

Mixed mode surveys:

a review for the NHS acute patient
survey programme

THE ACUTE CO-ORDINATION CENTRE FOR THE NHS
PATIENT SURVEY PROGRAMME

Chris Graham
Research Associate
Picker Institute Europe



Date published: 8th June 2007

Contacts

The Acute Co-ordination Centre for the NHS Patient Survey Programme¹
Picker Institute Europe
King's Mead House
Oxpens Road
Oxford
OX1 1RX

Tel: 01865 208127
Fax: 01865 208101
E-mail: acute@pickereurope.ac.uk
Website: www.nhssurveys.org

Key personnel

Sally Donovan
Esther Howell
Jason Boyd
Helen Sheldon
Julia Martin

¹ Previously the NHS Patient Survey Advice Centre

Contents

1	Executive summary.....	1
2	Conclusions and recommendations.....	2
3	Introduction.....	3
3.1	Context.....	3
3.2	Assumptions and constraints	3
3.3	What is mixed mode research?.....	4
3.4	Structure of this report	5
4	Modes of communication	6
4.1	Practical issues	7
4.2	Conclusions	9
5	Modes of data collection.....	10
5.1	Introduction.....	10
5.2	Current position – pen-and-paper.....	11
5.3	Interactive Voice Response (IVR)/Touchtone Data Entry (TDE).....	11
5.4	Internet and e-mail surveys	14
5.5	Bedside terminals.....	16
5.6	Electronic kiosks	18
5.7	Personal Digital Assistants (PDAs)	19
5.8	Handheld units	19
5.9	Mobile phone messaging	21
5.10	Conclusions	23
6	Other considerations relating to alternative modes of data collection	24
6.1	Timing of the survey.....	24
6.2	Linking responses to the survey sample.....	26
6.3	Conclusions	27
7	Response rates	28
7.1	Research background.....	28
7.1.1	Offering a choice of modes.....	28
7.1.2	Mode crossover.....	34
7.2	Conclusions	35
8	Mode effects.....	36
8.1	Mode effects in health surveys	36

8.2	Cognitive factors in mode effects	37
8.2.1	Audio versus visual presentation of questionnaires	37
8.2.2	Socially desirable responding	40
8.2.3	Acquiescence	41
8.3	Conclusions	42
9	Strategies for minimising mode effects.....	43
9.1	Unimode construction.....	43
9.2	Summary of questionnaire considerations relating to different modes of administration	46
9.3	Adjusting data to account for mode effects.....	47
9.4	Conclusions	48
10	Previous use of alternative data collection modes in surveys of NHS patients	49
10.1	Bedside terminals.....	49
10.2	Internet surveys	50
10.3	IVR.....	50
10.4	Conclusions	50
11	Cost implications	52
11.1	Research evidence.....	52
11.2	Conclusions	53
12	Bibliography.....	54

1 Executive summary

This report presents findings from an investigation into mixed-mode survey methodologies and their possible value with regards to the NHS national patient survey programme. It has been prepared by the Acute Co-ordination Centre for the NHS Patient Survey Programme at the Picker Institute, on behalf of the Healthcare Commission.

Whilst the current methodology for the patient survey programme generally produces high response rates and data quality, there is evidence that some groups of patients are not as well represented as others amongst the surveys' respondents. In addition, the past decade has witnessed a general pattern of decline in response rates to mail, telephone, and face-to-face surveys across all sectors, and health surveys are no exception. As such, any changes to the methodology that could improve response rates overall or within subgroups would be desirable.

Mixed mode surveys aim to balance the best possible and the best affordable methods for surveying populations and have been advocated as a way of increasing survey response rates. The rapid increase in the availability and accessibility of a range of new types of technologies for communication and data collection has fuelled the increased use of mixed mode approaches to surveys. In particular, advances in information technology allowing automated data collection – perhaps most notably the internet – have led to the possibility of vastly reducing unit costs for survey responses.

A range of alternative modes of data collection were investigated which could potentially be implemented alongside mailed surveys in the national patient survey programme. The strengths and weaknesses of each mode were considered in relation to NHS acute surveys. Similarly, we examined the possibilities of using different modes of communication to approach patients and encourage them to take part in these surveys. A range of existing and emerging technologies and approaches were considered, including automated telephones, web surveys, bedside terminals, hand held survey units and computers, and mobile phones. We looked at the key factors that would need to be considered before introducing new modes into the survey programme, and considered possible barriers that may complicate or prevent such change.

This report also details relevant published research on mixed modes methods and their impact on response rates, mode effects, and costs. Evidence shows that well implemented approaches can yield significant benefits in response rates, generally by contacting sample members through more than one mode, but that such improvements are by no means guaranteed: simply offering a choice of modes to respond by, for instance, seldom increases responses. The available literature also demonstrates the need to take account of mode effects, where individuals' responses may be affected by the particular mode of data collection. Due to the importance of generating comparable data throughout the programme, minimising such effects is extremely important.

Overall, this report provides considerable reason for caution regarding mixed modes research. Whilst the potential benefits are attractive and merit further exploration, so too is it clear that implementing such a methodological change without adequate testing and preparation could have severe adverse consequences for response and data quality. Based on our research we currently recommend continued monitoring of developments in mixed-mode methodologies and, in particular, pilot testing of the use of mobile phone text messaging for survey reminders.

2 Conclusions and recommendations

Based on this work, we make the following recommendations regarding mixed mode methodologies in the NHS acute survey programme in England:

1. It is critical that a consistent methodology for data collection is employed as far as possible in all trusts conducting the survey. At present **we believe a principally mailed methodology with paper questionnaires as the sole means of data collection to be the only viable approach to conducting the main national surveys**. This is because of problems related to the coverage of different modes and the comparability of data obtained from them.
2. We believe that SMS technology may function well as a hitherto untapped channel for engaging with respondents who are presently considered 'hard-to-reach'. As such we believe that there may be potential in the use of SMS text message reminders for non-responders. Furthermore, the financial impact of such a system would be extremely low, and we do not foresee any likelihood of responses being biased by a contact through this mode. Despite existing coverage issues and the fact that not all trusts will be able to access patients' mobile phone numbers, **we believe that the use of SMS reminders should be piloted in the near future**.
3. **It would be beneficial to conduct an in-depth investigation of the types, coverage, and quality of patient contact details currently collected by acute trusts**. This information would be invaluable for considering potential future use of mixed mode approaches.
4. We anticipate that the viability of using alternative modes of data collection, particularly the internet, will increase as coverage grows and people become more familiar with new technologies. However, there is a danger that overuse of such technologies by market research and sales organisations may lead people to become less receptive to completing surveys conducted in certain modes. **We believe that there would be value in continuing to monitor the development of new methodologies and their applications in the NHS or similar health sectors**.
5. Should any mixed-mode approaches be deemed useful in the future, **we would emphasise the importance of thorough and carefully controlled pilot testing** in order to develop a detailed understanding of the specific differences in response rates and demographic characteristics of responders, as well as response biases that may exist between different modes. If, as is likely, significant differences were found then it would be important to develop corrective analytic strategies for handling these, and these in turn would need to be checked regularly to ensure their continued validity.

3 Introduction

This paper presents the findings of an initial investigation conducted by the Acute Coordination Centre for the NHS Patient Survey Programme at Picker Institute Europe into the effects of using mixed mode methodologies in surveys. It is part of a wider agenda of development work being undertaken as part of the acute patient survey programme, which is funded by the Healthcare Commission.

3.1 Context

Since 2002, NHS trusts in England have been required to take part in national surveys of their patients as part of a commitment to collect the views of patients outlined in the NHS Plan². In the acute sector, these studies have looked at the experiences of adult inpatients (in 2002, 2004, 2005, and 2006), young patients (2004), of those seen in emergency (2003, 2004) and outpatients (2003, 2004) departments, and of emergency and urgent patients of ambulance trusts (2004). Whilst different surveys have targeted different groups of patients, the methodology employed from sampling through to data collection has remained largely constant. It has generally been possible to achieve good response rates and high data quality through the use of postal surveys with targeted reminders sent to patients who have recently been in contact with health services.

The Healthcare Commission is interested in looking at ways in which the survey process can be developed and improved for 2007 and 2008. Of particular interest are methodological variations that may lead to increases in response rates; inclusiveness (especially regarding younger people and black and minority ethnic groups, who are known to be underrepresented amongst respondents to the surveys); accessibility; timeliness; data quality; precision; or those that may reduce survey costs. Traditionally, the patient surveys have relied upon self-administered postal questionnaires, and it is not envisaged that this approach will be abandoned in the near future; the Healthcare Commission is interested, however, in the possibility of augmenting this approach with the use of further data collection modes to achieve some or all of the benefits described above.

3.2 Assumptions and constraints

Some basic assumptions are made throughout this document based on the current context of the survey programme. Certain features of the current methodology are unlikely to be changed for a number of reasons. Firstly, it is assumed that the surveys will continue to operate in a devolved fashion, with individual NHS acute trusts retaining responsibility for administering the bulk of the survey. No assumptions are made, however, regarding the possibility of using additional centralised samples, or providing trusts with access to centralised data collection systems. Secondly, financial constraints and the devolved nature of the surveys mean that methods of survey administration and data collection principally reliant on the use of human interviewers have to be ruled out for the main survey process – thus face-to-face and interviewer administered telephone surveying are not considered as viable alternatives to postal surveys in this document, although they are sometimes discussed in terms of comparison to other modes. Thirdly, for

² Department of Health. (2000) *The NHS Plan: A Plan for Investment. A Plan for Reform*. Department of Health: London.

consistency with previous surveys, and to ensure appropriate coverage of patients in the population of interest, it would not be advisable to make significant changes to the sampling frame or approach for surveys. Since contact information on these may be limited in some areas, some mediums of approach and contact may have to be ruled out: this could potentially apply to telephone contacts and would most definitely apply to e-mail contacts.

3.3 What is mixed mode research?

Mixed mode research is a broad term for research that utilises more than one mode of communication to contact, and/or more than one mode of data collection to obtain responses, from participants. Although there are a number of different ways in which mixed mode methodologies can be implemented, they may be best thought of in terms of two organising principles: modes of communication and modes of data collection^{following 3}. For instance, a survey with an initial mailed questionnaire phase may feature telephone reminders to increase response rates – mixed modes of communication. Similarly, a primarily telephone-based survey may utilise a dual-frame design such that a questionnaire can be administered to those without telephones in order to reduce the sampling error caused by their otherwise inevitable exclusion – mixed modes of data collection. Different modes may be used simultaneously or successively, and research designs may involve ‘crossover’ from one mode to another. Such survey designs have become more commonplace in recent years as survey developers seek to find cost effective solutions for minimising non-sampling error, especially in the light of a general pattern of decline in response rates to mail, telephone, and face-to-face surveys over the last decade^{eg 4,5}.

The purpose of using more than one mode of data collection for a survey, then, is to attempt to reach a compromise between the best affordable method and the optimal method were cost to be no object. In other words, “mixing modes gives an opportunity to compensate for the weaknesses of each individual mode at *affordable* cost”⁶. Clearly, any opportunity to reduce survey error whilst also lowering survey costs is bound to be attractive, but mixed mode methodologies still represent a relatively new approach and there is limited research on exactly how mode of data collection may influence responses, response rates, and survey costs. This is especially true of newer data collection modes, such as the internet and other electronic means of survey administration.

The research that has been conducted has shown that a range of potential ‘mode effects’ may occur. Factors such as the psycho-social context of the interview, the wording, response options, and the visual and/or verbal presentation of the stimuli, as well as other cognitive factors associated with respondents’ comprehension, evaluation, and reporting may all influence survey results. The effect of each of these factors is likely to vary across different modes, meaning that careful consideration must be given to survey design in order to minimise any instability in results caused by changes in mode.

³ de Leeuw, E.D. (2005) To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21 (2), pp. 233-255.

⁴ Atrostic, B.K., & Burt, G. (1999) Household non-response: what we have learned and a framework for the future. *Statistical Policy Working Paper 28, Federal Committee on Statistical Methodology, Office of Management and Budget*, pp.153-180.

⁵ Tourangeau, R. (2004) Survey research and societal change. *Annual Review of Psychology*, 55:775-801

⁶ de Leeuw, E.D. (2005) *op cit.* p. 235 (italics author’s own).

3.4 Structure of this report

This report is structured to provide a comprehensive introduction to and review of the range of different modes that could be used in the national survey programme and the existing research evidence regarding these and their likely effects. We begin by describing in detail the features of the current postal mode and of a range of alternative modes, incorporating a consideration of how the nature of each mode may relate to the survey programme in terms of both communication (4) and data collection (5). This is followed by a discussion of general considerations relating to the timing of data collection and the use of survey data (6).

In the subsequent sections we describe current research evidence regarding mixed mode surveys, looking particularly at how such approaches impact on response rates (7) and produce mode effects (8). Given the importance of ensuring comparability of data from different modes, we then focus on reported strategies for minimising and adjusting for mode effects (9). Finally, we describe the few previous attempts to use alternative modes for surveys of NHS patients (10) before looking at published evidence on the cost implications of mixing modes to determine whether such an approach could yield benefits in terms of cost-efficiency (11).

4 Modes of communication

Researchers routinely communicate with sample members at several points during a survey, providing information about the study, inviting people to take part, and issuing reminders to non-respondents. In the national patient experience survey programme, initial contact with patients is usually made by post, with up to two subsequent contacts also being mailed to non-responders. A number of studies have shown that switching between different modes of communication through the course of a survey is an effective means of reducing non-response and bias attributable to this⁷⁻⁸. In this section, we look at research evidence on the effect of using mixed modes of communication with sample members, and discuss the practicalities of employing such strategies in the survey programme.

Siemiatycki & Campbell (1984) conducted a health survey where patients were initially contacted and invited to complete a questionnaire either via mail or via telephone. Non-responders in either mode were subsequently approached via the alternate mode to participate in the same survey. In the initially-mailed arm of the survey, a response rate of 68.5% was attained in the first wave, rising to 80.9% after the telephone wave. Similarly, the telephone survey achieved an initial 72.7% response rate, rising to 88.3% after mailed follow-up. Comparing responses obtained from the two stages in both arms of the study showed that achieved samples were “adequately representative” of the entire sample on a range of socio-economic and demographic distributions. The authors concluded that there was no evidence of substantial non-response bias at either stage of either arm of the survey, and, comparing initial to final achieved samples, they argued that the mode crossover had little effect on the representativeness of the overall sample.

These results and the analytic approach used have, however, been questioned. Brambilla & McKinlay (1987) argued that comparing initial to overall samples lacked power, and that independently comparing samples achieved from the mailed survey to those obtained from the phone survey provides a better test of their comparability and representativeness. Doing this, and comparing results to independent population estimates from census data, showed that the use of a follow-up in an alternative modality reduced the degree of bias in the sample relative to the overall population of interest. In their own survey of 8,050 women aged between 45 and 55 in Massachusetts, Brambilla & McKinlay used a mode crossover from mail to telephone. Comparing samples achieved from the two waves of the survey, respondents to the initial mailed phase were found to be more likely to be from professional backgrounds, had greater average incomes, and were generally more educated than respondents to the subsequent phone wave. Reporting of health outcomes also varied by mode, with telephone respondents generally reporting better health but also more physical symptoms, and more frequent contact with and greater utilisation of health services. This variation was not eliminated by controlling for socioeconomic status and demographic variables: thus, differences in reports of health status did not appear to be the result of confounding by such traits. This indicates a potential for modality to impact upon responses (discussed further in section 5).

⁷ Brambilla, D.J. & McKinlay, S.M. (1987) A comparison of response to mailed questionnaires and telephone interviews in a mixed-mode health survey. *American Journal of Epidemiology*, 126: 962-971.

⁸ Siemiatycki, J., & Campbell, S. (1984) Nonresponse Bias and Early Versus All Responders in Mail and Telephone Surveys. *American Journal of Epidemiology*, 120: 291-301.

4.1 Practical issues

As discussed earlier, there may be limitations to the feasibility of switching from the current postal approach to patients used in the acute patient survey programme due to the sampling frames utilised. Firstly, contacting patients by e-mail would definitely have to be ruled out: despite consulting with staff at a number of acute trusts we are presently unaware of any that routinely record e-mail addresses for admitted patients. Generally, trusts will only have patients e-mail addresses if the patient has specifically requested that this be recorded. Any trusts that do attempt to record e-mail data are likely to have sporadic and inaccurate records. Coverage of e-mail addresses is low, so many patients will have none listed, whilst those that are recorded run the risk of inaccuracy – e-mail, unlike postal addresses, must be exactly correct and have no resilience to incorrectly entered digits – or being obsolete, as people change e-mails relatively frequently. Secondly, there is a risk that some trusts may be unable to provide consistently accurate telephone numbers for their patients.

As part of this report, the Acute Co-ordination Centre has undertaken some preliminary investigations into the recording of telephone contact information by acute trusts. The picture emerging is that there are significant variations in the extent to which trusts seek to record telephone contact information for inpatients. At one extreme, some trusts routinely seek to collect and record multiple telephone numbers – including, for example, work, home, mobile, and next-of-kin numbers – from all inpatients and report that at least one of these numbers is recorded for up to 100% of patients. One trust contacted, for example, stated that all inpatients are asked to provide as many contact phone numbers as possible, and that they were “yet to see any instances where a patient had refused to provide at least one” such number⁹. At the other end of the spectrum, however, we also spoke to a number of trusts who believed that whilst their PAS (Patient Administration System) contained fields for entering telephone numbers, patients were not routinely asked to give these, with the trust generally preferring to rely upon postal addresses. These trusts suggested that they would have telephone numbers for only a limited number of patients. Other trusts reported that recording of telephone numbers for inpatients was sporadic. One trust, for example, reported that “the majority of patient records contain details of a contact telephone number. However... the number [that] do not... is quite significant.”¹⁰

On balance, this suggests that it may be difficult to implement telephone contact of patients in a consistent way across the country. Whilst there are trusts that would be capable of providing at least one telephone number for all patients, there are many others who would be able to do this only in a minority of cases, and trust staff were generally unable to explain why these details may be obtained for some patients but not for others. A possible explanation may be that the severity of a patient’s illness or injury may be a related factor; another that their expected care and discharge arrangements may be connected, especially where their eventual care pathway may have a bearing on the trust’s need to contact the patient in different ways. One trust contact, for instance, stated that inpatients were much more likely to have their telephone numbers recorded if they had previously experienced, or were scheduled to receive, outpatient care at the trust. This

⁹ Hall H., Administrative Co-ordinator, Medical Records, Northumbria Healthcare NHS Trust. *Personal correspondence (telephone)*

¹⁰ Lowe, V., Patient Survey Co-ordinator, Clinical Audit Department, Shrewsbury and Telford Hospital NHS Trust. *Personal correspondence (via e-mail)*.

was because being able to ring patients to inform them about appointment times and details was seen as a higher priority in an outpatients setting.¹¹

Some evidence on the nature of record keeping relating to telephone numbers in acute providers has been generated by survey development work carried out for the Healthcare Commission's pilot survey of NHS and independent sector treatment centres¹². This work again gives a mixed impression regarding the availability of such data. None of the treatment centres were able to provide complete telephone records, and one could not provide any, although in total phone numbers were available for 88% of day case patients. These numbers were not recorded in a consistent format, however, with dialling codes sometimes missing. In centres with many missing telephone numbers, it was necessary to employ commercial organisations to use other patient data to determine contact numbers. Significant problems were found with this approach, however, with phone numbers generally obtainable for less than half of all patients¹³. Of the 88% of day case patients for whom numbers were eventually available, 5% were found not to work – suggesting an overall coverage error of around 16%. Given the different contexts in which such investigations have been undertaken, further investigation of the types, coverage, and quality of patient contact details routinely recorded by acute trusts is needed to determine clearly whether there is any possibility of using mixed modes for communicating with patients.

If the national surveys continue to use a sampling frame that provides only one consistently reliable route of contacting patients from all trusts (eg postal), then – assuming that interview based data collection methods are considered prohibitively expensive – this would significantly limit our freedom to utilise different technologies. Two options remain: adhering to the use of only postal contacts to maintain a strictly consistent approach across all trusts, or varying the methodology depending on the ability of each trust to provide systematically accurate contact details for different modes. Continuing to use a completely consistent approach, with all trusts applying an identical methodology, has clear advantages in terms of overall transparency and comparability of results, but these benefits do not necessarily constitute *a priori* grounds to dismiss any alternative approach. Allowing some variation, so those trusts that do have access to alternative means of contacting patients may use these, could be desirable if clear improvements in the survey process or quality can be demonstrated. In particular, it seems reasonable to hypothesise that alternative means of contact may be particularly useful in boosting response rates for certain sub-groups within the survey; for example, patients from black and minority ethnic groups¹⁴. This could come to represent an important improvement in data quality in some areas, and may be seen as a sufficiently attractive benefit to offset the potential negative consequences of a corresponding reduction in the overall consistency of the survey process across all trusts.

¹¹ Higgs, S., Clinical Effectiveness Manager, Royal West Sussex NHS Trust. *Personal correspondence (telephone)*.

¹² Healthcare Commission. (2006). Development report for the survey of NHS and independent sector treatment centres . Unpublished.

¹³ This generally resulted from patients being ex-directory or not the bill-payer for the telephone line. It is interesting to note that a similar approach of collecting patients' phone numbers from the telephone directory was used by Brambilla & McKinlay (1987), who were able to obtain numbers for 96.8% of their sample. Whilst this notable drop in the number of traceable patients partly reflects demographic differences between the samples, it also suggests that manually identifying phone numbers may now be more difficult than in the past (see also: Tourangeau, R. (2004) Survey Research and Societal Change. *Annual Review of Psychology*, 55: 775-801).

¹⁴ Eg Sheldon, H., Graham, C., Potheary, N., & Rasul, F. (2007) Increasing response rates amongst black and minority ethnic and seldom heard groups – a review of literature relevant to the national acute patients' survey. Oxford, UK: Picker Institute Europe.

The difficulties of using a mixed-mode approach with only one means of communication with sample members may be significant. If the surveys continued to operate based only on postal addresses then the only mixed-mode approach that would be possible without utilising a dual sampling frame design would be to make an initial mailed approach with invitation to respond by a choice of modes. Whether such a strategy could increase response rates, improve data quality, or be implemented without increasing survey costs would depend heavily on precisely how the initial contact was made and how respondents reacted to the choice of response mode. For example, if letters offering the choice of response by post or internet were sent to sample members *with a questionnaire enclosed*, then the cost of administration would only decrease if the overall response to the initial wave of the survey increased (as fewer reminders would need to be sent). Conversely, if response rates to the initial mailing did not increase, then it would be likely that overall survey costs would increase due to greater deployment costs unless a high proportion of respondents chose to use alternative responding methods with a significantly lower unit cost for processing.

4.2 Conclusions

Using multiple modes of communication in a survey, either simultaneously or sequentially in a crossover approach, may help to make achieved samples more representative and decrease nonresponse bias. This is especially true where certain groups are more likely to acknowledge communications via certain modes than via others: in other words, to show a preference to one mode over others. As such, using mixed modes of communication could help increase response rates to the survey, particularly amongst hard to reach groups who are less likely to respond to postal surveys.

At present, however, obstacles remain as to the feasibility of contacting patients by alternative modes. Contact methods that rely on details other than postal addresses are likely to be characterised by relatively high coverage error. This is especially notable for e-mail, where contact details would be hard to obtain and coverage poor, but also for telephone, where record keeping varies substantially between trusts. Thus whilst it might be possible to employ telephone contacts for patients of some trusts, it would not be universally possible, and mail remains the only mode that can be used consistently and reliably throughout England. Allowing variations between trusts in the approaches used to contact patients may, however, be acceptable if the benefits could be shown to outweigh the corresponding drawbacks of such methodological inconsistency. More research is required to determine the impact of specific approaches.

5 Modes of data collection

As new means of collecting survey data continue to emerge, so too do further potential combinations of methods. Mixing data collection methods may be desirable where there are known limitations with any given method relative to the aim of the survey. These limitations may include issues associated with unit or item non-response, where certain groups are known to be unlikely to respond using particular modes, or where cost constraints limit the use of a preferred data collection method. In this section we consider a range of data collection modes that are presently available, looking in detail at their characteristics and suitability for the NHS patient survey programme.

5.1 Introduction

Different data collection methods vary in a number of important ways. At a broad level, questions may be presented to participants either visually, in audio, or in a combination of these approaches. Furthermore, questionnaires may be static, such as a paper questionnaire which provides all the required information in a fixed format and where any routing must be 'opaque' (that is, explicitly presented to all respondents in the form of instructions), or questionnaires may be dynamic. Dynamic questionnaire design may utilise multimedia to further illustrate questions or concepts, incorporate advanced administrative features such as question order randomisation, – or apply 'transparent' routing, where questionnaires are effectively 'tailored', dependent on previous responses, to feature only items relevant to the individual respondent. These basic characteristics of different modes influence the kinds of questions that can be asked through them and the ways in which participants may be expected to respond, with certain question styles generally being favoured in particular types of questionnaire but avoided in others. Even such basic considerations as the number of questions that can feasibly be asked must be balanced against the choice of mode, taking into account the varying levels of burden that completion will place on respondents and the resultant acceptability of the overall survey. This is related to the means by which participants respond to survey questions, which may vary considerably from mode to mode. This can lead to survey participants utilising different modes having to perform different tasks from one another. Once again the distinction between visual and audio information is important, but even where visually based responding is used there are important differences in the kinds of knowledge and interaction involved in responding by ticking boxes, using a computer mouse and/or keyboard, pressing buttons, or inputting responses on a touch-screen. These differences, and the differences in the presentation of questionnaires discussed above, are likely to have effects on overall data because they affect the kinds of social exchange taking place in the survey interaction as well as the degree of cognitive load it produces, and these seemingly distal factors may influence people's reports of their experiences¹⁵¹⁶¹⁷¹⁸.

¹⁵ eg Tourangeau, R. (1984). "Cognitive Sciences and Survey Methods." In Jabine, T., Loftus, E., Straf, M., Tanur, J., and Tourangeau, R., eds., *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington, D.C.: National Academy of Science.

¹⁶ Schechter, S., Beatty, P., & Block, A. (1994) Cognitive issues and methodological implications in the development and testing of a traffic safety questionnaire. *Paper Presented at the 49th Annual Conference of the American Association for Public Opinion Research. May 11-15, 1994. Proceedings of the Survey Research Methods Section, American Statistical Association.*

5.2 Current position – pen-and-paper

The use of self-administered paper questionnaires is a robust and generally a cost-effective method of collecting feedback from patients, but it is not without its limitations. Questionnaires returned by patients must either be manually data entered or scanned. This can be a time consuming and laborious process, limiting the speed at which feedback can be obtained through postal surveys. In order to ensure an acceptable standard of data quality, it is then necessary for all data entered to be centrally cleaned and edited to ensure that questionnaires have been completed appropriately. Furthermore, self-administered paper questionnaires are generally considered to carry a relatively high cognitive burden, since they require respondents to be able to read the language used in the survey (which may be limited both by linguistic barriers and/or visual impairments), to have sufficient physical dexterity to complete the questionnaire (not necessarily a trivial concern in surveys featuring high proportions of older people recently discharged from hospital), and to be able to follow routing instructions¹⁹. As such, it is possible that the use of alternative forms of data collection, most likely in addition to paper questionnaires rather than instead of, could produce benefits both for those using survey data and for respondents completing the survey.

5.3 Interactive Voice Response (IVR)/Touchtone Data Entry (TDE)

Telephone surveying with an interviewer asking questions of participants selected either from a pre-existing sample list or by using random-digit dialling (RDD) has been a common research methodology since the 1970s and continues to be used extensively today. However, its feasibility for very large studies such as the acute surveys is limited because of the costliness of using interviewers. Interactive Voice Response (IVR) offers an alternative that is similar in form to a telephone interview but the human interviewer is replaced with a synthetic or recorded voice, making it considerably cheaper to run at scale. With this method, data is captured automatically by a computer system that recognises either vocal responses (IVR) and/or responses entered on the keypad of a touchtone phone (TDE), without the need for these to be keyed by an interviewer or operator.

One issue in the use of IVR that may raise important design considerations is that of 'humanisation' or 'personalisation' of the voice used in the IVR system. As mentioned above, either a synthetic or recorded human voice may be used to relay instructions and read questions to survey participants, but the choice between these may not be trivial and may lead to distinct mode effects. Underlying this possibility is the way participants react and relate to the IVR system; whether they think of the voice they hear as human, or as a machine. Whilst it remains impossible to engage IVR systems in anything approaching normal human conversation, research has shown that respondents favour IVR systems with anthropomorphic qualities over 'less human' ones. One experiment, for instance, demonstrated that respondents preferred an IVR system with 'humanistic' qualities – the use of the pronoun 'I' in its dialogue – to one without, even though 80% of those

¹⁷ Cacioppo, J.T., & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

¹⁸ Roberts, D. (2005). Recreating experiences. *Evaluation Journal of Australia*, 4(new series), 1&2, p44-51.

¹⁹ Eg Bowling, A. (2005) Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, 27(3):281-291

respondents reported not noticing the crucial difference between the two systems²⁰. As a result of this, some researchers have made efforts to make an IVR voice seem more human in an attempt to make using the system more satisfying to respondents. For example, studies have attempted through dialogue and tone to “convey animacy... and informality” in the IVR voice: yet at the same time, it has been necessary for these studies to note the possibility of undesirable effects arising from this approach, such as respondents being “less likely to report potentially sensitive or illegal information to a system that has humanlike qualities”²¹. This potential drawback is at odds with one frequently stated advantage of electronic data collection modes: that they ensure high degrees of perceived confidentiality²². Taking this argument even further, it has also been claimed that it may be normal for people to treat computers and similar media as if they were real social actors, even in the absence of humanisation or anthropomorphism²³. Given the technical limitations on the responsiveness of IVR systems, this could in some cases “raise respondents’ expectations about the system’s capabilities to unrealistic levels”²⁴, particularly where humanising features are used. This could raise levels of break-offs if peoples’ experiences with the system fail to meet their expectations and they become frustrated with its apparent inflexibility, inaccuracy, or unresponsiveness.

Another fundamental question associated with the use of IVR systems for surveys is whether or not this technology really does anything to reach otherwise inaccessible groups within the population. It is not clear whether any hard-to-reach group would prefer, or would be more inclined, to respond via an IVR system as opposed to any other survey method²⁵. That said, evidence regarding this issue is mixed and it has sometimes been suggested that response rates for certain groups, notably low income groups, may be increased by the use of IVR response options²⁶.

Length of the questionnaire instrument may also be an important issue when considering the use of IVR for surveys. Whilst many people are willing to complete lengthy telephone interviews with human interviewers, it is not necessarily clear that questionnaires of similar length could be employed via IVR. Regardless of whether or not respondents treat IVR systems as social actors it would appear that people are considerably more likely to hang up on an IVR system than on a real person; this is reflected in the higher break-off rates commonly observed in IVR studies versus comparable telephone studies. Providers of IVR systems strive to limit the number of such break offs. By attempting to make systems more interesting, responsive, and efficient, at least one provider claims a mean break-off rate of under 5% for questionnaires of up to 20 items:

²⁰ Boyce, S.J. (1999) Spoken natural language dialogue systems: User interface issues for the future. In: Gardner-Bonneau, D. (ed.) *Human factors and voice interactive systems*. Boston: Kluwer Academic.

²¹ Conrad, F., & Couper, M. (2004). Usability, comparability, and data quality across modes and technologies in census data collection. US Census Bureau, p18.

²² eg Tourangeau, R., & Smith, T. (1998) Collecting sensitive information with different modes of data collection. In Couper, M., Baker, R., Bethlehem, J., Clark, C., Martin, J., Nichols, W., & O’Reilly, J. (eds.) *Computer assisted survey information collection*. New York: Wiley.

²³ Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Massachusetts: Cambridge University Press.

²⁴ Conrad, F., & Couper, M. (2004), p18 *op cit*.

²⁵ Conrad, F., & Couper, M. (2004) *op cit*.

²⁶ Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., & Berck, J. (2001) Response rate and measurement differences in mixed mode surveys: Using mail, telephone, interactive voice response, and the internet. *Draft paper, last accessed 11/10/06, available at:* http://www.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20_with%20Callup_%20POQ.pdf

nonetheless, they do acknowledge the difficulty of getting respondents to complete longer instruments via this mode²⁷. Published research tends to draw similar conclusions about the fickleness and propensity to break-off of respondents to IVR surveys. Rodriguez et al (2006), for example, report a 5% break-off rate in a 16 item IVR health survey, compared with much lower rates of abandonment in identical web (1.2%) and mail (0.5%) survey²⁸, and Schneider et al (2002) comment that “IVR/ASR respondents who do not like what they hear will quickly become nonrespondents”²⁹. Similarly, in an analysis of results from the 2000 US Census test, Caspar (2003), used a questionnaire of considerably varying length (depending on the number of people within each household) and, noted that “respondents appeared to dislike lengthy surveys with [IVR]”³⁰.

In terms of the distinction between Interactive Voice Response (IVR) and Touchtone Data Entry (TDE) systems, IVR or mixed IVR and TDE systems have clear advantages over TDE only systems. This suggests that there is little merit in pursuing a TDE only data collection method. Firstly, TDE systems cannot be used by people with rotary or pulse-dialling phones; this would have a significant impact on the coverage afforded by such systems, particularly amongst lower income and older groups of patients. Secondly, a small majority of respondents tend to prefer speech recognition systems to TDE. One early study found 60% of respondents to prefer IVR over TDE, though this study was conducted in 1992 and it is likely that this figure will have since increased with growing public familiarity with responding to such systems³¹. Thirdly, IVR would be expected to be considerably more acceptable to mobile phone users, for whom the inconvenience of entering data from the phone handset should be apparent.

It should be noted that as well as the problems relating to humanisation and animacy, which were discussed earlier, another major issue in the use of IVR in mixed-mode surveys also featuring self-administered paper questionnaires or similar is the distinction between audio and visual presentation and responding. There are a number of reasons why the differences between these two kinds of presenting and responding to data may lead to divergences in responses achieved. These issues are discussed in depth in Section 8.

²⁷ Livock, T. (2006, Voxgen). Personal communication.

²⁸ Rodriguez, H.P., von Glahn, T., Rogers, W.H., Change, H., Fanjiang, G., & Safran, D.G. (2006) Evaluating patients' experiences with individual physicians: a randomised trial of mail, internet, and interactive voice response telephone administration of surveys. *Medical Care*, 44(2), p167-174. Note that the percentage 'abandoning' mailed surveys is likely to be an underestimate, since it represents only those who return partially complete questionnaires: it is not possible to numerate how many people begin to complete a questionnaire, abandon it, and never return it.

²⁹ Schneider, S.J., Cantor, D., Heller, T.H., & Brick, P.D. (2002) *Pretesting Interactive Voice Response/Automated Speech Recognition Surveys*. Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing Methods, Charleston, South Carolina, November 14-17, 2002.

³⁰ Caspar, R. (2003) Census 2000 Testing, Experimentation, and Evaluation Program Synthesis Report No. 18, TR-18, Results from the Response Mode and Incentive Experiment in 2000. U.S. Census Bureau, Washington DC 20233.

³¹ Clayton, R.L., & Winter, D. (1992) Speech data entry: Results of a test of voice recognition for survey data collection. *Journal of Official Statistics*, 8, 377-388.

5.4 Internet and e-mail surveys

Public access to and familiarity with the internet has risen sharply over the last decade and this has led to an explosion in the use of the internet for surveys, particularly in the market research sector. Some industry figures have gone as far as to suggest that internet surveys will become the most commonly used of all survey modes in the not too distant future, but for now concerns regarding their coverage and accessibility means that their deployment for healthcare surveys has been extremely limited. In the UK, 60% of people report having used the internet in the last three months, with 80% of these people reporting that they used e-mail in that period³². This, however, still means that more than half of all people have no e-mail access at all. In the NHS, collection of patients' e-mail addresses is extremely rare. Furthermore, it is likely that e-mail address records that do exist will suffer from the same kinds of problems experienced in other industries; notably, that errors in transcription and data entry of e-mail addresses are common, especially around the use of symbols such as underscores, meaning that many addresses on databases generated from written or verbal records are inaccurate. Additionally, people may change e-mail address with considerably greater ease than they may change their home address, so many addresses may quickly become out-of-date. Due to these numerous problems, it is believed that contacting patients via e-mail is not currently feasible, despite the low costs and administrative workloads associated with this process.

Given this, any attempt to use the internet within the NHS patient survey programme would most likely have to rely upon inviting participants, via another mode, to log on to a given URL and use a web based questionnaire to complete the survey. In practice this could be achieved in a number of ways, although the most obvious of these, and the one that bears the most similarity to the current methodology, would be to send patients a letter offering them the choice to complete the survey on the internet. Since around 40% of people are generally unable, unwilling, or unlikely to access the internet, however, this would have to be optional with patients also free to choose other, more universally accessible, modes in which to respond.

There are a number of obvious advantages to using the internet for data collection suggesting that, if it could be implemented effectively, it would be an extremely useful tool. Firstly, data collection via the internet is extremely timely and cost effective. Once a set of responses is submitted on the internet, it is transmitted virtually instantly and that data may be used straight away. Secondly, since an internet survey does not require the printing of paper questionnaires, the mailing back of responses, or any form of data entry, the actual unit cost of collecting a set of responses on the internet is negligible. The only real costs associated with internet surveys are development of the instrument, purchase and set up of software applications, and server hosting costs (eg, the price of keeping the survey available on the web, including hardware and network costs).

The internet also offers a host of benefits over traditional pen and paper questionnaires in terms of the range of ways in which questions of different types can be presented. Modern computers allow the possibility of multimedia presentation of questions, with the ability to accompany text with images, sounds, and videos. Likewise response options can be illustrated, and can be set to vary dynamically depending on the selected option: a classic example of this from consumer satisfaction surveys being sliding scale response ranges illustrated by cartoon faces that frown when the scale is low but move to a smile when the scale is high. Such potential to visually illustrate the meaning of responses may be especially useful in making surveys more accessible to

³² National Statistics. (2006). First release: internet access, households and individuals. 23rd August 2006. National Statistics: London.

groups with low literacy or learning disabilities. Similarly, more innovative and complex response systems are possible. Computer software allows the use of visual sorting exercises (like card sorts, as may be used in an interview), click maps (where an image is presented and locations of respondents' clicks on the image recorded), and makes it extremely easy to collect data from visual analogue scales, which are infamously time consuming to translate to data in paper-based questionnaires.

Electronic presentation also gives rise to a number of challenges. Whilst the presentation of questionnaires on paper has a long history and, consequently, has been thoroughly researched to a point where a good quality questionnaire can be informed through observing established canons, there are many more variables to be considered in web surveys. For instance, should the entire questionnaire be presented on a single web page, increasing the initial loading time and forcing the respondent to scroll through as they answer questions, or across a range of smaller pages, increasing incremental loading time? If the latter option is chosen, respondents may need to be made aware of their progress through the questionnaire as a whole so that they do not become frustrated and fail to complete the survey, but even the exact presentation of such indication of progress needs to be considered as it can have significant effects on data quality, and, in some cases, progress bars may even have a negative effect on completion rates³³. These are by no means the only kinds of design issues that need to be considered, but they are indicative of the increased complexity of designing good web versus postal surveys.

Not only are there not always clearly preferable options amongst the choices that the web offers questionnaire designers, but the actual values of variables will often be hidden from the researcher: unlike paper surveys, the final appearance of the questionnaire when viewed by an individual respondent will never be dependent solely on choices made by the researchers who devised the instrument. Factors related to the hardware – eg monitors, graphics cards, and other components – and software – particularly operating system, web browser, and display settings such as colour depth, resolution, and accessibility aids – will partly determine the exact appearance of the questionnaire. The effects of this may be significant. For example, a survey designed to display at full window height on a computer with a screen resolution of 1,024*768 pixels may obscure a bottom row of questions when viewed on an otherwise identical computer with a screen resolution of 800*600 pixels, forcing the latter respondent to scroll down to complete the page. Likewise, such seemingly innocuous differences as switching from one web browser to another can lead to significant changes in the appearance of the website as whole. Depending on the particular design chosen, such factors could have an impact on data quality and reliability. As such, web questionnaires need extensive testing on a range of systems before being deployed; even then, though, it is always likely to be hard to predict the sheer range of setups that the instrument will be confronted with when accessed by a variety of different users³⁴.

Internet surveys should, at least theoretically, allow generally increased levels of data quality. The survey software can keep track of respondents' answers and perform dynamic functions such as transparently routing participants past irrelevant questions and alerting them via pop-up messages to missing, inconsistent, or inaccurate responses. At an extreme, respondents can even be 'forced' to give a response to each item on a page before continuing – although this possibility again raises new design questions as it is often advisable to allow respondents to completely ignore a question if they wish. Whether or not web surveys do lead to improved data quality,

³³ van der Horst, W., Snijders, C., & Matzat, U. (2006). The effects of different kinds of progress indicators on online survey compliance and data quality. *Presentation from the 8th International General Online Research (GOR) Conference, March 21 2006.*

³⁴ Eg Dillman, D.A. (2000) *op cit.*

however, remains questionable and seems to depend heavily on the particular design and set up of the instrument and the backgrounds of responders. Several studies have explicitly looked at data quality in mailed versus internet surveys by comparing item nonresponse on matched instruments, but results have been mixed. For example, McMahon et al (2003)³⁵ report significantly lower levels of item nonresponse – and therefore higher data quality – in internet versus paper-based questionnaires, but other studies have reported an opposite effect, with higher item nonresponse observed in web than in mail surveys³⁶. Overall, it remains unclear as to whether web administration does, in practice, improve data quality.

As well as improvements to data quality on closed questions, there are benefits to the quality of data recorded from open-ended questions when using internet as opposed to interview surveys or pen-and-paper self-administered questionnaires. Because respondents are able to type in their open-ended responses, it can be assured that their comments are entered verbatim. This is in contrast to interviewer administered methods, where even the best interviewers will tend to paraphrase and interpret to some extent; and also to paper questionnaires, where the quality of comments can suffer in data entry, especially where legibility of hand-writing is an issue. How mode of administration affects the *quantity* of open-ended comments, however, is less clear. Schaefer & Dillman (1998) report 12% more responses to a final 'additional comments' question in the e-mail than in the postal arm of a mixed-mode survey, with comments collected via the internet four times more verbose on average, but results from this may not be easily generalisable as respondents were sampled from a university faculty and may have been particularly at-ease with the use of e-mail³⁷. Furthermore, this study was conducted some time ago when 'spam' e-mail was far less prevalent, and e-mail messages were more likely to be read and attended to. A more recent study, by contrast, reported fewer respondents giving open-ended comments over the internet than by mail – although responses were still more verbose when collected via the internet³⁸.

5.5 Bedside terminals

As part of the NHS plan, all large acute trusts are now expected to provide inpatients with access to personal bedside TV, telephone, and radio services³⁹. A number of private companies such as PatientLine, InTouchWithHealth, Trident (Zevo), and HTS (Hospital Telephone Services Limited)

³⁵ McMahon, S.R., Iwamoto, M., Massoudi, M.S., Yusuf, H.R., Stevenson, J.M., David, F., Chu, S.Y., & Pickering, L.K. (2003). Comparison of E-mail, Fax, and Postal Surveys of Pediatricians. *Pediatrics*, 111, 299-303.

³⁶ eg Kwak, N., & Radler, B. (2002) A comparison between mail and web surveys: response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18 (2), pp. 257-273.

Mi Kyung, J. (2005). Effects of survey mode, gender, and perceived sensitivity on the quality of data regarding sensitive health behaviours. Indiana University.

³⁷ Schaeffer, D.R., & Dillman, D.A. (1998) Development of a standard e-mail methodology: results of an experiment. *Public Opinion Quarterly*, 62(3), 378-397.

³⁸ Ballantyne, C. (2004) Online or on paper: An examination of the differences in response

and respondents to a survey administered in two modes. *Paper presented at the Australasian Evaluation Society 2004 International Conference 13-15 October-Adelaide, South Australia.*

<http://www.aes.asn.au/conferences/2004/FR21-Ballantyne,%20C.pdf> (last accessed 27/03/07)

³⁹ Department of Health. (2000) The NHS Plan: A Plan for Investment. A Plan for Reform

have developed bedside terminals. The most common of these systems, PatientLine, is currently installed in over 155 NHS hospitals across roughly 54% of English trusts⁴⁰.

Surveys conducted on bedside terminals would generally be similar to internet surveys in that the manner of presentation and data collection could vary and that they should be fairly flexible as a survey tool overall. As with the internet, audio, visual, and multimedia presentation could be used, and patients could use a keypad to respond to a variety of different types of question, although opened ended questions would generally not be well suited to this type of surveying since not all bedside terminals are equipped with keyboards for typing in responses. Questionnaire length is theoretically unlimited, but it remains to be seen how long a questionnaire patients would be willing to complete on one of these systems, and what kind of effects questionnaire length may have on response rates and user satisfaction with the survey process.

As with many electronic systems, it has sometimes been suggested that conducting surveys using bedside terminals may create a barrier to access for certain groups who may be intimidated by or lack confidence in using such systems. In particular, it has been suggested that older patients may sometimes have difficulty using bedside terminals. There is some evidence to support this idea; in a survey of users of bedside terminals conducted by BMRB on behalf of NHS Estates/Patient Power Review Group, patients aged 75 or older were significantly less likely than younger patients to report that systems were 'very easy to use' (38% v ~57%, $p \sim 0.005$). Likewise, patients over 75 years old were significantly more likely than all other patients to report that their bedside terminal was *not* easy to use (18% v ~4%, $p < 0.001$)⁴¹.

A common complaint from patients about bedside terminals is that they are expensive to use. This complaint may however represent an advantage of the unit for encouraging people to complete surveys, since we have discovered that it is possible to set terminals up such that patients completing surveys are given free credit for doing so; thus, credit for the service, which has high perceived value for inpatients and which is salient to the care experience, can be used as an incentive to encourage responses. Presumably this would be beneficial for survey operators as well, as it should be possible to arrange discounted purchase of such credit in bulk; this, however, would require further investigation with suppliers. A potential accessibility problem may, however arise from the use of bedside terminals to complete patient surveys, for it is not clear what the net effect would be should the terminal crash during the process of a patient completing a survey. If it became necessary, for example, for a patients terminal to be reset or have software reloaded in some way, they could become appreciably annoyed at the process.

One problem with assessing the potential usefulness of bedside terminals for patient surveys is that, although their use is gradually becoming more popular, we are not yet aware of any published research to explicitly look at their application in this kind of context. Whilst we have been in contact with a number of organisations who have attempted to use the terminals to collect patient feedback, none have conducted a rigorous investigation into factors such as response rates, mode effects, or validity of this approach. This reflects the fact that their key benefit tends to be seen as their immediacy – results can be delivered instantly without the need for manual collation or analysis. Consequently, it would seem that this kind of survey is typically thought of as being best

⁴⁰ Source: PatientLine

⁴¹ BMRB Social Research (2005). *Patient Power (Bedside TVs and telephones): Survey of users, staff, providers, and patient representatives*. London: BMRB. Note: Significance test not reported in published data – figures are calculated from data in tables within publication and should be taken as close estimates rather than exact; nevertheless, it is clear from the absolute sizes of the differences observed, and the low estimated p values, that they are significant.

suited to 'on the fly' research looking at a limited number of questions centred on the key issues at any given time.

The difficulty of testing the systems properly is further enforced by the default outputs, which do not link patients' responses between questions. Instead, in order to help preserve confidentiality in this high-speed point-of-care system, bedside terminals have tended to provide weekly averages for all questions, but have not allowed breakdown of responses by demographic characteristics or the crosstabulation of different questions. This makes it impossible to properly assess mode effects and so on, because there is no way of 'drilling down' into the data or picking apart trends; only top-line figures exist, and, since they are stripped of the variables that help to explain results, these are of little use. This is not, however, a fundamental limitation of the system – the makers of the most commonly used terminals, PatientLine, have indicated that their terminals do allow for more complex data collection, but this is something that has not been previously implemented simply because the survey application has not been designed with research in mind. This does though imply that the applications currently used would require significant overhaul before being useful for the national survey programme.

5.6 Electronic kiosks

Electronic kiosks, which offer touch screen or keypad based data entry, are becoming a familiar sight in England and are beginning to appear more frequently in health care settings as NHS providers seek to utilise these as a modern resource for disseminating information to patients and the public. Electronic kiosks used for collecting survey data vary massively in terms of their complexity and design, from very simple units with an ATM-like numeric keypad and a printed sheet of paper listing questions, to elaborate touch screen systems that allow the computerisation of survey presentation and the use of multimedia content. Different systems may have their own strengths and weaknesses, so the selection of a particular system would be of critical importance to the effectiveness of its use in terms of cost, response rate, and data quality.

For more complex systems, it is likely that many issues of survey administration and accessibility would be similar to those for bedside terminals. Whilst little or no work has been done to test this hypothesis, it would seem probable that, as with bedside terminals, older patients may experience greater degrees of difficulty in using these systems, and may report lower satisfaction with the systems than younger patients. This idea needs testing, however, to be considered reliable. Should these effects occur, though, it is unlikely that they would persist or be as pronounced for the far simpler styles of machines featuring only numeric keypads and a set list of questions, since people tend to be considerably more familiar with this type of technology - the earlier comparison to ATMs is illustrative of how widespread such data entry systems are. As with all electronic systems, however, the question of how to best handle system crashes remains open.

The greater accessibility of the simpler machines comes at a cost, however, since their simplicity also leads to their being relatively inflexible. Multimedia and, generally, audio presentation of questions are not possible with these and the systems involved are relatively 'dumb' – they do little more, on their own, than encode the pattern of key presses. In analysing data from such units one must be explicitly aware of the questionnaire used and the ordering of questions and responses, as unlike more complex systems it is not possible for them to relate responses to a certain program or script. Even these kinds of kiosks do, however, allow some relatively sophisticated analysis if properly approached. Latency time between responses, for example, is usually recorded; so not only can 'problem' responders - eg, those simply pressing buttons at random - be removed, but mean evaluation times can be calculated for different questions, and outlying cases identified. This may make these kinds of machines useful in pilot, as their capacity to automatically collect paradata of this type allows interesting analysis of peoples' cognitive interactions with the questionnaire.

5.7 Personal Digital Assistants (PDAs)

In recent years, PDAs, originally conceived as simple personal organisers, have effectively developed into small 'palm-top' personal computers, capable of playing audio and video, running complex software applications, connecting to the internet and to other computer hardware, and being used for data collection. Generally, data is entered onto the PDA using a touch-screen and a small stylus, although there are some, less common, variations on this. PDAs have sometimes been used in a healthcare context, most commonly by doctors – for example to be able to wirelessly check main databases of prescriptions and drug schedules during ward rounds – but also sometimes by patients, whom, it has been suggested, may in some situations be better able to communicate physical symptoms with medical staff by recording them on a PDA (for example by using on-screen sliders to measure and report their levels of pain). Their use in healthcare surveys has so far been relatively limited but they are becoming more commonplace in the market research industry, where they are now frequently used for data collection in face-to-face (F2F), computer-assisted personal interviews (CAPI), and sometimes CASA (computer assisted self-administration) studies. Given the earlier assumption in this document, however, that the use of a field force of trained interviewers to conduct face-to-face studies in hospitals would be prohibitively expensive for the national patient survey programme, self-administration would be necessary. In practice, some caution would be required here, as scenarios could be envisaged where patients would be unable to complete such surveys by themselves. If, for example, a patient were to get stuck completing a survey, or if the PDA were to crash or its battery to become empty, it would seem clear that intervention by ward staff would be required – creating burden on staff and, potentially, undermining data quality.

This being the case, it's most likely that NHS staff would have to pass PDAs to patients during their hospital stay and invite them to take part in the survey. This in itself may pose significant problems (see section 6). Ignoring these issues for a moment though, the potential benefits of using PDAs stem from their similarity to web surveys. PDAs are very similar in function to modern computers and, as such, they are highly flexible in terms of the ways in which they can present information. Audio, visual, or video stimuli could be used, as could complex transparent routing, and there would be little limitation on questionnaire design – except that imposed by screen size and resolution – in terms of either item or response length, and number of response options. It would also be possible to use more complex response formats; for example, visual analogue scales, click maps, illustrated responses, and even, through optimal character recognition (OCR) software, to allow patients to write in responses on a touchscreen tablet with a stylus.

5.8 Handheld units

Handheld units are simply electronic data collection systems that are small, portable, and that can be given to individual patients to allow them to give responses that are then stored on the unit until being uploaded to a main computer or server at a later point. We are aware of at least two distinctly different kinds of such devices: one using visual presentation of questions, the other using audio presentation. Given the major functional differences in these systems, they are considered separately.

The most basic kind of handheld unit are devices such as the Patient Experience Tracker (PET) developed by Patient Focus and presently being marketed by Dr Foster Intelligence (see Figure 1, below). These present up to five questions, allowing a range of up to four response options for each. Respondents are simply required to read the questions and respond by pressing a button to the right of these. The advantages of this system are its relatively low cost and its ability to very quickly and easily record feedback from patients on a number of key issues. People generally have little difficulty using the systems, which are distinctly 'low-tech' and unthreatening in nature,

and take very little time to complete surveys on them. Obvious disadvantages of this system, however, are its inflexibility in terms of the number and types of questions and responses that can be offered; changing any of the questions or response options requires the printing of a new cover for the unit and then manually swapping it. Regardless of its clear usefulness in very narrowly focused and localised data collection efforts, it is difficult to see what benefits this kind of system can add to large surveys.

Figure 1: Illustrative image of a PET device



An alternative to this kind of device are electronic data collection units relying on audio presentation of questions to respondents via headphones. These have been developed to facilitate the collection of responses from people with poor or no written literacy. Responses can be collected in a range of different ways depending on the particular system – some use buttons or sliders mounted on the device, whilst others use a microphone to collect vocal responses that can either be manually transcribed or analysed with voice recognition software later. Such methods have been used in research on pain management with a sliding scale used to record responses⁴², and are presently being developed to allow for responses to be made to multiple choice questions (currently with four response options permitted)⁴³. These devices are small portable units that connect to computers via USB cable and that store audio files in MP3 format. Scripts in different languages can be selected at the start of presentation, facilitating their use in surveys with a diverse range of respondents. Furthermore, the units can be programmed to ask and record responses to as many questions as necessary. Changing the set of questions on the unit can be achieved simply and quickly via USB connection to a PC. Whilst these units currently offer no greater flexibility than the PET in terms of the number of permissible response options, they are much more adaptable in terms of being used for different surveys and for multi-language communities. Again, since the technology in use is relatively unsophisticated, costs for such units should be relatively low.

One obvious disadvantage, though, is that reliance on audio presentation does increase the potential for mode effects to occur when comparing results from these devices to those from paper surveys. Early research (eg Naik et al, 2005) has attempted to test this by asking chronic pain patients to complete the paper based Short Form Brief Pain Inventory (SF-BPI) and then subsequently to complete an audio presentation version of this (EPICVox) using the slider based handheld system. Mean levels of agreement between results from the SF-BPI and EPICVox were

⁴² Naik, M., Pallett, E., Thompson, J.P., Watson, P., Maslowski, P., Toogood, L., & Rowbotham, D. (2005) Comparison of an electronic speaking data recorder with the short form brief pain inventory in chronic pain patients. *British Journal of Anaesthesia* 95 (4): 566–79P [ABSTRACT]

⁴³ Patel, N., Pallett, E., Ali, S., & Stone, M.A. (2006) Development of a method of collecting questionnaire data from people with mixed language and literacy skills: a tool for use in diabetes research [ABSTRACT]

shown to be relatively high, but standard deviations of mean differences were also high and limits of agreement were seen to “vary considerably”. A further problem is that the sequential use of the paper followed by audio versions of the questionnaire suggests that high levels of agreement could be attributable to memory effects rather than genuine responses patterns to the two questionnaires, and Naik et al’s study provides insufficient control to rule this out.

5.9 Mobile phone messaging

In the past decade the prevalence of mobile phones in the UK has increased dramatically. At present, estimates from rigorously conducted research on the proportion of adults in the UK who own or use a mobile phone range from 78%⁴⁴ to 85%⁴⁵ (although some reports, based on unclear methodology, have suggested the figure is as high as 95%⁴⁶), with this figure rising further for younger people. It has recently been reported that 90% of 12 year olds now own a mobile phone⁴⁷, and research from 2003 indicated then that 88% of under 35s compared with 75% of adults as a whole owned mobile phones⁴⁸.

One possibility that has received surprisingly little coverage in the academic and market research press is that of using mobile phones’ Short Message Service (SMS) as part of surveys. SMS or ‘text messaging’ services are more popular⁴⁹ and generally cheaper than mobile phone calls, and are especially widely used by younger people. These messages are limited to including only raw text (eg, they cannot include graphics, links, or dynamic content) and have a maximum length of 160 characters, which severely restricts their potential for questionnaire administration. As such, the deployment of this technology as a mode of data collection could not be seriously recommended for the national survey programme. Nonetheless, SMS also has a range of possible benefits which may make it a useful means of contacting patients. Firstly, this manner of communication is usually virtually instant both in transmission and receipt, unlike, for example, e-mails, which are transmitted instantly but which are received only when users actively check their inbox. Secondly, SMS is relatively unobtrusive compared with telephone calls but messages sent in this way are rarely ignored and, because of their short length, generally read in full: they are less easy to ignore than mail or e-mail communications, and not as readily refused as phone calls. For instance, Townsend (2005) argues that “SMS can cut through the clutter of too many e-mails which are easily confused with spam [*unsolicited junk e-mail*], even though the respondent may not utilize their wireless device to complete the research itself” (p57)⁵⁰. The concise nature of SMS messages could help overcome the problem encountered in telephone surveys where a majority of refusals occur before the (non) respondent has heard the arguments for participation⁵¹.

⁴⁴ Ofcom. (2006) *Mobile Call Termination: Report of Market Research Findings*. Office of Communication; London, England. 13th September 2006.

⁴⁵ Ipsos MORI Technology Tracker, September 2006. Ipsos MORI; London.

⁴⁶ NetSize Guide 2005 report for 2004

⁴⁷ LSE & Carphone Warehouse

⁴⁸ Ofcom Residential Survey, May 2003.

⁴⁹ 3.6 texts per day compared to 2.8 calls average: Mobile Life Report – Carphone Warehouse/LSE

⁵⁰ Townsend, L. (2005) The status of wireless survey solutions: the emerging “power of the thumb”. *Journal of Interactive Advertising*, 6, 1, p.52-58

⁵¹ Groves, R.M. (1990) Theories and methods of telephone surveys. *Annual Review of Sociology*, 16, 1 pp. 221-240.

Furthermore, SMS is a relatively inexpensive medium compared with more traditional methods – sending a text message currently costs between 5p and 10p, much cheaper than even the cost of postage for a mailed reminder, and discounted rates are available for bulk mailings. Using SMS in surveys also featuring telephone contacts can also produce additional benefits, as sending these messages proves a cost-effective way of identifying whether or not numbers are in service⁵².

Whilst mobile phones are now used by a significant majority of adults in the UK, their overall penetration is still some way short of being total and it is unlikely that this will ever reach 100%, especially given recent evidence that the bulk of those people who did not own a mobile phone in 2006 had no intention of owning one in the future⁵³. Furthermore, the proportion of patients for whom acute trusts have a mobile phone contact number will vary greatly from organisation to organisation, and in many cases mobile phone records will not be explicitly coded as such (although given UK phone numbering conventions, this is unlikely to be a problem as the prefix code allows identification between mobile and land lines). Despite this, there remains good, although largely untested, reason to believe that the use of SMS/text message communications during the course of a survey may help increase response rates amongst some 'seldom heard' groups who are typically under-represented. Response rates for younger people are generally poorer than for other age groups. This has been attributed partly to factors such as their relatively high mobility, disengagement, and, in telephone surveys, unpredictable working hours. The difficulty of reaching younger people relative to the general population is exemplified in findings from the 2005 national survey of adult inpatients: a significantly higher proportion of questionnaires were sent to patients aged 35 years or under were returned undelivered compared to those sent to patients aged 36 and above [1.82% vs 0.76%; $t(138,921)=15.202$, $p<0.001$]⁵⁴. Similar problems, in addition to language barriers, are encountered with immigrant communities. In both of these groups, however, as well as in other 'seldom heard' populations such as the homeless, mobile phone ownership remains extremely high⁵⁵; this has been attributed to the increased importance of mobile communication for groups who are unlikely to be able to access fixed telephone lines. For each of these groups, the use of SMS may actually serve as a much more reliable means of contact than a mailed approach.

One study focusing on the use of SMS invitations in a web/phone survey sent text messages to sample members selected by random digit dialling, asking them to respond with basic demographic information and indicating a preferred mode of completing the survey⁵⁶. The study used a 2x2 factorial design manipulating salience of the survey (with a high saliency topic on 'asylum seekers' and a low saliency topic on 'grocery shopping') as well as the nature of the sponsoring organisation (a university or a market research organisation), with two SMS reminders being sent to non-responders. Overall the study achieved a response rate of 36.5%, with significantly higher response rates in the market research (48%) rather than university sponsored (25%) conditions. The extremely small sample size of 200 ($n=50$ for each condition) does,

⁵² Steeh, C., Buskirk, T.D., & Callegaro, M. (2007) Using text messages in U.S. mobile phone surveys. *Field Methods*, 19 (1), 59-75.

⁵³ Office of Communications. (2006) *Consumers and the communications market: 2006*. Ofcom: London. http://www.ofcomconsumerpanel.org.uk/publications/consumer_panel_report06.pdf

⁵⁴ Graham, C. (2006) Unpublished analysis from 2005 inpatient survey.

⁵⁵ eg Thompson, M., & Crush, D. (2005) *Inclusion through innovation: tackling social exclusion through new technologies*. Social Exclusion Unit, Office of the Deputy Prime Minister. <http://www.aes.asn.au/conferences/2004/FR21-Ballantyne,%20C.pdf> last accessed 27/03/07.

⁵⁶ Balabanis, G., Mitchell, V-W., Heinonen-Mavrovouniotis, S. (2005). SMS-based surveys: strategies to improve participation. *Proceedings of the European Marketing Association Conference (EMAC), Milan*.

however, draw into question the robustness of this finding, as does the lack of any discussion of why zero responses were received from the second contact in the university conditions: compared with a 14% increase in responses at this point in the market research conditions, it does not seem unreasonable to speculate that technical problems may have led to some or all of the messages not being delivered (the study does not appear to have included any testing of the actual administration). One finding from this work that is interesting, however, is the predomination of young males in the achieved sample – 48.5% of respondents were under 30, and 64% were male. Whilst Balabanis et al suggest that this evidence that “SMS-based surveys... might appeal to a certain type of respondent” may be a drawback to using such technology, it seems apparent that in the context of a mixed-mode methodology they could potentially provide an innovative way of increasing responses from some traditionally under-represented groups.

Currently we are only aware of one experimental test of the use of SMS reminders in postal surveys. This was conducted in Finland, another country characterised by extremely high mobile phone penetration, and included three surveys with a total sample of over 16,000. Whilst this study has yet to be published⁵⁷, initial results are apparently “encouraging” and show significant increases in response rates when using SMS text reminders⁵⁸.

5.10 Conclusions

A wide range of alternative modes for communication and data collection that could potentially be used for surveys are now available. These typically rely on relatively new electronic technologies, but vary in a range of important ways that affect their viability for use in the patient survey programme. In this section we have looked at the key attributes of a selection of different modes, and this should clearly demonstrate the conceptual differences that exist in terms of the methods of presentation (eg, audio/visual) and responding (eg, by voice, button pressing, or other feedback devices such as sliders). The different characteristics of each of these modes have important implications, and this is further discussed in section 8.

⁵⁷ Forthcoming as Virtanen, V., Sirkiä, T., & Jokirinta, V. (2007) Reducing nonresponse by SMS reminders in mail surveys. *Social Science Computer Review*, 25 (3) – to be published Autumn 2007.

⁵⁸ Virtanen, V., Sirkiä, T., & Nurmela, J. (2005) Reducing nonresponse by SMS reminders in three sample surveys. *Paper presented at the conference of the European Association of Survey Research, 18-22 July 2005, Barcelona*. <http://sqp.nl/easr/programme/abstracts.pdf> (p71; last accessed 28/03/07).

6 Other considerations relating to alternative modes of data collection

Whilst the unique features of each specific mode of data collection force their own particular considerations upon survey design, there are some general considerations that must be borne in mind when considering how *any* alternative mode could be used in the national patient survey programme. These are related to the existing features of the surveys, and how future changes could be integrated to collect comparable data. Here we look at two particularly important issues: the timing of the survey in relationship to the care experience, and sampling and the linking of information about selected individuals to the final dataset.

6.1 Timing of the survey

Currently, patients included in the sample for any of the main acute surveys will normally receive a questionnaire between two weeks and three months after their discharge from hospital. Evidence suggests that the specific length of time within this period does not generally have a significant effect either on propensity to respond or on the overall positivity of responses⁵⁹. It seems likely, however, that the difference between a patient receiving an invitation two or 12 weeks after their discharge should be small compared to the difference between a patient receiving an invitation before or after their discharge from hospital. That is, we would anticipate that the differences in perception of care of a user recently discharged from a hospital would differ far more from the views of that user *whilst they were still in hospital* than from that same user some greater length of time post discharge.

A notable feature of using electronic kiosks, hand-held units, PDAs or bedside terminals to collect survey data is that patients would have to be asked to evaluate their care whilst still in hospital. This would preclude the asking of any questions about discharge procedures and would necessitate significant changes to the sampling procedure used for the surveys. Currently, most acute patient surveys are based on flow samples of patients leaving the care system – eg, being discharged – and actively exclude anyone known to be a current inpatient in a hospital. Thus, collecting the views of patients currently in hospital represents a fundamental shift in approach, and would present a challenge in terms of the overall co-ordination of the survey. In order to use a single sample for the survey, trusts would need to have a good idea in advance of what time period of discharges or incidents they would be sampling, and would need to be able to predict, whilst people were in care, whether or not they would be included in this sample. They would further need to be able to track whether or not patients had responded whilst in care in order to exclude these people from a mailed survey post-discharge, and this raises a number of issues regarding administrative feasibility and confidentiality. Alternatively, a dual-frame sample could be used, with a sample of discharged patients being drawn from a different time period to the fieldwork period for an ‘in-hospital’ or ‘point of care’ arm of the survey. This would reduce the administrative difficulties in conducting a survey with people at significantly different points in the patient journey, but establishing a robust sampling procedure for patients in hospital would most likely be extremely

⁵⁹ Graham, C. (2006). unpublished research based on results from the National Survey of Adult Inpatients 2005; Graham, C. (2007). unpublished research based on results from the National Survey of Adult Inpatients 2006

difficult (especially at trusts with multiple sites). A third alternative would be to dispense with any strict sampling for the collection of views of people currently in hospital – for example, a touch screen kiosk or similar could be placed on wards and all patients in hospital during the fieldwork period made able to access it if they wished – but this would limit the comparability of results from different sites and trusts as it would be impossible to ensure that demographically similar groups of patients were responding in different areas.

It may be that there are specific areas where surveying patients at the point of experience may be advantageous because asking them at this time removes the cognitive burden of recall and allows a more direct linkage between evaluation and report. This possibility, however, needs to be balanced against possible corresponding timing effects: it could be, for instance, that patients may be more reluctant to report shortcomings in their care whilst still in hospital, either because of fears that their future care may be contingent upon their responses or because they have yet to come to a clear conclusion about the quality of services. This may be especially true where clinical staff are involved in the administration process⁶⁰. Social desirability may also influence responses given in surveys where clinical staff are visibly involved in the administration process. For example, Dillman (2000) describes a hospital telephone survey producing overwhelmingly positive results as a result of nurses conducting the phone interviews: asked questions about their satisfaction with care by one of the very nurses who may have cared for them, patients were heavily inclined to offer the responses that they felt their interviewer wished to hear⁶¹. A further, more complex, possible source of timing effect may emerge if patients or subsets of patients show a reluctance to criticise care whilst in hospital, not because of fears about contingency of future care upon responses, but because of cognitive processes relating to peoples' actual experiences of illness and treatment. For instance, it may be that patients unconsciously seek to adopt a highly optimistic view of the care they are currently receiving because they wish for the care episode to pass as successfully and efficiently as possible and are therefore unwilling to acknowledge shortcomings in care in 'real-time'⁶². Clearly, thorough piloting would be necessary to resolve such questions about possible timing effects, and to develop statistical adjustments to correct for them if necessary.

A further concern regarding the use of most 'point of care' data collection modes is how patients' attention would be drawn to these. In a postal survey it is very easy to minimize or eliminate variations in the form and content of communication used to invite patients to participate in the survey. This is unlikely to be the case when surveying patients in hospital and it may be that this could be a source of instability in response rates and data quality. Whilst electronic kiosks and bedside terminals can be set to carry advertising type messages inviting patients to complete the survey, it is far from clear to what extent these would be noticed by patients or would encourage them to take part. Likewise, methods such as using posters situated on wards could be used, but it remains questionable whether these would have a sufficiently high impact to be worthwhile, and how a consistency of approach across trusts and sites could be centrally audited.

Further problems in approaching patients exist as regards the potential use of hand-held devices and PDAs, because by definition these have to be taken and given to the patient⁶³. Since we have

⁶⁰ eg Uttaro, T., Leahy, V., Gonzalez, A., & Henri, W.F. (2004) Effect of type of survey administrator on Consumer Assessment of Care. *Psychology Report*, 94 (3 Pt 2), p1279-82.

⁶¹ Dillman, D.A., (2000) *op cit.* p.226

⁶² Gribble, R.K., & Haupt, C. (2005) Quantitative and qualitative differences between handout and mailed patient satisfaction surveys. *Medical Care*, 43(3), 276-281.

⁶³ Although it might be worth noting that apparently 30% of PETs are in fixed locations...

already ruled out the possibility of using a specifically trained field force for administering national surveys, such responsibilities would then have to fall to ordinary NHS staff – most likely nurses. Aside from issues around asking already stretched clinical staff to devote time to surveys in this way, this could again limit data quality. One provider of hand-held units notes that “ward nursing staff must take the device to patients at their beds, before they are discharged, and prompt them to give feedback... [but] threatened staff avoid negative feedback from patients by not asking them to answer the questions on the devices.”⁶⁴ This should obviously raise serious concerns regarding the possibility of gaming of survey results. A further potential problem is whether using such devices, which would have to be used by many different patients and passed back and forth by staff, might have an impact on infection control. This may raise real concerns in some hospitals and wards, and one cannot rule out the possibility that some staff, highly conscious of the risk of spreading infections such as MRSA, would be reluctant to participate in administering such surveys.

6.2 Linking responses to the survey sample

Several of the data collection modes discussed above, particularly those where the patient could be asked to respond to the survey some time after their discharge from hospital, are well suited to integration with the main survey sample as, for reasons of coverage and achieving high response rates, it would be most likely that a postal response option would also be offered alongside these. In these cases, it would be necessary for responses to be able to be linked to sample information, so as to ensure that respondents using these methods were not then sent further reminders, to prevent people from completing the survey multiple times, and for the purposes of statistical analysis of responses and response rates.

In order to achieve this, respondents logging on to an ‘online’ system - be it internet, IVR, or TDE - would need to be able to identify themselves with respect to the sample. The logical way of doing this would be for participants accessing the system to enter a unique code: for example, the patient’s unique reference number (URN) as given on the questionnaire. This would require some changes to the sampling process as a standardised system of numbering would need to be used to ensure that every sample member had a genuinely unique reference, but, theoretically at least, this task should be trivial (for example, URNs could be structured by trust code and sequence number). This has in fact been done in past surveys, such as 2006 national survey of people with diabetes. To increase the accuracy of such a system, it would also be desirable to assign sample members with a separate password or personal identification number (PIN), or ask them to confirm a detail of their inpatient stay, such as their date of admission or discharge. This may be more problematic however, as using arbitrary codes or passwords would be difficult to co-ordinate in a devolved survey, and asking for personal information would be likely to raise questions regarding confidentiality. Furthermore, depending on the choices of modes on offer the nature of the codes to be used would have to be balanced with the means by which responses were to be provided. For example, alphanumeric codes could be used without difficulty on the internet, but may confuse some respondents using a TDE system, as, despite the growth of text messaging, using a phone keypad to enter letters is not necessarily a simple task.

Whatever approach may be adopted for linking of sample and response data, its accuracy would need to be assured as the success of this process has important implications for the administration and data quality of the survey as a whole. Because the tracking of responders (or, more accurately, non-responders) is so integral to a survey process designed to ensure good response

⁶⁴ Patient Focus, http://www.patientfocus.net/case_study_erewash.htm, last accessed 14/09/06

rates, errors in this linking would lead to numerous problems, including decreased validity of response rate analyses, reduced data quality in demographic analysis of responses, and increased costs arising from resultant errors in sending reminders. Another serious problem that would arise from a breakdown in linking of sample and response information would be that reminder letters would be sent to the wrong people, such that some nonresponders would not receive additional contacts whilst some responders would receive some in error. This could reduce response rates overall in the first instance and also cause a great deal of annoyance to people who had already responded in the latter. As such, careful consideration and testing of such a system would be critical to ensuring the smooth running of the survey as a whole.

6.3 Conclusions

Timing of data collection and the ability to connect responses to members of the survey sample are critically important issues for the national patient survey programme and transcend many of the mode specific issues described in section 5. Should any alteration be made to the survey methodology, it would be extremely important to give consideration to its impact on these two factors. At present there remain a number of questions regarding the administrative or analytic processes that might be required to deal with any changes that had bearing on these two factors, and thorough testing would therefore be necessary to ensure such changes could be implemented without damaging overall data quality.

7 Response rates

In this section we look at research evidence regarding the impact of mixed mode strategies on survey response rates. Potential effects on response rates, both overall and within different sub-groups, are highly important when considering the possibility of moving to a mixed-mode approach. Presently, NHS acute surveys tend to achieve reasonably high response rates: for example, the 2005 National Survey of Adult Inpatients had a final response rate of 59%, with response rates of 44% and 59% respectively for the 2005 Emergency Departments and Outpatients surveys. Variations in response rates for different sub-groups have been noted, however, with younger people, men, those living in relatively deprived areas and inner cities, and people from black and minority ethnic groups less likely to respond⁶⁵. Thus, whilst any increase in response rates is obviously a desirable target, methodological modifications that could result in improved response rates from these 'low response' groups may be of special interest.

7.1 Research background

A number of studies have looked at the effects on response rates of offering multiple modes of response to participants. Most often, these studies feature 'crossover' designs for the application of mixed modes; for example, having an initial data collection phase in a single mode and then switching to or offering an alternative mode for non-respondents. Far fewer studies have looked specifically at the effects of offering a range of different modes of response from the start of the survey. Given the significant methodological differences between these two kinds of approaches, it is best to consider these two types of designs separately.

7.1.1 Offering a choice of modes

The US National Census Bureau has conducted a number of studies looking at the possibility of using mixed modes of data collection to improve response rates and cut costs in their decennial census. Like the NHS Patient Survey Programme, only address information is available for sample members, so both surveys are similar in that they are constrained to only contacting participants by mail. The Census Bureau conducted the first experiment to test whether mixing modes can lead to "completion rates [being] improved when only addresses and not names and/or telephone numbers are available (ie when the respondent is invited to call in the response)",⁶⁶ p.568. The issue of whether completion rates could be improved by giving respondents the option of completing the (normally mailed) questionnaire over the phone was investigated using an experimental design involving five treatment groups of 4,500 households each, stratified for low and high response areas such that 2,250 households within each treatment group were within a

⁶⁵ Eg Raleigh, V.S., et al (2005) Variations in the experiences of patients in England: Analysis of the Healthcare Commission's 2003/2004 national survey of patients. The Healthcare Commission: London.

⁶⁶ Dillman, D.A., West, K.K., & Clark, J.R. (1994) Influence of an invitation to answer by telephone on response to census questionnaires. *Public Opinion Quarterly*, 58:557-568

low and 2,250 within a high response rate area⁶⁷. The five treatment groups varied in the number and type of contacts issued and in the availability of a telephone response option: timings between phases were constant. Figure 2 shows the details of each group (group 1 is a control group, offering no option of telephone responding):

Figure 2: Treatment groups in Dillman, West and Clark (1994)

Mailing Piece	Panel				
	1	2	3	4	5
Prenotice letter (3/26/93)	M	M	M	M	M + T
Questionnaire (3/29/93)	M	M	M	M	M + T
Reminder card (4/2/93)	M	M + T	M + T	M + T	M + T
Follow-up letter (4/6/93)	M + T	M + T	...
Replacement questionnaire (4/23/93)	M	M + T	M + T
Closeout (5/21/93)

NOTE. - M denotes mail piece and M + T denotes mail piece and telephone invitation.

Response rates to the survey by both mail and telephone within each group and stratum were calculated for each phase of the survey as well as for the whole survey duration. For the key comparison of group 1 and group 5 (which offered telephone responding at every stage), response rates within both strata of group 5 were slightly, although not significantly, *lower* than for the control group, despite 5.6% of households in group 5 choosing to complete the survey over the phone. Likewise, comparisons of responses obtained in treatment groups 1 and 2 before the sending of a replacement questionnaire showed no significant differences in response rates due to offering a choice to respond by telephone. Together, these results suggest that offering the choice to respond by telephone does not increase overall response rates but does lead to some degree of *substitution*, wherein people who would have completed the paper questionnaire anyway instead utilise the phone service. Significant increases in response rates were achieved, however, by using an additional follow-up letter with details of how to respond by phone but with no replacement questionnaire (group 2 compared with groups 3 and 4; the follow-up letter added 2.8% to the response rate). Presumably these findings could be generalised to an IVR system where respondents were simply told to call a given number without any indication that the process would be automated, but it is unclear whether the same results would be produced with other alternative data collection methods being used.

In terms of uptake, Dillman et al found that there was “a direct correlation between the number of telephone invitations [to individuals] and the telephone response rate. One invitation resulted in 1.7 percent of the sample responding by telephone, while two, three, and four contacts resulted in 4.1, 5.2, and 5.6 percent, respectively, responding by telephone” (p566-7). Since there was no treatment group in which sample members were offered a choice of telephone at every point in the survey *from, but not before*, the point of receiving the paper questionnaire, it’s difficult to estimate uptake levels for telephone responding under conditions where both modes are offered simultaneously. Furthermore, the number of responses achieved *before* sending the questionnaire in group 5 is not presented, which means that no conclusions can be drawn about the possibility of

⁶⁷ In practice this stratification was achieved simply by looking at the proportions of people from black and minority ethnic groups in different areas – because of the high correlation between proportions of black and minority ethnic groups within areas and area response rates, areas with high black and minority ethnic proportions were selected as low response areas. This led to the low response stratum having 64% black and minority ethnic population, and the high response stratum, whose response rates were ~10% higher in the 1990 census, having 15% black and minority ethnic populations.

lowering the costs by using an advance notification of a telephone response option to reduce the size of the first questionnaire mailing. Overall, though, the relatively modest 5.6% telephone response rate for group 5 over the course of the survey does indicate that when presented with the choice of a paper questionnaire or telephone completion, the vast majority of survey recipients will opt for the former.

The US Census Bureau looked again at the potential of mixed mode surveys in 2003, with technological developments in the intervening ten years seeing their alternative modes switch from telephone interview to IVR and internet⁶⁸. Both ‘choice’ and ‘push’ conditions were used: in the former, individuals were sent paper questionnaires along with instructions on how to access web/IVR⁶⁹ versions of the survey, whilst in ‘push’ conditions paper questionnaires were withheld until a final ‘crossover’ stage. Letters sent to the ‘push’ groups “used motivational language about the IVR and/or Internet systems, explaining the benefits of the electronic systems such as ease, convenience, and cost savings to taxpayers” – by contrast, no such motivational language was used in letters sent to the ‘choice’ groups. A summary of the groups of interest are shown in Table 1 (below).

Table 1: Treatment groups in Treat et al (2004)

Event	Date	C	M1	M2	M3	M4	M5	(CS2)
		Control	Push IVR	IVR choice	Internet choice	Push IVR & internet	Internet & IVR choice	IVR reminder
Pre-notice	21-22/01	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Initial Q'nnaire	30-31/01	Q'nnaire only	IVR only	IVR & q'nnaire	Internet & q'nnaire	Internet & IVR	Net, IVR, & q'nnaire	Q'nnaire only
Reminder postcard	7-8/01	Yes	Yes (inc IVR)	Yes (inc IVR)	Yes (inc internet)	Yes (inc IVR & net)	Yes (inc IVR & net)	No - IVR phonecall
2 nd Q'nnaire	18-19/01	Q'nnaire only	Q'nnaire & IVR	IVR & q'nnaire	Internet & q'nnaire	Internet, IVR, q'nnaire	Net, IVR, & q'nnaire	Q'nnaire only
Close	31/03							

As in the 1993 survey, each group was stratified by low or high response areas, although here the stratification was based on response rates from the 2000 national census rather than black and minority ethnic populations. Each treatment group contained a total of 10,000 households, with the exception of groups C (control) and CS2 (IVR reminder) which contained 20,000 households in

⁶⁸ Treat, J.B., Brady, S., Bouffard, A., & Stapleton, C. (2003) *2003 National Census Test: The Impact of Alternative Modes and Contact Strategies on Self-Response*. Unpublished report. Washington, DC: U.S. Census Bureau.

OR Treat, J.B. (2005) Research into the Use of Electronic Data Collection Modes for the 2010 United States Decennial Census. *Statistika*, 6/2005, p.449-478.

⁶⁹ It should be noted that instructions for accessing the IVR system *did not* state the it was an automated system – although Treat et al do not give details of precisely how the IVR response option was advertised, it seems reasonable to assume that it was described simply as a telephone responding system.

order to provide sufficient data for a separate item level analysis. Once the survey was completed, response rates for each of the different treatment groups were analysed to compare the effects of offering different modes for response: the original summary of results is reproduced below (Table 2).

Table 2: Co-operation Rates at the National Level and by Response Strata (reproduced from Treat et al, 2004)

Treatment Group	National		High Response Stratum		Low Response Stratum	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Control	67.1%	0.39%	72.0%	0.49%	53.5%	0.53%
Push IVR (M1)	62.2%	0.67%	67.9%	0.71%	46.4%	0.74%
IVR Choice (M2)	65.9%	0.58%	70.4%	0.63%	53.3%	0.69%
Internet Choice (M3)	67.3%	0.62%	72.9%	0.62%	51.6%	0.76%
Push IVR & Internet (M4)	61.4%	0.70%	67.1%	0.75%	45.4%	0.69%
IVR & Internet Choice (M5)	66.4%	0.66%	71.7%	0.71%	51.8%	0.77%
Telephone (IVR) Reminder Group (CS2)	73.0%	0.59%	76.6%	0.68%	60.3%	0.98%

Differences in rates between groups were investigated, and several findings were apparent from this. Firstly, significantly lower overall response rates were found in both of the ‘push’ treatment groups (M1 and M4; respectively 4.9 and 5.7 percentage points lower than the control group): these differences remained significant within both strata. By contrast, no significant differences were found, either within strata or overall, between the control group and any of the ‘choice’ groups (M2, M3, and M5). These findings suggest that offering a choice of modes does not lead to an increase in response rates, and that offering alternative modes *instead of* a paper questionnaire (the “push” groups) significantly reduces response rates, even when a crossover to mail is employed for the final phase of the survey.

In terms of uptake, more than two thirds of respondents used the IVR/internet systems in the two push conditions – for M1 (push IVR) >67% of respondents used the IVR system, with >38% and >30% of respondents respectively opting to respond by IVR and internet in M4 (push internet and IVR) – but the fact that such a relatively high proportion of achieved responses to these conditions were from mailed questionnaires firmly demonstrates the necessity of a mail crossover in sample surveys initially relying on self-administered electronic methods of data collection. In the choice conditions, the vast majority of respondents – between 88.4% and 93.2% of respondents within each group – chose to use the paper questionnaire. The largest uptake of alternative responses options was, as one might expect, in M5 (internet/IVR choice), where 11.6% of respondents used one of the electronic data collection methods. This still represents a relatively small share of responses, however, and this finding has implications for the potential of mixed-mode approaches to reduce survey costs⁷⁰. Like Dillman et al in 1994, though, Treat et al conclude that “offering a choice of alternative modes of response did NOT increase or decrease the [response] rates. Instead it simply shifted response to the alternative modes. However, this shift was relatively small” (p468-9).

It is also worth noting that response rates were significantly higher in the telephone (IVR) reminder condition. Overall, the response rate here was 5.9% points higher than in the control condition,

⁷⁰ Whilst unit costs are generally much lower for IVR and internet responses, set up costs may be considerably higher – more on this in section 11 (costs)

with larger (although not significantly larger) effects in the low response stratum (6.9% vs 4.6% points increase compared with control). A major limitation here, however, is that response rates for group CS2 were based only on households for which a telephone number was available – only 31% of the sample of 20,000. This confounds methodological and sample level variables for comparisons involving group CS2, since availability of telephone numbers may be linked to likelihood of responding (Treat et al suggest that those with unlisted telephone numbers may, generally speaking, be less inclined to act cooperatively). Consequently, findings from this section of the experiment should be treated with caution: although nonetheless it again raises the possibility that switching modes for re-approaching non-responders may be a particularly effective way of boosting response rates in low response areas and amongst hard-to-reach groups.

Whilst Treat et al (2004) did not include a ‘push internet’ condition in their survey design, we are aware of three studies which have looked experimentally at the feasibility of selecting probability samples of people without known e-mail addresses and using postal methods to invite them to complete a web survey. An early attempt at this was based on a large sample of military personnel (n=36,293) and compared the effects of offering web and mail responding either simultaneously (to 60% of the total sample) or sequentially (20%) against a control condition where only paper surveys were issued (20%)⁷¹. A notable feature of this survey was that it used a relatively long questionnaire – including 118 items and, in the print version, running to twenty pages: this is considerably longer than most questionnaires used in mixed-modes research and comparable to the questionnaires used in the acute patient survey programme. In all conditions Quigley et al utilised five mailings – these are summarised in Table 3 below:

Table 3: Mailing conditions in Quigley et al (2000)

Mailing	Condition		
	Mailed only (control)	Simultaneous mail & web	Sequential mail then web
1	Pre-notification letter	As for mailed only, but each of the 5 letters including instructions on how complete the survey over the internet.	Invitation to web survey
2	Questionnaire pack		Invitation to web survey
3	Reminder		Invitation to web survey
4	Second questionnaire		Invitation to web survey
5	Third questionnaire		Invitation + paper q'nnaire

Overall, the response rate in the control group who received only postal questionnaires was 40%. Similar response rates were recorded in both experimental conditions, with 42% responding in the simultaneous and 37% responding in the sequential mail and web condition. There were notable differences in the proportions of responses via each mode, and these are summarised in the table below:

⁷¹ Quigley, B., Riemer, R.A., Cruzen, D.E., & Rosen, S. (2000) Internet versus paper survey administration: preliminary findings on response rates. *Proceedings of the 42nd Annual Conference of the International Military Testing Association, Edinburgh, Scotland*, p125-128. Available at <http://www.internationalmta.org/Documents/2000/Proceedings2000.pdf> (last accessed 16/02/07).

Table 4: Response rates in Quigley et al (2000)

	Condition		
	Mailed only (control)	Simultaneous mail & web	Sequential mail then web
Overall response rate	40%	42%	37%
% of all responses by mail	100%	75%	27%
% of all responses by internet	N/A	25%	73%

Based on this, the authors of the above study suggest that offering a choice of completing mailed surveys over the internet may “slightly increase response rates” compared with a pure mailed condition. This finding, however, must be treated with caution: Quigley et al describe their results as “very preliminary” (p128) and do not present significance tests or detailed analysis of demographics, actual responses, or cost-benefits of different methods, so their outcomes and conclusions cannot be considered rigorous. Whilst it is not discussed in their report, the data recording proportions of responses by each mode does, though, reinforce the other findings reported here – where a choice of an alternative mode of completion is offered in a preliminary mailed survey, the majority of responders will still respond by post and, furthermore, the majority of people completing the survey by that alternative mode are likely to represent substitution from the primary mode, rather than additional responses.

Schonlau, Asch, & Du (2003) conducted a similar study following up the third condition in Quigley et al (2000)⁷². They approached sample members (n=1,750 college-bound high school students) by mail in two stages: first, with an invitation to complete an internet survey and in the second stage employing a crossover to a postal questionnaire. Two further methodological variations were also tested amongst randomly selected subsets of non-responders – a phone reminder in stage 1, and incentives (\$3 gift certificates for McDonalds) in stage 2. Despite sampling what was assumed to be a highly “internet savvy” audience and withholding the mail response option until later in the survey, responses collected on the internet still represented only 35.2% of those achieved overall. Use of a phone reminder in stage 1 did significantly increase response rates – from 17.8% amongst those who received no phone call to 30.3% amongst those who did – but the use of an incentive at stage 2 had no effect on web response rate, despite significantly increasing the likelihood of recipients returning the paper version of the questionnaire.

A third paper reporting the effects of a mixed-mode survey design intended to ‘push’ a web responding option found little benefits to this kind of strategy. From a random sample of 3,000 members of the general public in New Zealand, Brennan (2005) randomly assigned 2,000 to the experimental condition of a mail/web mixed mode survey on new product adoption, with the remaining 1,000 being assigned to a mail only control of the same survey⁷³. This study is of particular interest for two reasons. Firstly, sample members in the experimental condition were never automatically sent a paper copy of the questionnaire – rather they had to return a postcard, enclosed with each of the mailings, to request one if they so wished. This meant that there was no overt mode crossover: respondents had to undertake two steps to express a preference for and

⁷² Schonlau, M., Asch, B.J., & Du, C. (2003) Web surveys as part of a mixed mode strategy for populations that cannot be contacted by e-mail. *Social Science Computer Review*, 21(2), 218-222.

⁷³ Brennan, M. (2005) The effect of a simultaneous mixed-mode (mail and web) survey on respondent characteristics and survey responses. *Paper presented at Australia and New Zealand Marketing Academy Conference 2005, Perth, Australia.*

respond by mail, and web results can be considered 'purer' for this. Secondly, the survey took place in New Zealand, which has the highest level of internet penetration of any nation in the world at 79.3%⁷⁴. Given the unusually high level of coverage, then, and the relative immediacy of internet responding compared with responding by mail, this could be considered an ideal situation in which to test the feasibility of using web surveys based on probability samples contacted by post.

Brennan found low response rates for the mixed-mode study when compared with the mailed control (25.4% vs 40.0%), and of all those in the experimental group only 16.6% (65.6% of responders) responded by web. The use of a postcard to allow sample members to request a questionnaire was also noted to be ineffectual – only 12.0% of the sample requested a questionnaire, and of these 27.2% still did not return it completed. Furthermore, it was noted that responses from the experimental group were not actually received any sooner after the initial invitation was sent than were those from the mailed control survey. Strong differences were observed in responder characteristics between modes, with web responders generally younger, wealthier, more likely to be male, and showing greater adoption of new technologies. Polarisation of the mail and web responders to the mixed-mode survey compared to all responders in the control group was also apparent: for example, the mean age in the control group was 49.8 years, but significantly higher in the mail section of the mixed-mode response (55.4 years) and significantly lower in the web responses (44.6 years).

Overall, then, this paints a rather discouraging picture of the current potential for using internet surveys as part of mixed-mode strategies where sample members are to be contacted by mail. As Brennan reports from his study, "the purported benefits of using the web simply did not materialise... Unless there is compelling reason for using the web mode, it may be better simply to conduct the whole survey by mail." (p10-11). Some caution is needed in interpreting these findings, however, since only a smattering of research has so far been conducted on the use of this kind of approach and this has been spread thinly across a divergent range of sectors. Since factors such as survey salience and sponsorship are known to have major bearing on sample members' decisions to participate and, therefore, on overall achieved response rates^{eg 75, 76}, it must be recognised that the potential exists for interactive effects to emerge between modes and across different studies when such factors are not controlled for. As such, it would be unreasonable to presuppose that previous results from surveys of the military, high-school leavers, or the general public would transpose direct onto samples of recent patients of acute hospitals. Hence it is difficult to make firm statements about the probable effects of simultaneous mixed-mode designs on survey response rates, yet, nonetheless, the extant research evidence is such that some prudence is required when considering the potential benefits of such an approach.

7.1.2 Mode crossover

A number of studies have looked at the use of mode crossover in data collection, where non-respondents to an initial mode are subsequently invited to complete it in an alternative way. Often the intention of such approaches has been to use a relatively cheap mode to administer the initial wave of the study, then to use a more expensive mode to 'track down' non-responders and

⁷⁴ Source: ITU cf. www.nationmaster.com, 2004; last accessed 16/02/07.

⁷⁵ Groves, R.M., & Couper, M.P. (1998) *Nonresponse in household interview surveys*. New York: Wiley-Interscience.

⁷⁶ Groves, R.M., Singer, E., & Corning, A. (2000) Leverage-salience theory of survey participation: description and an illustration. *Public Opinion Quarterly*, 64(3), p298-308.

decrease non-response bias to the study as whole; thus achieving the common aim of mixed mode research, balancing affordable and optimal designs.

A good example of an attempt being made to use mode crossover to balance costs against low levels of non-response bias is a health survey of over 10,000 women aged 45-55 years that used a crossover from mail to telephone to improve response rates⁷⁷. The methodology for this experiment was based on Dillman's (1978) Total Design Method⁷⁸ and the earlier stages of it are therefore similar to the approach employed in the NHS national patient survey programme: an initial questionnaire is sent to all sample members, followed by a reminder a week later, then a second questionnaire after another week. Women who were yet to respond after three weeks were then telephoned *if* their phone number was available⁷⁹, and an attempt to complete the questionnaire over the phone was made. Alternatively, a 4th mailing was sent out if no phone number was present. The use of the 4th telephone/mail stage was shown to increase response rates by around 30% overall, with about seven in eight of those responding after this point doing so by telephone. Thus, this study showed clear benefits to using a crossover to telephone in terms of response rate, although it should be noted that the already considerable increase in costs of using a telephone contact must have been aggravated greatly by the laborious – and these days impractical – process used to identify people's phone numbers.

7.2 Conclusions

Published trials of the effect of mixed mode approaches robustly show that simply offering participants multiple options as to how to complete a survey is not, in itself, sufficient to yield improvements in response rates. Where a simple choice is offered, overall response rates tend to be similar to equivalent unimodal surveys except that some responses are “substituted” from one mode to another. Where people are sent paper questionnaires and also given the option to respond via another means, it is also likely that the vast majority of respondents will still respond by mail.

More encouraging results have been obtained when testing the sequential use of mixed modes. Following up nonresponders via an alternative method – ‘crossover’ approaches – seems to be the most successful approach for improving response rates, and there is a need for caution in any attempt to offer mixed modes of data collection via a unimodal contact strategy.

⁷⁷ Brambilla, D.J., & McKinlay, S.M. (1987) *op cit*.

⁷⁸ Dillman, D.A. (1978) *Mail and telephone surveys: the total design method*. New York: Wiley-Interscience.

⁷⁹ Phone numbers were obtained simply by looking people up in the published telephone directory; thus, numbers were not available for women without a telephone or whose numbers were unlisted.

8 Mode effects

One of the key drawbacks to the use of mixed mode approaches is the potential for “mode effects” – systematic changes to participants’ responses – to occur between different modes of data collections. There is a wealth of available evidence to demonstrate that asking the same question, with the same available response options and to similar groups of respondents, will not necessarily yield the same responses across all modes when multiple forms of self-administered surveys are employed^{eg 80,81}. Furthermore, the differences across modes that are unrelated to any other factors may often be substantial and have significant impact upon results and their interpretation.

This is clearly of particular importance where results are to be used for comparisons between organisations or groups or for the purposes of assessment. In this section, then, we consider the likelihood of such effects occurring in the NHS patient survey programme by looking at evidence from similar published health surveys. Following this we discuss the cognitive basis for mode effects with a view to understanding how and why they occur. We then summarise the salient characteristics of each of the modes discussed earlier before reporting published strategies for minimising and accounting for mode effects.

8.1 Mode effects in health surveys

Mode effects have been demonstrated in health surveys across a range of sectors and methodologies. For example, a study of recent female hospital inpatients in Florida surveyed parallel samples either by mail or telephone and found that responses made over the telephone tended to be markedly more positive and utilised less of the full range of response options⁸². Similar differences have been reported between phone and mail responders in satisfaction-type surveys of hospital attendees⁸³ and also in the CAHPS survey of health plan members in America, which uses similar questionnaires to the NHS Acute Survey Programme⁸⁴ and likewise. Walker and Restuccia (1984) argued that mail surveys were likely to be less biased than phone surveys because they were more often answered by the patients themselves – proxy respondents were used in 27.3% of mailed responses compared with 35.2% of telephone interviews – and significant differences were found between the responses offered by patients and their friends or relatives. In primary care settings, significant differences have also been observed between surveys mailed to patients compared with those handed out at the practice, the latter producing higher satisfaction scores and lower variation in ratings compared with mailed questionnaires⁸⁵.

⁸⁰ Dillman, D.A., (2000) *op cit.* Chapter 6.

⁸¹ Tourangeau, R., Rips, L.J., & Rasinski, K. (2000) *Op cit.* Chapter 10.

⁸² Hall, M.F. (1995) Patient satisfaction or acquiescence? Comparing mail and telephone survey results. *Journal of Healthcare Marketing*, 15, 54–61.

⁸³ Walker, A.H., & Restuccia, J.D. (1984) Obtaining information on patient satisfaction with care: mail versus telephone. *Health Services Research*, 19(3), 291-306.

⁸⁴ Fowler, F.J. Jr., Gallagher, P.M., Nederend, S. (1999). Comparing telephone and mail responses to the CAHPS survey instrument: Consumer Assessment of Health Plans Study. *Medical Care* 37(3):41-49.

⁸⁵ Gribble, R.K., & Haupt, C. (2005) Quantitative and qualitative differences between handout and mailed patient satisfaction surveys. *Medical Care*, 43(3), 276-281.

Presenting questions in a spoken format necessitates a different approach to the design of questions and response categories. Where closed questions are to be used, and particularly where respondents are unlikely to have a 'ready-made' answer to the question, it is generally essential that all permissible response options are explicitly read out to participants. Visual information cannot be used to help provide additional meaning to questions, so whilst factors such as equal spacing of response options in visual formats may imply linearly stepped intervals of values, respondents need to bring their own meaning to the terms listed in an IVR survey⁸⁸. Similarly, there is evidence that numbering response options, which is common in visually presented questionnaires but which seldom occurs in auditory presentation (with the exception of surveys utilising TDE for data collection) may affect the perceived meaning of scales⁸⁹. These factors may have implications for distributions of responses if participants define scales differently depending on the mode of presentation.

Response order effects may be particularly important when considering the mode effects that may arise between data collection modes relying on auditory and visual presentation of questions and responses. Such order effects are generally hypothesised as purely cognitive effects, based on psychological research on the nature of memory and attention, and there is evidence that the way in which these effects manifest may be a function of the specific ways in which information is presented⁹⁰. Traditional logic is that mail surveys have a tendency to exhibit primacy effects – where the first category offered is more likely to be chosen – and telephone surveys, conversely, are more likely to show recency effects – where the last category offered is more likely to be chosen. By contrast, it has been reasonably argued that there should be little or no reason to anticipate order effects occurring between different forms of visual self-administered questionnaires – such as paper questionnaires and web or e-mail surveys, for example – at least so long as the layout retains broadly equivalent: this is because the kinds of comprehension required are generally extremely similar⁹¹.

It has been argued that the primacy effect in mail surveys may exist due to greater and earlier attention being given to earlier response options: consequently they are selected immediately if acceptable or otherwise used as base-line comparators for later response options. The recency effect in telephone surveys may derive from respondents being given insufficient time to register and remember all categories, leading to the earlier ones being forgotten (particularly where the auditory space between question and first response option is small, such that the entire question and list of categories may be thought of as one block of speech or information)⁹².

A limited amount of evidence also exists to suggest that demographic factors may influence the extent of order effects in peoples' responses. Since such effects may be related to the need for

⁸⁸ A further interesting question here that thus far remains unaddressed relates to the *pacing* of the auditory relay of response options in IVR surveys and similar: if differences in the physical spacing on paper questionnaires may impact on interpretation of the value of response options, then it seems plausible that differences in the *temporal* spacing of spoken response options may have a similar effect in auditory presentation.

⁸⁹ Schwarz, N., Knauper, B., Hippler, H.J., Noelle-Neuman, E., & Clark, L. (1991) Rating scales: numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, 618-630.

⁹⁰ eg Miller, N., & Campbell, D. (1959) Recency and Primacy in Persuasion as Function of the Timing of Speeches and Measurements, *Journal of Abnormal and Social Psychology*, 54, 1-9.

⁹¹ eg Schaeffer, D.R., & Dillman, D.A. (1998) Development of a standard e-mail methodology: results of an experiment. *Public Opinion Quarterly*, 62(3), 378-397.

⁹² Eg Krosnick, J., & Alwin, D.F. (1987) An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201-219.

cognition required to provide an optimal response, it is perhaps not surprising that demographic factors related to cognitive ability have been found to be related to propensity for order effects, with such effects more likely amongst less educated respondents and those with more limited vocabularies⁹³. Likewise, recency effects in self-administered questionnaires are more likely to occur amongst those aged 65 or over⁹⁴, and this finding is likely related to normal declines in cognitive function associated with aging⁹⁵. More recent research, however, places severe doubt on conventional wisdom regarding the likelihood and direction of order effects, with evidence from a range of surveys showing that primacy and recency effects may occur both in telephone and in self-administered surveys in unpredictable and inconsistent ways. A review of 82 experiments found that only four of 33 mailed surveys, five of 26 telephone surveys, and three of 23 mixed-mode experimental comparisons produced significant primacy/recency effects in the expected manner⁹⁶. The implication of this is that any mixed-mode survey needs to be carefully piloted to determine whether and to what extent any such effects are an issue.

Question order effects may also vary between visual and auditory administration of surveys. Broadly speaking, these are effects that occur due to interactions between the responses given to separate questions on a survey, and a range of such effects have been described in existing literature. Some appear to be value based, activating social norms, whilst others may be rooted in cognitive processing and memory. By and large, it would appear that the potential for differences to emerge between modes of administration exists only in norm-based question order effects⁹⁷, and so others – such as ‘consistency and contrast’⁹⁸ and ‘additive and subtractive effects’⁹⁹ – are not discussed here. Rather, we focus on question order effects that are known to have the potential to vary across modes.

One well-known order effect is based on what has been described as the ‘norm of even-handedness’, where question order evokes a demonstration of fairness across responses to related questions. For example, a study conducted in America during World War Two asked respondents whether citizens of the United States should be allowed to enlist in allied (British or French) or enemy (German) armies. When the question about the German army was asked first, only 23% of respondents answered yes, with 43% answering ‘yes’ to the subsequent question on joining British or French armies. Reversing the question order, however, led to considerably more people answering yes to both the question on friendly (49%) and opposing (34%) armies¹⁰⁰. The differences in responses depending on the order of the questions can be explained thus: when the

⁹³ Krosnick, J., & Alwin, D.F. (1987) An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201-219.

⁹⁴ Knäuper, B., & Schwarz, N. (2004) Why your research may be out of order. *The Psychologist*, 17, 28-31.

⁹⁵ Park, D.C. (2000) The basic mechanisms accounting for age-related decline in cognitive function. In Park, D.C., & Schwarz, N. (eds.) *Cognitive aging: a primer*. Philadelphia: Psychology Press.

⁹⁶ Dillman, D.A., Brown, T.L., Carlson, J., Carpenter, E.H., Lorenz, F.O., Mason, R., Saltiel, J., & Sangster, R.L. (1995) Effects of category order on answers to mail and telephone surveys. *Rural Sociology*, 60, 674-687.

⁹⁷ Dillman, D.A. (2000) *op cit*.

⁹⁸ Schuman, H., & Presser, S. (1981) Questions and answers in attitude surveys: experiments on question form, wording, and context. New York: Academic Press.

⁹⁹ Moore, D.W. (2002) Measuring new types of question-order effects: additive and subtractive. *Public Opinion Quarterly*, 66(1), 80-91.

¹⁰⁰ Rugg, D., & Cantril, H. (1944) The wording of questions. In Cantril, H. (Ed). *Gauging public opinion*. Princeton: Princeton University Press.

questions are asked independently – as is the case when the first question is asked – respondents were more likely to answer affirmatively to the question about joining a friendly army. When, however, the question was asked second, respondents compared their response to that given to the first question. Since the question of allowing people to join armies, regardless of their alignment, is related, a social tendency towards fairness was activated causing people to be more likely to give similar responses to the second question. This effect has shown in other surveys, sometimes with enormous effects on substantive results¹⁰¹.

Some evidence does suggest that the effects of the ‘norm of even-handedness’ may vary across modes of survey administration as a result of the differences in the pacing and control of the survey. Whilst in interviewer administered surveys respondents are generally unable to go back and amend responses, or to read ahead and examine later questions before answering previous ones, this is not the case in pen-and-paper questionnaire studies. One comparative study of the norm of even-handedness found that whilst the effect did appear in telephone surveys it was eliminated in self-administered versions¹⁰². Other surveys have failed to reproduce this effect, however, and it has been argued that the context in which the respondent completes the questionnaire may also be a factor¹⁰³. As with response order effects, then, the exact nature of question-order effects in mixed-mode surveys is not entirely clear and some caution must be exercised around this issue when designing mixed-mode surveys.

8.2.2 Socially desirable responding

The phenomenon of socially desirable responding – where people offer responses that they believe are the most socially acceptable or that they feel their surveyor wants to hear – is one that is well known in the social sciences, although the range of its implications is often underappreciated. The effect of this kind of responding bias can vary across different modes of data collection, and can emerge in a range of different contexts.

The kinds of questions most obviously associated with socially desirable responses are ‘sensitive’ questions which ask respondents to divulge information of a personal or embarrassing nature, or that may be related to behaviours that transgress laws or social conventions. It has frequently been observed that sensitive questions are particularly susceptible to exhibit mode effects when switching between different modes, with those involving a weaker social presence – eg pen and paper self-administration – typically producing greater reporting of sensitive issues compared with those that involve an interviewer^{eg 104, 105, 106}.

¹⁰¹ Hyman, H.H., & Sheatsley, P.B. (1950) The current status of American public opinion. In Payne, J.C. (Ed.) *The teaching of contemporary affairs. Twenty-first yearbook of the national council of social studies*, p11-34.

¹⁰² Bishop, G., Hippler, H.J., Schwarz, N., & Strack, F. (1988). A comparison of response effects in self-administered and telephone surveys. In Groves, R.M., & Biemer, P.P., Lysberg, L.E., Massey, J.T., Nicholls, W.L. II, & Wakesberg, J (Eds). *Telephone survey methodology* p321-340. New York: Wiley.

¹⁰³ Sangster, R.L. (1993) *Question order effects: are they really less prevalent in mail surveys?* Unpublished doctoral dissertation, Department of Sociology, Washing State University, Pullma, WA. cf Dillman, D.A. (2000) *op cit* p 228

¹⁰⁴ Tourangeau, R., & Smith, T.W. (1996) Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60, 275-304.

¹⁰⁵ Kraus, L., & Augustin, R. (2001) Measuring alcohol consumption and alcohol-related problems: comparison of responses from self-administered questionnaires and telephone interviews. *Addiction*, 96(3): 459-71.

It is not just when asking sensitive question, however, that surveys need to take account of the possible effects of socially desirable responding. Rather, the tendency for our responses to all kinds of interactions to be influenced by social norms and conventions appears to be deeply ingrained in our social behaviour, and it has been noted that “social desirability operates at a threshold far below what one thinks of as anti-social behaviour”¹⁰⁷. A classic and highly relevant example of this occurs when asking people to report on their own health status: people consistently give better ratings of their own health when asked in a face-to-face interview compared with on the telephone, and here in turn better responses are given compared with in mailed surveys¹⁰⁸. This effect is highly robust and cannot be attributed to nonresponse error resulting from differences in mode preferences, as markedly more positive responses to health status questions have been observed in repeated measures design studies, wherein a cohort initially surveyed by mail is followed up via a personal interview a short time later¹⁰⁹. For the purposes of the NHS Acute Patient Survey Programme, the effect of socially desirable responding on questions about health status may be particularly important in terms of data analysis, because self-reported health status has been shown to be related to the overall positivity of responses¹¹⁰. It is not unfeasible that this may, therefore, eventually be incorporated into the weighting scheme for the survey results.

8.2.3 Acquiescence

‘Acquiescence bias’ in surveys is a predisposition towards offering affirmative responses, which Bowling (2005) describes as “a culturally based tendency to agree with others because it is perceived to be ‘easier’ to agree than disagree” (p.286)¹¹¹. Superficially it may seem reasonable to assume that such a responding characteristic may be more a trait of specific respondents than of specific instruments or methods, but research indicates otherwise: the extent to which individuals acquiesce in surveys has been shown to vary inconsistently across different types of questions and methods¹¹². Consequently, it is generally held that choice of mode and the way in which questions are asked is a more important consideration than the individual responding in terms of the likelihood of producing acquiescent responses¹¹³.

In keeping with the idea that acquiescence is based on the ‘ease’ of offering such responses, it has been argued that it results from applying simplified cognitive representations to give ‘satisfactory’

¹⁰⁶ Beebe, T.J., McRae, J.A. Jr., Harrison, P.A., Davern, M.E., & Quinlan, K.B. (2005) Mail surveys resulted in more reports of substance abuse than telephone surveys. *Journal of Clinical Epidemiology*, 58(4): 421-4.

¹⁰⁷ Dillman, D.A. (2000) *op cit*. p227

¹⁰⁸ Hochstim, J.R. (1967) A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62, 976-989.

¹⁰⁹ Biemer, P.P. (1997) Unpublished data. Research Triangle, NC: Research Triangle Institute. CF Dillman, D.A. (2000) *op cit* p217.

¹¹⁰ Raleigh, V. S., Scobie, S., Cook, A., Jones, S., Irons, R., & Hallt, K. (2004) *Unpacking the patients' perspective: Variations in NHS patient experience in England*. Healthcare Commission: London.

¹¹¹ Bowling, A. (2005) Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, 27(3), 281-291.

¹¹² Schuman, H. & Presser, S. (1981) *Questions and answers in attitude surveys*. New York: Academic Press

¹¹³ Groves, R.M. (1989) *Survey errors and survey costs*. New York: Wiley.

rather than 'optimal' answers¹¹⁴. As with order effects, then, the cognitive demands of different survey modes are related to the likelihood of acquiescence – and in particular, the cognitive demands of the mode have been related to the amount of time given to a respondent in which they may consider the question and response options. Self-administered survey modes where the respondent is free to consider and complete the questions in their own time should therefore exhibit less acquiescence than modes where participants have less perceived control over the pacing of the survey, such as face-to-face and, more so, telephone and IVR¹¹⁵. Such mode effects have been reported^{116 117 118}, but the existing evidence is inconsistent and other studies have reported no mode effects on acquiescence¹¹⁹. Part of the problem in firmly establishing the nature of acquiescence effects across modes is the difficulty of disentangling them from response order effects – Bowling, for example, argues that acquiescence in self-administered modes may often simply reflect the presence of primacy effects, since options indicating agreement tend to be presented first¹²⁰.

8.3 Conclusions

A range of existing research makes it abundantly clear that there is a potential for changes of modes of data collection to affect responses to surveys. These changes may be based on a number of different factors that vary between different modes, such as the mode of presentation, the mode of responding, and the overall degree of social presence perceived by respondents. Responses may differ because of issues relating to our social behaviour, such as acquiescence or social desirability, or to cognitive issues connected to the way respondents comprehend the questionnaire itself.

A significant concern related to the possibility of mode effects is that it is apparent from the research that the particular effects that might be observed in any given study or change of modes may not necessarily be predictable. Whilst some effects are typically related to certain modes or characteristics of modes – for example, acquiescence in the presence of an interviewer – others, such as order effects, may occur less consistently. This highlights the need for detailed pilot testing of any mixed mode survey to ensure that mode effects are properly understood and may be appropriately compensated for.

¹¹⁴ Krosnick, J., & Alwin, D.F. (1987) An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201-219.

¹¹⁵ de Leeuw, E.D. (1992) *Data quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT Publications. pp66-67.

¹¹⁶ Jordan, L.A., Marcus, A.C., & Reeder, L.G. (1980) Response styles in telephone and household interviewing: a field experiment. *Public Opinion Quarterly*, 44, 210-222.

¹¹⁷ De Leeuw, E.D., & van der Zouwen, J. (1988) Data quality in telephone and face-to-face surveys: a comparative meta-analysis. In: Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nichols II, W.L., & Waksberg, J. (eds) *Telephone survey methodology*. New York: Wiley.

¹¹⁸ Hall, M.F. (1995) Patient satisfaction or acquiescence? Comparing mail and telephone survey results. *Journal of Healthcare Marketing*, 15, 54–61.

¹¹⁹ De Leeuw, E.D. (1992) *op cit*. Note that De Leeuw does argue that “differences between methods do exist in self-selection of respondents, and therefore also in acquiescence” – p67.

¹²⁰ Bowling, A. (2005) *op cit*, p286

9 Strategies for minimising mode effects

Given the lack of clear, consistent evidence regarding mode effects, and given their complex, multi-faceted nature, it is not possible to put forward a single solution for eliminating them when mixing modes. Whilst much research has looked at the *kind* of mode effects that may occur, relatively little has looked at how to prevent or minimise them. To date, only one systematic approach has seen significant development – this is the ‘unimode construction’ approach proposed by Dillman¹²¹ – although more work has detailed the general design issues that may arise when constructing questionnaires for administration in different modes (much of this is described earlier in section 5).

Here, the unimode construction approach is discussed in detail as it is considered the best existing attempt to tackle the problem of mode effects. Following this, general design issues and considerations relating to different modes of administration are summarised.

9.1 Unimode construction

The aim of unimode construction is to maximise the similarities between the stimuli presented across different survey modes and seek to obtain an optimal questionnaire design *across* rather than *within* individual modes. This will generally involve some degree of compromise between choosing the optimal design for each mode and the design that can best translate across all modes.

Dillman sets out nine principles for unimode design:

1. **Response options should be constant in all modes, and if audio presentation is to be used, all options should be incorporated into the stem of the question.** Any option that can be used in any form of the survey must be explicitly apparent in all and must be initially offered in the same way – the most notable example of this being ‘no opinion’ or ‘don’t know’ responses that are often available, but not explicitly offered, in telephone surveys.
2. **Question structures should not change in ways that affect the stimulus.** For example, questionnaire designers must not switch from ‘tick all that apply’ structures to ‘yes/no’ lists when moving from self-administered questionnaires to telephone or IVR interviews or vice versa. In these instances, using the interview style structure (eg ‘yes/no’ lists) is generally preferable, both as it can be used in any mode and because respondents are less prone to satisficing (eg, select what they believe to be an appropriate number of responses to a ‘tick all’ style question and then simply move on to the next item without reading further response options)¹²².
3. **The number of response categories should be kept low so as to function well in all modes.** Whilst it is possible to use long scales in self-administered and face-to-face

¹²¹ Dillman, D.A. (2000) Mail and internet surveys: the tailored design method. New York: Wiley (Chapter 6)

¹²² Krosnick, J., Narayan, S., & Smith, W.R. (1996) Satisficing in surveys: Initial evidence. In Braverman, M.T., & Slater, J.K. (Eds.), *New directions for evaluation*, 70 (Advances in survey research), (pp. 29-44). San Francisco: Jossey-Bass. cf Dillman (2000), op cit.

surveys, these generally perform poorly in telephone, IVR, and TDE surveys. As such it is “desirable to find a common ground that will work for all methods” (p.235) by reducing the length of response categories.

4. **Use descriptive labels rather than visual scales.** Visual presentation of questions and response options provides respondents with additional information not present in surveys presented in audio. Consequently, the use of visual scales, especially where not all categories are labelled, must be avoided.
5. **Ranking questions should be preceded by rating questions on the same subject.** Questions where respondents are required to rank several different items are sometimes used in self-administered questionnaires and in face-to-face interviewing, often via card-sort exercises, but are generally avoided in telephone surveys because of the cognitive load they impose on respondents. Answering such questions over the phone is complex because it involves simultaneously memorising all available options and keeping track of these whilst evaluating and giving responses. Preceding such items with a rating question may make them more usable in all modes as the task of rating the responses encourages respondents to give individual cognitive focus to each response in advance, and therefore reduces the overall cognitive load of the ranking question. [NB: because of the high load and resultant poor data quality of ranking questions, we would generally advise against using these at all unless absolutely unavoidable.]
6. **Complex skip patterns should utilise equivalent instructions for respondents.** Whilst self-administered surveys generally seek to avoid complex routing instructions that are dependant on respondents’ answers to not one but several different questions, transparent routing in some survey modes enables these to be used much more easily. Where such routing is employed in self-administered postal surveys, designers generally attempt to simplify the task facing respondents by making reference to the nature of the questions that determine the skips (for example, “if you own your home (asked in Question 4) and are currently employed (asked in Question 30), then you should...”). This forces respondents to attend to certain issues raised at different points in the survey, but does not happen when transparent routing is available. As such, it is desirable to re-order questions to ensure that in all survey modes the relevant questions have been recently attended to. [NB: we would strongly recommend never using complex routing instructions in paper surveys because of the negative effect these have on data quality.]
7. **“Avoid question structures that unfold”** (p238). Telephone surveys sometimes break lengthy lists of response categories up into separate questions. For example, a question with responses of ‘agree’/‘disagree’ may be followed by a request for further detail: ‘do you strongly agree, somewhat agree, or slightly agree?’ Such question structures are inappropriate in self-administered surveys and prove onerous for respondents: to achieve similar stimuli in different modes they should be avoided completely.
8. **If order effects are anticipated, the order in which response options are presented should be reversed in half of the questionnaires to be used.** Whilst evidence regarding order effects in different modes is mixed, it is clear that category order effects can occur in any kind of survey. If there is a particular reason to believe that order effects may occur, or may occur between modes, then two versions of the questionnaire should be used – participants being randomly allocated to either one or the other. Because of the difficulty of handling this in postal surveys – especially devolved surveys such as the National Patient Survey Programme – and the resultant possibility of an impact on data quality, this should not, however, be done if there is no special reason to be wary of order effects.
9. **Interviewer instructions must be carefully considered and evaluated to ensure that these do not produce response effects, and, where feasible, explicitly present such**

instructions to all participants in all modes. Telephone interviewers are commonly given preset answers to expected common questions from participants, and may be told to encourage or thank respondents throughout the survey. This may also be an issue in IVR and TDE surveys where designers often include statements of encouragement throughout the questionnaire to deter break-offs. Such additional instructions and comments may affect responses both to temporally adjacent items and to the rest of the survey, and so their use must be carefully considered in advance. If interviewers are able to clarify questions for respondents who specifically ask for clarification, then it may be desirable to provide this extra information as standard to all respondents in all forms of the questionnaire.

Since the questionnaires used in the NHS Patient Survey Programme have generally been designed to be as simple, comprehensible, and easy to follow as possible, in a manner consistent with Dillman's (2000) 'Tailored Design Method'¹²³, few changes would be required in order to be compatible with the guidelines suggested above. There remain some instances, however, where items may prove problematic if used in modes other than self-administered postal surveys. Some questions, for example, include a relatively large number of different response options that may not function well in interviewer or IVR administered modes. Furthermore, although the surveys have traditionally avoided the use of visual scales, it does seem likely that many 'rating' type questions employed in the surveys (such as those that give 'Excellent/Very good/Good/Fair/Poor' type response categories) may be particularly susceptible to mode effects.

Whilst unimode construction goes some way towards providing a method of limiting mode effects at the design stage, it is not without its limitations. That the approach can help reduce problems relating to the structuring of the questionnaire is apparent, but this does not address issues arising from the different social contexts respondents are placed in when completing a survey by different modes. Effects related to social presence – such as social desirability and acquiescence – will not be reduced as a result of unimode design, nor will some effects attributable to the differences between audio and visual presentation of questionnaires be suppressed¹²⁴.

A further significant potential problem with the unimode construction approach is that by seeking to develop a questionnaire format that can be consistently applied across modes, one is, as Dillman implies, deliberately compromising the design within any particular mode. Whilst this may help reduce mode effects between different methods of administration, little or no research has properly investigated the effects this has on data quality within modes. The danger here, then, is that it may not always be possible for tried and tested items and instruments to be successfully ported between modes so that the survey can be considered both readily understandable and not excessively burdensome for respondents in all modes. Minimising the differences between modes may, in effect, restrict the potential of alternative data collection modes. Considering the possibility of using an IVR response option for the US Census, Conrad & Couper (2004) note that "if IVR is to be successful, its design must acknowledge and exploit the fact that speech is inherently different from visually presented information"¹²⁵. While the unimode construction approach goes some way towards this by attempting to push away from the use of question designs that are *reliant* on the use of visual information, it nevertheless is limited because it seeks to prohibit the exploitation of the specific features and benefits of different administration modes relative to each other.

¹²³ Dillman (2000) *op cit*. Note this differs from Dillman's (1978) 'Total Design Method'.

¹²⁴ Dillman, D.A., & Christian, L.M. (2005) Survey mode as a source of instability in responses across surveys. *Field Methods*, 17, 30-52.

¹²⁵ Conrad & Couper (2004), p18, *op cit*

9.2 Summary of questionnaire considerations relating to different modes of administration

Mode	Stimulus	Overall length	Item length	Response categories	Questions	Routing/ branching
Mail	Visual	Long	Any	Longer scales, fully labelled	Closed, 'tick all that apply'	Little (opaque)
Telephone	Auditory	Normally no longer than 20 mins	Short	Scales shortened, anchored endpoints	Closed, forced choice/open, further probing	Much (transparent)
IVR	Auditory	Extremely short	Short	Scale shortened further	Brief, closed	Much (transparent)
Internet	Multimedia (1 st ly visual)	Theoretically unlimited, but danger of terminations if long	Any	Longer scales, fully labelled. Opportunity for novel approaches – eg visual analogue scales/click-maps	Any – can include multimedia	Much (transparent/opaque)
PDAs	Multimedia (1 st ly visual)	Any, but cost of individual machines is such that it would be desirable to keep average completion time low so as not to limit total number of completions	Any	As with internet.	Any – can include multimedia	Much (transparent/opaque)
Hand-held units	Visual	Very short – five questions	Relatively short	Limited to four per question	Closed, forced choice	None
	Audio	Theoretically any	Any	Limited to four per question	Closed, forced choice	Unknown – could be set transparent
Kiosks	Multimedia (1 st ly visual)	Relatively short; danger of terminations if long	Any	Depends on input mode – touchscreen, keypad	Closed	Much (transparent)
Panel	Multimedia (1 st ly visual)	Any, but likely to need relatively high incentives if using longer questionnaires	Any	Longer scales, fully labelled. Opportunity for novel approaches – eg visual analogue scales/click-maps	Any – can include multimedia	Much (transparent/opaque)

9.3 Adjusting data to account for mode effects

Since it must be acknowledged that mode effects can most likely only be minimised and not eliminated, potential strategies for accounting for these at the analysis stage should also be considered. Again, little research has detailed the processes by which mode effects can be corrected for. A prime reason for this is that the specific processes which must be undertaken will depend heavily on the particular research design and methodology used, and thus generalising such corrective strategies is problematic.

We are aware of only one published research paper specifically focusing on adjusting data to account for changes in mode of administration and responding¹²⁶. Although this paper deals with moving from a mailed to a telephone methodology, it is nevertheless relevant here as it highlights an important challenge associated with changes to survey methodology: ensuring that data can be compared with that obtained via a different methodology in the past. In order to prepare for the migration of their survey from mail to phone, the authors undertook parallel sample surveys using both modes in order to compare response rates, demographics of responders, and obtained satisfaction ratings.

Finding significant differences in both demographics and satisfaction ratings, Burroughs et al initially weighted the mail and phone samples independently to adjust for demographic differences, making the two samples equivalent in terms of these demographic factors. Following this, significant differences in satisfaction ratings and response distributions remained between the two samples, with phone respondents giving more positive evaluations, being more likely to give 'top-box' responses, and using a smaller range of responses. In order to compensate for these differences, weighted linear regression models were created for the composite satisfaction variables used in the survey – similar to the 'domains' employed in NHS patient surveys – with mode of administration as the independent variable. Following this, the value of the regression coefficient was added to the composite variable for mail respondents on all variables exhibiting significant differences in the regression equation: effectively adjusting the expected values of the mail data to the expected values of the phone data.

The approach described above represents a good attempt to adjust data for changes in modes, but the necessary complexity of the method and the experimental research required to enable it should be noted: developing statistical adjustments for mode effects is a considerable undertaking. Even once this has been done, the stability of the adjustments need to be tested over time to ensure that they are behaving as intended – Burroughs et al recommend a “comprehensive review... with and without correction... after 12 months of data collection” (p361), however this may be difficult with surveys that do not operate on a rolling basis.

¹²⁶ Burroughs, T.E., Waterman, B.M., Cira, J.C., Desikan, R., & Claiborne Dunagan, W. (2001) Patient satisfaction measurement strategies: a comparison of phone and mail methods. *Journal on Quality Improvement*, 27(7), 349-361.

9.4 Conclusions

Mode of administration and response to surveys may form an unpredictable confounding factor. Minimizing mode effects is an important target when planning a mixed mode survey. Since it is generally the case that such effects cannot be eliminated altogether, correcting or adjusting for them is also vital. As such a consideration of mode effects must pervade every aspect of a mixed-mode survey, from planning and design through to administration and, ultimately, analysis. Evidence shows that the kinds of adjustments that must be made – both administrative and statistical – are not trivial: again, thorough piloting would be enquired to ensure that these are dealt with appropriately.

10 Previous use of alternative data collection modes in surveys of NHS patients

To date, the vast majority of survey work conducted with NHS patients has used either a postal methodology or relied upon the placement of 'drop-boxes' where patients can deposit completed surveys whilst within services (for example, collection boxes may be left on hospital wards, in doctor's practices, or in ambulances). Undoubtedly the relative simplicity of deploying surveys in these ways has been an influential factor here. With electronic equipment suitable for data collection purposes becoming both more affordable and more commonplace in healthcare settings, however, some attempts have been made to use alternative modes of data collection for NHS surveys, although these remain limited at the present time.

10.1 Bedside terminals

Despite the widespread availability of bedside terminals in modern NHS hospitals, usage of these for obtaining patient feedback is presently very limited. Having spoken to the biggest provider of such terminals, Patientline, we are aware of only two examples of surveys of patients' satisfaction with healthcare having been conducted on their systems (although further surveys have been undertaken looking at user satisfaction with the systems themselves, and we are aware of at least two further uses of Patientline to survey patients on their satisfaction with hospital food¹²⁷).

The first of these was the Healthcare Commission's report 'A Snapshot of Hospital Cleanliness in England', which made relatively limited use of the technology. The published report made little reference to the methodology used for this, but we have discussed this with the study's organiser at the Healthcare Commission, who saw the use of Patientline in the study as an exploratory way of attempting to gain data rather than as a strictly controlled scientific investigation. Only two questions were asked in the survey with no information about demographics of responders being collected.

We are aware of several NHS trusts that are now using Patientline for surveys of patient satisfaction with. As part of our research we have spoken both with staff at these trusts and staff from Patientline to discuss how surveys using bedside terminals have been implemented and how results have been used. Most trusts have sought to play to what they perceive to be the key strengths and advantages of the technology by using it not for 'one-off' surveys but for on-going monitoring of trends in patient satisfaction. Whilst this is commendable, we have serious reservations about the quality of the survey implementation here, and several shortcomings lead us to see this as a missed opportunity for the trusts. The most serious problem is that the systems we have seen are set up to automatically generate very simplistic reports rather than provide case-level data; eg, they provide frequencies at ward level for each question, but give no way of linking responses to different questions. Thus, whilst some of these surveys do ask questions on demographics, it is impossible to then use data from this to determine, for example, what proportion of negative responses to any other question on the survey were from men and what proportion were from women. This is not, we understand, an inherent limitation of the system

¹²⁷ The first of these was conducted in Spring 2003; the second in 2005.

though, and so there remains potential for more sophisticated and effective implementation of bedside terminals in surveying patients than has presently been attempted.

10.2 Internet surveys

Until very recently, no reported surveys specifically targeting NHS patients had been conducted principally or partially via the internet, as exemplified by a 2005 review that was unable to identify any such patient surveys¹²⁸. This reflects common concerns about the coverage achievable with internet surveys, but may not be entirely justified. Since then, an option to respond online has been included in at least two postal surveys of NHS patients. Firstly, the postal survey version of the Healthcare Commission's NHS and independent sector treatment centres' survey gave participants the option to respond online, but in spite of an exceptionally high response rate overall (72%) none of the 1673 sample members opted to complete the questionnaire via the internet¹²⁹. The 2007 GP Patient Survey also offered participants the opportunity to complete the survey online: to do this respondents must navigate to a given URL and enter an eight digit numerical username as well as a five digit alphabetical password so that they can be identified with respect to the sample¹³⁰. Information on the outcomes of this survey and the effects of using an e-mail response option are not presently available, but we would recommend that developments with regards to this are followed.

10.3 IVR

We are not presently aware of any surveys specifically targeting NHS patients and having been conducted principally or partially via the use of IVR. This reflects both the relative newness of the technology, its perceived complexity, and the difficulty of directly approaching patients by telephone.

10.4 Conclusions

Up to now, the vast majority of surveys conducted in the NHS have used a pen-and-paper based methodology, either via mailings or via handing out questionnaires to patients at the point of care. Few organisations have attempted to use alternative modes of data collection to gather feedback on patients' experiences, although the use of electronic data collection mechanisms is becoming more common as trusts seek to supplement data from national surveys with more regularly collected results for the purposes of on-going monitoring. Typically, these have involved surveying patients at the point of care using either bedside terminals or small hand-held units. Such efforts have generally prioritised a quick turnaround of data ahead of methodological rigour, however, and we believe that there is some way to go before organisations are making the best possible use of these alternative modes. Two larger surveys organised over many sites have also sought to incorporate an internet response option, but at present the usefulness of doing this is far from

¹²⁸ Castle, N.G., Brown, J., Hepner, K.A., & Hays, R.D. (2005) Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Services Research*, 40:6 part 2, p1996-2017

¹²⁹ Healthcare Commission (2006) Development report for the survey of NHS and independent sector treatment centres. Unpublished.

¹³⁰ <http://www.gp-patient.co.uk>

clear. More work is needed with all of the above modes of data collection to optimise their value for conducting surveys of NHS patients.

11 Cost implications

As noted earlier, one of the key reasons for adopting a mixed-mode approach is to ensure an optimal balance between minimising costs and maximising data quality¹³¹. How mixed and single mode approaches compare in terms of cost-effectiveness, however, is a complex question that depends heavily on the specific natures of the designs in question. In this section, we report on published evidence regarding the cost effects of mixed-mode approaches in different contexts.

11.1 Research evidence

Very few studies have looked explicitly at the cost-effectiveness of using mixed mode methodologies or the relative costs of different approaches, and of these the majority have included at least some interviewer administered methodologies. Presently, we are only aware of one study that explicitly looks at the relative costs of different mixed-mode approaches in the context of a health survey.

Rodriguez et al (2006)¹³² conducted a large scale randomised trial in the USA surveying patients on their experiences with individual physicians, using a range of methods including mail, internet, and IVR for data collection. Because no e-mail or telephone contact details were available for the patients in the survey, this involved using a unimodal approach by mail, with invitation letters either containing a paper questionnaire (mail arm) or details of how to complete the survey by accessing a given web address (web arm) or dialling a toll-free number (IVR arm) and entering a unique login code. Following this, non-respondents were sent a second letter one week after the first approach, and in the web and IVR arms a third letter containing a paper copy of the questionnaire was sent two weeks after the initial approach (mail crossover).

Rodriguez et al looked at the various costs of both the pure mode (eg, employing only one mode of data collection – ignoring responses from the mail crossover stage of the web and IVR arms) and mixed mode arms of the study. Unit completion costs were derived by dividing the total costs for each arm by the number of unique responses for that arm, meaning that response rates were – correctly – treated as an important aspect of cost efficiency. Comparable response rates (~50%) were obtained in the mailed arm and in both web and IVR arms after mail crossover, but prior to the use of the mail crossover response rates from both of the pure mode arms were poor – only 18.4% by web and 34.7% by IVR. Overall, costs were found to be much lower in the mail arm than in either the web or the IVR arm, both before and after crossover to mail. Recognising the potential for economies of scale to vastly reduce the marginal costs of internet and IVR surveying when used more regularly via reductions in “programming, testing, and production support costs” (p170), Rodriguez et al also tested the sensitivity of their findings by reducing these costs first by 50% and then by 100%, leaving only fieldwork costs. Mail remained the cheapest option per unique response (see table 5).

¹³¹ de Leeuw, E.D. (2005) To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21 (2), p. 235; italics author's own.

¹³² Rodriguez, H.P., von Glahnt, T., Rogers, W.H., Change, H., Fanjiang, G., & Safran, D.G. (2006) Evaluating patients' experiences with individual physicians: a randomised trial of mail, internet, and interactive voice response telephone administration of surveys. *Medical Care*, 44(2), p167-174.

Table 5: Survey costs by unique response for different approaches (reproduced from Rodriguez et al, 2006)

Arm	Mail arm	Web arm		IVR arm	
	Mail only	Web only	+ mail crossover	IVR only	+ mail crossover
Cost/unique response (\$)	5.19	13.94	8.01	9.04	8.06
-50% setup & production support costs	-	12.13	7.32	7.41	7.04
-100% setup & production support costs	-	10.33	6.64	5.77	6.01

Rodriguez et al's findings are difficult to reconcile with arguments extolling mixed mode methodologies as a way of cutting survey costs; rather, the evidence here indicates that mailed surveys are currently still likely to provide the greatest response rates at the greatest cost efficacy. It was argued, however, that the higher costs seen in this study simply reflect the need to support their use with mailings, since neither telephone numbers nor e-mail addresses were available. This, the authors argued, "affected [their] ability to capitalize on any cost savings associated with using these alternative survey modes" (p170). Nonetheless, they concede that "the cost advantage of traditional mail would remain so long as organisations remain faced with the need to rely on mail for initial outreach and/or to boost overall response rates"(p172).

11.2 Conclusions

Methodological details such as sampling strategies, means of approach, the possible routes of responding available, and likely outcomes such as response rates to different modes, as well as the specific development and implementation costs associated with these modes will all have an influence on overall costs. Since outcome related variables affect costs, it is reasonable to suggest that any estimates for unit costs of different modes would be unreliable without appropriate evidence from rigorous pilot work.

Rodriguez et al's (2006) study gives cause for concern regarding the costs of mixed-mode approaches, as it shows that cost benefits associated with some alternative modes will be undermined as long as it is necessary to support them with mailings. Since this would certainly be the case for NHS patient surveys, it may be difficult to implement a mixed mode design without raising survey costs.

12 Bibliography

- Atrostic, B.K., & Burt, G. (1999) Household non-response: what we have learned and a framework for the future. *Statistical Policy Working Paper 28, Federal Committee on Statistical Methodology, Office of Management and Budget*, pp.153-180.
- Balabanis, G., Mitchell, V-W., Heinonen-Mavrovouniotis, S. (2005) SMS-based surveys: strategies to improve participation. *Proceedings of the European Marketing Association Conference (EMAC)*, Milan.
- Beebe, T.J., McRae, J.A. Jr., Harrison, P.A., Davern, M.E., & Quinlan, K.B. (2005) Mail surveys resulted in more reports of substance abuse than telephone surveys. *Journal of Clinical Epidemiology*, 58(4), 421-4.
- Biemer, P.P. (1997) Unpublished data. Research Triangle, NC: Research Triangle Institute. cf Dillman, D.A. (2000) op cit p217.
- BMRB Social Research (2005). Patient Power (Bedside TVs and telephones): Survey of users, staff, providers, and patient representatives. London: BMRB.
- Bowling, A. (2005) Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, 27(3), 281-291.
- Boyce, S.J. (1999) Spoken natural language dialogue systems: User interface issues for the future. In: Gardner-Bonneau, D. (ed.) *Human factors and voice interactive systems*. Boston: Kluwer Academic.
- Brambilla, D.J. & McKinlay, S.M. (1987) A comparison of response to mailed questionnaires and telephone interviews in a mixed-mode health survey. *American Journal of Epidemiology*, 126: 962-971.
- Burroughs, T.E., Waterman, B.M., Cira, J.C., Desikan, R., & Dunagan, W.C. (2001) Patient satisfaction measurement strategies: a comparison of mail and phone methods. *Journal on Quality Improvement*, 27(7), 349-361.
- Cacioppo, J.T., & Petty, R.E. (1982) The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131;
- Caspar, R. (2003) *Census 2000 Testing, Experimentation, and Evaluation Program Synthesis Report No. 18, TR-18, Results from the Response Mode and Incentive Experiment in 2000*. U.S. Census Bureau, Washington DC 20233.
- Caspar, R.A. 2003 Synthesis of Results from the Response Mode and Incentives Experiment. National Census Test Response Mode Analysis report, p.19. cf Conrad & Couper, 2004.
- Castle, N.G., Brown, J., Hepner, K.A., & Hays, R.D. (2005) Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Services Research*, 40:6 part 2, p1996-2017
- Clayton, R.L., & Winter, D. (1992) Speech data entry: Results of a test of voice recognition for survey data collection. *Journal of Official Statistics*, 8, 377-388.

Conrad, F., & Couper, M. (2004) Usability, comparability, and data quality across modes and technologies in census data collection. US Census Bureau, p18.

De Leeuw, E.D. (1992) *Data quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT Publications. pp66-67.

De Leeuw, E.D. (2005) To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21 (2), pp. 233-255.

De Leeuw, E.D., & van der Zouwen, J. (1988) Data quality in telephone and face-to-face surveys: a comparative meta-analysis. In: Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nichols II,

W.L., & Waksberg, J. (eds) *Telephone survey methodology*. New York: Wiley.

Department of Health. (2000) *The NHS Plan: A Plan for Investment. A Plan for Reform*. Department of Health: London.

Dillman, D.A. (1978) *Mail and telephone surveys: the total design method*. New York: Wiley-Interscience.

Dillman, D.A. (2000) *Mail and internet surveys: the tailored design method*. New York: Wiley (Chapter 6)

Dillman, D.A. (2002) Presidential Address: Navigating the rapids of change: Some observations on survey methodology in the early twenty-first century. *Public Opinion Quarterly*, 66(3): 473-494.

Dillman, D.A., Brown, T.L., Carlson, J., Carpenter, E.H., Lorenz, F.O., Mason, R., Saltiel, J., & Sangster, R.L. (1995) Effects of category order on answers to mail and telephone surveys. *Rural Sociology*, 60, 674-687.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., & Berck, J. (2001) Response rate and measurement differences in mixed mode surveys: Using mail, telephone, interactive voice response, and the internet. Draft paper, last accessed 11/10/06, available at: http://www.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20_with%20Gallup_%20POQ.pdf

Dillman, D.A., West, K.K., & Clark, J.R. (1994) Influence of an invitation to answer by telephone on response to census questionnaires. *Public Opinion Quarterly*, 58:557-568

Fowler, F.J. Jr., Gallagher, P.M., Nederend, S. (1999) Comparing telephone and mail responses to the CAHPS survey instrument: Consumer Assessment of Health Plans Study. *Medical Care* 37(3):41-49.

Graham, C. (2006) Unpublished analysis from 2005 inpatient survey.

Graham, C. (2006) unpublished research based on results from the National Survey of Adult Inpatients 2005

Graham, C. (2007) unpublished research based on results from the National Survey of Adult Inpatients 2006

Gribble, R.K., & Haupt, C. (2005) Quantitative and qualitative differences between handout and mailed patient satisfaction surveys. *Medical Care*, 43(3), 276-281.

Groves, R.M. (1989) *Survey errors and survey costs*. New York: Wiley.

Groves, R.M. (1990) Theories and methods of telephone surveys. *Annual Review of Sociology*, 16, 1 pp. 221-240.

Groves, R.M., & Couper, M.P. (1998) *Nonresponse in household interview surveys*. New York: Wiley-Interscience.

Groves, R.M., Singer, E., & Corning, A. (2000) Leverage-salience theory of survey participation: description and an illustration. *Public Opinion Quarterly*, 64(3), p298-308.

Hall, M. (1995) Patient satisfaction or acquiescence? Comparing mail and telephone survey results. *Journal of Healthcare Marketing*, 15, 54–61.

Healthcare Commission (2006) Development report for the survey of independent and NHS treatment centres. Unpublished.

Hochstim, J.R. (1967) A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62, 976-989.

Ipsos MORI. (2006) Ipsos MORI Technology Tracker, September 2006. Ipsos MORI; London.

Jenkinson, C., Coulter, A., Reeves, R., Bruster, S., & Richards, N. (2003) Properties of the Picker Patient Experience questionnaire in a randomized controlled trial of long versus short form survey instruments. *Journal of Public Health Medicine*;25(3):197-201

Jordan, L.A., Marcus, A.C., & Reeder, L.G. (1980) Response styles in telephone and household interviewing: a field experiment. *Public Opinion Quarterly*, 44, 210-222.

Knäuper, B., & Schwarz, N. (2004) Why your research may be out of order. *The Psychologist*, 17, 28-31.

Kraus, L., & Augustin, R. (2001) Measuring alcohol consumption and alcohol-related problems: comparison of responses from self-administered questionnaires and telephone interviews. *Addiction*, 96(3): 459-71.

Krosnick, J., & Alwin, D.F. (1987) An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201-219.

Krosnick, J., Narayan, S., & Smith, W.R. (1996) Satisficing in surveys: Initial evidence. In

Braverman, M.T., & Slater, J.K. (Eds.), *New directions for evaluation*, 70 (*Advances in survey research*), (pp. 29-44). San Francisco: Jossey-Bass. cf Dillman (2000), op cit.

Kwak, N., & Radler, B. (2002) A comparison between mail and web surveys: response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18 (2), pp. 257-273.

Livock, T. (2006, Voxgen), Personal communication.

McMahon, S.R., Iwamoto, M., Massoudi, M.S., Yusuf, H.R., Stevenson, J.M., David, F., Chu, S.Y., & Pickering, L.K. (2003). Comparison of E-mail, Fax, and Postal Surveys of Pediatricians. *Pediatrics*, 111, 299-303.

Mi Kyung, J. (2005). Effects of survey mode, gender, and perceived sensitivity on the quality of data regarding sensitive health behaviours. Indiana University.

Miller, N., & Campbell, D. (1959) Recency and Primacy in Persuasion as Function of the Timing of Speeches and Measurements, *Journal of Abnormal and Social Psychology*, 54, 1-9.

Naik, M., Pallett, E., Thompson, J.P., Watson, P. Maslowski, P., Toogood, L., & Rowbotham, D. (2005) Comparison of an electronic speaking data recorder with the short form brief pain inventory in chronic pain patients. *British Journal of Anaesthesia* 95 (4): 566–79P [ABSTRACT]

National Statistics. (2006) First release: internet access, households and individuals. 23rd August 2006. National Statistics: London.

NetSize Guide 2005 report for 2004

Ofcom. (2006) *Mobile Call Termination: Report of Market Research Findings*. Office of Communication; London, England. 13th September 2006.

Office of Communications. (2006) *Consumers and the communications markets: 2006*. Ofcom: London. Available at:
http://www.ofcomconsumerpanel.org.uk/publications/consumer_panel_report06.pdf

Oftel Residential Survey, May 2003.

OR Treat, J.B. (2005) Research into the Use of Electronic Data Collection Modes for the 2010 United States Decennial Census. *Statistika*, 6/2005, p.449-478.

Park, D.C. (2000) The basic mechanisms accounting for age-related decline in cognitive function. In Park, D.C., & Schwarz, N. (eds.) *Cognitive aging: a primer*. Philadelphia: Psychology Press.

Patel, N., Pallett, E., Ali, S., & Stone, M.A. (2006) Development of a method of collecting questionnaire data from people with mixed language and literacy skills: a tool for use in diabetes research [ABSTRACT]

Quigley, B., Riemer, R.A., Cruzen, D.E., & Rosen, S. (2000) Internet versus paper survey administration: preliminary findings on response rates. *Proceedings of the 42nd Annual Conference of the International Military Testing Association, Edinburgh, Scotland*. Available at
<http://www.internationalmta.org/Documents/2000/Proceedings2000.pdf> (last accessed 16/02/07).

Raleigh, V. S., Scobie, S., Cook, A., Jones, S., Irons, R., & Hallt, K. (2004) *Unpacking the patients' perspective: Variations in NHS patient experience in England*. Healthcare Commission: London.

Raleigh, V.S., et al (2005) Variations in the experiences of patients in England: Analysis of the Healthcare Commission's 2003/2004 national survey of patients. The Healthcare Commission: London.

Reeves, B., & Nass, C. (1996) *The media equation: How people treat computers, television, and new media like real people and places*. Massachusetts: Cambridge University Press.

Roberts, D. (2005) Recreating experiences. *Evaluation Journal of Australia*, 4(new series), 1&2, p44-51.

Rodriguez, H.P., von Glahnt, T., Rogers, W.H., Change, H., Fanjiang, G., & Safran, D.G. (2006)

Evaluating patients' experiences with individual physicians: a randomised trial of mail, internet, and interactive voice response telephone administration of surveys. *Medical Care*, 44(2), p167-174.

Schaeffer, D.R., & Dillman, D.A. (1998) Development of a standard e-mail methodology: results of an experiment. *Public Opinion Quarterly*, 62(3), 378-397.

Schechter, S., Beatty, P., & Block, A. (1994) Cognitive issues and methodological implications in the development and testing of a traffic safety questionnaire. *Paper Presented at the 49th Annual Conference of the American Association for Public Opinion Research. May 11-15, 1994. Proceedings of the Survey Research Methods Section, American Statistical Association.*

Schneider, S.J., Cantor, D., Heller, T.H., & Brick, P.D. (2002) Pretesting Interactive Voice Response/Automated Speech Recognition Surveys. Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing Methods, Charleston, South Carolina, November 14-17, 2002.

Schonlau, M., Asch, B.J., & Du, C. (2003) Web surveys as part of a mixed mode strategy for populations that cannot be contacted by e-mail. *Social Science Computer Review*, 21(2), 218-222.

Schuman, H. & Presser, S. (1981) *Questions and answers in attitude surveys*. New York: Academic Press

Schwarz, N., Knauper, B., Hippler, H.J., Noelle-Neuman, E., & Clark, L. (1991) Rating scales: numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, 618-630.

Sheldon, H., Graham, C., Potheary, N., & Rasul, F. (2007) Increasing response rates amongst black and minority ethnic and seldom heard groups – a review of literature relevant to the national acute patients' survey. Oxford, UK: Picker Institute Europe.

Siemiatycki, J., & Campbell, S. (1984) Nonresponse Bias and Early Versus All Responders in Mail and Telephone Surveys. *American Journal of Epidemiology*, 120: 291-301.

The Carphone Warehouse. (2006) *The Mobile Life Report 2006: the impact of the mobile phone on the lives of young people*. Mobile Life. Available at <http://www.mobilelife2006.co.uk/PDF/Mobile%20Life%20Youth%20Report%202006%20Colour.pdf>

Tourangeau, R. (1984) "Cognitive Sciences and Survey Methods." In Jabine, T., Loftus, E., Straf, M., Tanur, J., and Tourangeau, R., eds., *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington, D.C.: National Academy of Science;

Tourangeau, R. (2004) *Survey research and societal change*. Annual Review of Psychology, 55:775-801

Tourangeau, R., & Smith, T. (1998) Collecting sensitive information with different modes of data collection. In Couper, M., Baker, R., Bethlehem, J., Clark, C., Martin, J., Nichols, W., & O'Reilly, J. (eds.) *Computer assisted survey information collection*. New York: Wiley.

Tourangeau, R., & Smith, T.W. (1996) Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60, 275-304.

Tourangeau, R., Rips, L.J., & Rasinski, K. (2000) *The Psychology of Survey Response*. Cambridge University Press: Cambridge, UK.

Townsend, L. (2005) The status of wireless survey solutions: the emerging "power of the thumb". *Journal of Interactive Advertising*, 6, 1, p.52-58

Treat, J.B., Brady, S., Bouffard, A., & Stapleton, C. (2003) 2003 National Census Test: The Impact of Alternative Modes and Contact Strategies on Self-Response. Unpublished report. Washington, DC: U.S. Census Bureau.

Uttaro, T., Leahy, V., Gonzalez, A., & Henri, W.F. (2004) Effect of type of survey administrator on Consumer Assessment of Care. *Psychology Report, 94 (3 Pt 2)*, p1279-82.

Van der Horst, W., Snijders, C., & Matzat, U. (2006). The effects of different kinds of progress indicators on online survey compliance and data quality. Presentation from the 8th International General Online Research (GOR) Conference, March 21 2006.

Virtanen, V., Sirkiä, T., & Jokirinta, V. (2007) Reducing nonresponse by SMS reminders in mail surveys. *Social Science Computer Review, 25 (3)*.

Virtanen, V., Sirkiä, T., & Nurmela, J. (2005) Reducing nonresponse by SMS reminders in three sample surveys. *Paper presented at the conference of the European Association of Survey Research, 18-22 July 2005, Barcelona*. <http://sqp.nl/easr/programme/abstracts.pdf> (p71; last accessed 28/03/07).

Walker, A.H., & Restuccia, J.D. (1984) Obtaining information on patient satisfaction with care: mail versus telephone. *Health Services Research, 19(3)*, 291-306.