



Is it safe to combine methodologies in survey research?

Is it safe to combine methodologies in survey research?

Steph Gray, MORI Research Methods Unit
October 2003

Overview

As consumers and employees we increasingly use a range of communication channels in our daily lives, and in a growing proportion of cases, are less accessible by traditional modes such as post or landline telephone. From a research point of view, this presents a challenge: how to maintain response rates to consumer and business surveys and maximise the reach of existing methodologies.

One common approach is to use different modes of research for different audiences, combining the results from telephone, face-to-face, or self-completion postal/online questionnaires among the different survey populations. Consequently, there is growing interest from researchers and their clients in exploiting these combined techniques in order to boost response rates and cut costs. But what are the implications of combining different modes of research on the validity of the data that is gathered? How comparable are telephone and online surveys, and can the results of one ever be reliably combined with the other?

This document presents some of the implications for researchers of mixing quantitative survey methodologies. It is based on a selection of published articles which address these issues, and is intended to be a brief overview rather than a comprehensive literature review. There are a number of leading academics who have conducted extensive research in this field, of which Don A. Dillman Washington State University is probably the most prolific. A bibliography is included at the end of this document.

The main conclusion to be drawn from most of the academic studies reported in the literature is that the major modes of research (telephone, face-to-face, self-completion web/postal) *can* often be combined effectively to boost response rate and reduce costs, as long as researchers seek to minimise differences between the questionnaires and samples and monitor the profile of responses.

Mixing modes: potential issues

A common observation among academics is that mixed-mode surveys are an increasingly popular way to maximise response while reducing costs. However, sample surveys are open to four main sources of error (Groves, 1989 cited in Dillman, Tortora and Bowker, 1999) which can affect the comparability of their outputs – and since the problems affect different modes to differing extents, this can have an impact on the comparability of the results:

Is it safe to combine methodologies in survey research?

1. **Coverage error:** the sample drawn does not give all members of the population an equal, non-zero chance of being drawn – therefore making it impossible to generalise the results to a wider population who have different characteristics.
2. **Sampling error:** the effect of surveying a sample rather than entire population – normally accounted for by ensuring results are published with accompanying data on tolerances.
3. **Measurement error:** inaccurate responses as a result of design, mode or interviewer effects (set out in more detail below, including issues of question wording and layout, or respondent reaction to the presence of a human interviewer by amending their answers).
4. **Nonresponse error:** the effect of ‘missing’ responses from those who did not take part in the survey which who would have differed from those who actually did take part, again, making the results hard to generalise to the entire population.

By way of example, Green, Krosnick and Holbrook (2001) explored respondent reaction to telephone surveys as compared with face-to-face interviews and found respondents significantly more likely to give socially-desirable answers, satisfice and express dissatisfaction with the length of the interview. By addressing all these sources of error and seeking to minimise potential issues within each, researchers can hope to improve the quality of the data they gather. Dillman devised the ‘Total Design Method’ (1978) as a means of controlling these errors, particularly in the case of postal questionnaires. This approach to survey design aims to boost response rate and minimise nonresponse errors through survey design and layout, enclosing a personalised covering letter, and reminding non-respondents at specific intervals. His updated book (1999) addresses the case of the internet too. His view appears to be that the approach is a good way to minimise error in surveys deployed across different modes, and to enhance their comparability.

Mode effects

All methodologies can be affected by mode effects, caused by the effect of the format in which the survey is administered. When methodologies are combined or mixed, the differential impact of these effects can introduce problems when trying to compare the resulting data.

Cobanoglu, Warde and Moreo, 2001 (citing Dillman, 1999) assert the main problem with mixed mode surveys as being four key measurement issues between modes:

Is it safe to combine methodologies in survey research?

- **Social desirability:** Interviewer-administered survey modes have been shown to influence respondents' answers by increasing 'socially-desirable' responses, or those which respondents feel the interviewer is most likely to approve of (Green, Krosnick and Holbrook, 2001). Naturally, this can introduce a measurement error into the responses, as respondents lie about their true opinions or behaviour (e.g. 'Bashful Conservative' supporters before 1992 election - de Leeuw, 1992 cited in Dillman et al 2000).
- **Acquiescence:** A general inclination by respondents to agree with statements, or say 'Yes' at bipolar 'Yes/No' questions. (de Leeuw, 1992, and Schuman and Presser, 1981, cited in Dillman et al 2000).
- **Question order:** The sequence of questions within the questionnaire, which can affect the type of responses given. For example, a respondent may give shorter answers towards the end of a long survey, give answers based on information learnt from previous questions, or give less information on subsequent questions out of a feeling of having 'made their point' at a previous question ('satisficing').
- **Primacy/recency effects:** Where scales or lists are used, there is a tendency to record higher scores for those items at the start of the list in the case of visual modes (self-completion postal, online) or those at the end of the list for aural modes (telephone, face-to-face without showcards).

There is evidence that different modes are affected by these issues to different extents. Green, Krosnick and Holbrook (2001) explored respondent reaction to telephone surveys as compared with face-to-face interviews and found respondents significantly more likely to give socially-desirable answers, satisfice (i.e. answer questions only partially, only to the extent they feel is necessary) and express dissatisfaction with the length of the interview.

Mode effects are more marked with attitudinal questions than factual or demographic questions. Dillman et al (2000) attribute this to the fact that respondents have a preformed answer to simple questions such as age or work status, and need only scan to see the appropriate category or wait for it to be read out. Attitudinal questions requiring selection from a scale or list are more susceptible to the problems outlined above.

Oosterveld and Willems (2002) argue that previous studies fail to distinguish between population effects and mode effects, by succumbing to differential drop-out rates between modes, or using otherwise self-selecting samples. Their

Is it safe to combine methodologies in survey research?

'crossover' design incorporates a two-part questionnaire administered to all respondents (students), of which one part is CATI-based, and the other web-based. They conclude that there are no significant differences between responses, and that differences found in other studies should therefore be attributed to population differences in the samples tested. However, it is hard to believe that these findings can be safely generalised, and it is likely that design effects between interviewer-administered CATI and self-completion web are still an issue for many surveys and audiences.

Impact of Design

While design issues apply to modes in isolation, it is clear that researchers trying to combine methodologies will need to bear in mind the different environments in which respondents are interacting with a questionnaire, and that questionnaires designed for one mode (e.g. telephone - aural, interviewer-administered) need to be comparable with those in relevant others (e.g. postal – visual, self-completion).

In any case, there are aspects of the survey design that have to be changed when switching modes, and this can have an important effect on the results. In the case of attitudinal scale questions, it has been found that respondents to interviewer-administered questionnaires (Telephone, Face-to-face without showcards) where only the endpoints of scales are labelled tend to offer more extreme responses than respondents to self-administered questionnaires (self-completion, online) where scales are fully-labelled (Dillman, 2000). The tendency to pick extreme points can be reduced in telephone interviews where a numerical scale is used instead of verbal labels (Burke, 2002). However, this can introduce confusion as respondents are required to translate their view into arbitrary numerical terms and leads to them distributing their answers less widely across the scale (Dillman et al, 2000).

Similarly, 'Don't know' options are rarely included on showcards or read out by telephone interviewers. Presenting a don't know option within a self-completion online survey leads to comparatively higher proportions selecting it (Burke, 2002).

Numerous aspects of layout in the case of self-administered questionnaires can affect response patterns by reducing respondent errors and encouraging different types of responses (Dillman & Christian, 2002):

Is it safe to combine methodologies in survey research?

- Fuller responses in larger open-ended boxes
- Reversing categories has a major effect in the pattern and number of those selected; e.g. where the initial categories are positive attributes more from the full list tend to be selected
- Changing 'check all that apply' to 'Yes/No' questions significantly increases the proportion coded 'Yes'
- Providing an ordered scale of responses makes it easier for respondents to complete a question than a written description of end-points (e.g. 'Where 1 is excellent and 5 is terrible') and a blank box in which to write a number
- Arrows to clarify routing or follow-up questions reduce respondent errors

Another example of these modal differences is the impact on results of 'tick all that apply' checkbox questions in online or self-completion questionnaires, which are conventionally structured in Yes/No terms within telephone or face-to-face interviews (Dillman, 2002)

Response rates

Clearly, respondents have preferences about the modes in which they are contacted to take part in research. Groves and Kahn, 1979 (cited in Dillman et al, 2000) found that of respondents to a national telephone interview, 39% indicated they would have preferred to be surveyed by telephone, 23% face-to-face and 28% by post. There is evidence that respondents to one mode of survey would not have agreed to take part in the same survey via another mode. Swoboda et al, (cited in Woong Yun and Trumbo, 2000) found that 41% of respondents to their web-based survey would not have completed it via telephone if asked.

Results of studies have been somewhat equivocal – some have shown that response rates can be improved by offering respondents a choice of modes of response (Dillman, Phelps, Tortora, Swift, Kohrell and Berck, 2001); others that it makes little difference to overall response rates, but does encourage switching to a preferred mode (Dilman, Clark and West, 1995). The effect is likely to depend on the audience: common sense suggests a young, mobile audience may prefer a web-based survey on technology issues to one conducted by post and may be accordingly more likely to respond.

Research by Dillman (1995, cited in Dillman, 2000) found that offering respondents a choice of modes of response did not actually improve the overall response rate, though a relatively small proportion of respondents did use the alternative modes offered. However, where non-respondents to one mode are *recontacted* using a different mode, response rates *are* substantially improved

Is it safe to combine methodologies in survey research?

(Shettle and Mooney, 1999 cited in Dillman et al, 2000). In an experiment testing this sequential approach, Dillman’s team found that response rates reached 80-83% when a combination of telephone followed by mail reminder were used, regardless of the order in which these modes were administered (see following table). However, this boosting of response rates did not make significant inroads into the hard-to-reach groups of non-respondents.

Mode	Phase 1 %	Phase 2 % (among non-respondents to Phase 1)	Combined response rate %
Postal, then CATI	75	32	83
CATI, then Postal	43	66	80
Interactive Voice Recognition (tel) then CATI	29	36	50
Web, then phone	13	45	48

Source: Dillman et al, 2000

McCabe et al, 2002, conducted an experimental survey using postal and web self-completion questionnaires among a population of students. They found a significantly higher response rate for the web questionnaire, (63% vs 40%). In addition, a follow-up survey showed that non-respondents to the web mode were more likely to have been unaware of the survey than postal non-respondents, and were also more willing to take part when reminded by telephone. They found that email records were more up to date than postal records among this audience, and hypothesise that this may be because the student respondents were more aware of the need to keep email details up to date to avoid missing out on important contact. A further finding was that around a third of non-respondents to the postal survey thought they had completed it – from which McCabe et al conclude that the greater complexity in completing, preparing to send, and sending a postal questionnaire leads to higher non-response rates than web-based questionnaires. No real differences existed between modes regarding concerns about confidentiality (slightly higher among postal non-respondents), though the survey subject matter concerned drinking and drug usage.

Is it safe to combine methodologies in survey research?

Cost implications

The academic explorations of mixed-mode research touch only briefly on issues of cost. Generally, they start from the argument that cost saving is the primary factor driving the popularity of mixed mode research (for example, in the US, where respondents in telephone surveys are routinely diverted to interactive voice recognition systems (IVR) to complete interviews to free up the interviewer to make more calls).

Those who do touch on cost issues (Comley, 1997) predictably report that online surveys are much cheaper to administer than alternative modes. Poulsen (2001) uses the example of a market segmentation study for European food labelling to demonstrate how an initial segmentation conducted by telephone can be followed up with online research among internet-accessible segments whose characteristics are known.

While this may be true, cautious researchers considering mixed-mode projects might incur additional costs, however, such as:

- extra time in questionnaire design to ensure survey designs are 'portable' between modes and can be repurposed without introducing significant design effects
- potentially, the time and cost of drawing and formatting different types of sample
- cost of conducting research with split samples to determine skews in online samples compared with alternative offline modes
- additional time developing quotas and weighting
- additional complexity in data processing and synthesis of datasets collected using different modes

Conclusions: issues to bear in mind

There is an assumption in the literature that mixed-mode surveys are here to stay and that given that all survey research is likely to be imperfect, they should continue to be used, but with caution. Issues to consider include:

- **Is the population in question accessible** by the modes of research proposed?
- **How much is known about the population** in question? (Adequate knowledge of demographic and attitudinal characteristics may enable a researcher to use a non-probability sample and correct for skews and non-response error).
- If **response rate** is an issue, consider sequential mode alternatives (reminders by telephone/email/post, following initial contact by another mode) rather than simultaneous mode choice. Does common sense suggest that the population is likely to respond better to one mode than another (e.g. business audiences, young internet users etc.)?
- Consider modes at the outset of the project, and ensure the **questionnaire** is 'portable' between modes in terms of its design and format
- Finally, use the following checklist of sources of error and minimise where possible:

Coverage error: is the whole population included in the sampling approach? If some people are more likely than others to be included, can you adjust for this in some way?

Sampling error: do normal sampling tolerances apply? Check with the Research Methods Unit and Sampling & Statistics Department.

Measurement error: is the interview format the same for each mode? Try to design questionnaires as similar as possible in layout and question type, and factor design/interviewer effects into your interpretation of the results

Nonresponse error: who has not taken part in the research? Is anything known about them and has the use of multiple modes affected the profile of non-respondents?

Is it safe to combine methodologies in survey research?

Bibliography & Further Reading

Backstrom C and Nilsson, C. 2002. 'Mixed mode: Handling method-differences between paper and web questionnaires'. Dept of Information Technology and Media, Mid-Sweden University.

Backstrom, C and Nilsson, C. 2002. 'Mixed mode surveying: a comparison of paper-questionnaires and web-questionnaires'. Dept of Information Technology and Media, Mid-Sweden University.

Baker, K., Curtice, J. and Sparrow, N. 2003. 'Internet poll trial: Research report'. ICM & The Guardian.

Bradley, N. 1999. 'Sampling for internet surveys. An examination of respondent selection for Internet research'. University of Westminster.

Cobanoglu, C, Warde, B, and Moreo, P. 2001. 'A comparison of Mail, Fax, and web-based survey methods'. Journal of the MRS.

Comley, P. 1997. 'The use of the internet for opinion polls'. ESOMAR.

Comley, P. 1997. 'The use of the internet as a data collection method'. ESOMAR paper.

Crawford, S., McCabe, S., Couper, M. and Boyd, C. 2002. 'From mail to web: improving response rates and data collection efficiencies'. International Conference on Improving Surveys, Copenhagen.

Curasi, C. F. 2001. 'A critical exploration of face-to-face interviewing vs. computer-mediated interviewing'. Journal of the MRS.

Dillman, D, Tortora, R and Bowker, D. 1999. 'Principles for Constructing Web Surveys'.

Dillman, D., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. 2000. 'Response rate and measurement differences in mixed mode surveys using mail, telephone, Interactive Voice Response and the Internet'.

Dillman, D. 2002. 'Navigating the rapids of change: some observations on survey methodology in the early 21st century'. Paper from Dillman website.

Is it safe to combine methodologies in survey research?

Dillman, D and Christian, L. 2002. 'The influence of words, symbols, numbers, and graphics on answers to self-administered questionnaires: results from 18 experimental conditions'

Green, M., Krosnick, J. A. and Holbrook, A. L. 2001. 'The Survey Response Process in Telephone and Face-to-Face Surveys: Differences in Respondent Satisficing and Social Desirability Response Bias'. Universities of Pennsylvania and Ohio.

Hogg, A. and Maszta, J. J. 2001. 'A practical learning about online research – Attribute rating scale type effects'. Advertising Research Foundation Workshop.

Hogg, A. 2002. 'Conducting online research'. Burke White Paper.

Hoppe, M and Lamp, R. 2001. 'The quality of online panels'. ESOMAR.

Mitchellhill, G. and Love, C. 2001. 'Going the extra distance: a case study on the setup and implementation of an online continuous customer experience monitor'. ESOMAR.

Moon, Y. 2000. 'Intimate exchanges: using computers to elicit self-disclosure from consumers'. MRS Journal.

Oosterveld, P and Willems, P. 2002. 'Two modalities: one answer? Combining Internet and CATI surveys effectively in market research'.

Poulsen, C. S. 2001. 'Efficiency vs effectiveness in Online Research'. ESOMAR.

Rosenbaum, P. R. and Rubin, D. B. 'The central role of the propensity score in observational studies for causal effects'. *Biometrika*, 70.

Schillewaert, N., Langerak, F. and Duhamel, T. 1999. 'Non-probability sampling for WWW surveys: a comparison of methods'. *Journal of the MRS*.

Strauss, J. 1996. 'Early survey research on the Internet: review, illustration and evaluation'. American Marketing Association Conference.

Terhanian, G. and Black, G. S. 1999. 'Understanding the online population: Lessons from the Harris Poll and the Harris Poll Online'. Advertising Research Foundation Workshop.

Terhanian, G. 2003. 'The unfulfilled promise of internet research'. Harris Interactive Europe.



MORI

Research at the heart of business and public life

79-81 Borough Road
London
SE1 1FY
United Kingdom
Tel: +44 (0) 20 7347 3000
Fax: +44 (0) 20 7347 3800
email: sri@mori.com
Visit our website at www.mori.com

44 Albany Street
Edinburgh
EH1 3QR
United Kingdom
Tel: +44 (0) 131 558 1515
Fax: +44 (0) 131 558 1717
email: scotland@mori.com
Visit our website at www.mori.com/scotland