

Effects of Survey Data Collection Mode on Response Quality:

Implications for Mixing Modes in Cross-National Studies

Jon A. Krosnick
Stanford University

September 2005

Copyright Krosnick

American National Election Study

- Face-to-face national surveys since 1948.
- Sharply rising costs.
- Consumed an increasingly large portion of the NSF Political Science.
- Recently declining response rates.
- Pressure from the community to switch to telephone to save money.

American National Election Study

- Shrinking grants from the National Science Foundation.
- Coped by reducing sample size.
- Community dissatisfaction with shrinking sample size.
- Conduct experiments with telephone.

American National Election Study

- NES commissioned an outside mode review committee.
- Conclusion: Changing Mode Ends the Time Series
- NSF commissioned an advisory committee.
- Conclusion: NES is an under-funded gold standard of quality – protect and nurture it.

American National Election Study

- Net Result:
Nearly tripling of the budget for the next 4 years.
- Maintain Super-Quality Face-to-Face
- Fieldwork: RTI, International.
- Minimum guaranteed response rate: 70%

General Social Survey

- Similar national face-to-face study for decades.
- Similar experiences with declining response rates.
- Similarly shrinking grants.
- Coped by selling questionnaire modules.
- Coped by double sampling in 2004.
- Coped by mixing modes: 16% telephone in 2004.

Reasonable Solutions for Increasing Costs

- Switch modes between waves
 - Shift from face-to-face to telephone
 - Shift from telephone or mail to Internet

Reasonable Solutions for Decreasing Response Rates

- Mix modes
 - Begin with face-to-face, end with telephone
 - Begin with telephone, end with face-to-face
 - Begin with mail, end with telephone

Even if cost is not an issue,

exhausting case recruitment with one mode and then adding a new mode increases the response rate.

So if high response rate is a goal, mixing modes is desirable.

Costs of Switching Modes Between or Within Studies?

- Is the response process different?
- Does accuracy change?
- Does the nature of response bias change?
- Do you get what you pay for?

- If reporting accuracy declines, is it offset by an improvement in sample representativeness ?

Realistically

- Increasing a response rate from 60% to 100% has the potential to change representativeness substantially and is guaranteed to eliminate unit non-response error.
- BUT ...
- Increasing a response rate from 60% to 68% might even *decrease* sample representativeness.
- That's the most we can hope for with mode mixing.

So what does mode switching do to response quality?

A Starting Point:
Some Stereotypes of Modes

- Face-to-face
 - Increasingly super-expensive
 - Very long field periods
 - Powerful social desirability pressures
- Telephone
 - Declining response rates
 - Increasing barriers to contact (call blocking)
 - Collapsing sampling frame (cell phones)
- Mailed Self-Administered
 - Low response rates
 - Slow turnaround
- Internet
 - Non-probability samples

Data Collection Mode

Face to face

Telephone

Paper and pencil

Computer/Internet

- Rapport and trust
- Confidentiality
- Modeling of commitment
- Accountability
- Pace
 - Discomfort with silence
 - One screen at a time
- Time of day
- Literacy
- Working memory burden

Two kinds of studies

- Comparisons of modes confounded with sampling methods:
 - E.g., telephone = RDD
 - Decide which mode to use first
- Comparisons of modes unconfounded with sampling methods:
 - E.g., random assignment experiments

Comparing Face-to-Face vs. Telephone

(Mode confounded with sampling method)

(Control for Sample Composition Differences in Analysis)

Table 1. Summary of Previous Studies Comparing Face-to-Face and Telephone Interviewing

Publication	The Same Respondents Were Interviewed Face-to-Face and by Telephone	Telephone and Face-to-Face Samples Were of Different Populations	Some People Assigned to a Mode Were Not Interviewed in That Mode	Face-to-Face Respondents Interviewed in Groups and Telephone Respondents Interviewed Individually	Respondents Were Given the Choice to Be Interviewed Face-to-Face or by Telephone	Respondents in One or Both Modes Were Interviewed Previously	Very Different Questionnaires	Questionnaire Not Described Adequately	Not an RDD Telephone Survey vs. an Area Probability Face-to-Face Survey	Not National Samples
Confounded mode comparisons:										
Henson, Roth, and Cannell 1978	X								X	
Herzog and Rodgers 1988	X								X	
Midanik, Rogers, and Greenfield 2001	X								X	
Rogers 1976	X								X	
Schmiedeskamp 1962	X								X	
Cahalan 1960	X								X	X
Larsen 1952		X							X	X
Siemiatycki 1979			X				X		X	
Mangione, Hingson, and Barrett 1982			X						X	
Hochstim 1962, 1967			X						X	
Herman 1977			X						X	X
Rosenstone, Petrella, and Kinder 1993			X						X	
Thornberry 1987				X						
Cannell, Groves, and Miller 1981				X						
Calsyn, Roades, and Calsyn 1992					X				X	
Yaffe et al. 1978						X				
Morchio, Sanchez, and Traugott 1985							X		X	
Essaiasson and Granberg 1993							X		X	
Herzog and Rodgers 1999							X		X	
Wolman, Turner, and Bushery 1980							X		X	
Aquilino 1992								X		

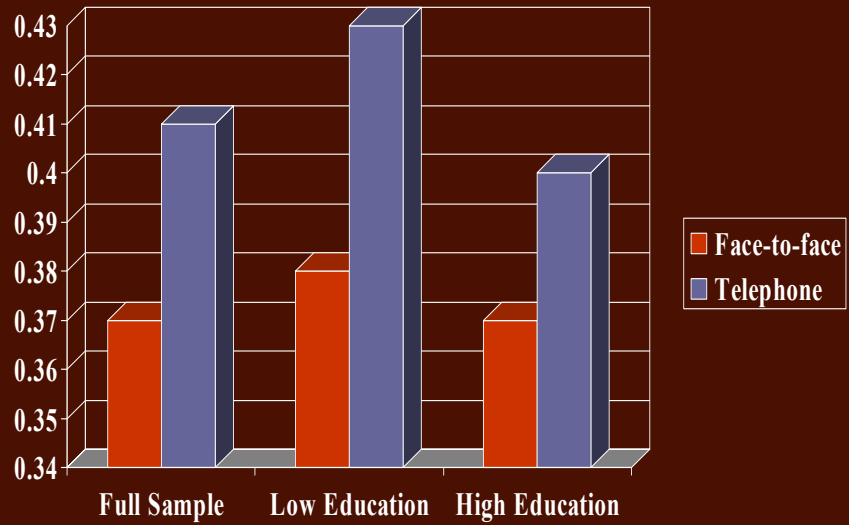
Comparing Face-to-Face vs. Telephone

- 1976 SRC Study (RR=72% vs. 62%)
- 1982 National Election Study Methods Comparison Project (RR=74 vs. 65%)
- 2000 National Election Study Experiment (RR=64% vs. 57%)

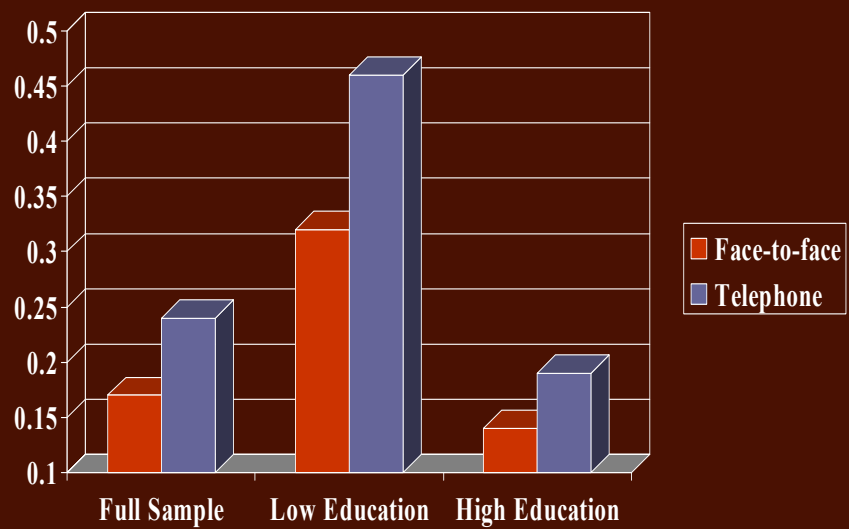
General Public Samples – 1 hour questionnaire

Satisficing

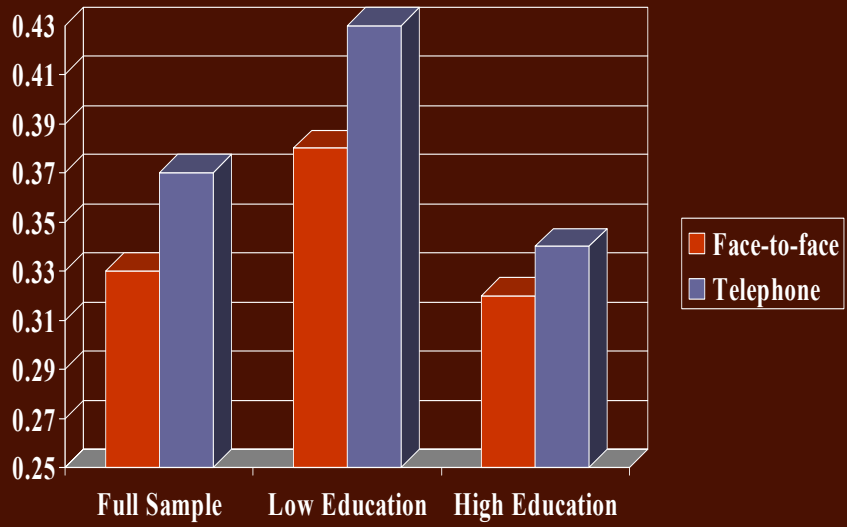
Non-Differentiation



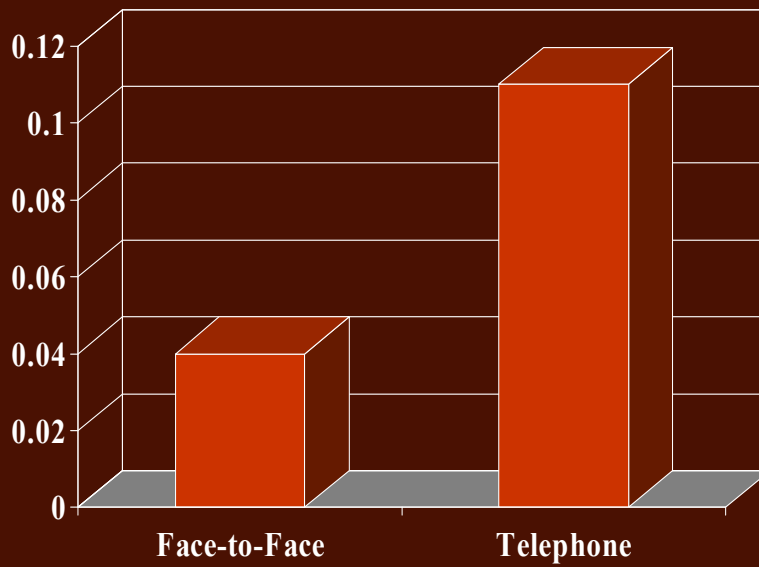
No-Opinion Responses



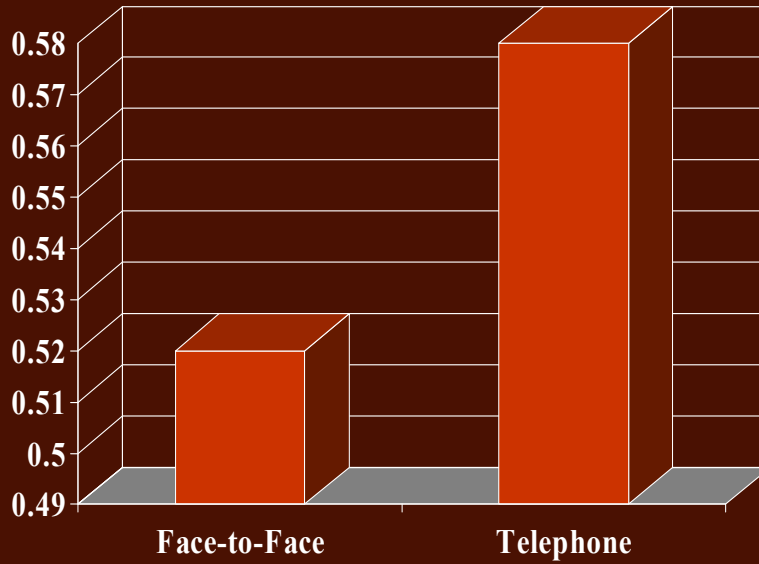
Acquiescence



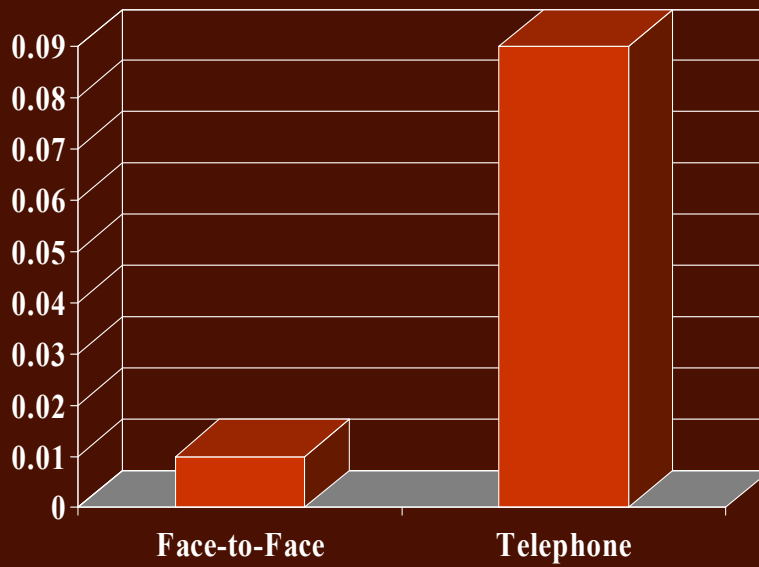
Asked How Much Longer the Interview Will Be

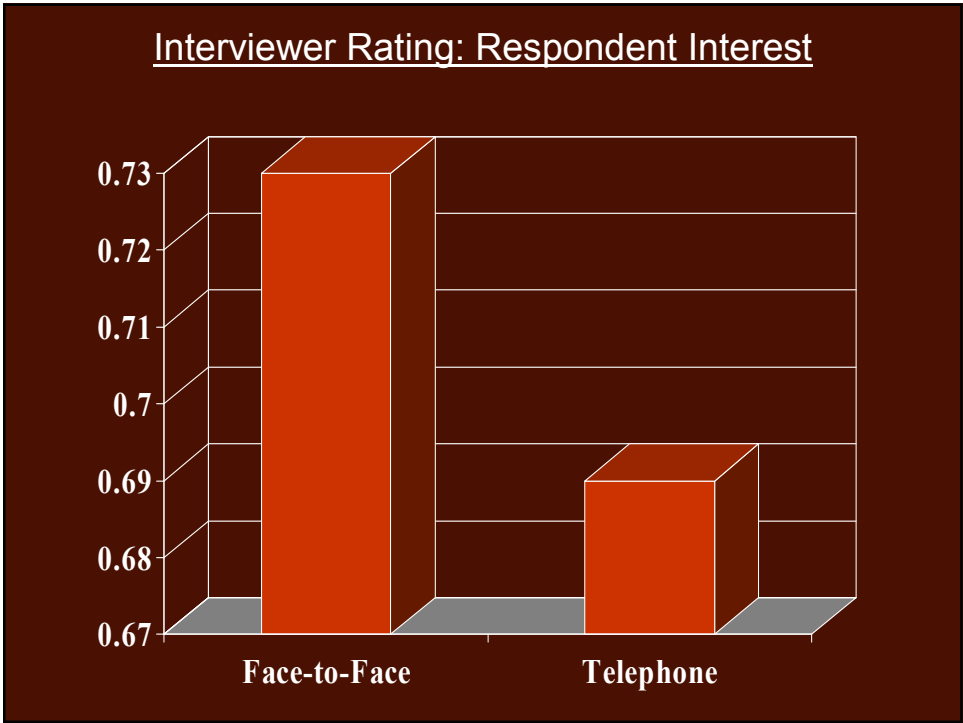
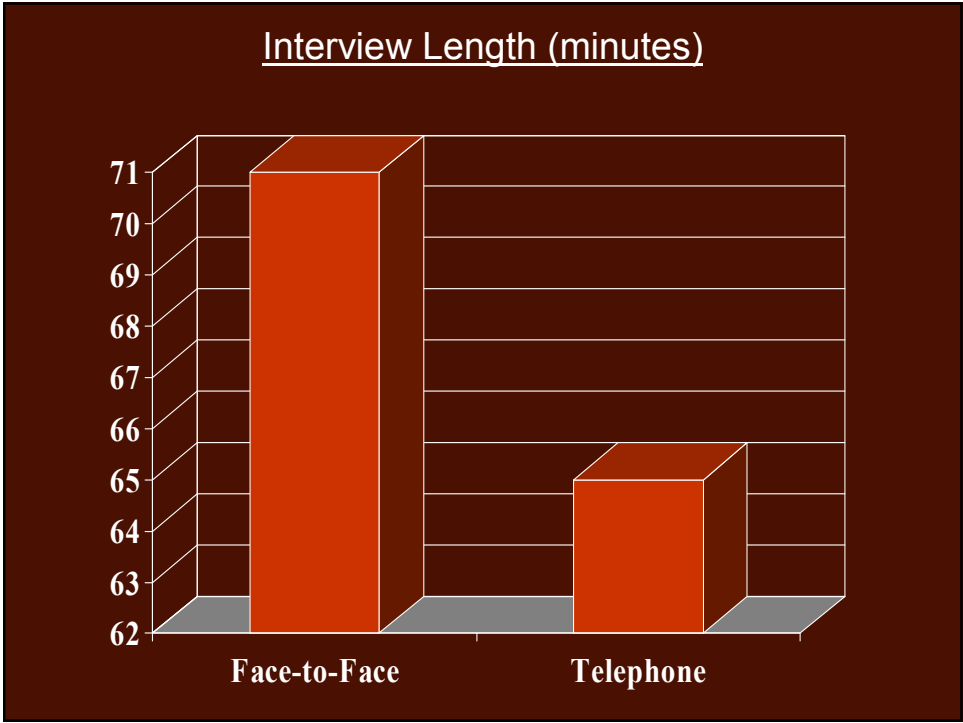


Expressed Dissatisfaction with Interview Length

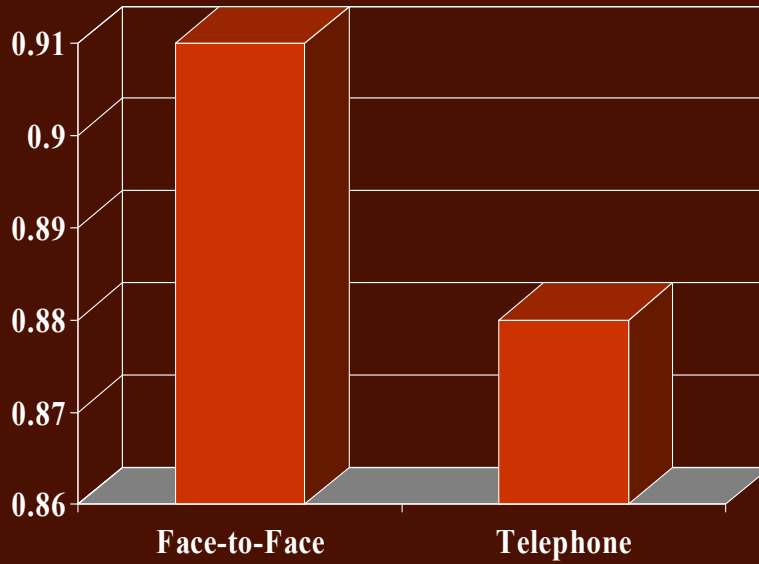


Said Interview Was Too Long



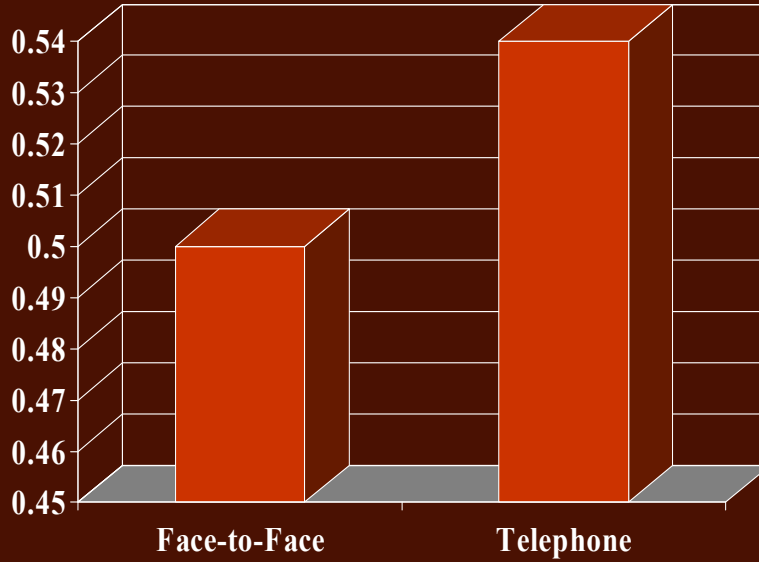


Interviewer Rating: Respondent Cooperation

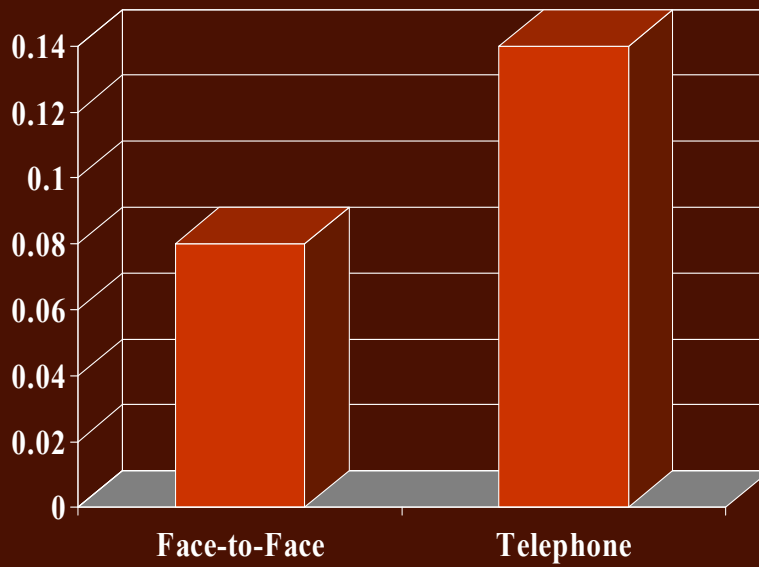


Social Desirability Response Bias

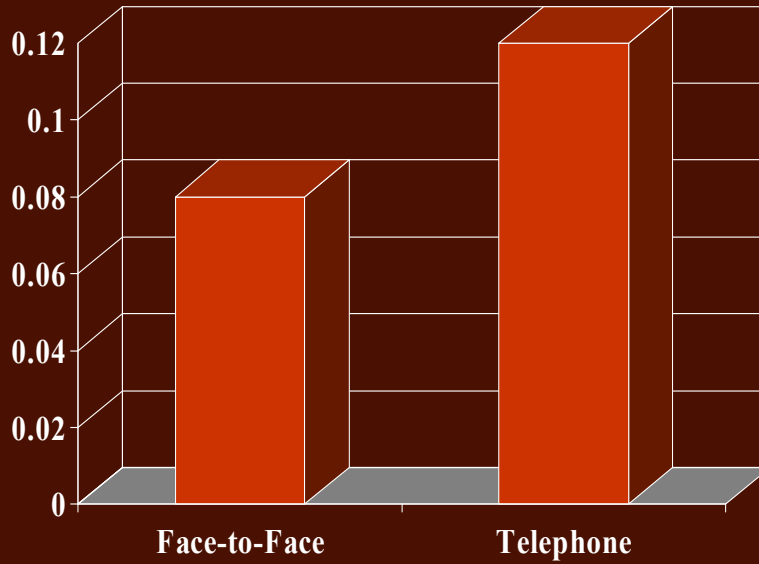
Socially Desirable Responding



Unease Discussing Sensitive Topics



Interviewer Rating: Respondent Suspicion



Telephone vs. Paper and Pencil

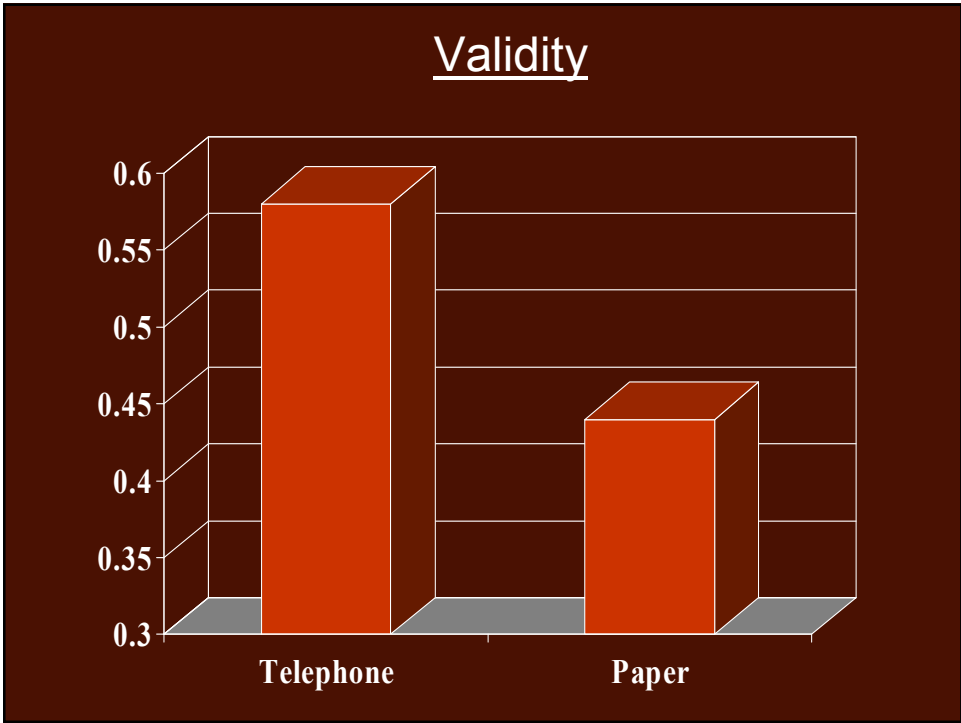
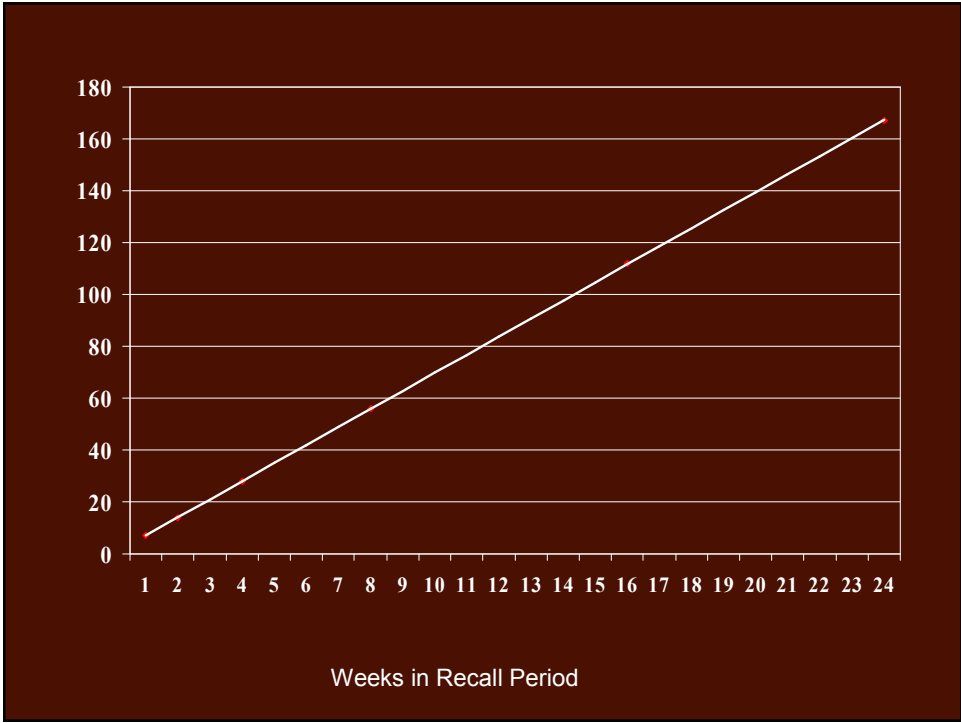
(Mode Unconfounded with Sampling)

National Aviation Operations Monitoring System

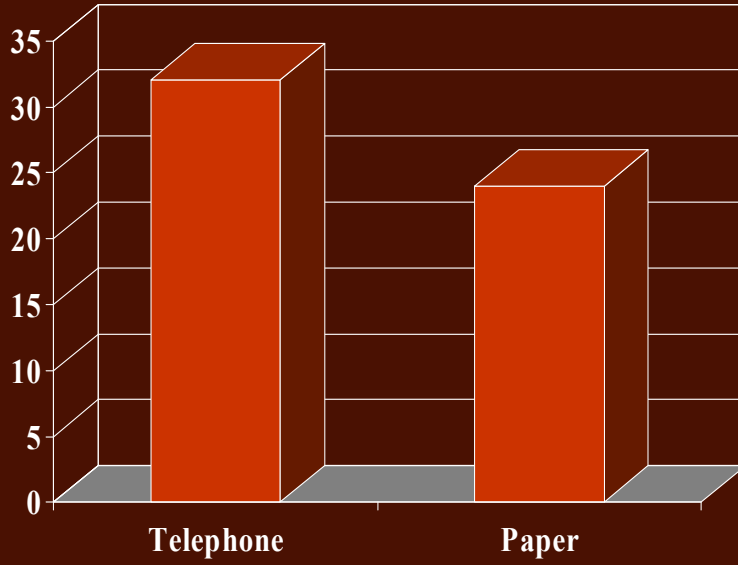
- Field Experiment
- Licensed Pilots
- Random assignment
 - Telephone (RR=73%)
 - Self-administered mailed questionnaires (RR=70%)
- Questionnaire
 - Measure number of safety-related events witnessed during a specified recall period.
 - FACTUAL QUESTIONS

Validity Measurement

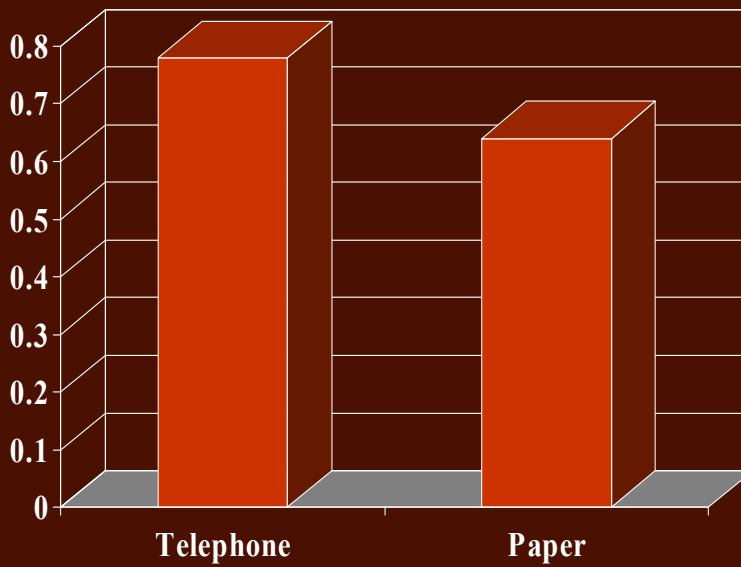
- Random assignment to recall period
 - 1 week
 - 2 weeks
 - 4 weeks
 - 8 weeks
 - 16 weeks
 - 24 weeks
- As recall period increases, so should number of events reported.



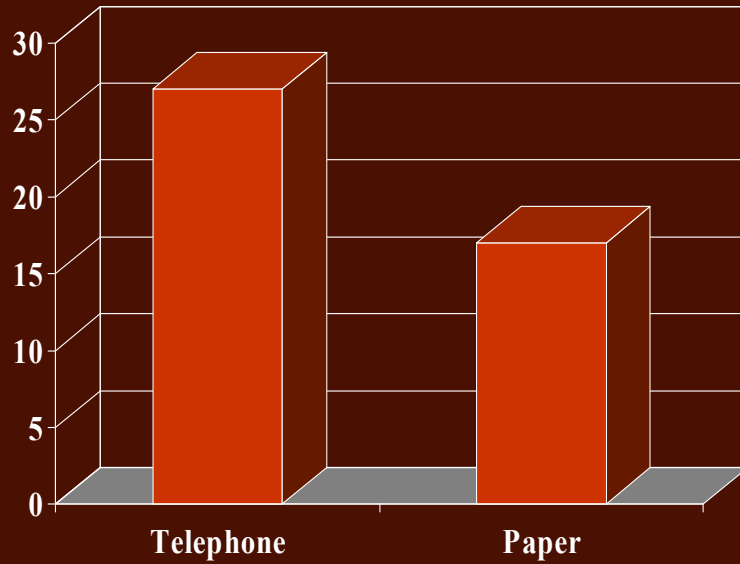
Total Number of Events Recalled



Perceived Accuracy



Completion Time



Telephone vs. the Internet

(Both: Mode Confounded AND Unconfounded with Sampling Method)

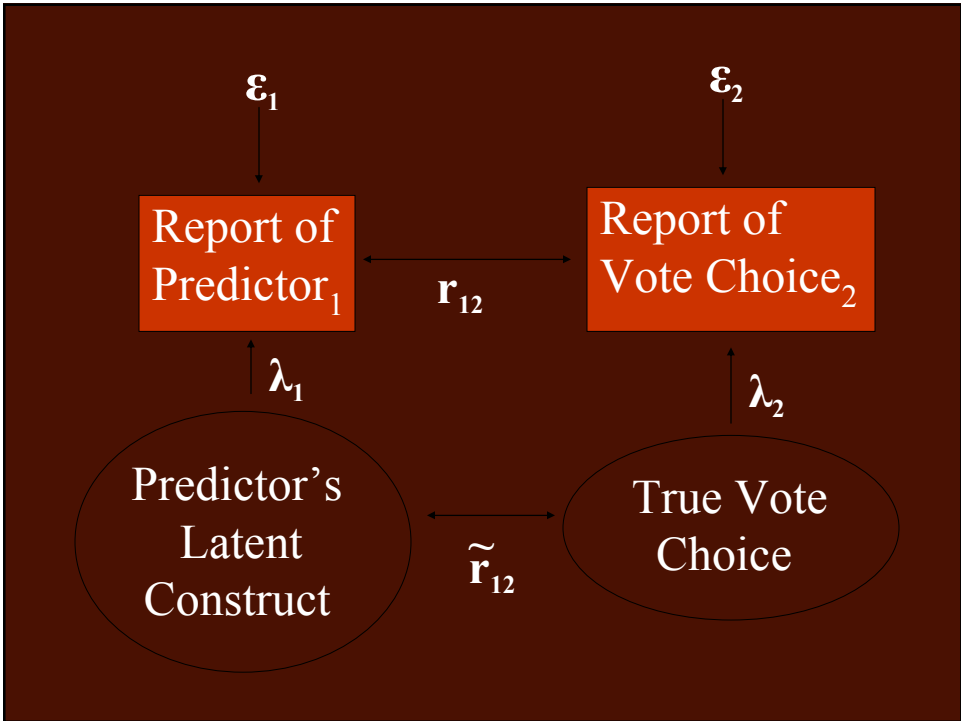
<u>Criterion</u>	<u>Which has the advantage?</u>	
	<u>Telephone</u>	<u>Internet</u>
Positive reinforcement	X	
Modeling of task involvement	X	
Accountability	X	
Literacy not required	X	
Computer proficiency not needed	X	
No long-term panel conditioning	X	
Practice effects		X
No interviewer effects		X
Honest responding		X
Self-pacing		X
Flexibility in completion time		X
Reduced memory burden		X

First Study

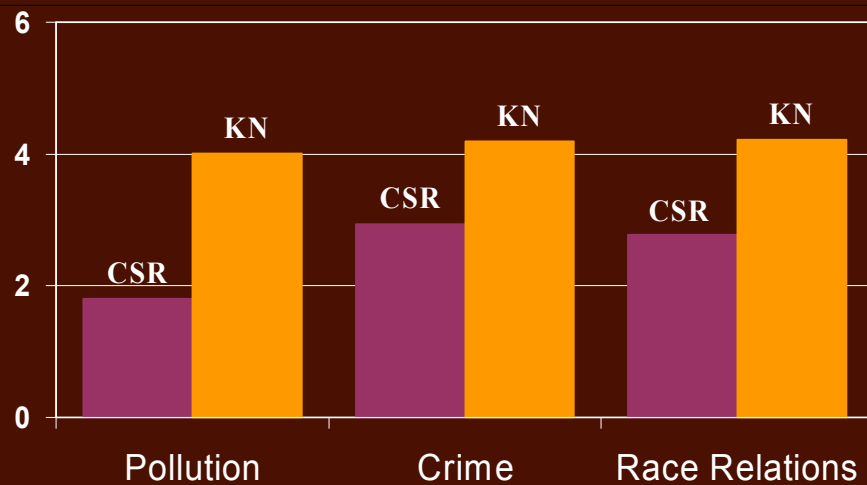
- Knowledge Networks (Internet)
- Ohio State University Center for Survey Research (RDD Telephone)

General Public Samples – 30 minute questionnaire

	<u>SRBI</u>	<u>Knowledge Networks</u>
Days in the Field	49	58
RR3	43%	
Completion Rate		70%



Logistic Regressions Predicting Vote Choice Using Clinton Approval Ratings



Differences Between KN and CSR





	KN > CSR (sig)	KN > CSR (non-sig)	KN < CSR (non-sig)	KN < CSR (sig)
Single Predictors	77%	16%	3%	0%
Model 1 (No control variables)	71%	23%	6%	0%
Model 2 (Controlling for Demographics and Knowledge)	71%	19%	6%	3%

Lab Experiment

- Random assignment to computer vs. intercom mode
- Interviewer training & supervision
- 332 respondents

Concurrent Validity

Significant Mode Differences

	Clinton	Gore	Bush
 > 	67%	71%	50%
 < 	0%	0%	0%

Total number of predictors = 39

Moderators

- Past experience
 - Respondents with no past experience completing surveys may benefit more from visual presentation and self-pacing
- Cognitive skills
 - Respondents with low cognitive skills may benefit more from visual presentation and self-pacing

DV = Clinton Thermometer

	Past Experience with Surveys	
	NO (n=83)	YES (n=246)
<i>Military Spending</i>	<i>-.08</i>	<i>.17*</i>
<i>Mode</i>	<i>-.06</i>	<i>.07</i>
<i>Military Spending x Mode</i>	<i>.35*</i>	<i>-.19</i>

Significant moderator effect present on 17% of predictors

DV = Bush Thermometer

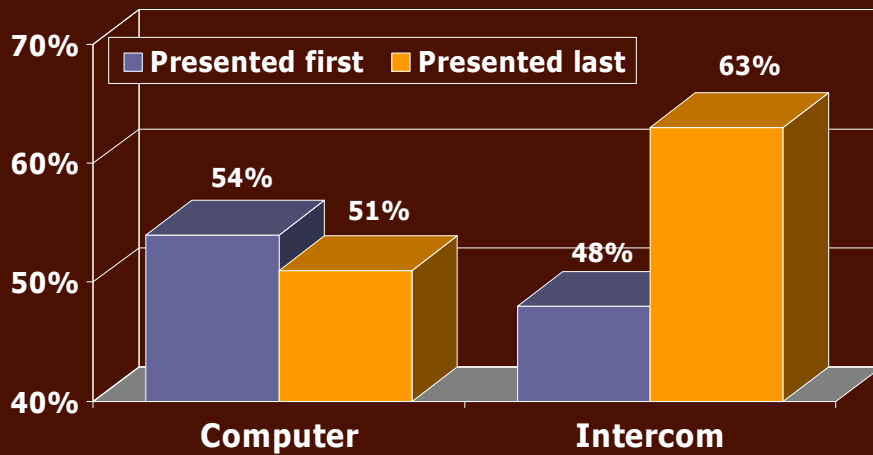
	Cognitive Skills	
	LOW (n=98)	HIGH (n=107)
<i>Expectations for Foreign Relations</i>	<i>.22**</i>	<i>.36**</i>
<i>Mode</i>	<i>.02</i>	<i>.02</i>
<i>Expectations x Mode</i>	<i>.23*</i>	<i>.06</i>

Significant moderator effect present on 25% of predictors

Response Order Effects

Response Order Effects

Percent Selecting "Crime & Violence" as the Most Important Problem Facing the Country



Response Order Studies Using Categorical Items

	<u>Primacy</u>	<u>Recency</u>	<u>Non-Significant</u>
Campbell & Mohr (1950)	1		
Becker (1954)			1
Kalton et al. (1978)			1
Schuman & Presser (1981)	2	5	8
McClendon (1986)		4	3
Bishop (1987)		3	3
Krosnick & Alwin (1987)	1		
Bishop et al. (1988)	1		1
Israel & Taylor (1990)	2		8
Ayidiya & McClendon (1990)	2		1
McClendon (1991)		7	9
Krosnick (1992)		1	3
Schwarz et al. (1992)	1	4	
Krosnick et al. (1996)		1	3

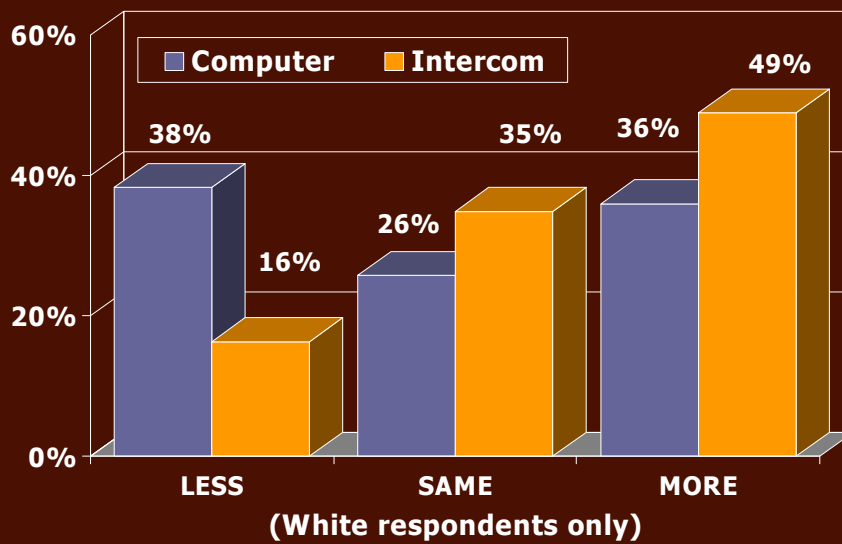
	<u>Primacy</u>	<u>Recency</u>	<u>Non-Significant</u>
<u>Visual</u>			
Campbell & Mohr (1950)	1		
Becker (1954)			1
Krosnick & Alwin (1987)	1		
Bishop et al. (1988)	1		1
Israel & Taylor (1990)	2		8
Ayidiya & McClendon (1990)	2		1
Schwarz et al. (1992)	1		
<u>Oral</u>			
Kalton et al. (1978)			1
Schuman & Presser (1981)	2	5	8
McClendon (1986)		4	3
Bishop (1987)		3	3
McClendon (1991)		7	9
Krosnick (1992)		1	3
Krosnick et al. (1996)		1	3
Schwarz et al. (1992)		4	

Voting in Elections

- 3% average primacy effect in the voting booth.
- 3% average recency effect in pre-election telephone surveys.
- Last 4 presidential election exit polls over-estimated Democratic share of the vote.
- All exit poll questionnaires listed the Democratic candidate first.

Social Desirability

Help for Black Americans



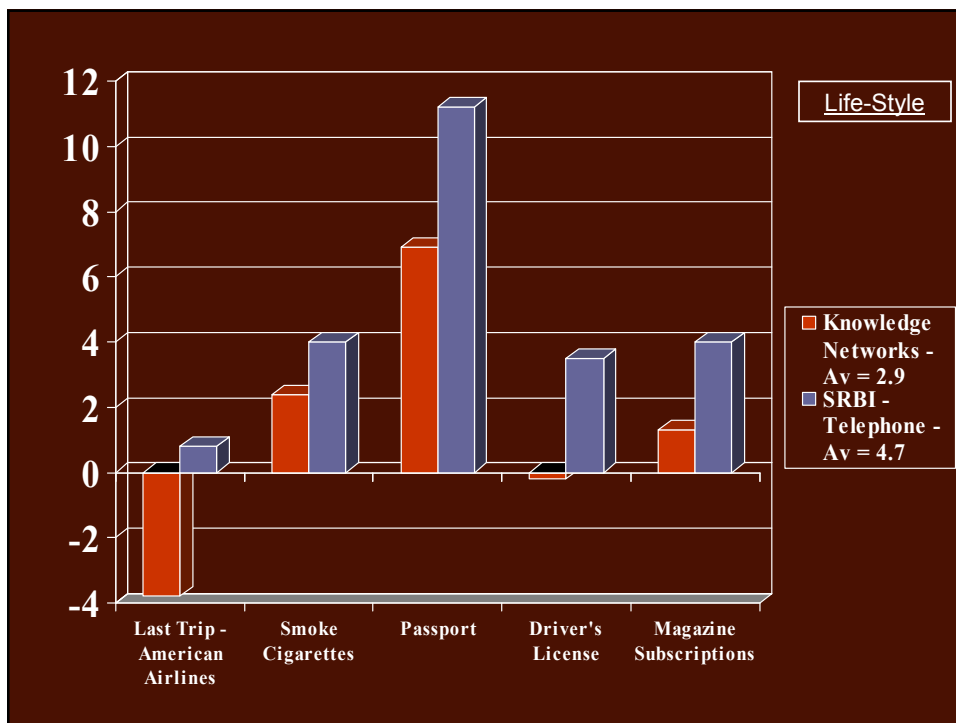
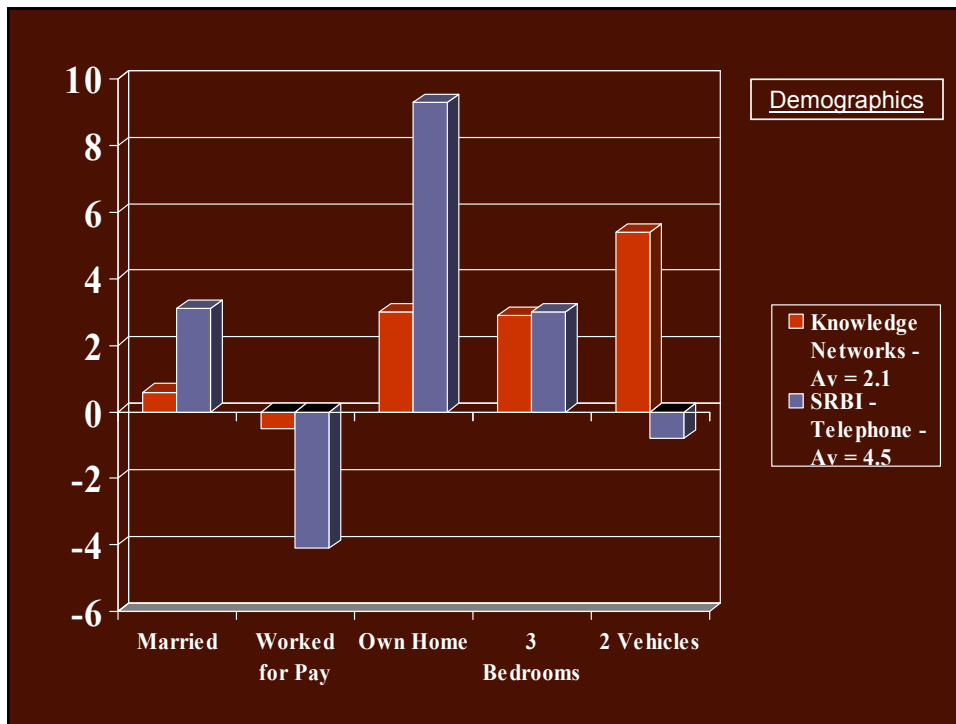
New Study

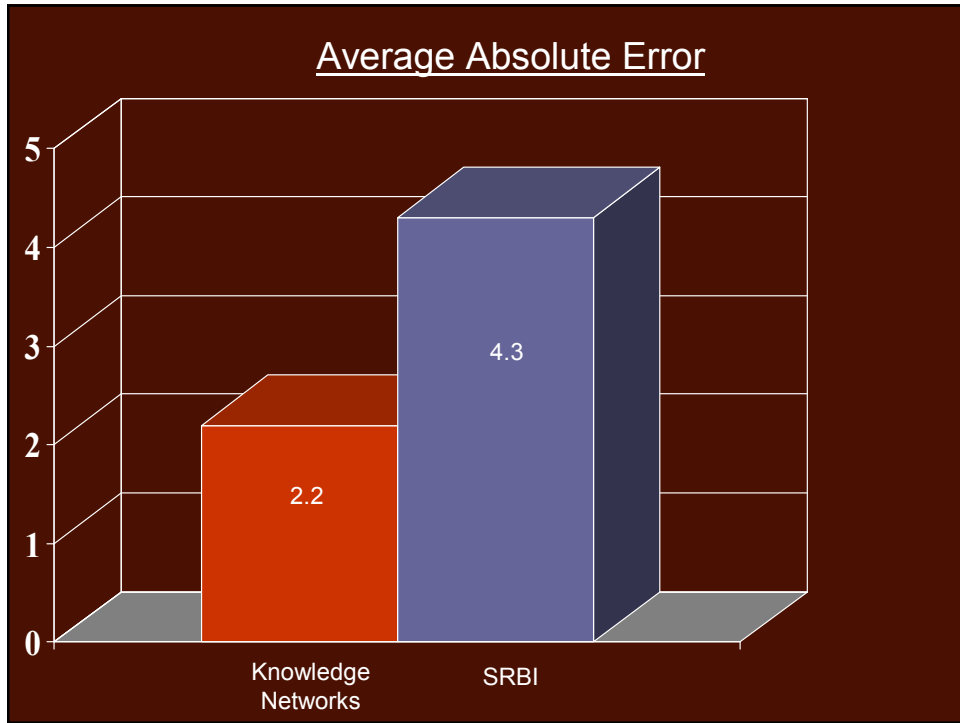
- Knowledge Networks (Internet)
- SRBI (RDD Telephone)

FACTUAL QUESTIONS

General Public Samples – 25 minute questionnaire

	<u>SRBI</u>	<u>Knowledge Networks</u>
Days in the Field	84	15
RR3	36%	
Completion Rate		73%





Conclusions

- Face-to-face beats telephone
- Internet beats telephone
- Telephone beats paper-and-pencil
- A guess:
 - F-to-F > Internet > Telephone > Paper-and-pencil

You (sort of) get what you pay for.

So ...

- Mixing modes can compromise response quality:
 - Satisficing
 - Social Desirability Response Bias
 - Inaccuracy
- Mixing modes can reverse response order effects:
 - Visual: primacy
 - Oral: recency

- Mixing modes can compromise correlational analyses:
 - Respondents self-select into second mode
 - Type of person is confounded with measurement artifacts
 - Correlation between two person attributes could be due to mode differences instead
 - Control for mode?
 - What if mode effects are interactive (different in different subpopulations)?

Cross-National Comparisons

If mode is confounded with country, we cannot simply control for mode during the analysis.

Test impact on sample representativeness of mixing modes intended to increase response rates.

Even if sample representativeness is improved, is response quality/comparability compromised?