Framework for Data Collection and Analysis

Week 1 Introduction, Classic Modes of Survey Data Collection

[Reading 1] THE FUTURE OF MODES OF DATA COLLECTION by MICK P. COUPER

• How the word "mode" has been used in survey methodology: Tests of paper and pencil versus computer-assisted interviewing are viewed as mode comparisons, although one could argue that it is the technology that changed, not the mode of data collection

More recent trends in what mode means…

• Proliferation of mode

➔ we have seen a rise in the number of different modes, and in the variations within a particular mode

➔ growing body of research evidence on the benefits of self-administration for sensitive questions or those subject to social desirability effects led to the development of new methods of self-administration as part of an interviewer-administered survey. Thus, self-administered questionnaires (SAQs) evolved into computer-assisted self-interviewing (CASI) and its variants (e.g., audio-CASI), and telephone versions (interactive voice response [IVR] or T-ACASI) (see Couper and Nicholls 1998).

• Increasing complexity

➔ computer-assisted self-interviewing can involve text, audio, or video presentation of questions or stimulus materials

➔ Internet / Web Surveys: Whereas Web surveys can be viewed as a single mode, there are many different ways that samples can be drawn for Web surveys (see Couper 2000 for an early review). These can include both nonprobability- and probability-based methods, and include such approaches as online access panels ++ There are many different ways one can implement Web surveys, sometimes resembling interviewerадministered surveys in some aspects of measurement, at other times replicating mail surveys.

• Mixed Modes

➔ Why? the threat to telephone surveys, for many decades a mainstay of much public opinion research, the feasibility of address-based sampling and the resurgence of interest in mail surveys, and the measurement opportunities offered by Internet surveys

• Dimensions of Mode

➔ Degree of Interviewer Involvement

Some modes fully administered by interviewer – Some modes are self administered but in the presence of interviewer (e.g. CASI, paper self-administered questionnaires SAQ) – drop off questionnaires – fully self-administered modes such as mail surveys, Web surveys

This has implications not only for costs, but also for errors of non-observation (e.g. sampling and non-response) and measurement errors (e.g. effects on sensitive questions, ability to motivate, probe, assist, etc.)

➔ Degree of contact with the respondent

Face to face survey – modes that involve pictures or videos of interviewers reading questions – telephone – mail, web surveys

➔ Channels of communication

Oral, Aural, visual etc. channels of communication

➔ Varying degree of privacy

High privacy modes: CASI, mail, web surveys // medium level privacy: interviewer administered survey // low privacy: exit interviews or intercept surveys (interviewed with others around them)

➔ Degree of computer technology used

• Research on Mode Differences

➔ Web v.s. Mail Comparisons

Response rate differences in Web surveys versus other modes (mostly mail) (see Lozar Manfreda et al. 2008; Shih and Fan 2007). Though it now seems clear that Web surveys generally get lower response rates than mail surveys, there are some exceptions. We don't yet know why this is the case, and under what conditions we are likely to get bigger or smaller differences in the rates ++ while we know quite a bit about nonresponse rates, we know considerably less about nonresponse bias differences between these two modes

➔ Mixing Modes

Modes are also mixed for a variety of different reasons, from reducing costs and increasing response rates, to addressing differential coverage, targeting specific subgroups, improving measurement, and so on. A key distinction is whether different modes are used for the contact phase, the response phase, or the follow-up phase.
The recent surge of interest has focused on mixed mode designs where some part of the sample are interviewed or provide data using one mode and another part do so using another mode.
e.g.1 dual frame design involving RDD and area probability samples (Groves and Lepkowski 1985), where the part of the sample with telephone access is surveyed by telephone whereas those without are interviewed in person.
e.g.2 where the entire sample is initially approached using one mode, but remaining nonrespondents are then followed up using a different mode.
Mixed mode can also be extended in time and space (1) across country mixed mode (2) within country mixed mode (3) across time mixed mode
The promise of mixed mode designs—especially those involving self-administration—is that the drawbacks of one mode can be compensated for by the strengths of another. A related hope is that by encouraging more people to use the cheaper mode, resources can be set aside to target particular subgroups with the more expensive mode.

Several studies offering respondents a choice of mode in **concurrent** mixed mode designs have found that doing so does not appear to increase response rates (e.g., Griffin, Fischer, and Morgan 2001; Tourkin et al. 2005; Gentry and Good 2008). However, **sequential** mixed mode approaches, in which sampled persons are offered first one mode and then the other, appear to be more promising

• Mixing Modes may add complications in terms of measurement error
➔ "prevention" strategy, attempts to minimize measurement error differences between the modes (Martin et al.'s (2007) "universal presentation" and Dillman's (2007) "unimode construction" approaches to questionnaire design exemplify this strategy)
➔ "correction" or "adjustment" approach: argues that measurement differences are fundamental features of the mode, and cannot be designed away. This approach argues for maximizing the design benefits of each mode (or optimizing the design for each mode), rather than compromising to produce the lowest common denominator that may characterize the prevention strategy. The measurement differences then have to be statistically adjusted to produce comparable measurements across the modes used (Vannieuwenhuyze, Loosveldt, and Molenberghs (2011)

• Another type of mixed mode design is one where the **modes are targeted at specific groups of questions,** rather than at groups of respondents.
➔ Here, the focus is on measurement error. The more we know about the characteristics of modes and the effects they have on certain types of respondents given certain types of questions, the more we can exploit the flexibility of modes to improve the quality of measurement.
For instance, research by Conrad and Schober (2000; see also Schober and Conrad 1997) suggests that conversational interviewing may be beneficial under certain circumstances, but not others. It is a challenge to expect interviewers to behave one way for one set of questions and a different way for another. But as we develop new methods and understand their properties, we will be better able to target particular approaches to specific sets of questions. This is already being done with CASI components of CAPI interviews for sensitive questions.

• How we REPORT modes
➔ We need better ways to describe the data collection methods employed in a survey.
e.g. when reporting on a mixed mode study, providing a single response rate is insufficient. Knowing the response rate to each mode—and the characteristics of those who responded in each mode—would be important information to disclose. Similarly, details of the protocol utilized would be needed to judge the value of the survey or its results.
➔ We also need better documentation for the user or analyst
e.g. In mixed mode designs, it may be important for the analyst to know in which mode a particular respondent responded to a particular item. In a similar fashion to the increasing recognition that nonresponse bias is a property of a statistic, the effects of mode may be best understood at the item and person level.

• What will the future look like…?
➔ interviewer-administered modes will continue to be an important part of the survey researcher's toolkit, albeit in increasingly limited roles. Relying on modes requiring some level of literacy, such as the Web or mail, may also limit generalizability for certain kinds of studies. Another related reason for the continued survival of interviewer-administered surveys is that Web surveys (in particular) have led to large increases in the number of surveys.

➔ survey takes longer than the transaction itself. **Surveys have become commodities** (Tourangeau 2007). There is much more competition for the attention of respondents, whom we should view as an increasingly scarce resource.

➔ Social Media… on the rise but…

First, despite the impressive number, and the fact that Facebook users are of interest in their own right, few would suggest that those who use Facebook are representative of the entire population of any particular country

Second, the set of Facebook users is not available to the outside community as a sampling frame. Researchers are restricted to using a variety of nonprobability methods (such as snowball sampling) to study Facebook users (see, e.g., Bhutta 2010).

Third, Facebook users are becoming increasingly aware of privacy issues, and new tools are being developed to give users control over who has access to what content. This limits the passive collection of data from users to those who are willing to publicly share the information, which is likely to be an increasingly selective subset of Facebook users. These rich data sources will add much to our understanding of public opinion, but will not replace surveys in the near future.

[Reading 2] To Mix or Not to Mix Data Collection Modes in Surveys by Edith D. de Leeuw

• The development of integrated programs stimulated the use of computer-assisted self-interviewing (CASI) in face-to-face interviews, and CAPI-CASI mixes became popular especially in interviews on sensitive topics. The procedure is straightforward: when sensitive questions have to be asked the interviewer hands over the computer to the respondent for a short period. The respondent can answer in all privacy and the interviewer remains at a respectful distance, but is available for instructions and assistance. This is the most common use of CASI and is equivalent to the traditional procedure where an interviewer might give a paper questionnaire to a respondent to fill in privately (cf. De Leeuw, 2002).

• it is a question of finding the **best affordable method**, and sometimes the best affordable method is a mixed-mode design.

• mixed-mode approach because mixing modes gives an opportunity to compensate for the weaknesses of each individual mode at affordable cost.

• Telephone survey coverage bias: **dual-frame mixed mode** employed

Coverage bias occurred because part of the population did not have a telephone and the no telephone households differed from the telephone households on socio-demographic variables such as age and social economic status. A dual-frame mixed-mode design has the advantage of the cost savings of telephone interviewing and the increased coverage of area probability sampling

• Web Survey coverage bias: mixed mode strategy employed as well

e.g. Parackal (2003) survey on mobile phones and interest in WAP tech ➔ used a mixed-mode or hybrid survey approach, in which **all sampled units were contacted by means of a paper letter and given the choice to either use the Internet or request a paper questionnaire.**

• To achieve higher response rates, while keeping the overall costs low, mixed-mode strategies are used, starting with the less costly method first.

➔ American Community Survey: mail survey with follow-up telephone interviews for non respondents, followed by face-to-face interviews for a subsample of the remaining non respondents

• One of the most consistent findings in mode comparisons is that self-administered forms of data collection perform better than interview-modes when sensitive questions are asked (for an overview, see De Leeuw 1992). Therefore, mixed-mode approaches using a paper self-administered form to elicit sensitive information in a face-to-face interview have been standard good practice for a long time

• It is important to realize that survey researchers communicate with sample members at different points in time and that they may use different modes of communication to do so.

• mixed-mode in Pre-contact phase: used for prenotification and recruitment.

➔ e.g. Nielson media research good example; use of paper advance letters in telephone surveys; it is easier to establish legitimacy and trust through an official letter that has a letterhead, contact information and a signature, than through a mere voice over the phone. That advance letters indeed reduce nonresponse in telephone surveys is shown by De Leeuw, Hox, Korendijk, Lensvelt-Mulders, and Callegaro (2004). Telephone precontact before a mail or web survey has been found to be effective as well.

➔ Another scenario: letting respondent to choose a mode // e.g. a paper mail advance letter with an invitation to complete a web survey, but also offering the opportunity to ask for a paper questionnaire

American Lung Association survey of asthma awareness among school nurses. In this survey postcards are sent to a random sample inviting nurses to participate online via an indicated web site or by telephone via a toll-free 800 number. A procedure like this is often used to reduce coverage error

BUT other errors come into play because there are now MULTIPLE modes
[1] self selection may cause differences in socio-demographic variables
[2] mode itself may cause measurement differences
Need to make choice… multiple-mode with reduced coverage error at the price of increased measurement error or a uni-mode approach with a larger coverage error component.
• **Concurrent mixed mode design**: paper mail survey with a web option; choice of web or telephone etc.
Aim is to reduce coverage bias and still complete survey at a reasonable cost; also assuming that giving sample member a choice of mode may reduce non response as some people may express reference over certain modes
➔ However, there is no firm empirical evidence for this. Dillman, Clark, and West (1995) did not detect any improvement in response rates when respondents were given a choice between sending in a mail questionnaire and phoning in their answers. A similar conclusion was reached by Lozar Manfreda et al. (2001), who offered a choice of web or paper mail to respondents. Balden (2004) also reports that in his experience providing respondents in market research with choices does not in general improve the overall response rates; this includes the choice combinations mail/web, mail/IVR, and phone/web.
• **sequential multi-mode system**: To reduce survey non response
Usually an inexpensive mode is used as the main mode for the whole sample and then a more expensive mode is used for the nonresponse follow-up to improve response rates
There is also evidence that a sequential mixed-mode raises the response rates in establishment surveys

• **One sample, one time point, but different modes for different parts of the questionnaire**
different modes are used for a subset of questions in the questionnaire during a single data collection period. Usually a mix of interview and self-administered forms is used to exploit the strong points of both methods.
e.g. Within an interview a self-administered form of data collection such as CASI or Audio-CASI is used for sensitive questions to reduce social desirability and enhance privacy as neither the interviewer nor any other person in the vicinity will know the answers given. This situation is basically positive and is not a case for concern.
• **One sample, multiple time points** mixed-mode design: longitudinal study or a panel in which the same respondents are surveyed at different time points, and different modes are used at those different time points.
➔ Together with the greater flexibility of an interviewer to gain cooperation at the doorstep and the opportunities for optimal screening, a face-to-face interview is often the preferred choice for the base-line study of a panel.
➔ When possible a less expensive method is used after the first wave to reduce costs. A combination of face to-face interviews for the first wave and telephone surveys for the second is used for labor force surveys in several countries.
➔ Another example of a mixed-mode panel combines an initial face-to-face interview with mail surveys in subsequent waves.
➔ Sometimes modes can even be switched back and forth. For instance, after an initial face-to-face survey, telephone and mail surveys are employed with an occasional face-to-face survey interspaced at crucial points. Also, in longitudinal health surveys and in growth studies of infants it may be necessary to include simple medical tests at regular times, which need a face-to-face contact. Internet panels, which are now becoming fashionable, are often formed after an initial telephone survey.
➔ Limitation: in sequential mixed mode, studies time and mode effects are confounded, and it is difficult to decide if a change over time is real or the result of a change of mode.
• **Different Samples, different modes**
➔ Different modes for different populations or subgroups
➔ Different countries may have different survey traditions and/or different practical constraints. In a densely populated country face-to-face surveys are feasible, but in sparsely populated areas that may not be the case. Some countries have detailed registers and address information to ensure successful mail surveys, while in other countries area probability based samples are the only option. Low literacy levels can preclude mail surveys, and in poor regions electronic equipment may be difficult to use. Furthermore, different data collection agencies may have different survey traditions and therefore may differ in availability of experienced and trained staff needed for specific modes and other required resources. To enhance comparability other design factors should be kept constant as far as possible.
➔ e.g. Behavioral Risk Factor Surveillance System (BRFSS) of the Centers for Disease Control and Prevention (CDC) in which 15 states participate in monthly data collections. A standard core questionnaire was developed by CDC for the states to provide data that could be compared across states. Data collection varies by state, which results in a uni-mode approach within one state but a mixed-mode design for the total study.

• **Follow up Phase**

➔ reminders: tool to increase response rates (either same or different mode of contact); sometimes a different mode of contact is used for the follow-up contact (costs and time constraints may prohibit in-person follow-ups in a face-to-face survey ++ different mode for the follow-up may also lead to additional information about the sampling frame or improved contact information and changing modes may improve the attention or novelty value of the reminder)
➔ persuasion letters
➔ If the mode change only involves the follow-up reminder, potential measurement errors due to mode effects are avoided, while the potential benefits of a second mode of communication can be exploited: a win – win situation, just as when using pre-notifications

• **Mode Effects and Data Quality**
➔ it is important to ascertain whether only the contact strategies should be mixed-mode and the data collection itself unimodal, or if multiple modes should be used for the actual data collection. Only in the latter case does the risk of mode effects on measurement error threaten the internal validity of the survey
➔ media-related factors: familiarity with a medium, use of medium, and locus of control (who has the most control over the question-answer process)
➔ factors influencing the information transmission: The way information is transmitted determines the cognitive stimulus people receive and this differs across modes of data collection.
The first mode difference is whether the information is presented visually or auditive (aurally), or both (Schwarz et al. 1991). When information is presented only aurally this demands more memory capacity of the respondent and may lead to recency effects in long lists (i.e., the last response category is chosen more often than earlier response categories on the list). availability of communication channels: verbal communication (words, text), nonverbal communication (e.g., gestures, expressions, body posture), and paralinguistic communication (e.g., tone, timing, emphasis).
➔ interviewer effects: It is important to realize that these factors are interrelated: in a face-to-face situation an interviewer can use more channels of communication than in a telephone situation, as information transmission and interviewer effects are related ++ The more control respondents have of the data collection, of their feelings of privacy and their willingness to disclose sensitive information increases and social desirability decreases (and vice versa)

• Empirical Mode Comparisons
➔ Comparing mail surveys with both telephone and face-to-face interviews, De Leeuw found that it is indeed somewhat harder to get people to answer questions in mail surveys. Both the overall nonresponse rates and the item nonresponse rates are higher in self-administered questionnaires than in interviews. However, when questions are answered, the resulting data tend to be of better quality. Especially with more sensitive questions, mail surveys performed better, with, in general, less social desirability bias in responses, more reporting of sensitive behavior like drinking, and less item non-response on income questions
➔ the structure of questions may be inadvertently changed during a mode change in a mixed mode desing. For instance, an unfolding procedure may be used in the telephone mode, while the full seven-point agree–disagree scale is being offered in a mail or Internet version. The resulting question wording effects will add to the measurement error and amplify mode differences.

[Reading 3] Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey by Scott Keeter, Courtney Kennedy, Michael Dimock, Jonathan Best and Peyton Craighill
• More Americans are refusing to participate in random digit dial (RDD) telephone surveys than was the case ➔ primary concern for survey researchers are the consequences of lower response rates for estimates of population characteristics. But measuring relationship between non-response and accuracy of survey statistics is complex and expensive
• Groves (2006) notes that the collective body of empirical work suggests no consistent relationship between response rates and nonresponse. Evidence that response rate is not necessarily an indicator of survey quality…?
BUT survey research environment has been evolving + what was true about the potential for bias may not be the case TODAY

• Smith (1983) demonstrates that extrapolation based on difficult cases has important limitations. He finds that respondents who are difficult to contact or refuse temporarily are not necessarily representative of final non-respondents. Comparisons between the Standard sample and the hardest-to-reach, therefore, only reflect the effect of the extra recruitment effort described above; they are not pure measures of non-response bias
• (1998 experiment: 5 day standard survey v.s. rigorous survey) Despite its higher response rate, the rigorous sample was not closer to the population parameters than the Standard sample in every comparison. It did reach more African-

Americans (12%) and Hispanics (11%) than the Standard survey, and the Rigorous sample was somewhat more reflective of the general public's education, as well as current levels of cigarette smoking. But it was no closer to the population in terms of the income and age distribution and was slightly less accurate than the Standard survey in its estimate of food stamp utilization, health insurance status, and home ownership. The Rigorous protocol also did not overcome a common geographic disparity in telephone surveys caused by lower response rates in urban areas. Counties with higher population density are systematically underrepresented in many RDD surveys. Whereas 40% of interviews should come from the two highest density quintiles of counties, just 34 and 35 percent of completed interviews in the Standard and Rigorous samples, respectively, came from these counties

• Survey non-response includes not just people who were never reached and those who refused to participate but also people who start the survey and then quit before the interview is complete

➔ The most notable trait of these people is their greater tendency to answer "don't know" to the attitude questions and in general to be more disengaged from public affairs

➔ Age 65+, high education or less, less likely to believe most people can be trusted, less likely to believe Islam doesn't encourage violence

1.1 What the course is and is not …

This Course is NOT ABOUT

• how to design questionnaires
• review of the entire survey process
• course based on opinions or feelings on the topic
• course in statistics, although statistical concepts and notation will be used

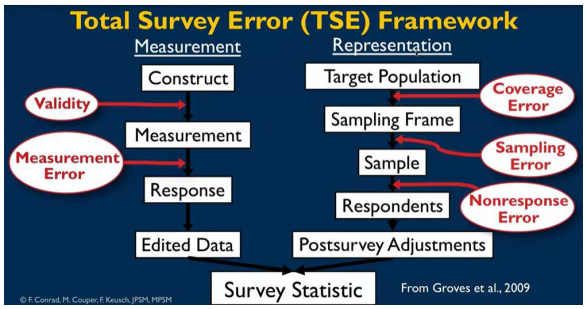
This Course is ABOUT ++ Topics Covered

• Mode of Data Collection
• Mode Decisions and Survey Error: FTF, Telephone, Mail
• Introduction to Non-Response
• Self Administration, Online Data Collection
  ➔ In Interviews (CASI, ACASI)
  ➔ Online (Web)
  ➔ Mixed Mode Surveys
  ➔ Respondent Mode Choice
• Interviewers and Interviewing
  ➔ Obtaining Interviews, selecting respondents
  ➔ Interviewer effects: bias and variance
  ➔ Interviewing Technique
• Emerging modes, new data sources
  ➔ Mobile web and SMS (text message) interviews
  ➔ Administrative records: statistical matching, consent
  ➔ Social Media

1.2 Survey Error

Introduction to Survey Errors

• Data Collection Design
➔ Affects the sources and types of surveys errors that can occur
➔ Affects other constraints, such as costs, time, mix of personnel
➔ Focus of this course: effect of various data collection design on survey error:
  Both variable errors and biases; both errors of observation and non-observation
• Total Survey Error (TSE) Framework

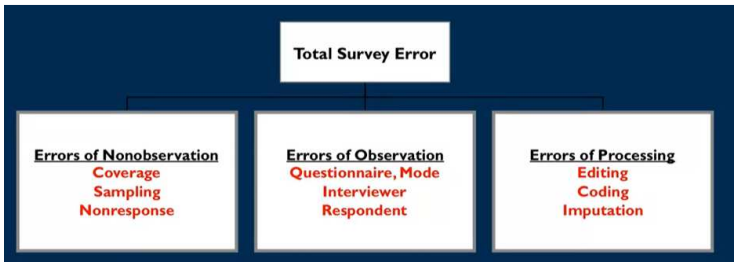


➔ TSE is a way of thinking about the various sources of error that may affect survey stats
➔ The goal of surveys is to make inference (to generalize, to represent) a broader population
- Survey quality is a measure of the success of that effort
- Error not equal to mistake, but rather reflects the uncertainty (or lack of confidence) of the inference
➔ The statistical notion of error is expressed as MSE
- MSE is sum of all variable errors and all biases for particular survey
- Errors are specific to a certain statistic or estimate
- In practice, MSE is rarely fully measured
- TSE approach is evaluation of survey design decisions on >= 1 error source: sampling, coverage, nonresponse measurement etc.

Variable Error and Bias

• True Value
➔ Actual value of what is being measured in the survey
- sometimes knowable e.g. from records, but often not (or we would not do the survey)
- sometimes defined in terms of measurement process e.g. IQ is the score on an IQ test and has predictive validity
➔ measurement errors is the diff between what is observed – an answer to a question – and the true value
• Variable Error
➔ Arises because answers vary over different units in the design (e.g. interviewers, sample people, questions)
➔ Measured by the variance of statistic i.e. responses to a Question
➔ can occur between or within participants
➔ Within participants requires:
- randomization
- essential survey conditions same across data collection waves
• Bias
➔ Directional Error e.g. bigger reports than are actually the case
➔ Concept of "True value" essential to measurement of bias
• Measuring variance and bias
➔ Variance can be estimated from the sample itself, using replication-based methods
➔ To estimate bias requires external data ("truth") or assumptions about direction of effects
• Interviewer bias (measurement error)
➔ Consistent under or overestimation e.g. of discouraged workers, because of failures to probe in looking-for-work Questions (low variance but high bias due to interview)
➔ possible that all interviewers collect similar but incorrect data
• Response variance (measurement error)
➔ Respondent answers same question differently across waves for a longitudinal survey
- for a question that should not change often e.g. person A lives at Address R in Wave I and address Q in Wave2
• Response bias (measurement error)
➔ Overestimation of sample mean for charitable contributions
➔ Can occur for reasons of social desirability, telescoping etc.

Total Survey Error



• Design each component of survey to minimize error inherent to that component
- use mail questionnaires or FTF interviews rather than web to reduce coverage error in survey of general population
- design shorter questions for telephone than mail administration to reduce measurement error; on telephone, Respondents need to remember entire question
• Involves explicit and informed trade-offs
• Success of TSE approach depends on good information on costs and errors (theoretical or empirical, past or current, direct or indirect)
• Survey errors can arise from many sources (survey topic, available funding, sampling frame data collection mode, interviewer training etc.)
• The data collection mode defines (to some extent) the types of levels of error that can occur (Mode chosen can significantly affect quality of survey results)
• Example) To reduce sampling error, want to increase sample size for fixed cost // to reduce non response we might opt for FTF instead of telephone interviews, but more expensive
• In sum, the notion of TSE guides our decisions about design
• Useful framework for understanding the potential impact of design decisions on survey errors
• Together with costs, should be an explicit part of design decisions
• One of the highest impact decisions to be made in implementing a research design
• Requires considerations of all sources of error and costs
• Decisions should be based on theory and empirical evidence

1.3 Survey Mode

What do we mean by mode?

• Combination of medium and agent involved in data collection
- medium = voice, text on screen, text on paper, video…
- agent = interviewer, respondent (self administration: can involve automated interview system)
- possible modes: interviewer voice(phone), self administrated voice(robo call), interviewer paper, self administrated paper…. Etc.
• Until recently, there were only a few main modes of data collection:
- face to face or personal visit surveys
- mail or postal surveys
- telephone (landline) surveys
• In the last few years, explosion of methods, most notably online ("web")
• variants of existing modes:
- mobile phone (as opposed to landline telephone) interviews
- mobile web: browser or app-based questionnaires on smart phone

	Medium	
Agent (Administered by)	*Paper*	*Computer*
Self	Mail-out Questionnaire, SAQ	Web, TDE, IVR,* CASI,* ACASI*
Interviewer	Telephone	CATI
	FTF/Personal Visit	CAPI

*Interviewer administers non-sensitive questions

CATI - Computer Ass't Telephone Interviewing
CAPI - Computer Ass't Personal Interviewing
TDE - Touchtone Data Entry
IVR - Interactive Voice Response (touchtone or speech)
CASI - Computer Assisted Self Interviewing
ACASI - Audio Computer Assisted Self Interviewing

• Some emerging modes
➔ Mobile web (browser or app-based, self administered questionnaires)
➔ Text (SMS)
- self administered (automated) or interviewer administered
- high coverage rates; may be well suited for studies in developing world
➔ Video
- Skype, Google Hangout, FaceTime, BlueJeans etc
- How similar to FTF
➔ Virtual interviewers
- Computer animated interviewing agents
- autonomous v.s. recorded agents
• Mode Proliferations
➔ Need to be more explicit about particular implementation of modes used in study
➔ Harder to make broad generalizations about mode
- Research literature does not (yet) cover all variations
- Need theory to inform expectations about effects
• Increasingly, combinations of modes are being used
- Mixed mode surveys or hybrid designs
• Dimensions of Data Collection Methods
➔ Degree of interviewer involvement (agency)
- Full involvement: CAPI
- Interviewer present, self-administered: CASI, ACASI
- None: Web
➔ Degree of contact with respondent
- Direct: FTF
- Indirect: telephone, mail
- None: direct observation, administrative records
➔ Medium of communication
- Aural only: telephone survey
- aural + visual: FTF (possibly with show cards)
- visual only: mail survey
• Two approaches to data collection method/mode
➔ Selection a mode: what is the most appropriate mode?
- constrained by resources, time, frame, topic etc.
- e.g. for a survey to explore the correlates of poverty, what mode is appropriate?
➔ Consequences of decision: what is the impact of a particular method of data collection on survey errors and costs?
- Given a particular mode, what effects to expect, and how to overcome them
- How to maximize the benefits of a particular mode
- e.g. given that one has decided on a web survey, what decisions should be made regarding the design of the questionnaire?
• Considerations when selecting a mode
➔ Sometimes the choice is dictated by the topic (e.g. what mode is needed for a literacy survey?)
➔ Sometimes the choice is driven by the frame (e.g. how do you survey voters on election day?)
➔ Sometimes the decisions is driven by cost, timeliness, resources etc. (e.g. not enough resources for interviewers, so self administered mode is best…)

More Choice (by respondent)
• Mode Choice by respondent (response rate down)
➔ Medway and Fulton(2012): **mode choice reduces participation**
Response rate was reliably lower overall in choice (mail or web) than mail only condition
Paradox of choice or costs of mode switching: Ironically, if you have more choices, you are paralyzed and become more satisfactory about the choice you make
• Mode choice by respondent (response quality up)
➔ Conrad (2013)
- Can choose 1 out of 4 interview modes (human voice, automated voice, human text, automated text)

- 50% assigned to a mode; remaining 50% given choice to choose mode
- Higher quality data from choice group: fewer rounded numerical responses, less straightlining (fewer breakoffs and higher completion); higher satisfaction from respondents who chose mode themselves; automated text chosen the most often
- **Choice reduce response rate** but only when switching from automated to human administered modes; switches in this direction unlikely to incur a delay; choosing a mode on single device minimizes mode switching costs

• Mode Preference (response rate down)

➔ Olson (2012): asked respondents preferred mode in telephone survey
- year later, same respondents were invited to participate in another survey to mail or web modes (or a sequence of mail-web or web-mail); some respondents answered in preferred mode and some did not
- **response rates higher if interviewed in preferred mode**

• Mode Preference (response quality up)

➔ Smith (2014)
- Examined (Olson 2012 data set) amount of "satisficing" (mental shortcuts) when responding in preferred and non-preferred mode
- Respondents randomly assigned to one of two questionnaire forms in which item formats differed
- Check all that apply v.s. Forced Choice
- Big text box v.s. small text box for open response

➔ Respondents not responding in preferred mode, endorsed fewer options in check-all-that-apply than forced choice and skipped open format question more with larger than smaller answer box // Respondents in preferred mode, no effect of format

Mode Summary

• Mode: combination of medium and agent (classic modes continue to be used but many new modes in use and many on the horizon)

• Many considerations when selecting mode for project or designing project around mode (topic, frame, resources, measurement goals)

• mode choice (for respondent) may reduce participation due to switching costs but can improve measurement

➔ responding in preferred mode – without having to choose it – increased participation and response quality

1.4 Mixed Mode Surveys

Mixed Mode Design

• Increasingly common to combine modes

➔ optimize trade-off between different costs and errors

➔ Increase RRs (goal to reduce non-response error)

➔ concern is creation of "mode effects"

• Different modes might not only be used for data collection but also for communicating with Respondents in different phases of study (e.g. pre notification, recruitment, reminders)

• Pre-notification

➔ Use different modes in pre-notification and data collection
- Pre notification alerts sample member to upcoming survey invitation (not actually an invitation)

➔ Examples
- telephone interviewing overall with ftf recruitment/pre-notification for some areas
- mail recruitment for web-only survey

➔ goals
- increase response rate by increasing study's salience
- no mode effects if data collection is unimodal

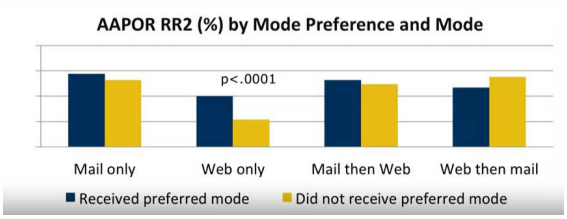
➔ Bosnjak (2008) test impact of web survey Response Rates of three pre-notification (SNS, email, none) and two invitation (SMS, email) modes
- SMS pre-notification most effective
- Email invitation most effective
- SMS pre-notification + email invitation most successful combination

• Data Collection Phase

➔ Concurrent mixed mode design: One sample, one time point, one questionnaire but different sample people receive different modes; choice

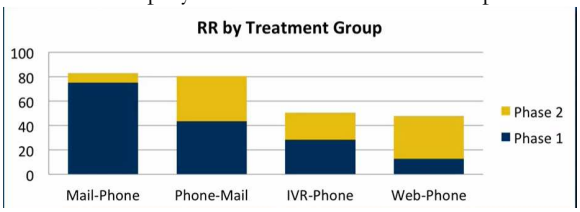
➔ sequential mixed mode: one sample, one time point, sample members recruited in increasingly effective and expensive – modes if previous contact attempts not successful
➔ switching modes within questionnaire: one sample, one time point; different modes for different parts of questionnaire for the same respondents
➔ longitudinal mixed mode: one sample, multiple time points; same people measured with different modes at different times
➔ parallel (separate) mixed mode: different samples, different modes combined altogether end the end (e.g. multinational survey)

Concurrent Mixed Mode
• Different modes for different subgroups of sample
  - e.g. web for those with internet access, mail for others
  - benefit: reduce coverage and non-response error
  - down side: mode effects on measurement confounded with subgroups
• Preference for Data Collection Mode
Effect especially strong for WEB ONLY



• Mode choice and non-response
➔ preferring a mode is not the same as choosing a mode
➔ fulton and medway(2012)
  - overall lower response rate for CHOICE
  - increases complexity making both options less attractive
  - creates break in response process reducing number of starts
  - may push respondents to web where connectivity problems deter start
• Mode choice and measurement error
➔ Conrad (2013): choice – especially when easy to implement – especially if Rs choose a convenient mode – should improve performance
  - contacted Rs on smartphones in one of 4 modes: human voice, human text, automated voice, automated text
  - required Rs to choose one of those modes, including contact mode
  - result: improved quality (less rounding, fewer breakoffs, greater satisfaction than an assigned mode)

Sequential (Follow up) Mixed Mode
• Goal1: reduce non response
  - use one mode as reminder for another mode
  - e.g. contact mail nonrespondents by phone to remind them to fill out paper questionnaire
• Goal2: reduce cost of non response follow-up
  - use increasingly more expensive modes for reluctant respondents
  - e.g. first mail, then telephone, then face-to-face
  - risk of mode effects on measurement
• Dillman (2009)
➔ followed up by Phone or Mail with non respondents in mail, phone, IVR and Web invitations



A bit unusual / not typical is the initial response rate (phase 1) is higher for Mail than phone (usually it's the other way around) ➔ think it can be attributed to the fact that mail had $2 gift card included already while that $2 reward was promised over the phone in the phone survey, so less incentive to respond

➔ Findings
- Offering modes sequentially can improve RR
- But does little to resolve demographic bias due to non-response
- mixing modes can mix mode effects
Aural mode (phone IVR) Rs tend to give more positive responses on bipolar scales than visual (mail, Web) modes

• Millar & Dillman (2011)
➔ In two experiments, concurrent choice (Mail or Web) leads to same RR as Mail only
➔ But web (or choice) then email follow-up (hence sequential) with link led to RR comparable to mail

• Mixed mode follow-up with initial incentive
➔ Beebe (2005) mailed questionnaires to a sample of Medicaid recipients; 50% received $2 incentive
- RR with incentive = 54%; no incentive = 45%
➔ telephone follow-up for minority/ethnic non respondents increased RR
- RR with incentive = 59%; no incentive = 64%
➔ Effect of initial incentive reduced in follow-up
➔ Overall, increased RR with telephone follow-up offset cost of incentive

• switching mode within questionnaire
➔ one sample, one time point but different modes for different parts of questionnaire for same Rs
- use different modes to collect different types of data from the same Rs
- improve data quality by increasing privacy of measurement
➔ example: FTF + audio-CASI

• Longitudinal Mixed Mode
➔ one sample, multiple time points but same people measured with different modes at different times
- use one mode for start of panel survey, another mode for later waves
- reduce costs, if start with more expensive mode
- confounds time and mode effects on measurement
➔ example: PSID and CPS (use ftf for initial wave, mostly telephone afterwards)

• Parallel (Separate samples) mixed mode
➔ mainly used for comparative studies to accommodate different regional survey traditions or practical constraints (coverage, literacy etc)
- sometimes even different time and questionnaires
- mode effects may cause differences by country, state, or region, although many differences beside mode
➔ examples: ISSP, ESS, BRFSS – different modes in different countries or states

Interview with David Weir (U. Michigan) on Mixed Mode Designs

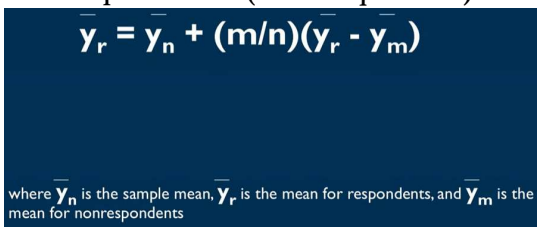
1.5 Non Response

Response Rates
• Response Rates very across modes
• Generally, FTF > Telephone > Mail > Web
- But many exceptions in the literature
- Affected by pre-notification, incentives
• Response Rates NOT EQUAL to Non response error
• No longer considered a measure, by itself, of survey quality
• But is a starting point for estimating non response error
• Multiple ways to calculate RR
• Two most widely used at AAPOR(American Association of Public Opinion Research) RRI and AAPOR RR2

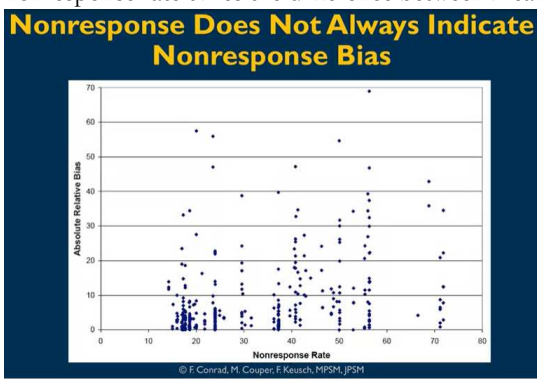
AAPOR Dispositions and RR calculations
• RR = Response Rate
• I = Complete Interview
• P = Partial Interview
• R = Refusal and break-off (agree & start survey / interview but terminate after some point)
• NC = Non-contact
• O = Other
• UH = Unknown if household/occupied
• UO = Unknown, other

• AAPOR RESPONSE RATES CALCULATION
RR1 = I / ((I+P) + (R+NC+O) + (UH+UO))
- Response Rate 1 (RR1), is the number of complete interviews divided by the number of interviews (complete plus partial) plus the number of non interviews (refusal and break-offs plus non-contacts plus others) plus all cases of unknown eligibility (unknown if housing unit, plus unknown, other)
• RR2 is similar to RR1, but includes partials in the numerator, i.e. if an interview is started, it counts
• RR2 generally > RR1 because considers includes completes and partials

Non Response Error (for each question!!)



• The mean for the respondents is a function of the sample mean (for a question) plus non response error, i.e. the nonresponse rate times the difference between means for respondents and non respondents



• Why does non-response sometimes produce bias and sometimes not (like the scatter plot above)?
➔ Sometimes likelihood of being contacted for survey or likelihood of cooperating given contact is related to topic in question, other times not
- e.g. Abraham, Helms & Presser (2009) found that estimates of volunteering are biased by higher response rates among those who volunteer
➔ RR is estimated for entire case (interview); non-response bias estimated per question
• Selecting a Mode: Non Response Error
➔ Modes differ in their response rates
- FTF interviews > Telephone > mail
- Low response rates do not guarantee non-response error
- e.g. Keeter et al. (2006); Groves (2006)
• But they do increase vulnerability of estimates to non-response error
• Nonresponse bias in Pre-election Polling (example)

White Rs	No Refusal	Refusal Conversion*
	%	%
Very Favorable opinion of Blacks	24	15
Blacks cannot get ahead because of racial discrimination	33	21
Blacks are responsible	54	70
Favor Proposed National Apology for Slavery	39	28

* initial nonrespondents

➔ Presumably a mode that increases response rate e.g. FTF, would include more of those who initially refuse a telephone interview, reducing non-response bias
- or additional monetary incentive or more active non-response follow-up conversion