

Mixing modes on the European Social Survey – Implications for data quality

Caroline Roberts, *City University London*
Peter Lynn, *University of Essex*
Annette Jäckle, *University of Essex*

One-day conference on mixed mode data collection in comparative social surveys - London, 15-9-2005

Current ESS Policy

- Face-to-face interviewing at data collection phase
 - Higher response rates
 - 'Better quality' data
 - Incomplete coverage of alternative modes
 - Problem of low literacy levels

The impetus for mixed modes

- Survey costs
- Response rates
- New technology
- National differences in survey practice
 - Experience and expertise in different modes
 - Penetration of different modes across countries
 - Social acceptability of different modes across countries

Options for mixing modes

- Different countries, different modes
- Sequential designs
- Respondent choice

Advantages & Disadvantages

- Advantages of a mixed mode future:
 - Reduce costs?
 - Improve response rates?
 - Respondent preferences?
 - National preferences?
- Disadvantages of a mixed mode future:
 - Mode effects
 - Continuity and quality of data

Mode effects

'Mode effects' –

1. Coverage Error: not all people can be contacted with all modes
2. Selection Bias: differential non-response because different modes 'attract' different people
3. Measurement Error: people respond differently to different modes

ESS-Gallup Research Questions

1. How well does the ESS questionnaire work in other modes?
2. Can we identify the types of ESS question most sensitive to mode?
3. Can we try to mitigate mode effects by modifying the design of questions and how they are administered in different modes?

ESS-Gallup Mixed Mode Methodology Project

- Phase 1 – pilot study in Hungary
 - 'Hall test' testing ESS questions in four modes
 - Repeated measures design
 - Findings:
 - face-to-face and telephone mode differed most from each other
 - Abstract and sensitive questions gave rise to biggest mode differences
- Peytcheva et. al (2004)

Face to face vs. Telephone

- What were the differences?

Design	N items	N sig.
Same	15	2
Showcard in f2f	11	8

Face-to-face vs. Telephone

- Sources of differences (de Leeuw, 2005)?
 - Media-related factors (e.g. control over pace & flow of interview)
 - Information transmission factors (e.g. visual vs. auditory stimuli; verbal vs. nonverbal communication, etc.)
 - Impact of interviewer
- Contribute to different types of response error, including:
 - Respondent satisficing
 - Social desirability bias

Design of Phase 2 Experiments

In each location (Budapest, Lisbon), sample selected from frame which includes both phone numbers and addresses.

Random allocation to 3 treatments:

1. f2f with showcards (standard ESS questions);
2. F2f, no showcards (adapted questions);
3. Phone (identical questions to treatment 2).

[Additional experiment within phone group involving a random subset of those with mobile phones being interviewed on their mobile: not analysed here.]

Response approx. 34% f2f, 37% phone (Budapest)

Analysis Set-Up

Some differences in sample composition between f2f and phone. F2f: more men, manual workers, less educated, slightly older.

All models include age, age², sex, occupation (3 groups), education (2 groups) as covariates.

Model response patterns that could be hypothesised to differ between particular treatments, to identify treatment effect.

Example dependent variables: Indicators of...

- ... satisficing (acquiescence, non-differentiation, no-opinion);
- ... social desirability bias (conformist responses, less extreme responses);
- ... primacy/recency effects.

Acquiescence Bias I

Six items used 5-point agree-disagree response scales (no explicit “don’t know” option).

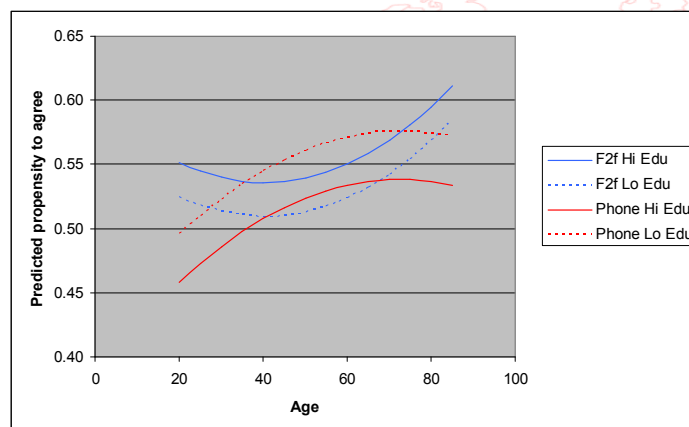
Measure of “tendency to agree” calculated as sum of ‘agree’ or ‘strongly agree’ responses, divided by 6 (thus range is 0 to 1). (cf. Holbrook et al, 2003)

Treatments 2 and 3 compared using regression model.

No treatment effect when only main effects considered.

But interactions of treatment with age and education.

Acquiescence Bias II



Social Desirability

Twenty-one items for which one or more response category had socially desirable connotations. (Note: sd not empirically proven.)

Measure of “tendency to give socially desirable responses” calculated as sum of ‘sd’ responses, divided by 21.

Treatments 2 and 3 compared using regression model. (Hypothesis: could go in either direction: distance vs. assurance)

Treatment main effect significant : more sd on phone. (Also, sex, occupation and education main effects.)

Interactions of age with treatment: sd associated with age only on phone.

Separate logit models for each of the 21 items showed 10 with significant treatment effect: 2 with more sd f2f and 8 with more sd on phone

Primacy/ Recency Effects

Fourteen items with labelled categories (including fully labelled scales).

Measure of “tendency to be subject to recency” calculated as sum of responses in second half of list (including mid-option), divided by 14.

Treatments 1 and 2 compared using regression model.

Significant and strong main effect of treatment (ME of show card -0.07).

No interactions of treatment with demographics.

Separate logit models for each of the 14 items showed 3 with significant treatment main effects and 2 others with significant interactions with treatment

Non-Differentiation

Four sets of items with identical response options (7, 3, 4 and 2 items per set).

For each set, index of non-differentiation calculated as maximum proportion of same responses.

Overall index calculated as sum of four indices divided by 4.

Treatments 1, 2 and 3 compared using regression model (2 as reference).

No treatment effect when only main effects considered.

No interactions of treatment with demographics.

No evidence of greater satisficing on phone or of any effect of visual response stimulus

Next Steps

- No-opinion responses (but not offered as explicit option)
- Item refusals, especially to sensitive questions such as income (sensitive questions to be identified by analysis of responses to Qs about sensitivity)
- Fixed vs. mobile phones; multi-tasking while on phone
- Consideration of specific items exhibiting effects and possible solutions
- Other ideas/ suggestions?

Mixing modes on the European Social Survey – Implications for data quality

Caroline Roberts, *City University London*
Peter Lynn, *University of Essex*
Annette Jäckle, *University of Essex*

One-day conference on mixed mode data collection in comparative social surveys - London, 15-9-2005

Sample Composition Differences

	F2f	Fixed	Mobile
% Men	40.5**	32.5	42.6
Mean age	56.1	55.3	48.5
% Manual	36.1***	25.4	26.5
% 'Low' Edu	55.1**	48.5	46.5