



Evaluation of Sample Designs for Telephone Surveys That Include Cell Phones

Charlotte Steeh

Contractor, Centers for
Disease Control and
Prevention

Background

- Worldwide major telecommunications transition.
- Communication is changing from a group based system centered on a precise location to a individual based system with no fixed location .
- Excruciating dilemma for telephone surveys.
- Struggle to preserve the accuracy and validity of data from telephone surveys.
- Must design general population samples that measure all telephone owning groups in their proper proportions.

The Challenge

- Difficult to include cellular and landline numbers in the same sample design.
- Overcoverage of dual users (those with both a cell phone and a landline phone).
- Weighting becomes complex.
- Several different sample designs try to eliminate complexities.
- Current research has not evaluated sample designs but has focused instead on determining the demographic uniqueness of residents of cellular-only households.

Types of Sample Designs

- Report from the AAPOR Cell Phone Task Force (2008) recognizes four designs that in some way take account of the overlap of dual users in a combined sample.
- Two are dual frame designs that actually include cellular numbers.
- Two are single frame designs that rely solely on weighting adjustments and do not include numbers from the cellular frame.

Dual Frame Designs

1. Independent Samples with Screening

Numbers are drawn from both frames but one sample is screened so that all units have nonzero probabilities of being selected from one and only one frame.

2. Independent samples without Screening

Numbers are selected from both frames without screening.

Requires a mixing parameter in the weights to take account of the overlap.

Single Frame Designs

1. RDD landline samples

Weights adjust for the segment of the target population without landline telephones.

2. RDD cell samples

Weights adjust for the segment of the target population without cellular telephones.

All four designs require post-stratification weights that reflect the correct distribution of telephone ownership in the general population.

Viability: Single Frame Designs

- RDD landline sample with standard weights is still accurate because cell-only households are only a small percentage of telephone households (Steeh 2003; Keeter 2005; Kennedy 2007; Keeter 2008)
- RDD cell phone sample appears too unrepresentative to be used alone. Landline-only percentage is larger than the cell-only percentage. Little research on this and percentages may be shifting.

Viability: Dual Frame Designs

- Without Screening

- Several studies have examined the accuracy of samples drawn using the dual frame design without screening (Brick et al. 2007a; Kennedy 2007).

- With Screening

- Other articles have looked at samples selected by the dual frame design with screening (Brick 2007; Fleeman 2007; Keeter 2006; Link et al. 2007).

Summary

- *There is, however, no consensus yet on which approach is the “best” design (AAPOR Task Force Report, p. 14).*

Evaluation Method One: Determine Bias

- Comparison with an independent gold standard from:
 - the U.S. Census,
 - the Current Population Survey, or
 - any other scientific survey with large sample sizes
- Data that will serve as a gold standard are:
 - Rare
 - Very difficult to obtain at the sub-national level
 - Nonexistent for attitudes

Evaluation Method Two: Compare Estimates

- Calculate differences between weighted survey estimates including and excluding the group screened out--for example, dual user households from the cellular frame.
- Kennedy (2007) compares differences of each sample from a benchmark estimate rather than from each other.
- Requires fully interviewing the samples from both frames.
- Needs complex sampling weights

Evaluation Method Three: Test for Significant Differences

- Test for significant differences between unweighted estimates calculated from each frame separately.
- Requires that interviews be conducted with the entire sample from each frame.
- Allows a wider range of variables to be tested including behaviors and attitudes.
- Sample sizes must be large.
- Will give us some idea about the variables needed for post-stratification weights.



Direction of Current Research

- Has not evaluated sample designs but has focused instead on the demographic uniqueness of residents of cellular-only households.
- Assumes that dual frame design with screening avoids many of the pitfalls involved in combining cellular and landline numbers in telephone samples.
- Has become a popular sample design.

Dual Frame Design with Screening

- Seems the simplest way to resolve problems.
- Basic frame is the landline frame supplemented with cell-only households screened from the cellular frame.
- Dual users are represented only by respondents contacted through the landline frame.
- The sample is weighted by the distribution of telephone ownership in the target population, the probabilities of selection, and post-stratification variables.

Possible Weaknesses

- Assumption that dual users are adequately represented by respondents obtained from the landline frame.
- Dual users in the cellular frame may be different from dual users in the landline frame.
- Differences can result from two sources:
 - Mode Effects
 - Nonresponse Bias
- This report focuses on the latter source but also addresses the first source indirectly.

Reasons for Differences

- Dual users contacted through the cellular frame may use their cell phones more often and thus be more likely to answer a call made to a cell phone.
- Dual users contacted through the landline frame may use their landline phones more often and thus be more likely to answer a call made to a landline phone.
- Differences in telephone behavior may indicate differences in other behaviors and in attitudes.



Hypothesis to Be Tested

- Residents of households with both types of telephone will show significant differences on a number of dimensions depending on how they are contacted—by cell phone or by landline.

Prior Empirical Evidence

- Studies have revealed substantial and statistically significant differences in telephone behavior depending on how dual users were interviewed (Steeh 2004; Brick et al. 2006).
- Direct comparison of dual users contacted by cell phone and by landline has yielded mixed results in other areas.
- An early PEW survey found differences between the two groups on political and social attitudes (Keeter and Kennedy 2006).
- Later analyses drew the opposite conclusion (Kennedy 2007; Keeter 2008).

Data

- 2003 National Mode Comparison Study conducted at Georgia State University
- Designed to test whether or not results from cellular phone surveys would differ significantly from results obtained in landline surveys.
- Questionnaire covered many different topics. (For more details about the study, see Callegaro et al. 2007 or Steeh and Piekarski, 2007.)

Table 1: 2003 Mode Comparison Study

	Cell	Landline
Response Rates (RR2)	21%	33%
Contact Rates	75%	78%
Refusal Rates	45%	38%
Refusal Conversions (% of Refusals)	8%	14%
Nonsample Rate	50%	53%
Number of Interviews	774	541
Length of Interview (minutes)	18	17
Length of Field Period (days)	181	63
Average Number of Calls	6	10
Total Number of Dialings	47,914	28,620

Corollary to Hypothesis

- The richness of the 2003 Mode Comparison Study allows us to explore differences along four dimensions.
 - telephone use
 - demographic characteristics
 - behavior
 - attitudes
- Predict that significant differences will be found across all of these dimensions.

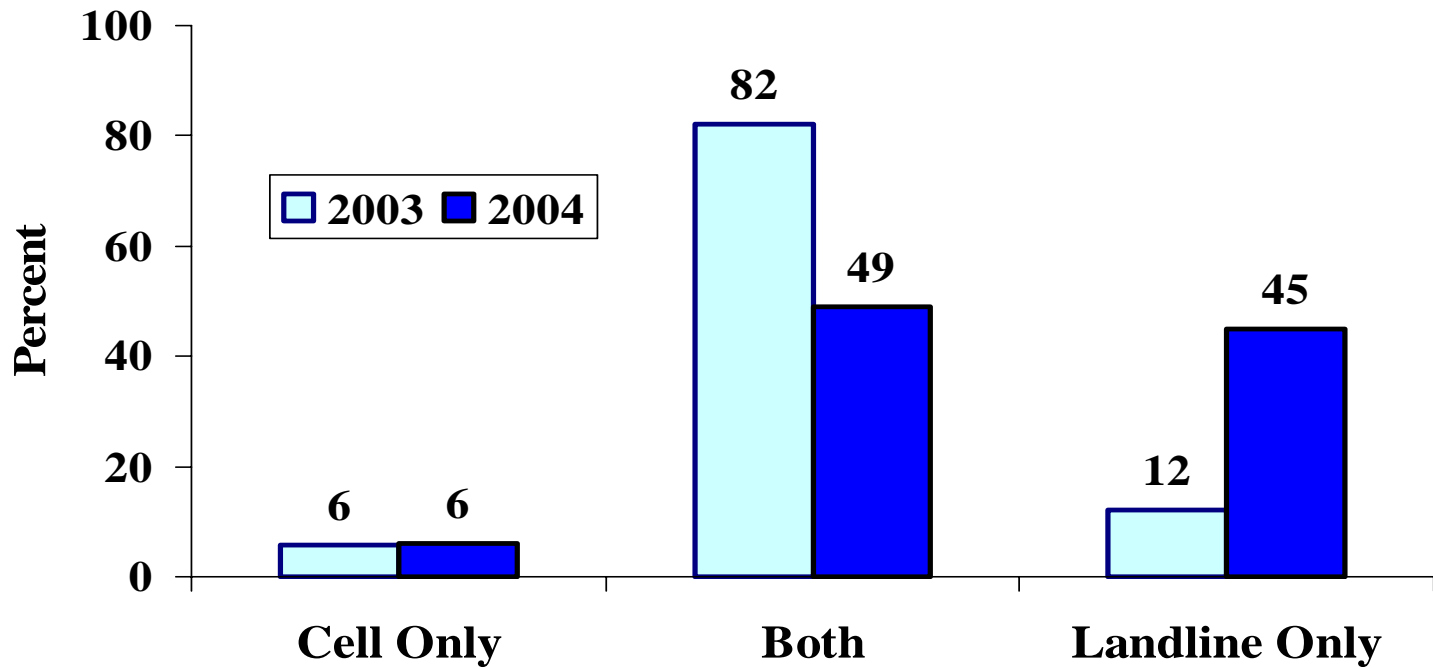


Figure 1: Distribution by telephone ownership, 2003 Mode Comparison Study versus 2004 CPS

Analysis Strategy

- Test for significant differences between unweighted estimates (Evaluation Method 3) calculated from each frame separately for dual users only.
 1. Dual users interviewed in the cellular survey, designated **CL**
 2. Dual users interviewed in the landline survey, designated as **LC**
- Emphasis on dual users gives research on cell phones new focus .

Method

- Calculate the percentage point difference between respondents from two surveys:

$$\Delta = CL\% - LC\%$$

- Have measures for **42 variables**.
- The cell phone survey contributes 688 dual users to the comparison and the landline survey contributes 386.

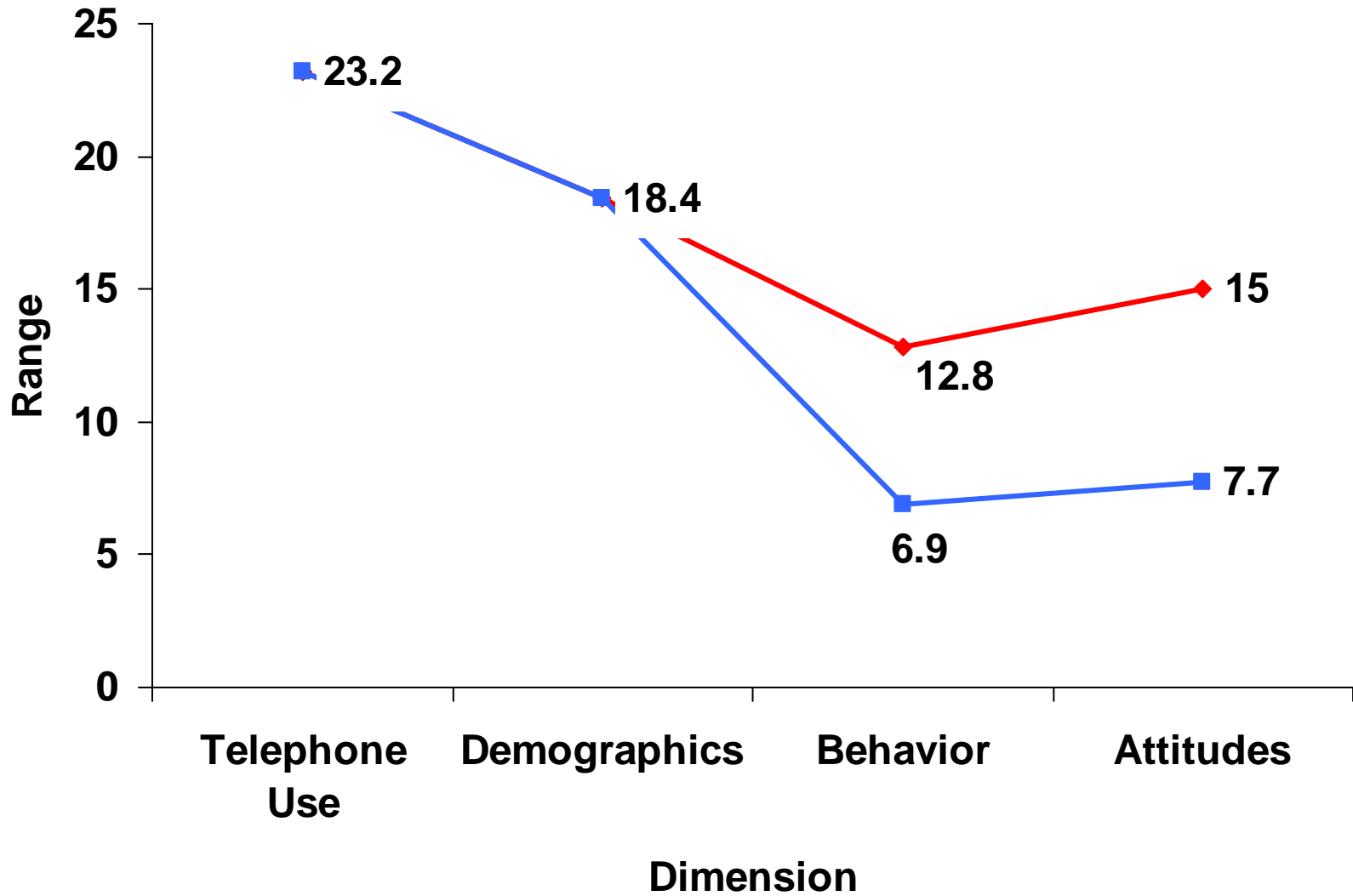
Table 2: Mean Difference by Item Category

Category	Number of Items	Average Absolute Difference	Number Significant beyond $p=.05$	Range of Differences in Percentage Points
Telephone Use	12	9.5	9	23.2
Demographics	11	4.7	5	18.4
Behavior	8	2.9	1	12.8
Attitudes	11	3.0	1	15.0

Outliers

- Only one significant difference for each of last two dimensions--behaviors and attitudes
- In testing multiple hypotheses, one result may be significant purely by chance.
- Dropping the outlier for each of the two categories reduces their ranges considerably.

Figure 1: Range of Differences by Item Category




Discussion


- Neither the primary hypothesis nor its corollary is confirmed.
- Significant differences in how LC and CL households use their telephones and in some demographic characteristics.
- Almost no significant differences in behaviors, or attitudes. One exception in each case is outlier.
- Similarity between LC and the CL respondents suggests the absence of mode effects proper.



Conclusions, Recommendations, and Warnings

- Cautious optimism about the viability of dual frame samples with screening.
- How important is it to include CLs in telephone samples?
- May not be critical at this time unless the research topic is telephone behavior.

- 
- In making the decision about which sample design to use, researchers should consider several factors.
 - purpose of the survey
 - definitions of the target population
 - estimates of cost.
 - The dual frame sample with screening would not be appropriate if the survey focus is measuring telephone use and behavior or measuring any form of telecommunications.

- 
- If the purpose is to measure attitudes and opinions, the dual frame sample with screening may be appropriate.
 - At present this design is costly because of the amount of screening.
 - As the cell phone only population grows, screening costs will diminish.
 - There are ways to ease the financial burden.
 - Screening during day time hours
 - Sending text messages to identify nonworking numbers

Another Finding

- For single frame sample designs, may need to redesign post stratification weighting procedures to include new variables, such as:
 - Household composition
 - Income
 - Extent of cell phone use
- Benchmarks for these kinds of variables are often not easily available.
- Help to keep the dual users contacted in landline surveys representative of all dual users.



Limitations of the Research

- Cell-only population nationally was three percent when surveys were conducted.
- Basic findings have been corroborated in subsequent analyses conducted after the cell-only group has grown substantially.
- Need more research on other behavioral and attitudinal measures.
- These analyses do not evaluate changes in bias as a result of using different weighting methods.



Thank you very much.