Survey Methodology Bulletin

Special Edition No. 58 August 2006

Contents

Editorial  Zeeshan Rahman  1

1. The impact of mode on the comparability of survey data  Zeeshan Rahman & Abigail Dewar  3

2. Developing a postal questionnaire on social capital  Abigail Dewar & Amanda Wilmot  11

3. Maximising response to a postal survey on social capital  Abigail Dewar  29

4. In search of a population sampling frame for UK postal surveys  Sarah Tipping & Gerry Nicolaas  34

5. Population coverage and response rates for a postal survey on social capital  Gerry Nicolaas  47

6. Mode effects in social capital surveys  Gerry Nicolaas & Sarah Tipping  56

7. Developing harmonised questions for use in a mixed-mode data collection environment  Amanda Wilmot & Abigail Dewar  75
The Survey Methodology Bulletin is produced primarily to inform staff in the Office for National Statistics (ONS) and the Government Statistical Service (GSS) about the work on survey methodology carried out by the ONS. It is produced by the ONS, and staff in the ONS are encouraged to write short articles about methodological projects or issues of general interest. Articles in the bulletin are not professionally refereed, as this would considerably increase the time and effort to produce the bulletin: they are working papers and should be viewed as such.

The bulletin is published twice a year, usually in March and September. This is a special interim edition of the Survey Methodology Bulletin. It only includes papers from one project; the background to which is given in the editorial on page 1.

The bulletin is free to staff within the ONS and the GSS. It is made available to others with an interest in survey methodology at a small cost to cover production and postage.

The Office for National Statistics works in partnership with others in the Government Statistical Service to provide Parliament, Government and the wider community with the statistical information, analysis and advice needed to improve decision-making, stimulate research and inform debate. It also registers key life events. It aims to provide an authoritative and impartial picture of society and a window on the work and performance of government, allowing the impact of government policies and actions to be assessed.

Prepared & Edited by: Zeeshan Rahman
Editorial

In order to address its commitment to promoting a consistent approach to the measurement of social capital, the Office for National Statistics (ONS) developed a harmonised question set (HQS) for use in interviewer-administered computer-assisted social surveys, primarily at the national level. The HQS comprises 50 questions based on the dimensions of social capital, and a shorter set of core questions for surveys which do not need to accommodate the full set. The core and full question sets, together with the dimensions to which they pertain, can be accessed at the following web address: http://www.statistics.gov.uk/socialcapital/downloads/Draft_UserGuide_v3.pdf.

There is however interest in collecting data on social capital at a local level, with a view to populating the community well being domain of Neighbourhood Statistics. Local authorities are primarily responsible for collecting local level data and many local authorities use postal self-completion as a more cost effective method of data collection.

Consequently, Social Inequalities Branch (SIB) commissioned Data Collection Methodology (DCM) at the ONS and the Survey Methods Unit (SMU) at the National Centre for Social Research (NatCen) to convert the core harmonised social capital questions from a face-to-face interviewer-administered to a postal self-administered format.

The main aim of this research was to design a postal questionnaire that would provide comparable outputs to the existing harmonised CAPI question set. Only the core question set was converted because the same principles can be applied to the full set and testing the full set would have been more resource intensive. (It should be noted that it was not in the remit of this project to improve or make changes to the wording of the HQS.)

Outputs from this project are in the form of written papers detailing the methodology employed and the results from each stage in the development and testing process. The papers report on: the conversion process itself and the qualitative methods used to inform and test the design of the paper questionnaire; population coverage and response in relation to postal surveys; an evaluation of a potential sampling frame for local authorities to use when conducting postal surveys; a quantitative experiment designed to assess the impact of mode of administration on the final results, and a summary of the entire project including a discussion about using harmonised questions in a mixed-mode data collection environment.

Despite the desire to produce a paper questionnaire that could provide comparable results to the CAPI instrument, significant differences were evident from the data pertaining to the use of certain question types. So, our conclusion to the study is that mode effects will reduce comparability of national and local data, for some question types, if the current set of harmonised social capital questions is used in local postal surveys.

This special edition of the ONS Survey Methodology Bulletin is intended to aid dissemination of the results from this project. The papers were written as the project progressed, and were intended to stand independently of one another; therefore there may be some repetition in the reporting.

---

1 Computer Assisted Personal Interviewing
Thanks are given to the research analysts who initiated the project and provided support and advice throughout, namely Jean Martin, Penny Babb, Figen Deviren and Lucy Haselden. Thanks are also due to the research and operations staff at ONS and NatCen. In particular Susan Purdon at NatCen who provided methodological advice; the staff at ONS who worked on the CAPI survey; and staff at NatCen who worked on the postal survey. Most importantly thanks are given to the respondents for completing the questionnaires.
The impact of mode on the comparability of survey data

Zeeshan Rahman & Abigail Dewar, UK Office for National Statistics

1. Introduction

The main aim of this project was to convert the core harmonised social capital question set from an interviewer-administered to a self-administered format for local authorities to use in postal surveys. The purpose of the conversion being to help local authorities to collect data on social capital at a local level that would produce comparable outputs to the CAPI\(^1\) question set used in national surveys. However, the comparability of local and national data could be affected by the different modes of data collection including any differences in population coverage, differential non-response bias and the mode of administration. Therefore these issues needed to be considered during the conversion process.

2. Differences in population coverage

Differences in population coverage could be introduced if different sampling frames are used for national and local surveys. The Postcode Address File (PAF) provides a more complete coverage of residential addresses than alternative sampling frames, for example the Electoral Register (ER) and as such, tends to be used by most face-to-face national surveys. However, it only provides a list of addresses and not individuals; making it unsuitable for postal surveys of the general population because mailings cannot be personalised. Mailing that is personalised is known to have a positive impact on response rates (Dillman, 2000; Krysan et al., 1994). Unlike the PAF, the ER includes the names of individuals living at each address on the register and does not include ineligible addresses (Foster, 1993; Foster 1994). Therefore, in the past, this has been the preferred sampling frame for local level postal surveys. However, the ER only includes adults aged 16 and over who are eligible to register and have actually registered to vote. Therefore the ER may not be a comprehensive representation of the general household population. Consequently, population coverage varies between local and national surveys because different sampling frames are used to draw samples. In order to maximise comparability, variation in population coverage needs to be minimised or controlled for.

Furthermore, the ER is no longer a viable option for postal surveys of the general population because of severe under-coverage on the edited version, which provides people with an option to opt out from being listed. Only the edited ER can now be used for survey purposes. The edited ER makes it more difficult to achieve comparable outputs between local and national surveys. Therefore, an alternative sampling frame to the edited ER was evaluated as part of this project. Issues related to coverage and sampling frames are discussed further in papers four (Tipping and Nicolaas, 2006) and five (Nicolaas, 2006) of this bulletin.

\(^1\) Computer Assisted Personal Interviewing
3. Differential non-response bias

The comparability of data used in local and national surveys, can also be affected by non-response bias. The extent of non-response bias is determined by both the survey response rate and the differences in the characteristics of respondents and non-respondents.

Response rates are often lower for postal surveys than face-to-face surveys. The sample composition is also likely to differ because some population groups, for example those with lower educational qualifications, are less likely to respond to a postal survey than to a face-to-face survey (Fowler, 2002; Dillman, 1978; Krysan et al, 1994). Hence, different response rates may reflect different sample compositions and, consequently the results may be skewed according to the characteristics of respondents in these sample compositions, thus potentially biasing survey estimates.

Measures need to be taken to increase the response rates of postal surveys to bring them in line with face-to-face surveys and remove or control for differences in the profiles of respondents between modes. Again, a more detailed discussion of non-response bias and the techniques used to minimise its effect can be found in papers four (Tipping and Nicolaas, 2006) and five (Nicolaas, 2006) of this bulletin.

4. Mode effects

One of the key components of this project was to investigate and minimise mode effects that may result from the conversion. The use of different data collection instruments is likely to produce mode effects, where the same questions asked in different modes produce different outputs.

The initial stage of the project entailed preliminary research to identify the likely mode effects that could result from converting the social capital question set in order to inform the design of the postal survey questionnaire. This consisted of quantitative analysis of data from the Health Survey for England (HSE) and a thorough review of current literature.

The quantitative analysis of the HSE data was conducted by the Survey Methods Unit (SMU) at the National Centre for Social Research (NatCen). This involved a comparison of answers to similar questions on social capital from the HSE 2000, which used an interviewer-administered CAPI-based data collection method, with the HSE 2001, which used a self-administered CASI-based method. Differences between the two sets of questions were investigated along with the effects these differences had on comparability. To ensure that the observed differences were due to survey mode, propensity score matching was used to control for differences in the sample profiles between modes.

The literature review was carried out by Data Collection Methods (DCM) at the Office for National Statistics (ONS). The review identified possible mode effects which could result from using different modes of data collection and provided recommendations on how to minimise such effects (Rahman et al, 2005).

---

2 Computer Assisted Self Interviewing
3 A description of propensity score matching can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp
Findings from the HSE analysis and best practice according to key survey methodological literature were used to inform the conversion process. This paper highlights the main factors which can influence comparability of data collected using different modes and summarises the combined findings from both the literature review and the HSE analysis. The results from the HSE analysis are used to support current research literature where appropriate and together illustrate the importance of considering mode effects in survey design.

For the purpose of this review, modes of data collection refer to interviewer-administered (face-to-face (CAPI) and telephone (CATI)) and self-administered (postal, CASI and web).

4.1 Visual and aural communication

The use of visual and aural communication varies between different modes. In a face-to-face survey, both visual and aural stimuli can be used to communicate information. Visual aids used within this mode mainly comprise of showcards, while aural communication refers to respondents listening to interviewers reading the survey questions, response options and instructions out loud. The absence of the interviewer in a postal self-administered survey means only visual stimuli can be used to convey information to the respondent. Dillman (2000) considers these distinctions between modes and suggests that the way information is communicated to respondents can influence the thought process involved in interpreting questions and response options. If respondents interpret and answer questions differently depending on the mode of administration, comparability of survey data between modes may be compromised.

Furthermore, category order effects, which can cause variations in the selection process of response categories, can occur as a result of these differences. The most common of these are primacy and recency effects. The primacy effect is the tendency to choose answer categories towards the beginning of a list of options. This is more common in self-administered surveys where respondents give greater cognitive consideration to earlier options, increasing the likelihood of their selection in comparison to options presented later. In contrast, the recency effect is the tendency for categories towards the end of a list to be selected. This is more common in interviewer-administered surveys, where in the absence of showcards, response options are read out by interviewers. Answer categories towards the end of a list tend to be selected more frequently because these are the ones respondents tend to remember, although sometimes not reflecting the respondents’ true circumstance or opinion (Dillman, 2000; Krosnick and Alwin, 1987). These effects occur as a direct consequence of the way information is presented in different modes.

In considering the differences between visual and aural communication, Dillman advises not to change the basic structure of questions when converting an interviewer-administered to a self-administered questionnaire. Even minor changes in the format or structure of a question can impact on respondents’ thought processes when answering. For example, descriptive information used in interviewer-administered surveys requiring the respondent to answer using a scale, such as ‘from a scale of one to ten’, is likely to be interpreted differently to a visual representation of the scale used in self-administered questionnaires. Therefore, attention is required in consistently using the same descriptive information in all modes.

---

4 Computer Assisted Telephone Interviewing
4.2 Question wording and response options

Similar to the effect of small changes in the format or structure of questions, many studies have found that minor changes to question wording and response options can have a major impact on respondents’ interpretation; resulting in differences in survey estimates (Schuman and Kalton, 1985 and Schuman and Presser, 1981 cited in Schwarz et al, 1991). Nicolaas and Tipping (2004) reported similar findings from their comparison of the HSE data, where differences in question wording and response options between HSE 2000 and HSE 2001 could have contributed to the variations in survey estimates between modes. For example, a question on ‘participation in group activities’ offered respondents a list of answer categories, one of which was slightly different between modes: ‘sports club’ in the interviewer-administered survey (using a showcard) and ‘sports club, gym, exercise or dance group’ in the self-administered version. A higher proportion (25.7%) of respondents selected this category in the self-administered version compared with the interviewer-administered version (18.8%). It seems likely that the provision of additional information could well have caused this variation.

Furthermore, Nicolaas and Tipping (2004) found that item non-response ranged from 0.9 to 8.9 percentage points higher in the self-administered mode than the interviewer-administered mode. In some cases this may also have been a result of differences in the question wording and response options between modes. For example, the inclusion of a ‘don’t know’ option for one of the questions in the self-administered questionnaire, which was not explicitly provided in the interviewer-administered version, is thought to be the likely cause of the greatest difference (8.9 percentage points) in item non-response. An explanation for this variation is the tendency for respondents to select such an option when explicitly presented to them. Krosnick et al (2002) claim that in most cases such an occurrence is due to the lack of time and effort invested by respondents in providing an answer rather than genuinely being unable to respond or not knowing the answer. Therefore, careful consideration should be given to the use of ‘don’t know’ response options when using a self-administered questionnaire.

However, not all differences in question wording and response options between modes influence survey estimates. Results from the HSE analysis showed that providing a spontaneous neutral option (an unprompted answer category), such as ‘neither’ or ‘depends’, in the interviewer-administered survey and not in the self-administered survey, did not necessarily affect the estimates. A question in the interviewer-administered HSE 2000 included the spontaneous option ‘neither’, which was not included in the self-administered equivalent in 2001, yet the estimates did not differ between the two modes.

In order to minimise the impact of mode on survey estimates in the conversion process, it was vital that the questions and answer categories in each mode were similarly interpreted by respondents. Dillman (2000) suggests a ‘unimode’ approach to the design; ensuring questions and response options are the same for all modes, thus developing one set of questions which can be applied to all modes of data collection. As Dillman’s approach relies on the simultaneous development of questions for all modes, rather than the conversion of those that already exist, it could not be used because the social capital harmonised question set had already been finalised. However, some of the principles of the unimode design could still be applied in order to try to minimise differences between modes.
4.3 Cognitive burden on respondents

Cognitive burden is the mental pressure on respondents to process survey information and provide answers. Question wording and response options may be the same across modes but if they are ambiguous then this is likely to cause additional cognitive burden. Likewise, unclear instructions and complicated routing will increase respondent difficulty in answering the questionnaire. The level of cognitive burden can vary between modes but is likely to be greater in a postal survey than an interviewer-administered survey as a result of ambiguous questions, unclear instructions or complicated routing.

In interviewer-administered surveys, interviewers usually give instructions and provide support to help respondents answer questions as intended. When converting an interviewer-administered survey to a postal format, instructions need to be incorporated into the postal questionnaire. However, converting verbal instructions into a visual format could be counter-productive, making them more confusing for respondents, and ultimately increasing cognitive burden.

Computer Assisted Interviewing permits the programming of a sophisticated instrument whilst remaining user-friendly. The computer is programmed to navigate interviewers and respondents through the questionnaire. However, postal self-administered surveys need to be as straightforward as possible because respondents must follow the routing unassisted (Martin et al, 1993; Tourangeau and Smith, 1996). If the design is not kept simple, cognitive burden is likely to increase.

Cognitive question testing is a research technique that can be used, prior to running a survey, to reduce the likelihood of respondent’s misinterpreting questions, terminology, instructions and nonverbal information, or making routing errors. This method of testing mainly consists of carrying out one-to-one interviews with respondents to explore the mental processes by which they reach and provide an answer to a question. This provides insight on whether the survey information is being interpreted as intended, and suggests unambiguous question wording, layout and routing. Once the questions have been converted to a postal format, it is advised that cognitive methods are used to investigate respondents’ interpretation of the information within the paper questionnaire.

4.4 Interviewer guidance and support

The interviewer-respondent interface is important and can be effective. Interviewers have a positive role in providing guidance to respondents and motivating them to complete a survey. This can reduce the overall burden on respondents.

A study of the UK International Passenger Survey (IPS), which compared answers to face-to-face interviews with those obtained from paper self-administered questionnaires, revealed that some respondents misinterpreted terms and did not follow instructions correctly in the paper version (Rowlands, 1994). If such misunderstandings result in respondents answering questions incorrectly, this could affect the quality of data collected. The presence of an interviewer can eliminate such problems by helping or probing where necessary (Schwarz et al, 1991).
4.5 Sensitive questions and social desirability

While the involvement of interviewers is often considered advantageous, they can have a negative impact on survey estimates, particularly when asking potentially sensitive questions; this is known as a social desirability effect. Social desirability refers to occasions when respondents give answers which they consider to be concurrent with the beliefs of the interviewer or the norms of society in general (Dillman 2000). This phenomenon is likely to be a major contributing factor influencing the results of the HSE study; such an occurrence was observed when answers to the questions about ‘problems in the neighbourhood’ were compared. Respondents were asked to rate the extent to which certain issues were a problem by using a scale, containing five categories ranging from ‘very big problem’ to ‘not a problem at all’. The results showed that a greater proportion of respondents admitted to problems in their neighbourhood in the self-administered survey than in the interviewer-administered version. Other research has found similar results. Dillman (2000) draws on three studies, Krysan et al (1994), Dillman and Mason (1984), and Tarnai and Dillman (1992) which looked at neighbourhood and community related questions. Each study found categories such as ‘not a problem’ and ‘never a problem’ were selected more frequently in interviewer-administered than in self-administered questionnaires. The difference was thought to be a consequence of respondents not wanting to give a negative view of their neighbourhood or community when interacting with an ‘outsider’, such as the interviewer.

Similarly, it is argued that self-administered modes tend to collect better quality data on sensitive issues (for example drug use, sexual behaviour, alcohol consumption) than interviewer-administered modes, as respondents feel they have more privacy in the absence of an interviewer and therefore feel more comfortable about giving honest answers (de Leeuw, 1992).

Hewitt (2002) studied whether women preferred to be interviewed face-to-face or using Audio-CASI5 when answering questions on numbers of sexual partners. Each respondent was interviewed using both modes. Forty-nine per cent of respondents said that the answers they had given using the Audio-CASI were the ‘most honest’ in comparison with 10.6% when interviewed face-to-face. The remaining respondents said they gave equally honest answers in both methods. The study also looked at how comfortable respondents felt in the two modes, while 52.7% of respondents were indifferent to the mode used, almost four times as many felt more comfortable with Audio-CASI than face-to-face interviews (37.7% and 9.6% respectively). Furthermore, the research found that some respondents reported having more sexual partners when using Audio-CASI. Hewitt concluded that Audio-CASI is better at capturing sensitive behaviour than face-to-face interviews. Similar findings were reported by Tourangeau and Smith (1996) who also found a higher level of reporting on both drug use and sexual partners in CASI and Audio-CASI compared with CAPI.

As well as the interviewer, the presence of other people can also affect the responses given. Evidence presented by Bajekal and Purdon (2001) shows that when other members of the household were present during an interview on social capital, respondents tended to give the consensus view of the household rather than their individual views.

5 Audio-CASI involves respondents hearing a recording of survey questions and providing their answers on a computer.
4.6 Question order

Finally, the order in which questions are presented to respondents can influence the answers given. This is known as a question order effect. The extent of such an effect would depend on the degree of interviewer and respondent control over the question-answer process as well as the time available for respondents to complete the survey. This type of order effect could, for example, occur if respondents are asked about problems in their neighbourhood before they are asked about their general opinion of the area they live in. In this situation, respondents are likely to be thinking about the problems referred to in the first set of questions when providing their opinion of the area.

In a postal survey, respondents are able to look through the entire question set before providing answers, whereas, in an interviewer-administered survey this flexibility does not exist because the order in which questions are asked is controlled. Furthermore, in a postal survey, respondents have more control over what specific parts of the questionnaire are read, which may result in part of a question or response categories being ignored (Dillman, 2000). It is therefore harder to control for question order effects in postal surveys where it is not possible to completely determine the order in which questions are seen or answered and whether all or part of a question is read (Krysan et al, 1994).

Research suggests that respondents have more time in self-administered surveys than in interviewer-administered surveys to understand and answer the questions (Schwarz et al, 1991). They also feel more at ease to go back and change or correct their answers (Bishop et al, 1988 cited in Dillman, 2000). In contrast, respondents participating in an interviewer-administered survey may hesitate in asking interviewers to change their responses to previous questions. De Leeuw (1992) states that having more time to consider questions and taking account of the survey context is beneficial to the question-answer process.

5. Conclusion

The evidence presented here illustrates the importance of considering mode effects and the extent to which they may impact on the comparability of survey estimates between modes. As one of the sponsor’s key objectives is to produce comparable outputs on social capital at the local and national level, it is important that it is addressed and ways in which to mitigate the mode effects described in this paper should be thoroughly reviewed and taken into account when the conversion takes place. While measures need to be taken to eliminate or minimise such effects wherever possible, the authors acknowledge that any recommendations to this effect can only be implemented within the cost and time constraints of the project.

References


Developing a postal questionnaire on social capital

Abigail Dewar & Amanda Wilmot, UK Office for National Statistics

1. Introduction

The second stage of the conversion involved the design, development and cognitive testing of a prototype questionnaire. This paper explains the rationale for the design of the postal questionnaire including the importance of numerical, symbolic and graphical language in helping to process and interpret the meaning of questions (Dillman and Redline, 2004), and discusses the cognitive research methods used to find out why and how respondents perceived the questionnaire based on this rationale. Key findings from testing the prototype and lessons learnt using these methods are reported. In addition, suggestions are made for techniques which might provide further insight into the ways respondents comprehend and complete the questionnaire, and as such aid questionnaire design improvement.

The paper therefore focuses on the conversion from a qualitative perspective, looking at how cognitive research techniques have informed design and development. However, while it was important to evaluate whether respondents perceived the postal questionnaire as the researchers’ intended and why they interpreted the questionnaire in a certain way, this approach does not provide a measure of the effect of using different modes. As part of the next stage of the conversion process, the CAPI-based core social capital harmonised question set were included as a module on the National Statistics Omnibus survey during April, May and June 2005. In order to field test the social capital questions and quantitatively assess any differences between their use in self-administered and interviewer-administered format, the former were included in a postal survey which was run in parallel with the social capital question module included in the Omnibus survey. The results of the quantitative analysis are presented in paper six of this bulletin (Nicolaas and Tipping, 2006).

2. Mixed-mode method of data collection

2.1 Design rationale

Using a mixed-mode method of collecting data on social capital could impact on the comparability of national and local data as different modes often produce different results. In order to minimise this risk, and produce a postal survey of good quality, it was necessary to consider how respondent requirements might differ when using different modes of data collection. This would ensure the self-administered questionnaire developed for the postal survey was similar as possible to the interviewer-administered version with regards to the data outputs.

Dillman, Sangster, Tarnai and Rockwood (1996) identify three main ways in which data collected by interviewer-administered surveys and similar data collected by self-administered surveys may be affected by mode differences: presence versus absence of an interviewer, visual versus aural communication, and whether the interviewer or

1 Computer Assisted Personal Interviewing
Abigail Dewar & Amanda Wilmot

Developing a postal questionnaire on social capital

the respondent controls delivery of stimulus. Accordingly, the CAPI-based social capital questions are reliant on interviewer delivery, while paper-based questions are reliant on respondent reception. The emphasis is placed either on how interviewers verbally deliver questions or how respondents see the questions, therefore appealing to the aural and visual senses respectively. Whichever mode of surveying is used, it is important that all respondents receive the information through common stimulus.

2.1.1 Unimode approach

In order to control for all mode effects the ideal would have been to take a unimode approach to the original questionnaire design. This is a method of questionnaire construction advocated by Dillman (2000b) as a way of attempting to limit response differences. Using this approach, survey questions are written and presented so that they are transferable across different modes. For the interviewer-administered and self-administered social capital questions, this would be achieved by formatting and re-wording the questions in such a way that the stimulus acquired from the visual appearance and layout of the self-administered questionnaire is equivalent to the aural stimulus when the questions are delivered in a face-to-face interview (Dillman and Christian, 2005).

In this way, following the unimode construction principles of commonality across modes requires the question styles usually adopted by specific modes to be replaced with a suitable compromise. This would involve optimising the question wording so that a generic question set can be developed that are acceptable regardless of survey mode. This would not be possible for the conversion of the harmonised social capital questions because the design of these questions has been finalised. As such, it was not intended that the wording of the questions would be tested, or the written style of the questions revised; instead researchers were required to find other ways to optimising question comprehension. Therefore the questionnaire design focussed on the non-verbal, visual channel of administering self-completion questionnaires as an alternative method of conveying the questions, that is, the paralanguage provided by visual aids.

The key challenges to the development of the self-administered questionnaire were firstly the ability of respondents to interpret the questions appropriately without an interviewer present, and secondly ensuring all respondents interpret the questionnaire in the same way. Without the aid of interviewers, and the freedom to adapt the questions (wording and ordering) themselves, it was necessary to concentrate on how numbering, symbolism (arrows and ticks) and graphics (changes in font size, shape, spacing and shading) could be manipulated to enable respondents to extract additional meaning from the questionnaire (Redline and Dillman, 2002).

3. Questionnaire testing

3.1 Methods used and lessons learned

Once a prototype had been developed, it was necessary to consider the best way of assessing whether the questionnaire was an effective operational research instrument, that it was able to provide researchers with accurate information.

Cognitive interviewing and observation techniques were used as the most proficient way to test if respondents were able to perceive all the information in the questionnaire, and also how they interpreted what they perceived. However, in
addition to understanding how respondents answer questions, without the presence of an interviewer in self-administered mode, it is important to gain some insight into how well motivated they are to complete the questionnaire. Therefore as well as cognitively analysing response behaviour, the techniques were also used as methods of assessing how the information given to respondents affects their motivation to complete or not complete the questionnaire. This approach to the development of self-administered questionnaires follows Jenkins and Dillman’s (1993) use of combined cognitive and motivational research methods to inform and improve self-administered questionnaire design.

3.1.1 Sample for cognitive interviews

A ‘purposive sampling’ method was used to select respondents as the aim of the sampling was to identify individuals chosen because they have particular characteristics rather than to achieve a statistically representative sample of the population.

Therefore, using the Office for National Statistics’ (ONS) qualitative respondent register (Wilmot, 2005) to recruit respondents, the sample was stratified by higher and lower education, age and sex: the aim being to recruit people with mixed reading and comprehension ability and with varied experience and attitudes towards completing questionnaires.

Seventeen cognitive interviews were carried out in London with seven males and ten females. Their ages ranged from 18 to 71 years old, seven were of higher educational level and ten lower educational level, representing a diverse range of respondents.

Further details on the sample specification including diagrammatic illustration of the composition of the achieved sample can be found in the technical report (Dewar, 2005).

3.1.2 The value of iteration

Testing was conducted using an iterative process whereby researchers reviewed interviews carried out in the first week of testing to identify any ways in which the methodology might be refined. The first iteration was a valuable learning experience for the cognitive researchers, with subtle amendments made to field documents and the progressive development of more systematic topic guides. This enabled cognitive researchers to take a more robust approach to the second stage due to greater understanding of testing required to thoroughly examine the complexities of visually communicating the social capital harmonised questions in self-administered format.

3.1.3 Concurrent think-aloud and observational cognitive assessment

Initially concurrent think-aloud and observational cognitive interviewing techniques were used to assess respondents’ understanding of the questionnaire. The concurrent think-aloud method required respondents to verbalise their thoughts while completing the questionnaire. Prior to each interview, cognitive interviewers trained respondents

2 Higher educational level was defined as above A’ level or equivalent, and lower educational level as A’ level or equivalent and below.
in how to think out loud using a test exercise as an example. The reason for using this method was that it enabled interviewers to gain insight into what was going through respondents’ minds as it occurred to them, rather than relying on the respondent to be able to correctly recall everything they had been thinking after they had completed the questionnaire.

However, after the first iteration, it became apparent that this technique was proving to be problematic. The researchers found that respondents were tending to comment on the wording of the questions and not the more numeric, symbolic or graphical features of the questionnaire. Although these comments were valuable and insightful, indicating that it might be necessary to revisit and improve the harmonised question set, as explained in section 2.1.1 above, the purpose of the study was not to evaluate the question wording. It also became apparent to researchers that conducting the interviews in this way inevitably resulted in a certain degree of interaction between the respondent and cognitive interviewer. Interviewer presence could be creating a false environment in which to answer the questionnaire, as ordinarily respondents would complete the questionnaire without an interviewer in attendance. Reading the questions out loud might also encourage respondents to focus on information they might otherwise have glanced over or possibly even skipped. Consequently, the think-aloud approach seemed not to be meeting the needs of the research objective.

3.1.4 Alternative cognitive interviewing technique: retrospective probing

Therefore a different method, retrospective probing, was used for the second iteration of interviews. This technique is often regarded as less obtrusive than think-aloud and less likely to change respondents’ behaviour, as cognitive interviewers are able to observe respondents while keeping their interaction to a minimum (Forsyth and Lessler, 1991). Using this method, respondents were asked to fill in the questionnaire as they would do at home, ignoring the fact that the cognitive interviewer was in the room. During questionnaire completion, the interviewers observed the respondents, noting anything that indicated respondents might be having difficulties answering the questionnaire, for example, facial expressions, hesitancies and missed routing instructions. Once respondents had completed the questionnaire, the interviewers used both prepared and spontaneous probing based on their observations, and specific information required about certain aspects of the questionnaire from a previously prepared topic guide (Dewar, 2005) to find out what respondents thought of the questionnaire.

3.2 Assessing all field documentation

It was important to explore respondents’ perception and interpretation of all the documentation they received, not just the questionnaire, as this might provide a richer understanding of the degree to which the information collectively affected whether or not they completed the questionnaire itself. Therefore research interviewers observed respondents reading the cover letter as well as completing the questionnaire. Furthermore, cognitive probing focussed on how likely the cover letter would affect respondents’ motivation to complete the questionnaire, as well as features of the questionnaire design.
3.2.1 Survey document order

The preliminary iteration revealed an important methodological finding regarding the order by which survey documents were tested. Researchers did not always give respondents the cover letter to read first. However, although the questionnaire was completed in an artificial test environment, it became clear from the comments made by respondents, that it was still essential for them to read the cover letter before completing the questionnaire. Some respondents who did not complete the task in the correct order queried why specific questions were being asked, for example question 4, which asks respondents how often they personally contact relatives, friends and neighbours. Others felt that people might be more inclined to fill in the questionnaire if they had a better understanding of what the survey was about; yet this information would have been clear to them had they been given the cover letter to read prior to completing the questionnaire.

3.3 Qualitative findings

The findings from the cognitive interviewing and observational assessment are solely a presentation of the opinions and experiences of the respondents themselves from completing the questionnaire, and are not the voice of the author.

The original (prototype) self-completion questionnaire and cover letter used during testing can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp.

4. Qualitative results - the role of the cover letter

Respondents’ comments highlighted the important role the cover letter plays as part of the survey package. In addition to providing an invitation to take part in the study and acting as an introductory medium, it also informs respondents’ first impressions of the questionnaire and can aid their motivation to complete it.

It was apparent from the findings of the cognitive testing that respondents had clear views of what they expected the letter content to be. For example, respondents wanted an informed title which enabled them to know exactly what the survey was about: it was commented respondents did not understand the title, ‘Community Perception Survey’. There seemed to be a great deal of uncertainty as to what the title was referring to, with suggestions that it was about the community, peoples’ role in the community, “what I am aware of in the community” and involvement in the community “beyond just living there”. The title was considered vague, and no-one perceived from the title that the survey was concerned with respondents’ views of their neighbourhood, indeed the word ‘neighbourhood’ was suggested as a better way of indicating what the questionnaire would be about. Respondents also wanted to know how their help would be of benefit: it was felt that more information was needed on how the data collected would be used and how similar surveys have helped in the past.

As such, there appeared to be a clear link between the letter and respondents’ motivation to complete the questionnaire, it seemed to be a crucial deciding factor as to whether respondents felt inclined to at least look at the questionnaire and find out what it entailed. Schwarz (1996) likens the way respondents attach meaning to the questionnaire based upon earlier information (such as the cover letter), as similar to everyday conversational exchange. The cognitive testing revealed that respondents...
were trying to gather as much information as possible to help them establish the context of the survey, which in turn seemed to help make the researchers’ request coherent to them.

4.1 Declaring confidentiality

As well as requests for more information in the cover letter, respondents also suggested where information provided could be kept to a minimum. In particular respondents felt that a brief confidentiality statement was sufficient to reassure people that their data would be protected.

5. Qualitative results - perceptions of the front page

5.1 Confidentiality: striking the balance between too much and too little

Respondents’ views on the confidentiality statement on the front page of the questionnaire were an extension of their views on the statement given in the cover letter.

While most respondents felt it was important to be clear of the research organisations’ commitment to protecting the confidentiality of the information they provide, some respondents felt that the detailed explanation given was unnecessary and others expressed no interest in the statement at all. In one instance it was felt that including a confidentiality statement on the front page would actually be detrimental as it felt like respondents were obligated to take part in the survey, as such it was suggested that the statement should be put on the back page.

It was important to strike a balance in the amount of information provided. Therefore, to ensure enough detailed information was still provided from an ethical perspective, while only a brief statement was included on the front page, a more detailed explanation was given at the back of the questionnaire. This meant that the confidentiality statement would no longer detract from other information on the front page, but for those with concerns and who required further detail, the statement would be available in full at the back of the questionnaire.

5.2 Providing instructions

Respondents also felt that the detailed confidentiality statement detracted attention from the instructions on the front page.

It is not common practice in postal social survey design to locate instructions on how to complete a questionnaire on the front page, and it can be argued that there is no requirement for them if individual instructions are also placed directly with questions. However the researchers’ intention was that the front page would provide key instructions believed to enhance the completion process of the overall survey rather than complicate it unnecessarily.

Furthermore while most self-administered paper questionnaires have a separate front cover with an illustrative graphical design identifying the survey prior to the start of the questionnaire, there is no evidence from research to-date recommending a definitive cover page. Therefore it was decided that the social capital questionnaire would cognitively test the placement of key instructions on the front page alongside
Abigail Dewar & Amanda Wilmot

Developing a postal questionnaire on social capital

minimal core identification details which both compliment information already provided in the cover letter, and are enough for the questionnaire to work as a stand-alone document if necessary. As such, this design approach favours launching the questionnaire with questions starting almost immediately rather than following an illustrative cover page (and possibly also a separate instructions page), with the view that this method would encourage respondents to effortlessly start completing the questionnaire.

Another option would be to start the actual questions on the front page, while there is empirical evidence supporting this concept (Dillman, 2000a) the cognitive researchers were concerned with the risk of launching respondents too quickly into the start of the questionnaire. Key instructions located on the front page seemed to be a good medium between having an illustrative cover page and having questions on the front page.

Cognitive findings revealed that while respondents had not necessarily come across this design concept before, it was generally felt that having instructions on the front page was an effective way to grasp peoples’ attention. Respondents felt that they would be more likely to read the instructions on the front page and consequently most felt that they should remain on there. Furthermore, they also felt that this prime positioning of the instructions motivated them to continue with the questionnaire. It seemed that reading the instructions on the front meant respondents were effectively starting the questionnaire; that they were inadvertently committing themselves to continue.

Locating the instructions on the front page acknowledges Nederhof’s (1988) recommendations to have a distinctive front cover which is both appealing to respondents and easily retrievable. The questionnaire will be recognisable and easy to distinguish from other mailings. Furthermore the instructions reinforce the ‘to do’ nature of the document; cognitive findings indicate that on seeing the front page, respondents immediately perceived the questionnaire as something requiring their attention and felt a sense of urgency for it to be completed and returned rather than read sometime in the future. A snapshot of the front cover of the prototype questionnaire is shown in Appendix A.

Therefore, as a follow-on from the cover letter, the front page of the questionnaire provided further scope to motivate respondents in actually carrying out the completion request.

6. Qualitative results – questionnaire comprehension

Subsequent to reaching a point where respondents are completing the questionnaire, it is important that they interpret the questions as the researchers’ intended.

The cognitive testing revealed various ways in which the questionnaire was affecting respondent comprehension and the possible impact this might have on the quality of survey data.

The following focuses on three of the five themes identified from the interviews: firstly navigation and the use of visual aids to guide respondents through the questionnaire; secondly how respondents interpret information on the questionnaire page as they follow the navigational path and thirdly, alternative questionnaire formats (portrait and landscape), and two different question types (successive and matrix), and discusses key elements within each of these themes.
6.1 Following a navigational path

Respondents must be able to navigate a clear route through the questionnaire in order to complete it efficiently. Therefore cognitive researchers wanted to ascertain whether information presented to respondents seemed to follow a logical order, it was important to find out if it was obvious to respondents what should be the correct way of proceeding through the questionnaire. Following elements of visual design (Dillman, 2000a) the questionnaire was developed using various visual aids and enhancements, the purpose being to map out a pathway which guided respondents to questions in the order intended by the researchers themselves.

6.1.1 Adhering to the order of information defined by the questionnaire designers

When a questionnaire is administered by interviewers, researchers have a degree of certainty that the questions will be asked and answered in the order they intended. In contrast, it is argued that there is more freedom in a self-administered questionnaire for respondents to look ahead or backtrack (Tourangeau, Rips and Rasinski, 2004).

Cognitive researchers were particularly interested to find out respondents’ views on the combination of numbers and letters used as visual aids to identify questions 2, 3, 4 and 5. For these questions, letters were used to distinguish between different problems, actions, types of contact and unpaid help (an illustrative example of the combination of numbers and letters is shown in Appendix B).

- Some respondents had difficulties following a combined number/letter style: it was felt that the problem statements in question 2 could be numbered consecutively, rather than using letters, although it was pointed out that this was preference only, it was not felt that this would make a difference to how the question was answered. Similarly, preference was made for numbers rather than letters in question 3. It was felt that using letters to identify each action was confusing as it made it hard to distinguish between the lettering and the start of the sentence:

  “Usually everything is put in numbers; the letters could be confusing for some respondents”.

- Other respondents liked the combined number and letter approach. It was felt that it helped to group together questions on a common theme, and it was considered particularly useful when the question is spread over two pages, such as in questions 2 and 4.

6.1.2 Intended use of accompanying questionnaire instructions

Instructions were designed with added visual elements to help respondents’ task of comprehending and answering questions as they follow through the questionnaire.

---

3 The full prototype self-completion questionnaire used during testing can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp
Abigail Dewar & Amanda Wilmot

Developing a postal questionnaire on social capital

• Most respondents read and understood the instructions. They were easily seen; the use of bold type to emphasise the colour, providing weight to the words and attracting respondents’ attention to the instructions seemed to be effective. It was commented that the bold text made you think “this is something I must read”. As such, variation in the weight of colour in text seemed to encourage respondents to follow the cognitive researchers’ desired navigational path (see Appendix B).

• Using a tick as a visual aid alongside the written instruction also proved effective. All respondents ticked within the boxes to answer the questions and it was commented that using the tick symbol surrounded by a white box to demonstrate what is required was a good idea; respondents liked the way this visual aid illustrated a verbal instruction. It was noticed that the tick boxes were white on a darker coloured background (see Appendix C for an example), and this contrast was viewed positively:

“it makes it easier to spot the boxes and easier on the eye”.

6.2 Visual interpretation of page layout

It was important to ensure key information in the questionnaire was made easily perceptible for respondents. Good design practice advises following the natural tendency to read from left to right and from top to bottom; using graphical features to alter the emphasis on different parts of the questionnaire also helps respondents understand and process information received in self-administered format.

6.2.1 Exploiting graphical design

Several graphical design features enhanced or diminished respondents’ perception and ability to process information on each page of the questionnaire.

• Respondents found question numbers in reverse print effective (see Appendix B); looking at the page, their eyes were drawn to the different print type making it easy to find the start of each question.

• Respondents liked the contrast between the shaded background field and white answer space. Most respondents understood that they should give their answers in the white answer space. However, some respondents had difficulties with the layout of the ethnicity question. The elongated white answer space used to invite respondents to ‘Please write below’ was problematic. The black outline surrounding the answer space seemed to be enhancing the white box. Respondents were regarding each section as separate questions, rather than answer options for sub-elements of the question as a whole. One reaction to this was that the categories seemed to be “all over the place” and that it was not clear where you should start from to find the answer you needed. The ethnicity question used in the prototype self-administered questionnaire is shown in Appendix D.

Therefore the black outline to the elongated boxes was removed so that sections of the question were presented on the page in a way that ensured respondents perceived all elements of the question as part of one unit; a method practised by Wallschlaeger and Busic-Snyder (1992).
6.2.2 Consistent use of graphical elements

In order to interpret the visual information on a questionnaire page, research has shown that respondents refer to their previous knowledge and what they expect to see based upon past experiences and visual context.

- Familiarity provided by consistent visual graphics (reverse print and regular use of square boxes to outline question numbers) enabled respondents to learn through experience what was expected of them as they answered the questions. Respondents commented that as they completed the questionnaire, they began not to notice the numbers themselves, but that this was because:

  “by this point (question 4) [they] assumed that a number type represents a new question”.

- Similarly, respondents commented that as they made their way through the questionnaire, the tick symbol and bold print became familiar visual graphics and it was clear to them what they should do. Many concluded that this was due to the fact that they were able to recognise the instruction was “the same each time” although the repetition was seen as a good idea, and a helpful reminder for some respondents about how to answer questions.

6.3 Alternative questionnaire formats

During the development of the questionnaire it became apparent that for two of the batteries of questions, namely question 2 and 4, duplicating the first part of the question before each statement might seem repetitive for respondents completing a paper questionnaire. To overcome potential risk of respondent fatigue, frustration, boredom or even non-response due to the repetitive nature of the successive questions, a simple, matrix question grid was designed whereby respondents were asked the main part of the question and then presented with a series of problems (question 2) and contacts (question 4) down the left-hand side of the page following on as elliptical sentences. Respondents were required to use categories arranged across the top of the page to provide answers. Appendix B illustrates the successive question style (question 4) and Appendix C illustrates the matrix question style (question 4).

The matrix grid was formatted in landscape to improve spacing as in portrait orientation this question style appeared crowded with the answer categories too close together.

Cognitive researchers investigated how respondents dealt with these two different types of question, addressing the navigational and visual perception issues discussed above and evaluating the impact of format and orientation in maintaining a common stimulus for all respondents as they answered the battery of questions.

All respondents saw both question styles, receiving the two different versions of the questionnaire in a controlled random order. It was intended that the ideal version would be selected based on the outcome of investigations into which design, the repeat successive question style printed in portrait orientation or the simple, matrix question style printed in landscape orientation, was preferred and worked better with respondents.
6.3.1 Controlling the navigational path

- Respondents found it difficult to answer the question in matrix style; the question layout meant it was easy to tick a box on the wrong row of the grid. For example a respondent was observed nearly ticking a box in the row for the previous question before realising their error and ticking the box aligned to the question they were answering, while others made mistakes which they corrected.

- Those respondents who preferred the successive questions felt that reading the question as an elliptical sentence in the matrix format might cause people to keep going back to the start to read the first part of the question again. As a result respondents found it difficult to map a direct route through the matrices; it was felt that there was too much room for error as the page was “covered in boxes”.

6.3.2 Interpreting the question style: in matrix format or repeated questions in successive spaces

- Respondents who preferred the matrix style tended to be those with previous experience of this type of question; they liked the fact that all the issues were presented on one page, the questions were perceived as being more “compact”, rather than questions spread over two pages. It was commented that the successive question type used wording which repeated itself for no particular reason.

- Persons with less experience of completing questionnaires perceived the questions provided in successive spaces to be easier to answer. The questions were considered to be less confusing. The repetition gave respondents more time to process what the questions were asking, and made them think more about their answers. As a result, it was felt that there was less chance of answering incorrectly.

6.3.3 Maintaining a common stimulus

A challenge in the design of the questionnaire itself, was managing to create a common stimulus for all respondents. Similar to interviewers who are trained to read questions verbatim, deliver them in the same order and follow routing instructions determining which questions they should ask, it was important that the self-administered questionnaire was visually designed so each respondent received information in the questionnaire in the same way.

However respondents seemed to read the matrix style questions in different ways. For example, in order to answer question 2, some respondents read the entire question set, then read the first part of the question again before giving each answer. Others read the first part of the question every time they provided their answers, while some read the first part of the question only once. Consequently, respondents didn’t use the same method to answer the question and there no longer seemed to be a common stimulus for respondents reading and answering the questionnaire:

- Respondents who did not refer back to the start of the question before providing an answer to each problem in question 2 found it difficult to remember and make a connection between the first part of the question and
the following parts. It was commented that the series of problems read “as though part of a sentence is missing”. Respondents who did read both the first part of the question and then the problem that followed, commented that they did not like having to refer back up to the start of the question each time and it was suggested that the question would be easier to understand as a series of statements rather than as an elliptical sentence.

- Respondents seemed more likely to read questions 2 and 4 in the same way when the information was presented to them as successive questions. Respondents were inclined to methodically work their way through each question. However some respondents found presenting the series of problems in question 2 as successive questions repetitive. It was commented that the question was a bit repetitious and that it kept saying the same thing, “how much of a problem, how much of a problem...” and as a result there was a tendency to skip this part of the question.

6.3.4 Perceiving the page layout: landscape or portrait orientation

Cognitive researchers investigated respondents’ ability to handle landscape orientated matrix questions while completing the questionnaire.

No respondents had difficulties turning the questionnaire around in order to complete the matrices for questions 2 and 4 in landscape orientation, however it was considered to be annoying and “odd that you have to flip the paper around” indicating the landscape orientation could be detrimental from a motivational perspective.

7. Outcome and conclusions

The successive question style was selected as a result of the overall finding from the analysis; the cognitive testing also generated a number of recommendations specifying how the questionnaire could be improved (Dewar, 2005).

Collectively these improvements are demonstrative of the key part visual survey design plays in the development of self-administered questionnaires. Paralanguage is a powerful tool in conveying information, enabling respondents to extract additional meaning from the questionnaire via appearance and layout. In particular, a combination of graphical features such as spacing, shading, print type, symbols and shapes guide respondents along an intended navigational path and also encourage respondents to focus on important information. When applied consistently to the survey design, this enables respondents to establish some familiarity with the task they are required to perform.

However, caution should be taken in how paralanguage is used, as graphical design features can diminish as well as enhance perception and consequently the ability to process information in the questionnaire. This is illustrated in the high potential for response error associated with the matrix question format. Using the grid style proved a considerable challenge to the basic question and answer process and created unnecessary respondent burden. It is important not to overcomplicate visual design; the use of successive questions individually spaced, although simple, was more effective in encouraging respondents to progress along the navigational path and made it easier for them to interpret questions. This has a cumulative effect as respondents gain confidence, feel more motivated and better able to provide clear answers with fewer errors; consequently improving the quality of survey data.
8. Future research

Retrospective probing was an effective way to capture the completion process however further insight might be provided by making better use of observational techniques.

The interviews carried out in the cognitive research laboratory were also video taped, with unobtrusive cameras angled to focus on respondents’ faces, hands and the questionnaire to capture any observational data which might be missed by the interviewer alone.

Currently this visual database is small, however it is intended that more respondents are video taped completing the questionnaire to build upon the observational data set. This will enable cognitive researchers to code and analyse the video recordings and provide an additional technique to inform self-administered questionnaire design.

Such insight will aid questionnaire comprehension and as a result, the accuracy of the postal survey data collected. Ultimately, this will provide further opportunity to enhance survey quality when measuring social capital at the local government level.

References


Appendix A: Front cover of the prototype self-administered questionnaire

Community Perceptions Survey 2005

Confidentiality

The information you give us will be strictly confidential. Your name will not be
linked to any report that results from the survey and no information that identifies you
will be passed on to anyone outside the research team.

Instructions: Please read before completing the questionnaire

♦ Most questions can be answered simply by putting a tick in the box next to the
answer that applies to you, like this:

Yes... ✓
No....

♦ Usually, after answering each question you will go on to the next question, unless a
box you have ticked has a ‘Go to’ instruction:

Yes... ✓ Go to 2
No....

♦ If you mark the wrong box, fill in the box and put a tick in the right one, like this:

Yes... ✓
No....

♦ If you change your mind, put a tick next to the right box, like this:

Yes... ✓
No....

♦ In answering the questions, please give your opinion only and not the collective
opinion of other people.

♦ When you have completed the questionnaire, please post it back in the prepaid
evelope provided.

Thank you for taking part in our survey, your contribution is important to us.
Please turn to page 2 to begin the questionnaire

Notes:
This front cover has been resized from its original size (A4) for illustrative purposes.
## Contact with relatives, friends and neighbours

The next set of questions is about how often you personally contact your relatives, friends and neighbours. Please do not count the people you live with.

For each of the following questions please tick one box only.

| 4a How often do you speak to relatives on the phone? |
|------------------|------------------|------------------|------------------|------------------|------------------|
| On most days     | Once or twice a  |
|                  | week             |                  |                  |                  |
|                  |                  |                  |                  |                  |
|                  |                  |                  |                  |                  |

| 4b How often do you speak to friends on the phone? |
|------------------|------------------|------------------|------------------|------------------|
| On most days     | Once or twice a  |
|                  | week             |                  |                  |                  |
|                  |                  |                  |                  |                  |
|                  |                  |                  |                  |                  |

| 4c How often do you speak to neighbours face to face? |
|------------------|------------------|------------------|------------------|------------------|
| On most days     | Once or twice a  |
|                  | week             |                  |                  |                  |
|                  |                  |                  |                  |                  |
|                  |                  |                  |                  |                  |

| 4d How often do you meet up with relatives who are not living with you? |
|------------------|------------------|------------------|------------------|------------------|
| On most days     | Once or twice a  |
|                  | week             |                  |                  |                  |
|                  |                  |                  |                  |                  |
|                  |                  |                  |                  |                  |

Notes:
This page has been resized from its original size (A4) for illustrative purposes.
Question 4e was on a different page of the paper questionnaire so has not been included here.
Appendix C: Question 4 in a matrix style landscape orientation from the prototype self-administered questionnaire

<table>
<thead>
<tr>
<th>Contact with friends, relatives and neighbours</th>
</tr>
</thead>
<tbody>
<tr>
<td>The next question is about how often you personally contact your relatives, friends and neighbours.</td>
</tr>
<tr>
<td>♦ For each of the following questions please tick <strong>one box only</strong></td>
</tr>
</tbody>
</table>

4 Not counting the people you live with, how often do you...

<table>
<thead>
<tr>
<th></th>
<th>On most days</th>
<th>Once or twice a week</th>
<th>Once or twice a month</th>
<th>Less often than once a month</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>...speak to relatives on the phone?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...speak to friends on the phone?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...speak to neighbours face-to-face?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...meet up with relatives who are not living with you?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...meet up with friends?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please turn over

Notes: This page has been resized from its original size (A4) for illustrative purposes.
Appendix D: Ethnicity question from the prototype self-administered questionnaire

What is your ethnic group?

Please choose one section from A to E, then tick ✓ the appropriate box to indicate your ethnic group.

A  White
   British .......................................................... ✓
   Irish ............................................................... ✓
   Any other White background, please write below .......................................................... ✓

B  Mixed
   White and Black Caribbean .......................................................... ✓
   White and Black African ............................................................... ✓
   White and Asian ................................................................. ✓
   Any other Mixed background, please write below .......................................................... ✓

C  Asian or Asian British
   Indian ............................................................... ✓
   Pakistani ............................................................. ✓
   Bangladeshi ......................................................... ✓
   Any other Asian background, please write below .......................................................... ✓

D  Black or Black British
   Caribbean .......................................................... ✓
   African ............................................................... ✓
   Any other Black background, please write below .......................................................... ✓

E  Chinese or other ethnic group
   Chinese ............................................................... ✓
   Any other, please write below .......................................................... ✓

Notes:
This page has been resized from its original size (A4) for illustrative purposes.
Maximising response to a postal survey on social capital

Abigail Dewar, UK Office for National Statistics

1. Introduction – limitations of postal surveys

In the design and development of the Social Capital Postal Survey, the limitations of such surveys became self-evident. Most noticeably, without an interviewer present, the ability for the researchers to communicate with respondents was restricted. This in turn meant that there was no means of ensuring that the survey was administered to the sampled respondent, and that there was no verbal means of explaining what the survey was about, what information was required, or any way of allaying respondents’ concerns about taking part.

The implications for survey design and implementation were twofold:

Firstly, significant efforts had to be put into practice to maximise response rates as without an interviewer to ensure that the sampled respondent completed the questionnaire or to persuade potential respondents to take part, the means of securing respondent co-operation was reduced.

A combination of different factors used to help improve response had a positive effect on the number of productive returns for the postal survey (see paper five of this bulletin: Nicolaas, 2006) and ultimately resulted in a very favourable co-operation rate\(^1\) of 69% compared with 74% for the interviewer-administered version run on the Omnibus survey (see paper six of this bulletin: Nicolaas and Tipping, 2006).

Secondly, it meant the questionnaire and survey materials such as the covering letter had to be of an optimum standard as there was a necessary reliance on the postal package as the sole method of minimising non-sampling error in order to collect quality data. Therefore development focussed on the content and wording used to explain the purpose of the survey in the covering letter, and the use of paralanguage and general visual design principles to help guide respondents through the questionnaire and aid comprehension (as discussed in paper two of this bulletin: Dewar and Wilmot, 2006).

This paper discusses how a high rate of co-operation was achieved using methods which can also be applied to postal surveys in general.

2. Methods used to achieve high response to a postal survey on social capital

2.1 Personalisation

Personalising correspondence has been proven to increase response rates to postal surveys (Dillman, 1991); using names from CACI’s Consumer Register and matching them to the addresses sampled from the Postcode Address File (see paper four of this bulletin: Tipping and Nicolaas, 2006), enabled this technique to be used in the Social Capital Postal Survey. This method also helped maintain data quality in that, to a

\(^1\) The co-operation rate indicates the number of achieved interviews as a proportion of those contacted during the fieldwork period (Lynn et al, 2004).
certain extent, it was possible to deal with the lack of control over who completed the postal questionnaire; by addressing the questionnaire to a named individual, it was more likely that the correct member of the household would complete it.

2.2 (Respondent-friendly) Questionnaire design

Undoubtedly, the questionnaire design was also an important factor encouraging overall unit response and reducing item non-response. For example, not only did the appearance (perceived legitimacy and length) influence respondents’ first impressions of the questionnaire and the decision as to whether or not to complete it, but for those respondents who had begun answering the questions, their motivation to continue was largely dependent on how difficult they perceived the task they were asked to carry out (see paper two: Dewar and Wilmot, 2006).

2.3 Token incentive payment

Likewise, the inclusion of an incentive payment was also influential. This method is often used as a way to maximise survey response and can be particularly effective when sent with the initial survey request (Dillman, 2000). To investigate this hypothesis, members of the sample were systematically assigned to one of two treatment groups: an incentive group who received a book of six first class stamps included with the first questionnaire mailing, and a control who did not receive an incentive; the result was a five percentage point increase in response for the incentive group (see paper five of this bulletin: Nicolaas, 2006). The findings support previous research which consistently reveals that a pre-paid incentive, albeit a token gesture, to complete the questionnaire, can increase survey response rates significantly (James and Bolstein, 1992; Church, 1993; Singer et al, 1999; Singer, 2002).

Therefore there is persuasive evidence as to the reasons why it is worth investing time and effort into many different aspects of survey design in order to achieve high response. However it is the researchers’ opinion that the prime factor contributing to the successful running of the survey was the implementation procedures; that is, the multiple contacts, which collectively had a greater influence on response rates than the previous attributes of the communication process (personalisation, questionnaire design, financial incentives) mentioned above.

2.4 Multiple contacts

Similar to the way interviewers carry out repeated visits to maximise the number of contacts with households in face-to-face surveys, various methodological studies have shown multiple contacts to be the most effective way of increasing response to postal surveys (Scott, 1961; Linsky, 1975; Dillman, 1991). Lynn and Thomas (forthcoming) cite the number of reminders sent as being the ‘single most powerful determinant of the response rate achieved on postal surveys’. As such, a schedule of five contacts was included in the fieldwork period for the postal survey, this consisted of:

- An advance letter to all sampled persons;
- The questionnaire with covering letter one week after the advance letter to all sampled persons;
• A thank you/reminder postcard one week after the questionnaire to all sampled persons;
• Another questionnaire with a different covering (reminder) letter two weeks after the thank you/reminder postcard to all non-responders;
• Another questionnaire with a different covering (reminder) letter two weeks after the previous contact to any remaining non-responders.

2.4.1 Contact objectives

This multiple contact approach is similar to the five-contact ‘system’ recommended by Dillman (2000) in the tailored design method for postal survey implementation. The overall objective was to encourage participation. However, each of the contacts varied in nature, and was purposively worded to reflect a different emphasis each time. This was because it was important that each contact should appear different from the last. The rationale was not only to put across a suggestion of renewed effort to communicate, but also to try to get the attention of non-responders who had clearly been indifferent to the request researchers had attempted to convey in the previous contact.

In this way, the first of the five contacts\(^2\), the advance letter, was used to notify sample members that they would be receiving a questionnaire in the post. In the second form of contact, the covering letter included with the initial questionnaire explained the nature and objective of the survey and a request to take part. The third contact was brief and to the point, used to both thank respondents who had taken part and to act as a brief reminder to those who had not yet responded. Changing the mode of contact from letter to postcard was deliberate; it was intended to provide a contrast in communication, following up the initial request. It was also hoped that the postcard would encourage non-responders to return the initial questionnaire thereby saving the cost of producing and posting out the questionnaire a second time. In the fourth type of contact, another questionnaire was sent out to replace the first, and the emphasis in the covering letter was altered slightly. For this contact, the focus was changed to try to appeal to peoples’ tendency to wish to have their opinions heard. As such, the letter stated that the views of people who had already responded had been wide-ranging and very useful, but that the success of the study depended on as many people as possible taking part; highlighting the particular importance of their response to ensure a representative cross-section of the population. In the last contact, another questionnaire was sent with a final reminder. The letter explained that the study was drawing to a close and expressed a final wish ‘to be able to include your views’. In addition to personalising the letter in this way, the importance of getting a wide response as possible was also reiterated. The letter pointed out that their views as respondents might be different from other peoples’, to imply that this was a chance for their opinion to be counted; the importance of gaining a representative sample of the population was also re-emphasised, and a deadline given for returning the questionnaire. Rather than the plain, white C4 envelopes used to send out the questionnaires in previous mailings, a brown, C4 envelope with the HMSO logo was used to ensure that the visual appearance of the final contact, as well as the content of the letter, was distinct from the other mailings. The purpose of changing the appearance of the final contact was also to demonstrate continued effort, and was a

\(^2\) Examples of the advance letter, covering letter, postcard and reminders are available at the following website address:
http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp
last attempt to convey the survey’s value, reflected in the researchers’ commitment to gaining response.

A business reply-paid, addressed envelope for respondents to use when returning completed questionnaires was included every time the questionnaire was sent out; this helped to ease respondent burden and lessen the risk that the returns would go astray in the post.

Sometimes, in the ultimate endeavour to encourage remaining non-responders to take part, sampled members are also contacted by some other mode of communication, for example, by telephone. Dillman (2000) recommends this ‘special’ form of contact as a final method of improving overall postal survey response rates. However for the Social Capital Postal Survey, the demands on resource such as time, funding and availability of telephone numbers, meant that it was not feasible in this instance.

2.4.2 Time intervals

The timing of correspondence was also a fundamental part of the contact process: the advance letter was sent one week prior to the initial questionnaire; if it had arrived too soon, recipients would have forgotten about it, and its contribution to the contact schedule would have been wasted. Likewise, careful timing of the reminders was essential, if they were sent too early it would not have allowed enough time for potential responders, for whom sending a reminder would be unnecessary, to return the questionnaire, and people may have felt harassed. Furthermore, sending out reminders too early is likely to put additional pressure on operations staff who need time to book-in completed questionnaires and identify non-responders requiring reminders. However, if reminders were sent a long time after the initial questionnaire, arguably the importance and urgency of the request would have been lost. Therefore, while it was considered appropriate to send the postcard as a (gentle) reminder a week after recipients would have received the initial questionnaire, an interval of two weeks was timetabled into the fieldwork period prior to sending each of the replacement questionnaires. As such, care was taken to only target reminders to members of the sample who had not replied. This not only reduced survey costs through savings made on postage and survey materials, but also helped to avoid annoying those respondents who had already sent in or intended on returning their questionnaires.

3. An appropriate response strategy

This postal survey strategy was effective in achieving a high co-operation rate to the Social Capital Postal Survey and its implementation is recommended for similar postal surveys. The methods identified are certainly not exclusive, however their selection was based upon the range of experience on the design and implementation of postal surveys provided in the survey literature. Other techniques, such as first class stamps for outgoing as well as return envelopes and material incentives are noted in the survey literature (Dillman, 2000; Lynn and Thomas, forthcoming) however many of these have no proven effect, while others are more likely to be used when it is not appropriate to incorporate the methods described above into the survey design. Varied techniques may well be more effective in particular survey situations; however the general concept is to tailor the methods according to circumstance.
4. Final recommendation

It is recommended that merely by incorporating multiple contacts into the survey implementation process will, alone, improve co-operation and response rates.

Furthermore, including other design factors: personalisation, respondent-friendly questionnaire design and a token financial incentive in addition to the five scheduled contacts, will provide an overall, resoundingly effective strategy for positive response which, when followed, will not only improve the quality of postal survey estimates, but will also create an efficiency saving as less resource and less time is required chasing non-returns.

References


In search of a population sampling frame for UK postal surveys

Sarah Tipping & Gerry Nicolaas, National Centre for Social Research, UK

1. Introduction

Ideally the sampling frame for a postal survey should include the whole population of interest as well as auxiliary information to allow each unit in the sample to be identified and contacted by post. For the Social Capital Postal Survey, the population of interest were all adults living in Great Britain. In order to identify and contact each adult in the sample, the sampling frame also had to include names and addresses.

Unfortunately a frame that includes the names and addresses of all adults living in Great Britain does not exist. We therefore needed to find a sampling frame that, although not ideal, would be adequate for our purposes. To assess the adequacy of sampling frames we needed to know the per cent of the population that was excluded from the frame and the characteristics of those who were excluded.

2. An overview of sampling frames

2.1 The Postcode Address File

In Great Britain, the most commonly used sampling frame for large-scale surveys of the general population is the Postcode Address File, also known as PAF.

PAF is a list of all addresses to which mail is sent, maintained by Royal Mail. The small-user residential file tends to be used for survey sampling. This file only includes addresses which receive on average less than 50 items of mail per day, thus excluding most non-residential addresses. Nevertheless, about 10% of the addresses in the small user file are non-residential; e.g. businesses, institutions, derelict, empty, second homes.

The coverage of PAF has not been assessed since the 1991 Census check (Foster, 1994). This check showed that PAF had almost complete coverage of households; 96.6% of all private households occupied on Census night and 96.9% of adults usually resident in these households, were at addresses listed on PAF.

PAF is, however, a list of addresses; it does not include names. For this reason, PAF tends to be used as a sampling frame for face-to-face surveys where an interviewer visits the sampled address, determines who lives there and then selects one or more residents for interview. The PAF is considered unsuitable for postal surveys because it is not possible to send the questionnaire to a named individual. Sending questionnaires to addresses without a name is likely to have a negative effect on response rates. Furthermore, among those who do respond, there is no control over who completes the questionnaire at the address which may introduce bias.
2.2 The Electoral Register

The Electoral Register (ER) lists the name and address of everyone in the United Kingdom who has registered to vote. Anyone aged 16 or over can register to vote if they are citizens of the United Kingdom, other European Union countries or Commonwealth countries. It is therefore an incomplete listing of all adults; it excludes adults who live in the UK but are not eligible to vote and adults who are eligible but fail to register. Despite the exclusion of these adults, the ER has until recently been considered the most suitable sampling frame for postal surveys because it was a fairly complete listing of adults with names and addresses.

This changed in October 2002 when electors were allowed to opt out of the publicly available ER, known as the edited ER. According to Experian, the proportion of adults registered on the full ER who opted out of the edited ER was 32% in September 2005. This proportion varied greatly by local authority, ranging from 4% to 75%.

In addition to the opt out rate of 32%, sampling from the edited ER will also exclude those adults who are not eligible to vote and those who are eligible but have failed to register. In 1991 this proportion was 12% according to the 1991 Electoral Register Check (Foster, 1993). Unfortunately a more up-to-date figure is not available.

It is very likely that the people who are not listed on the edited ER are different to those who are, thus introducing the risk of bias if sampling from the edited ER.

2.3 Commercial databases

Since the full ER was no longer available for marketing/sales purposes, some commercial companies in the data industry have been looking for alternatives to the edited ER; e.g. CACI’s Consumer Register, Experian’s National Canvesse, Equifax’ ConnectSelect, EuroDirect’s Data Exchange. These companies supplemented the edited ER with data pooled from other sources to produce commercial databases with greater coverage than the edited ER. For example, CACI’s Consumer Register (CR) starts with the edited ER which is then supplemented with data from Claritas, DataWorks, Bounty’s market leading mother and baby file, and large mail order databases. It includes names and addresses as well as information which could possibly be used for sampling, such as modelled age and income. CACI update the CR every three months. The CR has almost 40 million names and addresses of UK adults.

Although CACI and these other companies claim to have almost complete coverage of UK households, there is no empirical evidence that they have adequate coverage of adults which would make it suitable for probability sampling.

---

1 Although it is possible to register when aged 16 and 17, it is not possible to vote unless aged 18 or over.
3. Methodology

3.1 Introduction

For reasons outlined in sections 1 and 2, both the PAF and the Edited ER fall short of the requirements for a postal survey sampling frame. With no ideal alternative to previous postal survey practice using the ER, the Office for National Statistics (ONS) and National Centre for Social Research (NatCen) took the opportunity to investigate the feasibility of using the CR and comparing the two, using the PAF as the standard. This was carried out using a representative sample drawn from the PAF for the ONS Omnibus survey and checking the extent to which responding sample members were covered by the edited ER and the CR.

3.2 Design of ONS Omnibus survey

The ONS Omnibus survey is a multipurpose survey developed by the ONS for use by non-profit making organisations, including government departments, public bodies, charities and academics. Each month interviews are conducted with approximately 1,200 individuals living in private households in Great Britain.

Using the ‘small user’ PAF as a sampling frame, 67 postcode sectors are sampled and 30 addresses are selected from within them so that, in total, 2,010 addresses are issued to field each month.

If an address contains more than one household, the interviewer uses a standard ONS procedure to select just one household at random. The interviewer then selects one person aged 16 or over at random from each household and invites them to take part in the interview. A full Omnibus interview is achieved at 65% of eligible households on average.

Further details about the design of the survey can be found on the National Statistics website: http://www.statistics.gov.uk/about/services/omnibus/default.asp.

3.3 How the matching was carried out

All sampled Omnibus addresses for the months of April, May and June 2005 were sent to CACI (6,030 addresses) and matched to one or more names listed on the edited ER and the CR. The matching rate at the address level was 81% for the edited ER and 90% for the CR.
The next step was to collect the names of the adults actually living at the sampled addresses so that these could be compared with the matched names from the edited ER and the CR. Only the first name of the adults living at the sampled addresses was collected by ONS interviewers working on the Omnibus survey. However, it was not possible for the interviewers to collect this information from non-responding addresses; i.e. addresses where no contact was made during the fieldwork period and addresses where contact was made but the resident refused to take part in the survey. For the purpose of this study we have assumed that the accuracy of the matched names would be similar for responding and non-responding addresses.

ONS then compared the names found by the Omnibus interviewers at responding addresses with the names listed on the edited ER and the CR. The edited ER and the CR only contain people aged 18 and over; therefore any person aged under 18 was not included in the comparison.

The comparison was made on address and first name only. ONS Omnibus interviewers were not able to collect full names because of concerns about response rates and data confidentiality. For a small number of cases (263), the interviewers were not able to record the first name. Instead the interviewer coded household members as Mr, Mrs, sister, son, husband, wife, or by a number (1, 2, etc). These cases were excluded from the analysis.

As mentioned in section 3.2, if a sampled address contained more than one household, the interviewers were instructed to select one of these households at random and no information was collected about the other households. Consequently, a small number of names on the edited ER/CR that were coded as ineligible may in fact belong to people living in households that were not selected at the address. The
effect this will have on the matching rates is negligible because the proportion of multi-household addresses in Britain is less than 1%\(^2\).

### 3.4 Matching codes

A number of codes were used to record how well the names from the edited ER and the CR matched the names of adults actually found at the responding addresses by the Omnibus interviewers. These codes distinguish between identical names (exact matches), names with alternate spellings, truncated versions and abbreviations (likely matches) and names which were completely different (mis-matches). The full range of individual matching codes is given in Table 1.

#### Table 1 Individual matching outcome codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exact match on name</td>
</tr>
<tr>
<td>2</td>
<td>Likely match - typographical error (e.g. Caroline and Carline)</td>
</tr>
<tr>
<td>3</td>
<td>Likely match - alternate spelling (e.g. Caroline and Carolyn)</td>
</tr>
<tr>
<td>4</td>
<td>Likely match - prefix of full name (e.g. Alex and Alexander). Includes names where the first letter of the name is recorded in either the Omnibus or the CR/ER but the full name corresponding to that letter is recorded on the other (e.g. L and Louise) - between 20 to 30 cases were found each month.</td>
</tr>
<tr>
<td>5</td>
<td>Likely match - abbreviation which is not a prefix of full name (e.g. Richard and Dick)</td>
</tr>
<tr>
<td>6</td>
<td>Incorrect match - individual recorded in the CR/ER but not listed in the Omnibus household grid</td>
</tr>
<tr>
<td>7</td>
<td>Incorrect match - individual listed in Omnibus household grid but not recorded in the CR/ER</td>
</tr>
<tr>
<td>8</td>
<td>Incorrect match – CR/ER and Omnibus records different</td>
</tr>
</tbody>
</table>

The ONS team coded all adults within responding households. The team then checked for any codes which could not co-exist and amended any anomalies. Further quality assurance of the coding and matching exercise was then carried out by a member of ONS staff who had not been involved in the matching exercise.

\(^2\) According to ONS figures, 0.7% of households in Britain are multi-households.
4. Results

4.1 Introduction

There were 3,779 responding households, according to the Omnibus data. These households contained 6,786 adults aged 18 and over, of which 6,523 adults provided their first names and were therefore included in the analysis. The results of matching these individuals to the edited ER and the CR are shown in Figures 2 and 3 respectively.

Figure 2  Results of matching Omnibus data to the edited ER
4.2 Over-coverage

The matched records from the edited ER and the CR included the names of people who were not found at their listed addresses; 18% and 19% respectively (see Table 2). This means that about one in five postal questionnaires will not reach the named person if the sample is drawn from either the edited ER or the CR.

Table 2  Proportion of adults not found at listed addresses

<table>
<thead>
<tr>
<th>Adults matched to the edited ER</th>
<th>Adults matched to the CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not found at listed address</td>
<td>861</td>
</tr>
<tr>
<td></td>
<td>(18%)</td>
</tr>
<tr>
<td></td>
<td>1,050</td>
</tr>
<tr>
<td></td>
<td>(19%)</td>
</tr>
<tr>
<td>Found at listed address</td>
<td>3,857</td>
</tr>
<tr>
<td></td>
<td>(82%)</td>
</tr>
<tr>
<td></td>
<td>4,436</td>
</tr>
<tr>
<td></td>
<td>(81%)</td>
</tr>
<tr>
<td>Total adults aged 18+ on register</td>
<td>4,718</td>
</tr>
<tr>
<td></td>
<td>5,486</td>
</tr>
</tbody>
</table>
Possible causes for not finding these people at their listed address include delays in recording changes of address, delays in removing people who have died, and duplicate listings of people. Only those people who have moved address should be classified as eligible; those people who have died or whose names appear in duplicate listings should be classified as ineligible; i.e. over-coverage.

Unfortunately we do not know how many of these people are eligible. If we assume that they are all eligible (i.e. they are all movers), then the rate of over-coverage is zero. If we assume that none of them are eligible (i.e. they are all duplicates or deceased), then the rate of over-coverage is 18% for the edited ER and 19% for the CR.

Given the methods for producing the edited ER, we would expect a very large proportion of its cases to be eligible; i.e. people listed on the edited ER but not found at the listed address are more likely to be movers than duplicates and deceased. As described in section 2.3, the CR consists mainly of the edited ER but also includes additional cases found from other data sources. We would expect to find more over-coverage amongst these additional cases compared to the edited ER.

Therefore, for the purpose of illustration, it is assumed that the proportion of eligibles among the unmatched names is about ninety to ninety five per cent for the edited ER and about ten to fifty per cent for the additional cases on the CR. Thus the corresponding rates of over-coverage (shown in Table 3); range from one per cent to two per cent for the edited ER and about three per cent to five per cent for the CR.

### Table 3 Estimated rates of over-coverage

<table>
<thead>
<tr>
<th></th>
<th>Over-coverage on the edited ER</th>
<th>Over-coverage on the CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>0% - 18%</td>
<td>0% - 19%</td>
</tr>
</tbody>
</table>

**Proportion of people not found at listed address who are eligible:**

- 90% on edited ER: 2%
  - and 10% on CR++: 5%
  - and 25% on CR++: 4%
  - and 50% on CR++: 3%

- 95% on edited ER: 1%
  - and 10% on CR++: 4%
  - and 25% on CR++: 3%
  - and 50% on CR++: 3%

* CR+ stands for those cases on the CR that were not on the edited ER.
4.3 Under-coverage of adults

According to the results of the matching exercise, 41% of the adults found at the responding sampled addresses were not listed on the edited ER and 32% of them were not listed on the CR (see Table 4). These results imply that a sample of adults drawn from the edited ER would exclude two out of five adults in Great Britain, whereas a sample of adults drawn from the CR would exclude roughly one in three. Exclusion of this kind is regarded as under-coverage.

### Table 4  Adults in Omnibus survey and whether they were on the edited ER and the CR at the listed addresses

<table>
<thead>
<tr>
<th>Adults in Omnibus survey</th>
<th>On edited ER at listed address:</th>
<th>On CR at listed address:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>3,857 (59%)</td>
<td>2,666 (41%)</td>
</tr>
</tbody>
</table>

Possible causes for people not being on the edited ER at the listed address include delays in recording new addresses, not being eligible to vote, and failing to register to vote. These people still have a chance of being included on the CR if they are included in one or more of the data sources that are combined with the edited ER to make up the CR database.

4.4 Variation in the coverage of adults

Table 4 in section 4.3 indicates that using the CR to draw a sample of adults would provide slightly better coverage of adults than the edited ER; 68% compared to 59%. However, further analysis is required to assess whether the extra 9% on the CR would improve the sample composition compared to PAF. If so, the CR would be preferable to the edited ER as a sampling frame for postal surveys of the general population.

In order to compare sample compositions of the edited ER and the CR with PAF, we analysed interview data from the Omnibus survey. A wide range of socio-demographic characteristics were used in the comparison; i.e. age, sex, ethnicity, marital status, household type, employment status, occupation, tenure, highest education attainment and whether or not the respondent had a limiting long term illness.

Table 5 shows the differences in sample composition between the responding samples of the PAF, edited ER, and the CR. It can be seen that individuals listed on the edited ER and CR are less likely to be in the 18-24 age group, to rent their accommodation or have a degree. There were no differences by sex but there was some indication that
people belonging to minority ethnic groups were also less likely to appear on the edited ER and CR.

Table 5  Characteristics of individuals matched to the CR and edited ER

<table>
<thead>
<tr>
<th></th>
<th>PAF sample (%)</th>
<th>Edited ER (%)</th>
<th>CR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>45</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>Aged 18-24</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Non-white</td>
<td>7</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Single, never married</td>
<td>24</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Married</td>
<td>48</td>
<td>51</td>
<td>52</td>
</tr>
<tr>
<td>Lone parent</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Renting</td>
<td>27</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>Has degree</td>
<td>18</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Elderly retired household</td>
<td>27</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>Worked in reference week</td>
<td>53</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Routine/manual occupations</td>
<td>40</td>
<td>43</td>
<td>42</td>
</tr>
<tr>
<td>No car</td>
<td>23</td>
<td>22</td>
<td>21</td>
</tr>
</tbody>
</table>

Base 3,584 2,128 2,444

Some of these characteristics are related to each other. For example, it is possible that the observed relationship with renting is, in part, an age effect. In order to distinguish the separate influences of age and renting on edited ER registration and CR registration, it was necessary to carry out multivariate analysis, i.e. logistic regression.

Two separate logistic regression models were fitted, one for the edited ER and the other for the CR. The dependent variable for each model was coded 1 if the Omnibus respondent was listed on the register and 0 if not. The models were run using the social demographic variables as predictors.

The results confirm that both the edited ER and the CR under-represent younger people, especially those aged between 18 and 24 and people who rented their accommodation from a private landlord. According to the fitted models, the edited ER under-represents those in non-manual occupations whereas the CR under-represents those who continue with full-time education beyond the age of 16. It is very likely that these two variables are measuring the same effect; i.e. the better educated tend to have non-manual occupations. Finally, the results show that men were less likely than women to be listed on the CR. The full logistic regression models are shown in Tables A1 and A2 in Appendix A.

To conclude, the sample composition of the CR is not much closer to the PAF than that of the edited ER. Similar groups are under-represented by both registers.
5. Conclusion

The main purpose of this study was to select a sampling frame for the Social Capital Postal Survey which would minimise differences in population coverage when compared to an equivalent face-to-face survey that used the PAF as a sampling frame. We assessed two possible sampling frames for postal surveys: the edited ER and CACI’s CR.

Unfortunately the results show that both registers have inadequate coverage of adults aged 18 and over living in Great Britain (59% and 68% respectively) and both registers under-represent specific population sub-groups: young adults aged 18 to 24, people who rent their accommodation from a private landlord and people who had either left full-time education by the age of 16 (edited ER) or were in non-manual occupations (CR). These characteristics are likely to be related to social capital estimates. Consequently, sampling from either the edited ER or CR is likely to affect comparability of local authority postal survey data and national face-to-face survey data.

We only assessed one commercial database; CACI’s Consumer Register. However, we believe that the results would be similar for other commercial databases. The companies that produce these databases tend to supplement the edited ER with personal information from other data sources such as mail order databases. However, people who opt out from the edited ER are also less likely to allow these companies to include their personal information in these databases. Hence, the addition of these other data sources is unlikely to reduce bias.

In the absence of a single comprehensive sampling frame of named individuals for use in postal surveys of the general population, other solutions to this problem should be explored. If PAF is shown to be the only viable sampling frame for a postal survey, then methods for reducing bias need to be explored; e.g. methods for increasing response to unpersonalised mailings and self-administered techniques for respondent selection. The cost-effectiveness of mixed-mode designs for increasing access to all units selected from the PAF could also be explored; e.g. postal questionnaires for those addresses with matched names from the edited ER and face-to-face interviews for those addresses without (although mode effects would have to be considered; see paper six of this bulletin (Nicolaas and Tipping, 2006). Alternatively, the feasibility of combining two or more sampling frames to improve coverage for postal surveys could be investigated.

References


3 Permission to use personal information for commercial purposes is required under the Data Protection Act 1998.
### Appendix A: Results of logistic regression

**Table A1**  Results of logistic regression on Omnibus respondents matched to edited ER

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respondent age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 to 24</td>
<td>-1.31</td>
<td>0.193</td>
<td>46.2</td>
<td>1</td>
<td>0.00</td>
<td>0.27</td>
</tr>
<tr>
<td>25 to 44</td>
<td>-0.59</td>
<td>0.136</td>
<td>19.1</td>
<td>1</td>
<td>0.00</td>
<td>0.55</td>
</tr>
<tr>
<td>45 to 54</td>
<td>-0.35</td>
<td>0.147</td>
<td>5.7</td>
<td>1</td>
<td>0.02</td>
<td>0.70</td>
</tr>
<tr>
<td>55 to 64</td>
<td>-0.09</td>
<td>0.140</td>
<td>0.4</td>
<td>1</td>
<td>0.51</td>
<td>0.91</td>
</tr>
<tr>
<td>65 to 74</td>
<td>-0.15</td>
<td>0.144</td>
<td>1.1</td>
<td>1</td>
<td>0.29</td>
<td>0.86</td>
</tr>
<tr>
<td>75 and over (baseline)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NS-SEC (3 categories)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managerial and professional occupations</td>
<td>22.2</td>
<td>3</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate occupations</td>
<td>0.09</td>
<td>0.099</td>
<td>0.8</td>
<td>1</td>
<td>0.37</td>
<td>1.09</td>
</tr>
<tr>
<td>Routine and manual occupations</td>
<td>0.38</td>
<td>0.088</td>
<td>18.3</td>
<td>1</td>
<td>0.00</td>
<td>1.46</td>
</tr>
<tr>
<td>Not classified</td>
<td>-0.01</td>
<td>0.180</td>
<td>0.0</td>
<td>1</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Tenancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owns outright</td>
<td>1.44</td>
<td>0.153</td>
<td>89.0</td>
<td>1</td>
<td>0.00</td>
<td>4.23</td>
</tr>
<tr>
<td>Owns mortgage</td>
<td>1.59</td>
<td>0.143</td>
<td>123.3</td>
<td>1</td>
<td>0.00</td>
<td>4.88</td>
</tr>
<tr>
<td>Rents LA/HA</td>
<td>1.07</td>
<td>0.156</td>
<td>47.3</td>
<td>1</td>
<td>0.00</td>
<td>2.92</td>
</tr>
<tr>
<td>Rents privately</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-0.72</td>
<td>0.237</td>
<td>9.3</td>
<td>1</td>
<td>0.00</td>
<td>0.48</td>
</tr>
</tbody>
</table>

1. The response is 1=Omnibus respondent was matched to edited ER, 0=no match
2. The model R² is 0.080 (Cox and Snells)
3. B is the estimated co-efficient with standard error SE
4. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom. If the test is significant (sig<0.050) then the categorical variable is considered to be ‘significantly associated’ with the response variable.
5. The Wald level for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.
### Table A2  Results of logistic regression on Omnibus respondents matched to CR

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 to 24</td>
<td>-1.24</td>
<td>0.235</td>
<td>27.8</td>
<td>1</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>25 to 44</td>
<td>-0.61</td>
<td>0.173</td>
<td>12.4</td>
<td>1</td>
<td>0.00</td>
<td>0.54</td>
</tr>
<tr>
<td>45 to 54</td>
<td>-0.26</td>
<td>0.182</td>
<td>2.0</td>
<td>1</td>
<td>0.15</td>
<td>0.77</td>
</tr>
<tr>
<td>55 to 64</td>
<td>-0.02</td>
<td>0.174</td>
<td>0.0</td>
<td>1</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>65 to 74</td>
<td>0.01</td>
<td>0.171</td>
<td>0.0</td>
<td>1</td>
<td>0.94</td>
<td>1.01</td>
</tr>
<tr>
<td>75 and over</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>7.3</td>
<td>1</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.21</td>
<td>0.079</td>
<td>7.3</td>
<td>1</td>
<td>0.01</td>
<td>0.81</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age respondent left full time education</td>
<td>22.0</td>
<td>3</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 or under</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 - 20</td>
<td>-0.22</td>
<td>0.103</td>
<td>4.7</td>
<td>1</td>
<td>0.03</td>
<td>0.80</td>
</tr>
<tr>
<td>21+</td>
<td>-0.42</td>
<td>0.105</td>
<td>16.3</td>
<td>1</td>
<td>0.00</td>
<td>0.65</td>
</tr>
<tr>
<td>Missing</td>
<td>-0.87</td>
<td>0.299</td>
<td>8.9</td>
<td>1</td>
<td>0.00</td>
<td>0.41</td>
</tr>
<tr>
<td>Tenancy</td>
<td>171.5</td>
<td>3</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owns outright</td>
<td>1.76</td>
<td>0.156</td>
<td>127.4</td>
<td>1</td>
<td>0.00</td>
<td>5.83</td>
</tr>
<tr>
<td>Owns mortgage</td>
<td>1.72</td>
<td>0.143</td>
<td>145.1</td>
<td>1</td>
<td>0.00</td>
<td>5.60</td>
</tr>
<tr>
<td>Rents LA/HA</td>
<td>1.14</td>
<td>0.156</td>
<td>53.3</td>
<td>1</td>
<td>0.00</td>
<td>3.11</td>
</tr>
<tr>
<td>Rents privately</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.99</td>
<td>0.351</td>
<td>8.0</td>
<td>1</td>
<td>0.00</td>
<td>0.37</td>
</tr>
</tbody>
</table>

1. The response is 1=Omnibus respondent was matched to CR, 0=no match
2. The model R² is 0.126 (Cox and Snells)
3. B is the estimated co-efficient with standard error SE
4. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom. If the test is significant (sig<0.050) then the categorical variable is considered to be ‘significantly associated’ with the response variable.
5. The Wald level for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.
Population coverage and response rates for a postal survey on social capital

Gerry Nicolaas, National Centre for Social Research, UK

1. Introduction

This paper describes the population coverage and response rate for the Social Capital Postal Survey. This survey was carried out in parallel with the Office for National Statistics’ (ONS) Omnibus survey which included the same social capital questions as the postal survey. The purpose of the parallel run was to compare data collected using a self-administered questionnaire with an interviewer-administered instrument.

To aid comparison between the postal survey and the Omnibus survey, addresses for the postal survey were selected using the same sampling frame (‘small user’ Postcode Address File (PAF)) and the same 201 Primary Sampling Units (postal sectors) as the Omnibus survey for the months April, May and June 2005. Within each postal sector, 35 addresses were selected at random. In order to maximise response rates and to ensure that a random adult in the household completed the questionnaire, names of adults aged 18 and over were matched to the sampled PAF addresses using CACI’s Consumer Register (CR). Given this sample design, the target population for this survey can be defined as the adult population aged 18 and over living in private households in Great Britain (GB). (See paper six for a description of the sample design: Nicolaas and Tipping, 2006).

CACI matched the sample of 7,035 PAF addresses to the CR and found one or more names for 6,373 addresses. This matching rate of 91% was a lot higher than the anticipated rate of 78%. To keep costs within budget, we therefore sent questionnaires to a random selection of 5,460 addresses with matched names; i.e. a sub-sample of 85.67%. For those addresses with one or more matched names, one adult was selected at random.

The matching rate and the outcome of the mail-out are shown in Table 1.
### Table 1  Matching and outcome rates for the postal survey

<table>
<thead>
<tr>
<th>Outcome of matching names to addresses and mail-out</th>
<th>N</th>
<th>N</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial Sample of PAF addresses:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not matched to CR name(s)</td>
<td>662</td>
<td>567</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Matched to CR name(s)</td>
<td>6,373</td>
<td>5,460</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td><strong>Mail-out</strong> (86% sub-sample of addresses with matched name):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No reply</td>
<td>1,686</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replies</td>
<td>3,774</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Replies:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineligible PAF address</td>
<td>21</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person unknown (incorrect name matched to eligible address)</td>
<td>570</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusals and other non-response</td>
<td>377</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completed questionnaire</td>
<td>2,806</td>
<td>74</td>
<td>51</td>
<td></td>
</tr>
</tbody>
</table>

In order to simplify the presentation of results in the remainder of this paper, all calculations are based on the assumption that the starting sample was (7,035 * 85.67%), which is 6,027.

### 2. Eligibility

In order to calculate the response rate, it is necessary to determine the eligibility status of all sampled cases. This can be complicated for postal surveys because the status of no replies is difficult to ascertain. The additional step of matching names to the sample of addresses has complicated this even further for this particular postal survey.

However, the addresses for the postal survey were selected in the same postal sectors as the ONS Omnibus survey and the study also included a coverage check of the CACI Consumer Register which is described in paper four of this bulletin (Tipping and Nicolaas, 2006). Consequently the proportion of ineligible addresses in the sample of the postal survey can be assumed to be the same as the Omnibus sample and the results of the coverage check can be used to estimate the proportion of eligible cases among addresses with matched name(s), addresses without a matched name and addresses that did not reply to the mail-out.

The following information from the ONS Omnibus survey and the CR Coverage Check was used to estimate the proportion of eligible and ineligible cases for the postal survey (see Table 2):

- According to the ONS Omnibus survey, 9% of the sampled addresses were ineligible.
- Results from the coverage check showed that about 70% of addresses that had not been matched to a name on the CR were in fact eligible (i.e. residential address with at least one adult aged 18 and over).
• Results from the coverage check showed that just under 7% of the addresses that were matched to names on the CR were ineligible (e.g. non-residential addresses).

Having estimated the number of ineligible addresses in the mail-out (360) and knowing how many ineligible addresses there were among the replies (21), it is then possible to calculate the number of eligible and ineligible addresses among the ‘no replies’ (1,347 and 339 respectively).

### Table 2 Eligibility of sampled addresses

<table>
<thead>
<tr>
<th>Outcome</th>
<th>All</th>
<th>Eligible</th>
<th>Ineligible</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial Sample of PAF addresses:</strong></td>
<td>6,027</td>
<td>5,495</td>
<td>532</td>
</tr>
<tr>
<td>(100%)</td>
<td>(91%)</td>
<td>(9%)</td>
<td></td>
</tr>
<tr>
<td>Not matched to CR name(s)</td>
<td>567</td>
<td>395</td>
<td>172</td>
</tr>
<tr>
<td>(100%)</td>
<td>(70%)</td>
<td>(30%)</td>
<td></td>
</tr>
<tr>
<td>Matched to CR name(s)</td>
<td>5,460</td>
<td>5,100</td>
<td>360</td>
</tr>
<tr>
<td>(100%)</td>
<td>(93%)</td>
<td>(7%)</td>
<td></td>
</tr>
<tr>
<td><strong>Mail-out (issue sample):</strong></td>
<td>5,460</td>
<td>5,100</td>
<td>360</td>
</tr>
<tr>
<td>(100%)</td>
<td>(93%)</td>
<td>(7%)</td>
<td></td>
</tr>
<tr>
<td>No reply</td>
<td>1,686</td>
<td>1,347</td>
<td>339</td>
</tr>
<tr>
<td>(100%)</td>
<td>(80%)</td>
<td>(20%)</td>
<td></td>
</tr>
<tr>
<td>Replies</td>
<td>3,774</td>
<td>3,753</td>
<td>21</td>
</tr>
<tr>
<td>(100%)</td>
<td>(99%)</td>
<td>(1%)</td>
<td></td>
</tr>
</tbody>
</table>

### 3. Non-coverage

In section 1 we defined the target population as the adult population aged 18 and over living in private households in GB. However, some units in this target population had either no chance of selection for this survey or if their address was selected were not included in the survey for other reasons. This is referred to as ‘non-coverage’.

Given the sample design for this survey, there are three possible reasons for non-coverage:

- According to the last coverage check of the PAF which was carried out in 1991 (Foster, 1994), about 3% of private households in GB were not covered by the PAF. It is unknown whether this proportion has changed since 1991 but it is likely to be of a similar magnitude.

- According to the CR coverage check, just under 7% of eligible addresses did not have any names listed on the CR. Consequently no questionnaires were sent to adults living at eligible addresses which were not matched to the CR.

- According to the coverage check, 19% of the names on the CR that were matched to eligible PAF addresses did not belong to any individuals living at the address. Although questionnaires were sent to these addresses, they would have been either ignored or returned to National Centre for Social Research (NatCen) as ‘unknown person living at the address’.
For the purpose of this study, we have ignored the non-coverage introduced by using the PAF as a sampling frame because we do not have an up-to-date figure, the figure is likely to be very small and the last coverage check of the PAF (Foster, 1994) showed that the non-coverage bias of the PAF is very small. However, the amount of non-coverage caused by matching names from the CR to the PAF addresses is very high and cannot be disregarded. Table 3 shows that about 26% of the eligible addresses in the sample were not included in the survey. The main cause for this high rate is the large number of addresses matched to incorrect names on the CR. The consequences of this high non-coverage rate are discussed in section 5.

Table 3  Non-coverage in the initial sample of PAF addresses

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial sample of PAF addresses:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineligible addresses</td>
<td>532</td>
<td>9</td>
</tr>
<tr>
<td>Eligible addresses</td>
<td>5,495</td>
<td>91</td>
</tr>
<tr>
<td>Eligible addresses in initial sample:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not matched to CR name(s)</td>
<td>395</td>
<td>7</td>
</tr>
<tr>
<td>Matched to incorrect CR name(s)</td>
<td>1,052</td>
<td>19</td>
</tr>
<tr>
<td>Matched to correct CR name(s)</td>
<td>4,048</td>
<td>74</td>
</tr>
</tbody>
</table>

4. Response rates

The response rate should indicate how many interviews were achieved as a proportion of those eligible for the survey (Lynn et al, 2001). The calculation of this response rate for postal surveys can be complicated because the eligibility status of ‘no replies’ can be difficult to ascertain. For this reason, the response rates for postal surveys are often calculated by dividing the total number of completed questionnaires by the total number of questionnaires that were mailed out (i.e. the issued sample size), irrespective of whether these questionnaires were sent to eligible or ineligible cases. According to this calculation, the response rate for the Social Capital Postal Survey is 51% (see Table 1).

\[
\text{Response rate 1} = \frac{\text{Number of completed questionnaires (2,806)}}{\text{Issued sample size (5,460)}} \times 100\% = 51\%
\]

However, for this postal survey we have been able to estimate the total number of eligible cases using outcome data from the Omnibus survey. It is therefore possible to calculate a response rate that reflects the percentage of the intended sample, irrespective of whether the addresses were matched to a name on the CR or not. Using information from the Omnibus survey, we have estimated that the number of eligible cases in the initial PAF sample was 5,495 (see Table 2). These figures produce a response rate of 51% which, by coincidence, is the same as the response rate based on the issued sample size.

\[
\text{Response rate 2} = \frac{\text{Number of completed questionnaires (2,806)}}{\text{Number of eligible units in sample (5,495)}} \times 100\% = 51\%
\]

In addition to the response rate, useful information about the quality of the survey process is provided by the contact rate and the co-operation rate.
For the purpose of this postal survey, we have defined those contacted during the three month fieldwork period as those who are likely to have received a questionnaire because a correct name had been matched to an eligible address. The contact rate measures the proportion of all those contacted and is therefore an indicator of the quality of the matching of names from the CR to eligible addresses. According to this definition, the contact rate for this survey is:

\[
\text{Contact rate} = \frac{\text{Number of eligible addresses with correct name (4,048)}}{\text{Number of eligible units in sample (5,495)}} \times 100\% = 74\%
\]

The reasons for this low contact rate are (1) the failure to find names on the CR for a proportion of the eligible addresses and (2) the matching of incorrect names to a proportion of eligible addresses.

The co-operation rate indicates the number of achieved interviews as a proportion of those contacted during the fieldwork period (Lynn et al, 2001). This rate is useful for evaluating the effectiveness of various survey procedures such as multiple contacts, the use of incentives, the persuasiveness of the letters and user-friendly questionnaires. We estimated that the co-operation rate was 69% for the postal survey. This is quite high for a postal survey and compares well with the co-operation rate of 74% for the Omnibus survey which was carried out face-to-face in the same postcode sectors and at the same time as the postal survey.

\[
\text{Co-operation rate} = \frac{\text{Number of completed questionnaires (2,806)}}{\text{Number of eligible addresses with correct name (4,048)}} \times 100\% = 69\%
\]

Table 4  Outcome categories

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>%</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial sample of PAF addresses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineligible address</td>
<td>532</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligible address</td>
<td>5,495</td>
<td>91</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td><strong>Non-coverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligible address without name on CR</td>
<td>395</td>
<td>7</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>Eligible address with incorrect name on CR</td>
<td>1,052</td>
<td>1,447</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td><strong>Eligible address with correct name on CR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No reply</td>
<td>868</td>
<td>16</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Refusals and other non-response</td>
<td>377</td>
<td>7</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Completed questionnaire</td>
<td>2,806</td>
<td>51</td>
<td></td>
<td>69</td>
</tr>
</tbody>
</table>

5. The impact of non-coverage and non-response on sample representativeness

Table 5 compares the age and sex distribution of the postal sample with the corresponding distribution for the Omnibus survey and the mid-2004 population estimates for Great Britain. Looking first at the comparison with the mid-2004 population estimates, note that the contact rate is equivalent to the coverage rate of 74% calculated in section 3; Table 3.

---

1 Note that the contact rate is equivalent to the coverage rate of 74% calculated in section 3; Table 3.
population estimates, the postal survey under-represents men and young adults aged 18 to 34 and over-represents adults aged 45 and over. However, the age and sex distribution of the postal survey is no different than that of the Omnibus survey, except that the Omnibus survey appears to be slightly better at representing adults aged 25 to 34 and women in the older age groups.

Table 5  Age and sex distribution of the responding postal sample compared with population estimates for Great Britain and the Omnibus survey

<table>
<thead>
<tr>
<th></th>
<th>Postal sample April-June 2005 (%)</th>
<th>Omnibus sample April-June 2005 (%)</th>
<th>Mid-2004 population estimates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>25-34</td>
<td>5</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>35-44</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>45-54</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>55-64</td>
<td>9</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>65 and over</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>All men aged 18 or over</td>
<td>45</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>25-34</td>
<td>7</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>35-44</td>
<td>10</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>45-54</td>
<td>11</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>55-64</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>65 and over</td>
<td>13</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>All women aged 18 or over</td>
<td>55</td>
<td>54</td>
<td>52</td>
</tr>
<tr>
<td><strong>Base (weighted for unequal selection probabilities)</strong></td>
<td>2,740</td>
<td>3,679</td>
<td>45,340,000</td>
</tr>
</tbody>
</table>

Sources: Office for National Statistics, General Register Office for Scotland. The mid-2005 population estimates were not available at the time of writing the paper.

Up-to-date population estimates for other socio-demographic characteristics are not available. However, we can compare various distributions for the postal survey with the Omnibus survey. Table 6 shows that there are some significant differences in socio-demographics between the postal survey and the Omnibus survey. The postal sample respondents were more likely to be from a white ethnic background, to be married, to be educated to a higher level and to be owner-occupiers. They also were far more likely to be suffering from limiting long-term illness. A likely explanation for these differences is the differential non-coverage and non-response bias between the two surveys.
### Table 6  Distributions of socio-demographic characteristics for the postal survey and the Omnibus survey

<table>
<thead>
<tr>
<th></th>
<th>Postal sample April-June 2005 (%)</th>
<th>Omnibus sample April-June 2005 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-white</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Single, never married</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>Married</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>Higher education qualification</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td>Worked in reference week</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>Limiting long-term illness</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td>Owner-occupiers</td>
<td>79</td>
<td>76</td>
</tr>
<tr>
<td><strong>Base</strong></td>
<td>2,805</td>
<td>3,680</td>
</tr>
</tbody>
</table>

The main aim of running the postal survey and the Omnibus survey in parallel was to determine whether any differences in social capital estimates could be attributed to mode, and not to sample composition. For this reason, propensity score matching was used to eliminate the observed differences in sample composition.

6. The impact of incentives on the response rate of the Social Capital Postal Survey

It has become common place for government sponsored face-to-face surveys in GB to include a book of postage stamps with the advance letter. Studies have shown that this can increase response by a few percentage points: an experiment on the Family Resources Survey showed an increase in response from 67.0% to 70.4% (McConaghy and Beerten, 2003). It was therefore agreed that the postal survey should include an experiment to test whether a book of stamps would have a similar effect on its response.

The sample of names and addresses were systematically assigned to one of two treatment groups:

1. Control group - no incentive.
2. Incentive group – inclusion of a book of six first class stamps with the first questionnaire mailing.

The analysis was based on the sample of addresses that were sent a questionnaire. Table 7 shows that the incentive significantly increased the percentage of completed questionnaires from 49% to 54%. This increase of 5% was mainly due to a lower rate of non-returns (29% compared with 33%).

---

2 A description of propensity score matching can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp
Table 7  Outcome category by treatment group

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Treatment group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No incentive (%): Book of six 1st class stamps (%):</td>
</tr>
<tr>
<td>No reply</td>
<td>33 29</td>
</tr>
<tr>
<td>Replies:</td>
<td></td>
</tr>
<tr>
<td>Ineligible PAF address</td>
<td>0 0</td>
</tr>
<tr>
<td>Person unknown at address</td>
<td>10 11</td>
</tr>
<tr>
<td>Refusals and other non-response</td>
<td>7 6</td>
</tr>
<tr>
<td>Completed questionnaire</td>
<td>49 54</td>
</tr>
<tr>
<td>Base (issued sample)</td>
<td>2,730 2,730</td>
</tr>
</tbody>
</table>

Pearson Chi-Square=0.001

7. Conclusion

The response rate for the postal survey was much lower than that for the Omnibus survey; 51% compared to 69%. This was not unexpected because it is generally assumed that face-to-face surveys will achieve higher response rates than equivalent postal surveys because of the valuable role the interviewer plays in persuading target respondents to take part in the survey.3

However, the main reason for the difference in response rate between these two surveys was the high non-coverage rate for the postal survey rather than a higher refusal rate. If we combine the ‘no replies’ among the eligible addresses with the refusals, then the refusal rate for the postal survey was 23% (see Table 8). This compares with a refusal rate of 25% for the Omnibus survey.

The high non-coverage rate for the postal survey was caused by the matching of names from CACI’s Consumer Register to the sample of addresses. About one in four eligible addresses were either not sent a questionnaire because the address had not been matched to a name (8%) or a questionnaire was sent but this was addressed to someone not living at the address (19%). The non-coverage rate on the Omnibus survey was not measured but was assumed to be negligible.

3 It should be noted that the Omnibus survey is a multi-purpose survey comprising a number of different modules on different topics for different sponsors. The Social Capital or ‘Neighbourhood’ Survey focussed specifically on neighbourhood issues and was considerably shorter in length than the Omnibus.
In contrast, the co-operation rate of the postal survey compares extremely well with that of the Omnibus survey. Among those assumed to have received a postal questionnaire (i.e. questionnaire addressed to an individual living at an eligible address), 69% returned a completed questionnaire (see Table 8). This compares to 74% of Omnibus sample members contacted by an interviewer. This high cooperation rate for the postal survey probably reflects the salience of the survey topic and the efforts made to persuade people to complete the questionnaire.

These results suggest that if a sampling frame with good coverage of the study population is available, then it should be possible to achieve acceptable response rates for postal surveys provided the survey topic is salient to the population of interest, the documents are well-designed (e.g. user-friendly, not too long) and non-responding sample members are contacted several times to encourage participation (see paper three for further details: Dewar, 2006). Furthermore, we have demonstrated that it is possible to increase postal survey response rates by several percentage points using a relatively inexpensive incentive sent with the first questionnaire.

References


Mode effects in social capital surveys

Gerry Nicolaas & Sarah Tipping, National Centre for Social Research, UK

1. Introduction

The main aim of this project was to design a postal self-administered questionnaire that would produce comparable data to the Office for National Statistics’ (ONS) harmonised interviewer-administered questions on social capital. Therefore, the core question set was converted to a self-administered format suitable for use in a postal survey. A detailed discussion of the design, development and cognitive testing of a prototype questionnaire can be found in paper two of this bulletin (Dewar and Wilmot, 2006).

In order to assess any differences between the self-administered format and the interviewer-administered format, a postal survey was run in parallel with an interviewer-administered survey.

2. Methods

2.1 Data collection

The core set of harmonised social capital questions were included in the ONS Omnibus survey in April, May and June 2005. The Omnibus survey is a multi-purpose survey which uses Computer Assisted Personal Interviewing (CAPI). In addition to the core set of social capital questions, the Omnibus survey included several modules covering a range of different topics and a set of classificatory variables.

The self-completion questions were included in a postal questionnaire which was conducted by National Centre for Social Research (NatCen) from April 2005 to June 2005. In addition to the core set of social capital questions, the postal questionnaire included a set of classificatory variables which were also included in the Omnibus survey.

2.2 Sample design

The Omnibus survey and the postal survey used the same sampling strategy; random probability sampling stratified by region, the proportion of households with no car; the proportion of households where the household reference person is in the National Statistics Socio-economic Classification (NS-SEC) categories 1-3; and the proportion

---

For further details about the design of the Omnibus survey, see http://www.statistics.gov.uk/about/services/omnibus/default.asp
of people who are aged over 65 years. The sampling frame was the small user Postcode Address File (PAF).²

First of all, 201 postal sectors were selected, with probability of selection proportionate to size. Within each sector, 30 addresses for the Omnibus survey and 35 addresses for the postal survey were chosen at random. Samples of addresses were selected from the same postcode sectors for both surveys to avoid any area effects. A larger number of addresses per sector were selected for the postal survey than for the Omnibus survey, to take into account the expected lower response rate for the postal survey.³

Only one adult per address/household was selected. For the Omnibus survey, interviewers were instructed to determine the household composition and then randomly select the respondent from amongst all the over-16s using a Kish grid. For the postal survey, names of adults aged 18 and over were matched from CACI’s Consumer Register to the sample of addresses.⁴ If more than one name was matched to a sampled address, one name was selected at random. Because only one person per household is interviewed, the data for both surveys were subsequently weighted to correct for the unequal probability of selection that this causes.

2.3 Standardisation across modes

To ensure any differences identified between the two sets could be attributed to the mode of administration, it was necessary to control for other known factors which may cause differences between the two surveys:

1. The social capital questions were identical for both modes, apart from some differences that were necessitated by the particular mode (see Dewar, 2005).

2. As described above, addresses were selected from the same postcode sectors using the small user Postcode Address File to avoid any area effects.

3. The surveys were conducted at the same time to ensure there were no differences due to changes over time.

4. In order to minimise differential non-response bias between the two surveys, Dillman’s Tailored Design Method (Dillman, 2000) was applied to the postal survey, as a way of minimising the anticipated difference in response rates between the postal survey and the interviewer-administered survey (see paper three of this bulletin for more detail: Dewar, 2006).

Despite these standardisation efforts, there were a number of small yet significant differences between the demographic make-up of the Omnibus and postal samples. The postal sample respondents were more likely to be educated to a higher level, to be owner-occupiers and to be from a white ethnic background. They also were more likely to be suffering from limiting long-term illness.

² For a description of the sample design, see the sampling section on the website of the Omnibus survey - http://www.statistics.gov.uk/about/services/omnibus/sample.asp
³ The response rate for the Omnibus survey was 69% and 51% for the postal survey. See paper five for a full account of the coverage and response rate for the postal survey (Nicolaas, 2006).
⁴ For a description of the matching exercise, see paper four of this bulletin (Tipping and Nicolaas, 2006).
We wished to be able to attribute any differences between the two samples to survey mode, and not to sample composition. We therefore used propensity score matching\(^5\) to make the profile of the postal sample similar to that of the Omnibus sample.

\[\text{(1) Problem of people using or dealing drugs (10.4\% and 1.2\% respectively) and (2) Problem of people being racially attacked or harassed (4.5\% and 0.6\% respectively).}^6\]

\[\text{Both of these questions could be considered sensitive. If so, respondents would have been less likely to answer these questions when an interviewer was present.}\]

The results from the parallel run of the postal survey and the Omnibus survey are summarised in Table 1.

\[^5\] A description of propensity score matching can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp

\[^6\] There are two other questions in this series which show higher rates of item non-response in the Omnibus survey compared to the postal survey: (1) Problem with teenagers hanging around on the street and (2) Problem with troublesome neighbours. Although these differences are statistically significant, they are small.
Table 1 Item non-response* to social capital questions by mode

<table>
<thead>
<tr>
<th>The core set of harmonised social capital questions</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Postal (%)</td>
</tr>
<tr>
<td>Likelihood of lost purse/wallet being returned**</td>
<td>9.0</td>
</tr>
<tr>
<td>Problem of drunk/rowdy people in public places</td>
<td>1.2</td>
</tr>
<tr>
<td>Problem of rubbish/litter lying around</td>
<td>0.7</td>
</tr>
<tr>
<td>Problem of vandalism/graffiti/other damage</td>
<td>0.7</td>
</tr>
<tr>
<td>Problem of people using or dealing drugs</td>
<td>1.2</td>
</tr>
<tr>
<td>Problem of people being racially attacked or harassed</td>
<td>0.6</td>
</tr>
<tr>
<td>Problem of teenagers hanging around on the street</td>
<td>0.2</td>
</tr>
<tr>
<td>Problem of troublesome neighbours</td>
<td>0.1</td>
</tr>
<tr>
<td>Contacted a local radio station, TV station or newspaper</td>
<td>4.0</td>
</tr>
<tr>
<td>Contacted appropriate organisation to deal with problem</td>
<td>2.1</td>
</tr>
<tr>
<td>Contacted a local councillor or MP</td>
<td>2.7</td>
</tr>
<tr>
<td>Attended a public meeting or neighbourhood forum</td>
<td>3.2</td>
</tr>
<tr>
<td>Attended a tenants or local residents group</td>
<td>3.5</td>
</tr>
<tr>
<td>Attended a protest meeting or joined an action group</td>
<td>3.7</td>
</tr>
<tr>
<td>Helped organise a petition on a local issue</td>
<td>3.4</td>
</tr>
<tr>
<td>How often do you speak to relatives on the phone?</td>
<td>0.5</td>
</tr>
<tr>
<td>How often do you speak to friends on the phone?</td>
<td>0.5</td>
</tr>
<tr>
<td>How often do you speak to neighbours face-to-face?</td>
<td>0.5</td>
</tr>
<tr>
<td>How often do you meet relatives not living with you?</td>
<td>0.6</td>
</tr>
<tr>
<td>How often do you meet up with friends?</td>
<td>0.4</td>
</tr>
<tr>
<td>Raised/Handled money, taken part in sponsored events</td>
<td>3.0</td>
</tr>
<tr>
<td>Led a group or were a member of a committee</td>
<td>3.0</td>
</tr>
<tr>
<td>Organised or helped to run an activity or event</td>
<td>2.8</td>
</tr>
<tr>
<td>Visited people</td>
<td>2.9</td>
</tr>
<tr>
<td>Befriended or mentored people</td>
<td>3.7</td>
</tr>
<tr>
<td>Given advice, information or counselling</td>
<td>3.2</td>
</tr>
<tr>
<td>Provided secretarial, admin or clerical support</td>
<td>3.2</td>
</tr>
<tr>
<td>Provided transport or driving</td>
<td>3.2</td>
</tr>
<tr>
<td>Acted as a representative</td>
<td>3.4</td>
</tr>
<tr>
<td>Participated in campaigns</td>
<td>3.6</td>
</tr>
<tr>
<td>Given other practical help</td>
<td>2.8</td>
</tr>
<tr>
<td>Given any other help</td>
<td>3.3</td>
</tr>
</tbody>
</table>

* Base (weighted) 3,664 3,669  
* Base (unweighted) 2,806 3,680

* Item non-response includes ‘don’t know’, no answer given and more than one answer given for single coded questions.
** This question included an explicit ‘don’t know’ option in the postal questionnaire, whereas this option was not offered to respondents in the CAPI version but could be coded as such if the respondent spontaneously said ‘don’t know’.
*** Statistically significant at the 95% level.
3.2 Social capital estimates

The core set of harmonised social capital questions provides measures for each of the following dimensions of social capital:

- Reciprocity and trust
- Views about the local area
- Civic participation
- Social networks and support
- Social participation

3.2.1 Trust

The core set of harmonised social capital questions includes one question on trust. In the Omnibus survey, this question and its response options were read out by the interviewer. The only response option that was not read out, was the ‘don’t know’ option which could only be coded if the respondent spontaneously said ‘don’t know’. No showcard was used for this question in the Omnibus survey.

The question wording was exactly the same for the postal survey, except the ‘don’t know’ option was explicitly offered to respondents. The postal survey question was pre-tested without a ‘don’t know’ option, but the results of the pre-test suggested that it was required (Dewar, 2005).

Omnibus survey, face-to-face (CAPI)

Suppose you lost your (purse/wallet) containing your address details, and it was found in the street by someone living in this neighbourhood. How likely is it that it would be returned to you with nothing missing...

Running Prompt

(1) very likely,
(2) quite likely,
(3) not very likely,
(4) or not at all likely?
(5) don’t know(Spontaneous only)

Postal survey (self-administered)

Suppose you lost your purse/wallet containing your address details, and it was found in the street by someone living in your neighbourhood. How likely is it that it would be returned to you with nothing missing?

Please tick one box only

Very likely……………………………………... □
Quite likely……………………………………. □
Not very likely………………………………... □
Not at all likely…………………………….. □
Don’t know…………………………………… □
Table 2 shows statistically significant differences in responses to this question by mode. Item non-response rates were higher for the postal survey than for the face-to-face survey, as noted in the previous section. Furthermore, the results show that face-to-face respondents were more likely than postal respondents to select the most extreme response options, i.e. ‘very likely’ and ‘not at all likely’.

Table 2  Likelihood of purse/wallet being returned by mode

<table>
<thead>
<tr>
<th>Likelihood of purse/wallet being returned without anything missing</th>
<th>Postal (%)</th>
<th>Face-to-face (CAPI) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very likely</td>
<td>17.7</td>
<td>28.8</td>
</tr>
<tr>
<td>Quite likely</td>
<td>34.8</td>
<td>33.7</td>
</tr>
<tr>
<td>Not very likely</td>
<td>28.2</td>
<td>18.8</td>
</tr>
<tr>
<td>Not at all likely</td>
<td>10.2</td>
<td>14.2</td>
</tr>
<tr>
<td>Don’t know</td>
<td>6.9</td>
<td>4.3</td>
</tr>
<tr>
<td>Not answered</td>
<td>2.1</td>
<td>0.2</td>
</tr>
<tr>
<td><strong>Base (weighted)</strong></td>
<td>3,665</td>
<td>3,669</td>
</tr>
<tr>
<td><strong>Base (unweighted)</strong></td>
<td>2,806</td>
<td>3,680</td>
</tr>
</tbody>
</table>

Pearson Chi-Square=0.000

3.2.2 Views about the local area

The core set of harmonised social capital questions includes a series of six questions about potential problems in the area. In the Omnibus survey, these questions were read out by the interviewer and the response options were provided on a showcard. The only response option that was not provided on the showcard, was the ‘don’t know’ option which could only be coded if the respondent spontaneously said ‘don’t know’.
Omnibus survey, face-to-face (CAPI)

I am going to read out a list of problems which some people face in their neighbourhood. For each one, please tell me how much of a problem it is.

Showcard
(1) Very big problem
(2) Fairly big problem
(3) Not a very big problem
(4) It happens but it’s not a problem
(5) Not a problem at all
(6) Don’t know (SPONTANEOUS ONLY)

♦ Taking your answer from this card, how much of a problem are people being drunk or rowdy in public places?
♦ How much of a problem is rubbish or litter lying around?
♦ How much of a problem are vandalism, graffiti and other deliberate damage to property or vehicles?
♦ How much of a problem are people using or dealing drugs?
♦ How much of a problem are people being attacked or harassed because of their skin colour, ethnic origin or religion?
♦ How much of a problem are teenagers hanging around on the street?
♦ How much of a problem are troublesome neighbours?

The wording of these questions and response options in the postal survey was identical to the wording in the face-to-face survey. The response option ‘don’t know’ was not provided.7

7 NB The response options suggest that the question is measuring two different issues. First, whether the events occur at all. And if they do occur, whether the event is considered to be a problem. A redesign of this question should be considered.
Postal survey (self-administered)

Here is a list of problems which some people face in their neighbourhood.

For each of the following questions, please tick one box only

♦ How much of a problem are people being drunk or rowdy in public places?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem is rubbish or litter lying around?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem are vandalism, graffiti and other deliberate damage to property or vehicles?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem are people using or dealing drugs?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem are people being attacked or harassed because of their skin colour, ethnic origin or religion?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem are teenagers hanging around on the street?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

♦ How much of a problem are troublesome neighbours?

<table>
<thead>
<tr>
<th>Very big Problem</th>
<th>Fairly big problem</th>
<th>Not a very big problem</th>
<th>It happens but it’s not a problem</th>
<th>Not a problem at all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For each of these questions, face-to-face respondents were far more likely than postal respondents to report that these events were not a problem at all in their neighbourhood (see Table 3). This could be the result of a social desirability effect, i.e. respondents are more likely to admit to problems in a self-administered format than in an interviewer-administered format. The same result was found when analysing social capital data from the Health Survey for England (Nicolaas and Tipping, 2004).

### Table 3 Problems in the neighbourhood: ‘not a problem at all’ by mode

<table>
<thead>
<tr>
<th>Problem</th>
<th>Postal (%)</th>
<th>Face-to-face (CAPI) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>People being drunk/rowdy in public places</td>
<td>28.5</td>
<td>46.9</td>
</tr>
<tr>
<td>Rubbish/Litter lying around</td>
<td>16.0</td>
<td>34.1</td>
</tr>
<tr>
<td>Vandalism/Graffiti</td>
<td>21.8</td>
<td>39.5</td>
</tr>
<tr>
<td>People using/dealing drugs</td>
<td>38.0</td>
<td>52.2</td>
</tr>
<tr>
<td>People being racially attacked/harassed</td>
<td>60.7</td>
<td>74.5</td>
</tr>
<tr>
<td>Teenagers hanging around on the street</td>
<td>16.9</td>
<td>35.4</td>
</tr>
<tr>
<td>Troublesome neighbours</td>
<td>56.5</td>
<td>75.8</td>
</tr>
</tbody>
</table>

*Base (weighted)* 3,665 3,669  
*Base (unweighted)* 2,806 3,680  

All results are statistically significant at the 95% level.

### 3.2.3 Civic participation

In the Omnibus survey, face-to-face respondents were given a showcard listing seven types of activities that could be undertaken to address local issues. Respondents were asked to indicate which of those activities they had carried out in the last twelve months. This question format tends to be called ‘check all that apply’ in the survey literature.

A ‘check all that apply’ format is not recommended for postal surveys because it would not be possible to distinguish between genuine negative responses (i.e. the respondent had not undertaken that activity) and item non-response (i.e. the respondent had skipped that item). Consequently, the ‘check all that apply’ question was converted into seven separate questions for the postal survey, each requiring either a ‘yes’ or a ‘no’ answer.
Omnibus survey, face-to-face (CAPI)

In the last 12 months have you taken any of the following actions in an attempt to solve a problem affecting people in your local area?

Code all that apply

Showcard
1. Contacted a local radio station, television station or newspaper
2. Contacted the appropriate organisation to deal with the problem, such as the council
3. Contacted a local councillor or MP
4. Attended a public meeting or neighbourhood forum to discuss local issues
5. Attended a tenants’ or local residents’ group
6. Attended a protest meeting or joined an action group
7. Helped organise a petition on a local issue
8. No local problems
9. None of the above
10. Don’t know (Spontaneous only)

Postal survey (self-administered)

We would now like to ask you a question on your participation in local issues.

In the last 12 months, have you taken any of the following actions in an attempt to solve a problem affecting people in your local area?

<table>
<thead>
<tr>
<th>Action</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contacted a local radio station, television station or newspaper</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Contacted the appropriate organisation to deal with the problem, such as the council</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Contacted a local councillor or MP</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Attended a public meeting or neighbourhood forum to discuss local issues</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Attended a tenants’ or local residents’ group</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Attended a protest meeting or joined an action group</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Helped organise a petition on a local issue</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Table 4 shows that postal respondents were more likely than face-to-face respondents to report all seven activities in the last twelve months. It is generally accepted that a series of separate questions, each requesting a ‘yes’ or a ‘no’ answer, will produce more affirmative responses than a ‘check all that apply’ question. Smyth et al (2005) have recently shown that Yes/No questions appear to produce higher rates of positive
answers because they encourage respondents to process the question more deeply and to discourage survey satisficing.\(^8\)

However, there is an alternative explanation for this finding. We cannot rule out the possibility that the postal survey, which was described as a ‘Neighbourhood Survey’ on all fieldwork documents, might have been disproportionately attractive to people who were active in their neighbourhoods. The face-to-face survey would have been introduced to respondents as a multi-purpose survey, and the mention of the neighbourhood questions would have been left to the interviewer’s judgement. Furthermore, people who lack social involvement are known to have lower response propensities (Groves and Couper, 1998). This is likely to be more noticeable in a postal survey than a face-to-face survey because of the absence of an interviewer who might otherwise be able to persuade reluctant respondents.

We would have hoped that the propensity score matching would have removed the most obvious differences between the two samples but the matching variables (i.e. demographic and household characteristics) may not have correlated well enough with social involvement to correct for this difference. It is therefore possible that the large differences shown in Table 4 are jointly caused by mode effects and real differences in social involvement between the two responding samples.

Table 4 Civic participation by mode

<table>
<thead>
<tr>
<th>Participation in local issues</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Postal (%)</td>
</tr>
<tr>
<td>Contacted a local radio station, television station or newspaper</td>
<td>2.6</td>
</tr>
<tr>
<td>Contacted the appropriate organisation to deal with the problem, such as the council</td>
<td>27.0</td>
</tr>
<tr>
<td>Contacted a local councillor or MP</td>
<td>10.8</td>
</tr>
<tr>
<td>Attended a public meeting or neighbourhood forum to discuss local issues</td>
<td>12.0</td>
</tr>
<tr>
<td>Attended a tenants’ or local residents’ group</td>
<td>7.6</td>
</tr>
<tr>
<td>Attended a protest meeting or joined an action group</td>
<td>3.4</td>
</tr>
<tr>
<td>Helped organise a petition on a local issue</td>
<td>5.2</td>
</tr>
<tr>
<td><strong>Base (weighted)</strong></td>
<td><strong>3,665</strong></td>
</tr>
<tr>
<td><strong>Base (unweighted)</strong></td>
<td><strong>2,806</strong></td>
</tr>
</tbody>
</table>

All results are statistically significant at the 95% level.

\(^8\) Survey satisficing occurs when the respondent does not make the effort to answer the question optimally.
3.2.4 Social networks and support

The core set of harmonised social capital questions includes a series of five questions about contacts with friends, relatives and neighbours. In the Omnibus survey, these questions were read out by the interviewer and the response options were provided on a showcard. The only response option that was not provided on the showcard, was the ‘don’t know’ option which could only be coded if the respondent spontaneously said ‘don’t know’.

The wording of these questions and response options in the postal survey was identical to the wording in the face-to-face survey. The response option ‘don’t know’ was not provided.

**Omnibus survey, face-to-face (CAPI)**

The next set of questions are about how often you personally contact relatives, friends and neighbours. Not counting the people you live with, how often do you do any of the following?

Showcard
(1) On most days
(2) Once or twice a week
(3) Once or twice a month
(4) Less often than once a month
(5) Never
(6) Don’t know (SPONTANEOUS ONLY)

♦ How often do you speak to relatives on the phone?
♦ How often do you speak to friends on the phone?
♦ How often do you speak to neighbours face-to-face?
♦ How often do you meet up with relatives who are not living with you?
♦ How often do you meet up with friends?
The next set of questions are about how often you personally contact relatives, friends and neighbours. Not counting the people you live with, how often do you do any of the following?

For each of the following questions, please tick one box only

- **How often do you speak to relatives on the phone?**
  - On most days
  - Once or twice a week
  - Once or twice a month
  - Less often than once a month
  - Never

- **How often do you speak to friends on the phone?**
  - On most days
  - Once or twice a week
  - Once or twice a month
  - Less often than once a month
  - Never

- **How often do you speak to neighbours face-to-face?**
  - On most days
  - Once or twice a week
  - Once or twice a month
  - Less often than once a month
  - Never

- **How often do you meet up with relatives who are not living with you?**
  - On most days
  - Once or twice a week
  - Once or twice a month
  - Less often than once a month
  - Never

- **How often do you meet up with friends?**
  - On most days
  - Once or twice a week
  - Once or twice a month
  - Less often than once a month
  - Never

Although the results showed a number of differences in the frequency of contacts between the two modes, the only consistent finding for all of the questions in this series is the higher rate of ‘never’ responses among face-to-face respondents compared to postal respondents (summarised in Table 5). This could be the product of real differences in social involvement between the two samples, as suggested in the previous section. However, ultimately, we are talking about relatively small percentages.
Table 5  Contact with friends, relatives and neighbours by mode

<table>
<thead>
<tr>
<th>Contact with friends, relatives and neighbours</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Postal (%)</td>
</tr>
<tr>
<td>Never speak to relatives on the phone</td>
<td>0.8</td>
</tr>
<tr>
<td>Never speak to friends on the phone</td>
<td>1.0</td>
</tr>
<tr>
<td>Never speak to neighbours face-to-face</td>
<td>1.6</td>
</tr>
<tr>
<td>Never meet up with relatives</td>
<td>2.0</td>
</tr>
<tr>
<td>Never meet up with friends</td>
<td>2.0</td>
</tr>
<tr>
<td>Base (weighted)</td>
<td>3,665</td>
</tr>
<tr>
<td>Base (unweighted)</td>
<td>2,806</td>
</tr>
</tbody>
</table>

All results are statistically significant at the 95% level.

3.2.5 Social participation

The final dimension of social capital in the core set is social participation. These questions ask about respondents’ involvement in groups and voluntary activities.

In the Omnibus survey, face-to-face respondents were given a showcard listing 11 forms of unpaid help as well as a catch all category ‘given any other help’. Respondents were asked to indicate which forms of unpaid help they had given in the last twelve months.

As for the series of questions on civic participation, this ‘check all that apply’ question was converted into 11 separate questions for the postal survey, each requiring either a ‘yes’ or a ‘no’ answer.

**Omnibus survey, face-to-face (CAPI)**

During the last 12 months have you given any unpaid help to any groups, clubs or organisations in any of the ways shown on this card?

Code all that apply

Showcard

1. Raising or handling money/taking part in sponsored events
2. Leading the group/ member of a committee
3. Organising or helping to run an activity or event
4. Visiting people
5. Befriending or mentoring people
6. Giving advice/ information/ counselling
7. Secretarial, admin or clerical work
8. Providing transport/driving
9. Representing
10. Campaigning
11. Other practical help (e.g. helping out at school, religious group, shopping)
12. Any other help
13. None of the above
14. Don’t know (Spontaneous only)
Postal survey (self-administered)

During the last 12 months have you given any unpaid help to any groups, clubs or organisations in any of the ways shown on this card?

<table>
<thead>
<tr>
<th>Activity</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised or handled money, taken part in sponsored events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Led a group or were a member of a committee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organised or helped to run an activity or event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visited people</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Befriended or mentored people</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Given advice, information or counselling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided secretarial, admin or clerical support</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided transport or driving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acted as a representative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participated in campaigns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Given other practical help</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Given any other help</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similar to the results for civic participation, the series of separate Yes/No questions in the postal format produced a higher proportion of affirmative responses than the ‘check all that apply’ question in the face-to-face format (Table 6). As previously explained, it is possible that these differences are not only caused by mode effects but also by genuine differences in social involvement between the two samples.
Table 6  Social participation by mode

<table>
<thead>
<tr>
<th>Social participation</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Postal (%)</td>
</tr>
<tr>
<td>Raised/handled money, taken part in sponsored events</td>
<td>29.5</td>
</tr>
<tr>
<td>Led a group or were a member of a committee</td>
<td>16.4</td>
</tr>
<tr>
<td>Organised or helped to run an activity or event</td>
<td>22.5</td>
</tr>
<tr>
<td>Visited people</td>
<td>25.4</td>
</tr>
<tr>
<td>Befriended or mentored people</td>
<td>13.2</td>
</tr>
<tr>
<td>Given advice, information or counselling</td>
<td>20.2</td>
</tr>
<tr>
<td>Provided secretarial, admin or clerical support</td>
<td>10.0</td>
</tr>
<tr>
<td>Provided transport or driving</td>
<td>19.7</td>
</tr>
<tr>
<td>Acted as a representative</td>
<td>9.5</td>
</tr>
<tr>
<td>Participated in campaigns</td>
<td>5.4</td>
</tr>
<tr>
<td>Given other practical help</td>
<td>21.1</td>
</tr>
<tr>
<td>Given any other help</td>
<td>27.0</td>
</tr>
<tr>
<td>Base (weighted)</td>
<td>3,665</td>
</tr>
<tr>
<td>Base (unweighted)</td>
<td>2,806</td>
</tr>
</tbody>
</table>

All results are statistically significant at the 95% level.

4. Discussion

The analysis has revealed three broad types of mode effects:

1. Item non-response
2. Social desirability effect
3. Formatting

4.1 Item non-response

On the whole, item non-response rates tended to be higher in the postal survey than in the face-to-face survey. However, this is with regards to very small percentages for most questions; less than 1% on the face-to-face survey and less than 4% on the postal survey.

The notable exceptions to this general finding were the questions that could be considered sensitive; i.e. problems with people using/dealing drugs in the neighbourhood and problems with people being attacked/harassed because of skin colour, ethnic origin or religion. Levels of item non-response for these questions were much higher on the face-to-face survey than the postal survey.

The easiest way to reduce this mode effect would be to use a self-completion format for all modes. For CAPI surveys this would mean inclusion of the social capital questions in a CASI instrument (Computer Assisted Self Interviewing) or a paper and pencil self-completion booklet. The equivalent self-completion format for telephone surveys would be touchtone data entry (TDE) but this is not a common data collection method for social surveys in the UK and still requires further methodological research.
Another issue that has been raised, is how to cater for respondents who genuinely don’t know the answer to a question. Interviewer-administered questions tend to be designed without explicit ‘don’t know’ options but interviewers can code ‘don’t know’ if respondents spontaneously say this. The reason for this approach is to minimise survey satisficing; i.e. respondents are tempted to choose the offered ‘don’t know’ option rather than exerting cognitive effort to answer the question. However, it is generally considered good practice for postal surveys to provide a ‘don’t know’ response option. This is recommended so that researchers can distinguish between genuine ‘don’t knows’ and other forms of item non-response such as ‘refused to answer’ or missed the question. If we had adopted this practice for the postal survey, then the postal survey would have had much higher levels of item non-response than the interviewer-administered format. To maximise comparability between the postal and interviewer-administered surveys, ‘don’t know’ options were not included in the postal survey and it was assumed that postal respondents who left the question blank were equivalent to face-to-face respondents that spontaneously say ‘don’t know’.

If the harmonised questions are to be suitable for different modes, then the handling of ‘don’t know’ responses requires further consideration.

### 4.2 Social desirability effects

As noted in the previous section, questions that asked about drug or racial problems in the neighbourhood had much higher levels of item non-response in the face-to-face mode than in the postal mode. We believe that this may be a social desirability effect.

In addition to higher levels of item non-response for these two ‘sensitive’ questions, there were other indications of possible social desirability effects. The results showed that face-to-face respondents were more likely than postal respondents to report that there were no problems in the area. The same result was found when comparing Health Survey for England data on social capital which were collected using CAPI in 2000 and CASI in 2001 (Nicolaas and Tipping, 2004).

Although we suspect that this might be a social desirability effect, most of these questions do not appear to be sensitive (e.g. ‘How much of a problem is rubbish or litter lying around?’). It is possible that the cause of these effects is not the sensitivity of the question itself but the response options; i.e. very big problem, fairly big problem, not a very big problem, it happens but it’s not a problem, not a problem at all. Perhaps it is socially undesirable to admit to problems in face-to-face encounters. Further research into the underlying causes for this strong mode effect is required.

This mode effect could be minimised by using self-completion formats in all surveys, irrespective of the main mode of data collection. For example, paper and pencil self-completion booklets or CASI could be used alongside the main data collection instruments in face-to-face surveys.
4.3 Formatting

Some of the harmonised social capital questions had to be reformatted so that they could be used in a postal survey (i.e. civic participation and social participation).

‘Check all that apply’ with showcard is a common question type in face-to-face surveys. Interviewers and respondents tend to find this approach less repetitive and tedious than a series of separate questions, each requesting a ‘yes’ or a ‘no’ answer. It also helps to maintain the speed of the interview.

However, a ‘check all that apply’ question format is not recommended for postal surveys because it is not possible to distinguish between genuine negative responses (i.e. the respondent had not undertaken that activity) and item non-response (i.e. the respondent had skipped that item). Consequently, the two ‘check all that apply’ questions were converted into two series of questions for the postal survey (one series of seven questions and one series of 12 questions), each requiring either a ‘yes’ or a ‘no’ answer.

All of these questions had higher rates of affirmative responses in the postal survey compared to the face-to-face survey. It is generally accepted that series of questions requiring a ‘yes’ or a ‘no’ answer are likely to produce higher rates of affirmative responses than ‘check all that apply’ questions because of deeper processing and reduced satisficing (Smyth et al, 2005).

If harmonised questions are to be suitable for different modes, then ‘check all that apply’ formats should not be used.

4.4 Limitation of the study

In order to identify mode effects, the surveys were designed to minimise other possible differences (i.e. area differences, time differences, differences in question wording and differential non-response bias) and propensity score matching was used to remove any remaining socio-demographic differences between the two responding samples. However, some of the differences we observed may have been partly caused by differences in sample composition that had not been accounted for by the propensity score matching.

We hypothesised that the postal survey, which was described as a neighbourhood survey, may have been disproportionately attractive to people who were more socially involved in their neighbourhood. This effect was less likely to occur in the face-to-face survey because it was a multi-purpose survey, i.e. there was no mention of the neighbourhood module in the advance letter nor would interviewers necessarily have mentioned it in their survey introductions.

Furthermore, it is generally accepted that people who are more socially involved have higher response propensities (Groves and Couper, 1998). In other words, they are more likely to take part in a postal survey than people who are not socially involved. This effect would be less extreme in face-to-face surveys because the interviewer can persuade reluctant respondents.

Consequently, we cannot completely disentangle mode effects from possible differences in levels of social involvement between the two responding samples.
4.5 Conclusion and recommendations

Mode effects could undermine the comparability of national and local data if the current set of harmonised social capital questions are used in local postal surveys. It is therefore recommended that some of the questions are redesigned so that they are less predisposed to mode effects.

On the whole, the higher rate of item non-response on the postal survey compared to the face-to-face survey does not give reason for concern. The mode effects introduced by reformatting the ‘check all that apply’ questions into series of separate Yes/No questions can be overcome by adopting the latter approach for all modes. Social desirability effects can be minimised by recommending that the social capital questions are included in self-completion instruments for all modes; e.g. either CASI or paper and pencil self-completion booklets for face-to-face surveys.

This study has also highlighted two areas requiring further research:

- An approach for handling ‘don’t know’ responses that is acceptable for all modes.
- The underlying causes for the strong mode effects that were observed for the ‘views about the neighbourhood’ questions.

References


Developing harmonised questions for use in a mixed-mode data collection environment

Amanda Wilmot & Abigail Dewar, UK Office for National Statistics

1. Introduction

A major factor behind the creation of National Statistics in the UK was the desire to make statistics accessible to everyone, providing a more coherent picture of British society. In the UK Government White Paper ‘Building Trust in Statistics’ (October 1999), new arrangements were set in place for establishing a framework under which the national statistical service would be able to provide ‘better and more reliable official statistics that command public confidence’. Through the introduction of a National Statistics Code of Practice\(^1\), in effect, National Statistics could be interpreted as a quality marker for official statistics in the UK.

Supporting the concept of National Statistics is a progressive programme of harmonisation of survey questions, concepts and classifications, in an attempt to make survey results more comparable. This programme of harmonisation is considered as a ‘step towards achieving the ultimate goal of making the interpretation and analysis of data easier’\(^2\).

One of the intentions of the harmonisation programme is to also help data users plan surveys at a local level that can provide data comparable with national surveys. Furthermore, there is an intention to make greater use of administrative data to support local level, neighbourhood statistics. In order to exploit such data it is therefore considered important to also extend the principles of harmonisation beyond the national surveys (and the Census).

Further impetus includes the requirement to provide data to Eurostat\(^3\), which is comparable cross-nationally. Indeed Eurostat are also engaging in a programme of harmonisation with this aim in mind, and have recently established a European harmonisation task force.

However, establishing any programme of harmonisation is not straightforward. Indeed, the Office for National Statistics (ONS) acknowledges that certain practical issues arise.

Differences in survey design and methods used are acknowledged by the harmonisation team. Factors inherent in these different designs may lead to different estimates; one of these factors being mode effect - that is the impact that the mode of data collection (for example, interviewer-administered or self-administered) may have on the estimates.

Most ONS social surveys are primarily interviewer-administered using CAPI\(^4\), which is generally accepted as the preferred mode of data collection to ensure high response

\(^2\) [http://www.statistics.gov.uk/about/data/harmonisation/](http://www.statistics.gov.uk/about/data/harmonisation/)
\(^3\) The European Community Statistical Office
\(^4\) Computer Assisted Personal Interviewing
and data quality, and the harmonisation programme has focussed on developing questions for this single mode of administration.

However, with technological and socio-economic changes in society (Groves and Cooper, 1998) and an increased need to capture data in different ways, the survey research world has, in recent years, put greater emphasis on improving survey response and reducing cost through mixed-mode survey administration, where more than one mode of data collection are used in conjunction.

Accepting the desire for a harmonised approach to the design of survey questions to try to ensure equitability within a changing research environment, this paper highlights the need to consider the design of harmonised survey inputs in a mixed-mode data collection environment. The results of an experiment conducted in 2005 comparing data collected in a face-to-face interviewer-administered survey with data from a matched sample, postal self-completion survey, are used for illustrative purposes.

2. Evolution of mixed-mode data collection

The move towards mixed-mode data collection in national UK surveys is not a new phenomenon. A classic example of a survey designed for use in a mixed-mode environment is the Labour Force Survey (LFS), where all of the first wave cases are interviewed face-to-face and around 80% of interviews on waves 2-5 are carried out over the telephone. Surveys such as the General Household Survey have supplemented CAPI with modules of questions for self-completion, offering a paper or CASI\(^5\) alternative where questions are considered sensitive in nature. By doing this researchers have acknowledged, historically, the fact that the presence of an interviewer may influence people to provide more socially acceptable responses, for example, to questions about alcohol and tobacco consumption. Indeed, surveys with a diary keeping element such as the Expenditure and Food Survey and the National Diet and Nutrition programme of surveys have, by their very nature, meant that both CAPI and paper self-completion methods are used in combination - the face-to-face interview often used to supplement or validate information recorded in the diary.

With technological developments there are now many more alternative modes of data collection to choose from, the use of the Internet probably being the most important, along with Audio-CASI, and Interactive Voice Response surveys which can recognise spoken answers or where respondents can enter data using a telephone key pad. With these technological developments comes the increased difficulty of maintaining a consistent approach and the increased need to examine the effect of mode on data quality.

The development of the new UK Integrated Household Survey (IHS) has also highlighted the need to consider the impact of different modes. The IHS public consultation document states that a mixed-mode data collection instrument will be used and that all survey questions should be suitable for both personal and telephone interviewing modes.

Researchers managing the European Social Survey also have to consider the most complex issue of carrying out surveys in a mixed-mode environment, cross-nationally. As such, cultural preferences for different modes and the resulting effect on response and data quality will need to be addressed.

\(^5\) Computer Assisted Self Interviewing
So, there is a growing appreciation that operating in one mode only is not practical, especially with the increasing focus on response, the demand for collecting data of sufficient quality at a local as well as a national level, and for international comparisons.

3. Evidence of mode effect

It is accepted that the survey methodology will affect survey estimates, but what is less certain is the extent to which any differences in estimates can be attributed to the mode of administration.

Labour Force Survey investigative work supports the view that mode effect should not be disregarded as a decisive factor impacting on survey estimates. Research carried out by Elliot et al (2006), identified significant differences in some of the estimates between the first two waves (face-to-face and telephone) where differential non-response (attrition) bias could account for only some of the observed differences. The paper concluded that while the remaining differences may be caused by some form of panel conditioning, mode effects may also be a key contributing factor, although the extent of the effect was unknown.

It is widely acknowledged that the cognitive stimulus respondents receive differs across modes, and differences in the way that information is presented demands different cognitive processes and elicits differential respondent burden. Issues such as providing answers that are more socially acceptable when an interviewer is present are supported in the literature. As Dillman and Christian (2005) states ‘considerable evidence now exists that the choice of survey mode affects respondents’ answers to survey questions that are worded the same’.

Is one mode of administration ‘better’ than another? De Leeuw (1992) conducted a meta analysis of mode comparisons which concluded that, when questions are answered, data quality is better in self-administered surveys, especially with more sensitive questions whereby less social desirable responses are produced. De Leeuw’s own experiment (1992) replicated the finding that the main differences are found between postal and face-to-face surveys. The self-administered postal questionnaire, where respondents had more control and more time to read and answer questions, resulted in more reliable and consistent responses and less acquiescence than face-to-face interviews.

Where multiple modes are used for data collection, de Leeuw states that it is crucial to recognise that mode effects are an important design consideration and should be reduced as much as possible. However, she goes on to say that ‘hardly any theoretical or empirical knowledge is available on how to design optimal questionnaires for mixed-mode data collection. Empirical research is needed to estimate what constitutes the same stimulus across different modes’.

---

6 Acquiescence is the tendency to agree when interacting with another person because people generally find it easier to agree than disagree.
4. Further comparative research on mode effect

Although no comprehensive consensus exists on how to operationalise a common stimulus across modes, the demand for mixed-mode data collection continues to gather momentum. A key element of the ONS harmonisation programme has been the feasibility of collecting comparable local and national data, looking in particular at ways in which to produce more granular social capital estimates, that is, estimates available for smaller geographical areas. The provision of small area statistics would facilitate more in-depth data analysis.

Therefore as part of the harmonisation strategy, including the provision of small area statistics, the ONS commissioned the Data Collection Methodology branch at ONS and the National Centre for Social Research (NatCen) to convert the harmonised social capital question set into self-administered format that could be recommended for use by local authorities. The existing set of ONS harmonised social capital questions were developed for use in CAPI-based national surveys. The requirement for data to be available at the local level would necessitate a change in mode of administration. Local government tend not to have the resource to implement interviewer-administered face-to-face surveys and instead use postal self-completion as their main method of data collection.

The principal aim of this commissioned research was to design a self-administered paper questionnaire that would provide comparable outputs to the harmonised CAPI question set.

4.1 Health Survey for England mode comparisons study

To inform the design and development of a self-administered questionnaire NatCen carried out a comparison of data collected from the Health Survey for England (HSE) in two different modes (Nicolaas and Tipping, 2004). The HSE 2000 was conducted using CAPI, while for the HSE 2001 the question set on social capital was included as a CASI module. This comparative study was specifically salient to the conversion work as the HSE questions on social capital were similar to the ONS recommended harmonised set.

The HSE analysis provided valuable methodological insight into what does or doesn’t work in interviewer to self-administered conversion. To try and ensure that any differences found could be attributed to different data collection modes and not differences in sample composition, propensity score matching was used so that the sample profiles were as similar as possible. While it was recognised that the time frame differed between mode administrations, it was deemed unlikely that social capital would have changed significantly in a relatively short time period.

The findings from the analysis showed differences between modes with regard to five key question design issues: factual and opinion question types; question wording and response options; item non-response; problem scales; and spontaneous ‘don’t know’ and neutral response options in face-to-face mode.

While the literature suggests that factual questions are less susceptible to mode effects than opinion questions, differences between mode were identified for factual

---

7 The propensity score matching was carried out in Stata V8.2 using the psmatch2 command. A description of propensity score matching can be found at: http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp
questions in the HSE. However, it is possible that some of these could be explained by factors other than mode, such as changes in the question wording and response options used in each mode. For example, precodes used by interviewers but not used in the self-administered version; or where the interviewer-administered questions included a filter question, but the filter was not available in self-administered format. Other differences were an extra category ‘charity, voluntary, or community group’ included for the self-administered questions about involvement in organisations, and additional information provided for one of the categories, ‘sports club’ which included gym, exercise or dance group. The effect of these changes resulted in substantial differences in survey estimates (Nicolaas and Tipping, 2004).

The study confirmed findings from previous research that item non-response tends to be higher for the self-administered (CASI) mode compared with interviewer-administered (CAPI) mode; for example 2.3% and 0.3% respectively for a question asking respondents how long they have ‘lived in this local area’ (Nicolaas and Tipping, 2004). Indeed, even higher levels of non-response could be expected for a postal survey which does not have the benefit of the automatic routing available in computer assisted instruments, where the risk of item non-response for CASI-based questions caused by routing error and respondents skipping questions, is minimised. Furthermore, item non-response was generally higher for opinion questions in the self-administered format than in the interviewer-administered version, which was considered likely to be a consequence of changing the mode of data collection. However, changing the response options, for example from yes/no to an agree/disagree scale was also likely to cause respondent burden in the self-administered version and should not be ruled out as a contributing factor. Yet, supporting the research carried out by Dillman and Christian (2005), the analysis showed that using the same response scale does not necessarily improve comparability. The analysis indicated that even when questions were worded the same, differences still occurred. For example, the questions about teenagers hanging around on the streets and vandalism/graffiti/property damage were identically worded, and used the same 5-point response scale ranging from a ‘very big problem’ to ‘not a problem at all’ in both modes, yet large differences were observed. Only 19.8% of CASI respondents said that teenagers hanging around on the streets was not a problem compared with 49.2% of CAPI respondents (Nicolaas and Tipping, 2004).

Dillman and Christian (2005) observed that the reason why modal differences tend to occur for identically worded questions is because of context and order effects, acquiescence, primacy and recency effects and also social desirability. With regard to the HSE analysis, it would seem that the problem scales are sensitive to changes in mode. In-keeping with Dillman’s research, this suggests a social desirability effect was occurring whereby respondents were less likely to admit to problems in the face-to-face interview with an interviewer present than with the self-administered questionnaire.

The study also revealed potential problems with how to deal with spontaneous (hidden) ‘don’t know’ options used in an interviewer-administered question when converting to self-administered format. Results from the HSE analysis showed that item non-response increased substantially when an explicit ‘don’t know’ option was given in a self-administered question about trust in people generally; 9.7% compared with 0.8% for the interviewer-administered question with a spontaneous ‘don’t know’

8 Primacy is the tendency for respondents to choose from the answer categories offered first whereas recency is the tendency to choose from those offered last.
option. However, it should be noted that there was no evidence to suggest that these differences in response led to any differences in the survey estimates of trust; 39.6% of 2000 and 2001 HSE respondents reported that most people could be trusted (Nicolaas and Tipping, 2004).

What is clear from the HSE analysis was the paramount importance of keeping the way respondents receive information across modes as similar as possible in order to reduce the risk of item non-response, measurement error and susceptibility to mode differences.

Together with a preliminary literature review, the results from the HSE comparative analysis provided a basis for understanding possible sources of response differences in survey estimates of social capital and the extent to which they might be overcome or reduced.

4.2 Conversion of the CAPI social capital harmonised question set into postal format

Taking account of this preliminary analysis, the core set of ONS harmonised social capital questions were converted from interviewer-administered to a self-administered format. The core set included one or more questions from five social capital domains; trust, views about the local area, civic participation, social networks and support, and social participation. In general, the question wording and response options were kept the same in both modes.

The process of converting the questions from one mode of data collection to another emphasised the importance of assessing how modes differ in the way they operate in practice, as this ultimately affects the way researchers design questions. In face-to-face surveys, interviewers are a key medium of interaction with respondents; the benefits of oral communication and use of body language help question comprehension and reduce the cognitive burden placed on respondents. For the self-administered version, questionnaire design was necessarily required to focus on the non-verbal visual channel of administering the self-completion questionnaire (see paper two of this bulletin (Dewar and Wilmot, 2006) for more information on the design and development of the self-administered questionnaire). Without interviewers present, considerable reliance was placed on paralanguage provided by visual sources of information. As such, numbering, symbolism (ticks and arrows) and graphics (changes in font size, shape spacing and shading) were used as an alternative method to help respondents understand the questions and to guide them through the questionnaire.

Cognitive pre-testing revealed various recommendations for revisions to the postal questionnaire which could lead to improvements in the quality of the survey data; the results indicated the importance of visual design, using paralanguage as the key tool in the development of questionnaires in self-administered paper format (see paper two: Dewar and Wilmot, 2006).

Following the qualitative stage of the conversion, Nicolaas and Tipping carried out a quantitative evaluation, using the recommended postal survey questionnaire, to measure the effect of using different modes to collect social capital data. (Also see paper six of this bulletin (Nicolaas and Tipping, 2006) for the full comparison and results from this evaluation.)
4.3 Mode comparisons between the CAPI and Social Capital Postal Survey 2005

In order to assess any differences between the interviewer and self-administered questions, the postal survey was run in parallel with the CAPI-based questions included as a module in the National Statistics Omnibus survey, and the outcomes from the two samples compared.

To ensure any differences identified between the two sets could be attributed to the mode of administration, it was necessary to control for other known factors which may cause differences between the two surveys, such as differences in population coverage and differential non-response bias. Therefore, careful consideration was given to the sample design. Both surveys used the same sampling strategy: a stratified, clustered, random probability design; both samples of addresses were selected from the same postcode sectors using the small user Postcode Address File\textsuperscript{9} to avoid any area effects; the surveys were conducted at the same time to ensure there were no differences from changes over time, and Dillman’s Tailored Design Method (Dillman, 2000) was applied to the design of all documentation relating to the postal survey, as a way of further minimising differential non-response bias between the face-to-face and postal surveys. Propensity score matching was also used in order to control for socio-demographic characteristics and remove any remaining differences between the two responding samples. Overall, this made it possible to make more direct comparisons between the surveys and to be reasonably confident that any enduring differences found could be attributed to the use of different modes of data collection.

The results from the analysis revealed key differences between the surveys which could be attributable to three types of mode effect: item non-response; formatting and social desirability (as identified in paper six: Nicolaas and Tipping, 2006).

4.3.1 Item non-response

Similar to the findings for the HSE analysis and consistent with the survey literature, item non-response was generally higher for the self-administered questions than the interviewer-administered questions, although it was relatively low for both surveys; less than 4% and less than 1% respectively.

One exception was the level of item non-response for the trust question on the likelihood of a lost purse or wallet being returned, which was higher than the level observed overall (9% postal compared with 4.5% CAPI). As with the rest of the interviewer-administered question set, the question included a spontaneous ‘don’t know’ response option. Despite results from the HSE analysis indicating caution in its use, an explicit ‘don’t know’ was offered to respondents in the postal survey for this specific question. This was because findings from the cognitive pre-testing indicated that on this occasion, there was a genuine requirement for respondents to have the option to say ‘don’t know’; quite apart from other non-response scenarios such as refusing to answer or having missed the question (Dewar, 2005). This need to be able to provide a ‘don’t know’ response at this question is supported by the CAPI results. However, as also evident for the HSE questions, providing ‘don’t know’ response options led to higher levels of item non-response in self-completion mode. It thereby increased the differences in non-response between the postal and face-to-face surveys.

\textsuperscript{9} The Royal Mail’s list of addresses (delivery points). Considered the most comprehensive sampling frame for national face-to-face surveys of the general population in Great Britain.
and ultimately would risk reducing the comparability of local and national survey estimates.

Interviewer-administered and self-administered surveys have juxtaposing requirements with regard to the treatment of ‘don’t know’ responses. For interviewer-administered questions, spontaneous ‘don’t know’ options seem the most appropriate way to minimise the risk of survey satisficing\(^\text{10}\); it is generally assumed that if explicitly available to them, respondents will be tempted to choose the ‘don’t know’ option as it requires the least amount of cognitive effort. However, for postal surveys, providing visible ‘don’t know’ response options is the only way possible to distinguish between valid ‘don’t know’ answers, and refusals/missing responses to questions.

Therefore, in order for harmonised questions to be of optimum use in a mixed-mode data collection environment, a more appropriate way of managing ‘don’t know’ responses needs to be addressed.

However the most marked differences in item non-response were found for two questions on ‘problems in the area’. In contrast to trends highlighted in the survey literature, here, it was in fact the interviewer-administered questions with higher item non-response. The questions were concerned with people using/dealing drugs in the neighbourhood, where item non-response was 10.4% (face-to-face) and only 1.2% (postal), and people being attacked/harassed because of skin colour, ethnic origin or religion (4.5% and 0.6% respectively). Both questions could be considered sensitive in nature, and it might be that respondents were less inclined to answer such questions in a face-to-face interview, or with other household members present.

4.3.2 Formatting

When administering surveys in different modes, there is a tendency to simply adapt, or make adjustments to, questions because different question formats work better in different modes. Dillman and Christian (2005) attributes this to the fact that ‘a “preferred” mode of constructing survey questions has evolved for each of the different modes and tends to be shared across organisations’. The social capital conversion was no exception; some formatting changes were necessary to make questions in the civic and social participation domains suitable in postal survey format in order to be able to distinguish between genuine negative responses and item non-response.

Often in face-to-face interviews, questions are designed with a showcard element, and in order for the question to make sense, respondents are required to use this visual aid to help them answer. For example, in the face-to-face survey, the series of civic participation questions required respondents to tell interviewers which activities they had carried out to address local issues in the last twelve months; respondents were given a showcard listing seven types of activities from which they could choose which were relevant to them.

This ‘check all that apply’, multicoded approach, works well in face-to-face surveys as it avoids the need to ask respondents a series of separate questions requiring individual prompt ‘yes’ or ‘no’ answers, which can be repetitive and inevitably slows

\(^{10}\) Satisficing is the inclination to choose a response which requires the least cognitive effort in order to answer a question.
down the interview. However, it does require the respondent to be able to read, assimilate and comprehend all of the options listed on the card quickly.

However, it could be argued that the ‘check all that apply’ approach is inappropriate for use in self-administered modes. For example, in the postal survey it would not be possible to identify whether respondents had simply not taken part in a particular activity or whether they had skipped the item. Consequently, the civic and social participation multicoded questions used in the face-to-face survey were converted into a series of separate questions with ‘yes’ / ‘no’ response options for the postal survey.

Tailoring question construction in this way helps to maximise the effectiveness of specific questions in different modes; however it can also contribute to overall survey mode effects (Christian et al, 2006).

Following the change from a check-all question in the interviewer-administered survey to a Yes / No forced-choice in the postal survey, it was possible to assess the interaction between question format and survey mode. While the forced-choice format made it easier to identify item non-response, adjusting the multicoded questions in this way also increased reporting, with a significantly higher number of items selected when the questions were presented in forced-choice format. These results confirm previous research findings which compare these two question formats in various interviewer-administered and self-completion modes (Jordan et al, 1980; Rasinski et al, 1994), and have been reinforced by Smyth et al (2006) who observe a similar phenomenon when making comparisons across Internet and telephone modes.

The social capital reporting rates were higher for postal respondents than for face-to-face respondents for all seven activities in the civic participation domain and similarly for every item in the social participation domain (see paper six of this bulletin: Nicolaas and Tipping, 2006). Consistent with the research carried out by Smyth et al (2006), this implies that adapting questions in this way to accommodate postal mode limitations will alter the question stimulus to the extent that national and local estimates may not be comparable. The cognitive findings from the conversion process also supported the quantitative analysis as respondents stated that repetition gave them more time to process what each individual question was asking and made them think more about their answers (see paper two of this bulletin (Dewar and Wilmot, 2006)).

Similarly, Smyth et al (2005) surmise that the repetitive nature encourages respondents to process the question more deeply; it also discourages the survey satisficing behaviour some respondents reveal when answering questions in the check-all format (Smyth, Dillman, Christian and Stern 2005). Moreover, recent findings indicate that the forced-choice format transfers across interviewer-administered and self-completion modes quite well (Smyth et al, 2006). Therefore a series of ‘yes’ / ‘no’ rather than check-all questions would seem beneficial for all modes, including self-administered surveys and would certainly be an appropriate step towards harmonising the channels of communication when surveys are conducted in a mixed-mode environment. However, further research is required to evaluate the impact on respondents and interviewers in interviewer-administered surveys.

Another key characteristic, which may have influenced likely mode effects for the civic and social participation questions, is the tendency for response in postal surveys to be linked to the saliency of the topic and the variables of interest being measured. The interviewer-administered social capital questions were presented to respondents as part of a multipurpose Omnibus survey. However, the self-administered questions
were referred to in all fieldwork documentation as part of a ‘Neighbourhood Survey’. Direct mention of the term neighbourhood might have disproportionately appealed to people who were active in their neighbourhoods or took more of an interest in neighbourhood issues. Research indicates that people who lack social involvement are less inclined to respond to surveys (Groves and Couper, 1998) and this is exacerbated still further by the fact that it is more difficult to persuade reluctant respondents to take part in a postal survey than in a face-to-face survey where the interviewer is adept at encouraging participation.

As a result of this, and in view of the possibility that the propensity score matching did not sufficiently correct for all differences in the sample composition, large variations in the estimates could be due to both mode effects and real differences in social involvement between the two responding samples (see paper six of this bulletin: Nicolaas and Tipping, 2006).

4.3.3 Social Desirability

The final area where strong modal differences were identified between the two surveys was in relation to scalar questions. The analysis indicated that social desirability had an effect on the distribution of responses to a 5-point problem scale. When asked about potential problems in the area, face-to-face respondents were more likely than postal respondents to report that the possible scenarios described were ‘not a problem at all’ in their neighbourhood. For example, 52.2% of respondents interviewed face-to-face reported people using/dealing drugs were not a problem at all, compared with 38.0% in the postal survey (paper six of this bulletin: Nicolaas and Tipping, 2006). It seems feasible that when asked these potentially sensitive questions, respondents adapted their answers according to what they thought the interviewers wanted to hear or found acceptable, while in a self-completion environment, answers were more likely to reflect what respondents really thought. These results support Dillman and Christian’s (2005) and Nicolaas and Tipping’s (2004) findings discussed above concerning mode differences for identically worded questions influenced by social desirability.

However, the higher proportion of face-to-face respondents reporting no problems in their local area was not limited to sensitive questions. The pattern was observed for less contentious problems: 35.4% of respondents interviewed face-to-face said that teenagers hanging around on the street was not a problem at all compared with 16.9% of postal respondents. This pattern was replicated in the HSE analysis. For almost identical questions, 49.2% of respondents interviewed by CAPI compared with 19.8% of CASI respondents said that teenagers hanging around on the street were not a problem at all. Recently, similar findings have been identified by Christian et al. (2006) comparing scalar formats used in telephone and Internet modes. More positive ratings were observed amongst telephone respondents than Internet respondents to questions using scales with both 5 and 11 categories, and they were also likely to choose the most extreme positive category. Furthermore, the results were observed across various scale types; such as agree/disagree and satisfaction scales (Christian et al, 2006). Yet, again, the questions were not considered sensitive, and social desirability was not deemed to be a suitable explanation for the differences. These findings also confirm results from past research (Dillman and Mason, 1984; Tarnai and Dillman, 1992) which show a stronger likelihood amongst telephone and face-to-face survey respondents to select response categories at the positive end of scales than postal survey respondents. Overall, these results show the occurrence of this modal effect ranging across substantive scales and throughout the interviewer-administered versus self-completion mode spectrum.
With regard to the social capital research, this implies that it was not the nature of the questions that caused mode effects per se, but the response options themselves. It might be that people generally are reluctant to admit any type of problem in face-to-face encounters, if they consider it socially unacceptable to do so. It could also be that respondents needed more time to contemplate problem scales than that afforded to them in a face-to-face interview which was likely to be conducted at a faster pace than when completing a paper questionnaire; the speed of the interview commanding a much quicker response to the question/answer process. Further research is required to identify the underlying causes of this type of mode effect. However, what is clear is that there is a requirement for optimal scalar questions for surveys that provide the same stimulus in different modes, or that an alternative to the concept of problem scales is needed if estimates are to be truly comparable across modes.

Past research has shown that self-administered modes of data collection help to mitigate mode effects such as item non-response and social desirability. Evidence has shown that reporting levels for sensitive questions, ranging across a number of topics increased compared with interviewers asking the same questions (Tourangeau and Smith, 1996). Erdman, Klein and Greist (1983) observed that answering questions via a CASI instrument was ‘less embarrassing’ for respondents. It should be noted that comparative studies also show CASI instruments produce similar results to paper and pencil interviews. This implies that the key to increased response for sensitive questions is self-administration, regardless of whether it is computer assisted or not (Tourangeau et al, 1997). The privacy provided by self-administration makes it ideal when respondents might be inclined to provide more culturally acceptable or socially desirable answers. Although, as observed in the design and development of the Social Capital Postal Survey, an element of caution is required when the questionnaire is the sole means of conveying information because paralanguage becomes key to ensuring respondents comprehend the questions as the researchers’ intended.

Nevertheless the inclusion of the social capital questions in self-completion instruments for all modes would be an effective way to minimise item non-response and social desirability effects and improve comparability between modes. For example, they could be included in a CASI instrument or a paper and pencil self-completion document when included in face-to-face surveys; similar to current practice in the Health Survey for England.

The modes comparison study showed that despite efforts to minimise risk, for local authorities wanting to use the harmonised social capital questions in a postal survey, converting the current set of harmonised social capital questions, primarily designed to be interviewer-administered will predispose some of the survey estimates to mode effects. This could therefore undermine the comparability of national and local social capital data.

5. Limitations and methods of improving comparability between modes

One way to improve comparability would have been to reword the harmonised questions set on social capital in both the interviewer and self-administered formats. However, this was not possible because the interviewer-administered question set was already operationalised and therefore questions were simply adapted to suit the self-administered environment.

Even if a redesign of the question wording had taken place, as already discussed, fundamental differences between interviewer-administered and self-administered modes of data collection, with respect to the way in which respondents answer survey
questions, would have remained. Specifically, research has shown survey modes to differ as a result of; the presence versus absence of an interviewer (de Leeuw, 1992; Dillman, Sangster, Tarnai and Rockwood, 1996); aural or visual methods of conveying or transmitting information, and whether the interviewer or respondent commands delivery of stimulus (Schwarz et al, 1991; de Leeuw, 1992; Dillman 2000; de Leuuw 2005). Consequently, the choice of survey mode would still affect respondents’ answers to questions irrespective of any redevelopment to ensure the question wording was the same in both modes (Dillman and Christian, 2005). Indeed, there is considerable evidence to suggest that the words themselves are just one way of extracting meaning from questions; the use of paralanguage to aid question comprehension in the postal version of the social capital questionnaire is a prime example of an alternative method of fostering question comprehension.

Better ways of ensuring comparability when using mixed-mode survey designs is being reviewed in research to-date. One notion is to go a step beyond recommendations to adapt questions designed for use in the primary mode of collection to secondary modes, and to take a ‘unimode’ approach to question construction in an attempt to ‘further extend effort to obtain equivalent data across survey modes’ (Dillman, 2000) - an objective akin to that of the ONS harmonisation programme.

Dillman advocates this unimode approach which, amongst other things, focuses on visual principles to form the foundation of question design and which requires survey designers to consider question styles from a different perspective. Rather than developing questions which optimise performance in a single mode, the aim is to achieve functional equivalence across many different modes. For the social capital questions this would involve re-developing the questions for use in different modes at the same time so that the aural and visual stimulus provided by the interviewer-administered and self-administered modes respectively would elicit the same responses.

Currently, there is a requirement for further empirical research to find out what constitutes the same stimulus across different modes in addition to the work carried out for the social capital conversion study, which although not conclusive, does add to the international research on the subject. Cognitive pre-testing is of course valuable in systematically highlighting questions which do not transfer easily across modes, although it is important to also be able to quantify the impact, so questions and questionnaires would need to be pre-tested and piloted in all modes to ensure equivalence.

6. Conclusion

Currently different modes affect the way the researcher designs questions and there is a tendency to adjust questions because different question formats work better in different modes.

Developing questions in the primary mode and then testing them in other modes would help to develop more mode portable questions or questionnaires. The social capital conversion project has shown that questions and questionnaires can be adapted to accommodate the characteristics of different modes of administration but some fundamental differences will always remain. Designing questions and questionnaires for use in different modes at the same time through a unimode approach could further reduce any differences.
However, this concept of replacing questions developed for specific modes with a
generic question set suitable for use in any survey mode is something which may be
deemed counterintuitive for survey researchers and certainly represents a paradigm
shift in the way in which questions are designed, and further research is required to
establish exactly how it can be done.

It is worth noting that Data Collection Methodology Branch at ONS carried out an
investigation of key survey variables, such as income and education, on behalf of the
new Integrated Household Survey and adopted a ‘unimode’ approach to the design of
these questions. Recommendations for the redesign of the questions to operate in both
face-to-face and telephone modes were successfully applied. A combination of
qualitative question development and testing techniques were used together with
quantitative split sample experiments to check for validity, reliability and to monitor
the effects of any change to outputs (Harper et al, 2004; Mackown et al, 2005).

There is however scepticism as to whether it is in fact possible to account for survey
measurement error in this way. Developing entirely mode insensitive questionnaires
for use in a mixed-mode environment may not be wholly achievable. An alternative
approach is that further research should focus on developing and applying methods of
correcting for mode effects. For example, looking at ways to evaluate the validity of
responses in different modes or using imputation methods to adjust for mode
differences.

It is the authors’ view that it is better to attempt to prevent than correct for mode
effects and advise addressing the approach to the design of the harmonised question
sets, although realistically both approaches, prevention and post hoc adjustment, are
likely to be adopted.

References and further reading

on answers to scalar questions in telephone and web surveys. Paper presented at the
Second International Telephone Survey Methodology Conference, January 11-15
2006, Miami, Florida.

known as the Integrated Household Survey), available at:

Postal Questionnaire: The cognitive phase of the Social Capital Conversion project.


response. Paper presented at the annual meeting of the American Association for
Public Opinion Research, Delavan, Wisconsin.

differences in people’s answers to telephone and mail surveys’, in Braverman M.T.
and Slater J. K. (eds.) Advances in Survey Research: New Directions for Evaluation


Rasinski K.A., Mingay D. and Bradburn N.M. (1994) ‘Do respondents really ‘mark all that apply’ on self-administered questions?’, *Public Opinion Quarterly* 58, 400-408.


Subscriptions and Enquiries

The Survey Methodology Bulletin is priced £5 in the UK and EIRE or £6 elsewhere.

This edition, together with supporting documentation, can be found at:

http://www.statistics.gov.uk/about/services/dcm/reports_publications.asp

Paper copies of the current and previous editions are available from:

Survey Methodology Bulletin Orders
Data Collection Methodology
Office for National Statistics
Room A028
Cardiff Road
Newport
NP10 8XG

dcm.business.surveys@ons.gov.uk

Telephone: +44 (0) 1633 655641

www.statistics.gov.uk