et al. 2018)

Within-household selection may be problematic

- Approach #2: Rostering and selection step in the hands of the household via cover letter instructions and/or survey instrument questions



- Methods used reflect range used on telephone

One stage of selection methods

- Any adult, most knowledgeable person, or head of household (e.g., Biemer, et al. 2018)
 - Example: Survey of Consumers: “the head of the household or his or her partner complete the questionnaire” (Elkasabi, et al. 2014)
- All adults (e.g., Battaglia, et al. 2008; Brick, Andrews, and Mathiowetz 2016)
 - Example: BRFSS: “This survey should be completed by every adult, age 18 or older, living in your household.” (Battaglia, et al. 2008)
 - Example: Proxy reporters for entire household (Brick, Andrews, and Mathiowetz 2016)
- Age/Order selection (e.g., Bosa, Gagnon, and Caron 2017; DeBell, et al. 2017; Wells, et al. 2018; Smyth, Olson, and Stange forthcoming)
 - Requires multiple cover letters, each specifying a different combination of age/order position in the household (and sometimes gender)
 - Example: CHIS pilot (Wells, et al. 2018)

Step 1: Identify who should complete the survey

How many adults, 18 years of age or older, are in your household?

One adult: You should complete the survey.

Two adults: The **older adult** should complete the survey.

Three or more adults: List the three oldest adults in order from **oldest to youngest**. The **third person on the list** should complete the survey.

1. _____ 2. _____ 3. _____

If you are **not** the selected adult, please share this information with the selected adult and ask them to complete the survey by going to the website listed below. 25

Birthday methods

- Commonly used in telephone surveys; Often used when transitioning to self-administered surveys.
 - No consistent method for implementing the birthday method across surveys – most use “next” birthday
- Example: Bosa, Gagnon, and Caron (2017)
 - Who should complete this survey? The person in your household who had the most recent birthday, and is 18 years of age or older, has been selected to participate.
- Example: Hicks and Cantor (2012)
 1. Is there more than one person age 18 or older living in this household? Yes/No
 2. Including yourself, how many people age 18 or older live in this household? ___
 3. The adult with the next birthday should complete this questionnaire. This way, across all households, HINTS will include responses from adults of all ages.
 4. Please write the first name, nickname, or initials of the adult with the next birthday. This is the person who should complete the questionnaire. _____

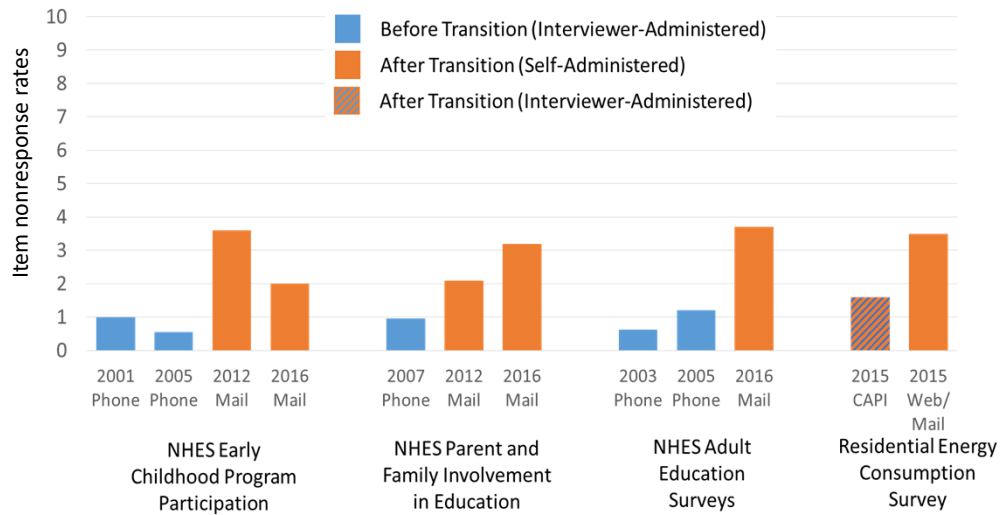
Long questionnaires may be shortened

- Many surveys that transitioned shorten the questionnaire
 - Examples: Residential Energy Consumption Survey shortened a 40 minute F2F survey to a 20-30 minute web and mail questionnaire by focusing on most critical content
 - 2007 Health Information National Trends Survey shortened a 40 minute telephone interview to a 30 minute questionnaire
 - National Household Transportation Survey –
 - “trip purpose” response categories reduced from 30 to 19.
 - “means of transportation response categories reduced from 24 to 20.
 - Reduced the burden of trip rostering.
 - If another household member reported they traveled together, the destination and many details (party size, vehicle, etc.) was automatically populated on the participant’s roster. The participant merely had to confirm and/or edit the trip and details.
- Some surveys have examined offering the questionnaire in multiple modules, although the efficacy of this approach is still under investigation
(e.g., Peytchev, et al. 2019; Liao, et al. 2019)

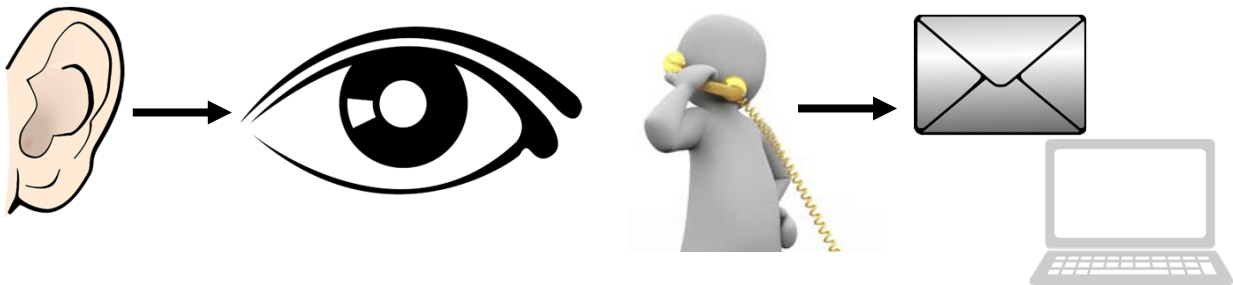
Computerization may affect data collection decisions

- Automation commonly used in CATI, CAPI, and Web is not available in mail.
- Mail questionnaire may need to be simplified, abbreviated, or redesigned to avoid complex skip patterns
 - Example: National Household Education Survey simplified skip patterns; moved set of questions applying only to homeschoolers (3% of population) to a separate questionnaire
(Chapman and Hagedorn 2009).
- Different modes may adopt different levels of personalization or question wording.
 - Example: National Survey of Children’s Health (U.S. Census Bureau 2018)
 - Web – skip patterns, range checks, “pick lists”, fills, and required responses for screening questions, soft edit prompts, online help screens.
 - Mail – child’s name, initials, or nickname; age; and sex were printed on topical questionnaires

Item nonresponse rates may increase when surveys transition to self-administered or mixed modes



Loss of benefits of interviewer administration



- Changes in survey estimates on items related to:
 - Knowledge questions
 - Socially (un)desirable issues
 - Autobiographical information obtained from records
 - Topics subject to acquiescence
 - Ordinal attitudinal items

Additional decisions related to language of administration

- Most surveys that transitioned were administered in English only or in English and Spanish only
 - Decision often based on the prevalence of different languages in the previous interviewer-administered version (e.g., Ghandour, et al. 2018)

- Surveys that transitioned in multiple languages often include cover letters and survey questionnaires in multiple languages from initial mailing, sometimes to a targeted set of households (e.g., Zuckerberg and Mamedova 2012; Montaquila, et al. 2013; Brick, et al. 2012; Brick, et al. 2013; Jans, et al. 2013; US Census Bureau 2018; Ghandour, et al. 2018)

Additional decisions related to language of administration

- Respondents may complete surveys in multiple languages
 - Web surveys have offered respondents the ability to toggle between languages
 - Respondents answer some questions in mail surveys in one language and other questions in a different language, depending on the design of the survey
 - Survey organizations then to make decisions about how to record “language of interview” – for example, as the language used on the last question of the questionnaire

- Asking respondents to call into a language line is less successful than including translated materials in the mailings (e.g., Cantor, et al. 2009; Wells, et al. 2018)

Increased difficulty in interviewing children and teens

- Is the child or teen being reported on by their parents or on themselves?
- Child-specific household roster may be embedded in a household-level questionnaire
 - Focal child selected from the household in a separate screening instrument or a web roster
 - Proxy reports by parents on children generally as successful in surveys that transitioned to self-administered or mixed modes as doing the same task in telephone
- Self-reports by teens very challenging in surveys that transitioned
 - Parental permission required
 - Separate contact information needed for teens
 - Contacting and then obtaining cooperation from teens themselves

Increased difficulty in collecting nonsurvey data

- Physical measurements and neighborhood observations
- Biological measurements
 - Some studies send a separate observational team (Harris 2018)
 - Some send respondents to clinics (Sakshaug, et al. 2015)
 - Others use self-administration, with success depending on how request is made and what biological samples are requested (Gatny, Couper, and Axinn 2013; Sakshaug, Couper, and Ofstedal 2010; Dykema, et al. 2017)
- Recording linkage consent rates and nonconsent bias higher in self-administered than interviewer-administered modes (Fulton 2012; Sakshaug, et al. 2017).

What hypotheses are of interest to test when transitioning?

- $H_0: \theta_{Phone} - \theta_{Self-admin} = 0$
- $H_A: \theta_{Phone} - \theta_{Self-admin} \neq 0$

Are the overall estimates the same or different?
(Both sets of estimates may have biases, but the biases are about the same size in each mode.)

- $H_0: (\theta_{Phone,t=1} - \theta_{Phone,t=2}) - (\theta_{Self-admin,t=1} - \theta_{Self-admin,t=2}) = 0$
- $H_A: (\theta_{Phone,t=1} - \theta_{Phone,t=2}) - (\theta_{Self-admin,t=1} - \theta_{Self-admin,t=2}) \neq 0$

Are the overall change estimates the same or different?
(Both sets of change estimates may have biases, but the biases are about the same size in each mode.)

No single approach to measuring or adjusting for mode differences

- Estimates from telephone and self-administered or mixed mode data collection efforts may differ because of
 - Selection errors (coverage and nonresponse)
 - Measurement errors
- Data are needed to evaluate differences between these modes
 - “Gold standard” or administrative record systems
 - Parallel surveys, conducted in different modes on different respondents, sometimes called “benchmark” or “bridge” surveys
 - Repeated measurements on same respondents in different modes
- Many statistical and analytic approaches exist
 - “Mode Bias Factors,” regression models, propensity score adjustments, imputation
- Different approaches may come to similar conclusions, but this is estimate specific

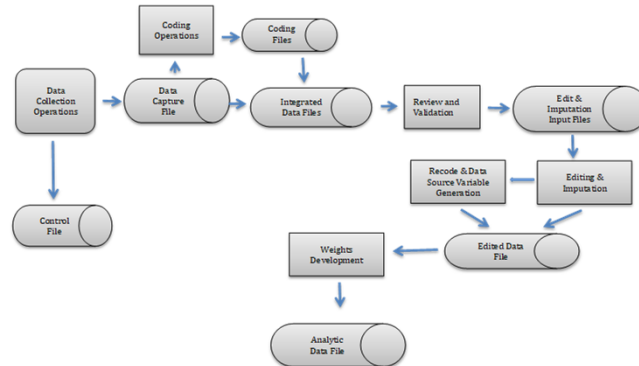
Deciding whether to transition using bridge surveys

- Some, but not all, organizations simultaneously fielded the survey in the old mode and the new mode for purposes of calibrating estimates.
- This is known as a bridge study
 - Expensive. May be necessary for estimation of error source changes. May not be relevant if survey landscape has changed (e.g., National Household Travel Survey)
- Transitions take time
 - Can take months to years to plan and implement the transition, leading to years without estimates
 - Estimates may differ so much in new survey that need to plan additional pilot and field tests; build in additional years of bridge studies.
 - New modes may lengthen field period or reduce it.
- To address these issues, some organizations decided to compare the new mode implementation with the most recently fielded interviewer-administered mode rather than fielding a bridge study.

What design decisions have received little empirical attention in surveys that transitioned?

Almost no empirical or experimental data about how processing changes when moving to self-administered and mixed mode surveys

- A few key decision: Should (and what amount of) data quality checks should be built into computer-assisted instruments that can't be built into a mail survey?



39

Processing decisions

- Transparency is key
 - Need to have clear documentation for each decision made as data are being combined/pooled across modes
- **Data entry:** Will single integrated system be used for all modes (including mail) or separate systems for different modes?
- **Deduplication:** May receive multiple copies of a completed survey in different modes. Deduplication rules may depend on (1) mode of completion, (2) level of completion, (3) date of completion. Should be preidentified before data collection.
 - Example: National Survey of Children's Health prioritized a completed questionnaire in any mode, but selected the completed web questionnaires if both web and mail questionnaires were returned and completely filled out.
- **Coding** of open-ended questions may differ in amount, detail, and characters used across modes
 - Example: American Community Survey needed to "clean" the text of characters entered in web survey that were not recognized by automated coding system (":") (US Census Bureau 2014)

40

We need research on informed consent design when transitioning modes

- 13 of 23 survey organizations reported that their informed consent process changed with the transition.
- Transition from interviewer-obtained consent to respondent self-administered consent
 - This process “may be challenging for youth assent and literacy issues”.
- Few studies that transitioned reported on comparing informed consent across interviewer- and self-administered modes. More research is needed here.
- More research is needed on the best design features for informed consent in self-administered modes.

Other human subjects issues

- Interviewer Administered Modes
 - **Informed consent:** Interviewers are trained to read all of the consent language, answer questions and provide clarification as needed, and watch for verbal and non-verbal cues for understanding/misunderstanding.
 - **Personally identifiable information:** Interviewers can remind respondents not to share this information and/or choose not to record it in the CAPI/CATI instrument.
 - **Mandatory reporting of abuse/harm:** Interviewers trained to tell respondents and understand what triggers mandatory reporting requirements. Reports can be made quickly.
 - **Handling respondent distress:** Interviewers trained to see R distress and start a distress protocol.
- Self Administered Modes
 - **Informed consent:** No clear way to check for literacy, reading of, or understanding of the consent document.
 - **Personally identifiable information:** Respondents can write PII on the questionnaire (open-boxes and/or margins), which will need to be redacted from all data sets. Web surveys also collect additional PII (e.g., IP addresses) that needs to be protected.
 - **Mandatory reporting of abuse/harm:** Mandatory reporting information can go into the informed consent or questionnaire, but there’s no guarantee respondents read or understand it. Considerable time can pass before data is processed and anyone sees what was reported. Sensitive topic surveys may want to anticipate ongoing review of open-ended responses, automated flags for key words triggering mandatory reporting.
 - **Handling respondent distress:** No way to engage a distress protocol in self-administered surveys. Can provide resources. More research is needed.

What new reporting requirements may be needed?

Surveys should transparently identify on data file or in documentation

- Surveys should transparently identify on data file or in documentation
 - Changes in design features from past data collection cycles to this data collection cycle
 - Mode of data collection (e.g., SRVMODE in NSCG)
 - Device of data collection (from self-report or paradata, if possible)
 - Clear documentation of all questionnaires, including screenshots from web and mobile questionnaires
- Weighting decisions
 - Decisions about calculating response rates may change when transitioning, and thus may change weighting decisions.
 - Different eligibility may be observed across modes (e.g., 2015 Residential Energy Consumption Survey)
 - Separate weights for different modes of data collection (e.g., American National Election Studies)

What organization-level shifts are occurring as surveys transition from telephone to self-administered and mixed modes?

Data Collection Systems are Changing

- Important to have data collection systems that effectively tracks what contacts cases have received
 - Allows for evaluation of the effectiveness of data collection strategies
 - Helps ensure interventions are properly employed
- In mixed-mode designs, systems need to talk across modes
- May be challenging or require significant infrastructure development at survey organizations
 - Large organizations may develop in-house data collection systems (e.g., Cheung and Maher 2015, Wernimont and Snowden 2015, Edwards, Maitland, and Connor 2017, Bonhomme 2018)
 - Small organizations may be doing this manually using existing spreadsheet or statistical software packages

Cost structures are changing

- Mail only vs. web only vs. mail and web modes
 - For small surveys, mail only may be more cost effective.
 - What the “tipping” point is for programming a web survey in addition to a mail survey varies across studies (ranges from n=620 to n=2000).
 - Costs may have been invested in programming a complicated telephone instrument that reduces the costs for programming a web survey.
 - More work is needed here to understand the cost vs. response rate and coverage tradeoffs (e.g., Fricker and Schonlau 2002; Griffis, Goldsby, and Cooper 2003; Lien 2015; Lesser, et al. 2017; Kaminska and Lynn 2017)
- Mail surveys require data entry, a task that some organizations may not be set up for

Conclusion

- The survey research field is incredibly adaptable.
- There is no single way that a survey is transitioned from telephone to self-administered or mixed modes of data collection.
- Clearly communicating the decisions made during the transition, and how they affect survey estimates is key.

Thank you!

kolson5@unl.edu @olson_km
jsmyth2@unl.edu @jolene_smyth