National Centre for Research Methods Review Paper

On the experience and evidence about mixing modes of data collection in large-scale surveys where the web is used as one of the modes in data collection

Shirley Dex, Centre for Longitudinal Studies, IoE

Julia Gumy, Magdalen College, University of Cambridge

NCRM
National Centre for Research Methods

E·S·R·C
ECONOMIC & SOCIAL RESEARCH COUNCIL

National Centre for Research Methods Review Paper

On the experience and evidence about mixing modes of data collection in large-scale surveys where the web is used as one of the modes in data collection

Shirley Dex

Centre for Longitudinal Studies
Institute of Education, University of London
s.dex@ioe.ac.uk (for communications)

and

Julia Gumy

Magdalen College
University of Cambridge

November 2011

Acknowledgements.
The authors wish to express their gratitude to a long list of people who have helped in identifying material relevant to this Review by responding to emails, sharing their knowledge, sending us papers and giving feedback on an earlier Draft. In alphabetical order: Sami Benyahia, Frédéric Berger, Helene Berard, Peter Betts; Bev Botting, Paul Bradshaw, Matthew Brown, Josef Brüderl, Brendan Burchell, Jon Burton; Lisa Calderwood, Debbie Collins, Mick Couper, Marcel Das, Fiona Dawe, Matthew DeBell, Don Dillman, Bobby Duffy, Jane Elliott, Bob Erens, Pablo de Pedraza Garcia, Debbie Griffin, Anders Holmberg, Nick Howat; Annette Jäckle; Dominique Joye, Joost Kappelhof, Antje Kirchner, Bengt Oscar Lagerstrom, Aat Liefbroer, Peter Lugtig, Peter Lynn; Enrique Fernadez Macias, Peter Martin, Bob Michael, William Mockovak, Nick Moon, Gerry Nicolaas, Donna J. Nordquist, Mary Beth Ofstedal, Richard Puttock, Debbie Reide, Ulf-Dietrich Reips, Jay Ryan, Hildegard Schaeper, Annette Scherpenzeel, Jackie Scott, Matti Simpanen, Patrick Sturgis; Brenda Taggart, Alison Thompson, Kea Tijdens, David Weir, Knut Wenzig, Boris Wernli, Hannah White, Joel Williams, Laura Wilson, and Natasha Wood.

ABSTRACT

This paper presents experience and evidence, from international sources, of the effects on survey quality of the use of mixed modes in data collection, where the web is used as one of the modes. It is recognised that this is an area which is changing fast as the spread of new technologies assist in widening the population who have access and familiarity with using the web. Surveys found to have used mixed modes (including web) were classified around the conventional schema in the literature; according to whether they used concurrent or sequential approaches in their mixed modes and according to whether they were longitudinal or cross-sectional data collections. In addition, surveys are divided according to whether they attempted to survey general populations or sub populations. . Existing longitudinal data collection is of particular interest, and especially where data collection modes have changed at some point into the tracking of individuals. Relatively few longitudinal surveys were identified that had attempted to use the web in data collection and no analyses of the consequence of the mode of response on subsequent wave attrition were found. Many of the examples found focused on sub populations of students and young people, who might call the web-savvy populations. This review showed that experience and familiarity with carrying out mixed mode surveys that include the web have grown considerably over the first decade of the 21st century. For experience of surveying general populations, one needs to look to Scandinavian countries, especially the Netherlands. These countries have also been ahead of most others in their prevalence of households with access to the internet.

The review of the literature on survey quality effects of mixed modes shows that use of sequential mixed modes, using the web first, followed by more expensive modes, can achieve response rates on a par with good response rates from high quality single mode studies. The different modes recruit samples with different characteristics. However, such approaches, in combining responses from different modes can also gain coverage of the general population that improves on single mode studies. The evidence from experiments also points out how to get the highest web responses at the first stage of a sequential mixed mode data collection series. This is by failing to mention that there are other mode options for the response. However, measurement errors, namely nonresponse errors and mode effects, are likely to be evident in the data collected. These mode effects are still being charted, and are seen to vary according to the type of question, the type of response codes and even the particular topic content; in some cases the mode effects are relatively minor and in other cases substantial. There is also a problem of confounding which is often present with non-response measurement errors or time effects for longitudinal data. There are well documented generalisations about the social desirability consequences and satisficing under different modes. Other research work is continuing to try and identify how to devise questions, by type, that will minimise mode effects. This involves painstaking attention often to the detail of individual questions. Most researchers think the unimode approach to mixed mode question construction is likely to dominate for some time yet, and UK survey fieldwork organisations also tend to adopt this approach.

Contents

Glossary of terms

CAPI/CATI/CASI/ - computer assisted Personal/Telephone/Self-Administered interviews
CAWI - computer assisted web interview
DK – Don't Know, as a potential answer to a survey question.
F2F – face to face
IVR – Interactive Voice Recognition
NEET – Not in employment, education or training.
RDD – Random digital dialling
RR – response rate (non-standardised)
Web – includes and covers terminology of online, internet and web as accessed by any route (smart phone, desktop, laptop, handheld device).

1. Introduction - This Review

The main purpose of this Review is to seek out and present the experience and evidence of the effects on survey quality of the use of mixed modes in data collection, where the web is used as one of the modes. This focus has grown out of the interest of the UK's major longitudinal social science resource datasets to consider new innovations and potentially adapt to new austerity funding environments. At the same time there is a need to maintain the highest quality to these data resources.

There have been other reviews of written material on mixed-modes of data collection in recent years (Betts and Lound, 2010a; Roberts, 2007; Ariel et al, 2008; and more on particular (sub) topics, for example Shih and Fan, 2007 comparing web and mail surveys; Watson, 2009 comparing web and paper). These reviews show that this is an area of great interest and rapid development. This review draws on, without repeating in detail, some of these earlier reviews, but also differs in a number of respects. Firstly, this review focuses on mixed methods that include web/online data collection as one of the modes whereas the existing general review papers cover all of the potential mixed mode options. In taking the wider specification, earlier reviews end up being less detailed on any one type of mixed mode approach. The narrower focus of this review lends itself to a more detailed examination on mixed mode data collection where one mode is the web. However, the speed of development in this area means that even within one year, things have moved on, so reviews face a fast-changing environment.

This review has one background factor as a major interest and constraint; that is the collection of longitudinal data where earlier waves have been done with the minimum of mixed mode approaches (largely limited to CAPI plus CASI within one interview, or CAPI-CATI options to avoid refusals). So experience and evidence about using the web in the context of collecting longitudinal data is of particular interest. This is especially the case when mixed mode data collection was adopted at some point into the longitudinal tracking of individuals. Nonetheless, experience of the use of mixed mode for cross-sectional data collection and the collection of longitudinal data planned as mixed mode from the outset may potentially have lessons to impart.

In particular we are interested in answering the questions:

- Are there longitudinal studies that we can learn from in terms of their experience of using the web as part of mixed modes in data collection?
- Are there specific issues to face in changing modes during the course of collecting data for a longitudinal study?
- What is the impact of mixed mode data collection which includes the web on data quality?
- What are the likely response rates to a web survey?
- What are the impacts of mixed mode approaches on total response rates?
- Do fieldwork agencies in the UK have the successful experiences and capacity to carry out mixed mode surveys of various kinds?
- Are mixed mode approaches likely to be cheaper?

We first layout the various types of mixed-mode approaches.

What type of mixed-mode approach?
A mode can, confusingly, be used to refer to the method of data collection, or even the entire survey process. As lead authors in this field have pointed out, there are a number of key dimensions that the data collection mode involves; these include, the degree of interviewer involvement; the degree of contact with the respondent; the channels of communication; the locus of control; the degree of privacy and the degree of technology use. The modes we use, and how we combine them, are likely to have implications for survey quality.

It is common to note De Leeuw's (2005) distinction between mixed mode systems which can operate at the three main phases of survey operations: the contact phase, the response phase and the follow up phase of fieldwork. This Review focuses on the Response Phase only and does not include reviews of the use of two or more modes across these three phases. So, for example, an advance letter sent by post for contact, followed by an web survey in the response phase, would not be included here as an example of mixed mode. De Leeuw (2005, Fig 1, p.238) makes further distinctions about the Response phase in Table 1 below, where other necessary distinctions have also been incorporated.

The main types into which mixed mode data can be introduced in Table 1 cover the standard cross-sectional survey (Type MM-1) with different modes being used to collect data from different subsamples. Injecting another mode within the one cross-sectional survey which all respondents complete (Type MM-2) is another variant type. Type MM-3 represents the longitudinal survey which has multiple time periods with different modes being used to collect the data at different time point waves. A set of more complex possibilities which would cover cross-national comparisons is the fourth type (Type MM-4). In the cases of Types MM-1 and MM-3 a further distinction is highlighted; this is the possibility of introducing mixed mode data collection as a concurrent or sequential set of choices. Concurrent choices are all presented at the same time, with respondents able to select the one they prefer. In sequential choices, respondents are steered to one option in the first instance, and only if they fail to respond, will they subsequently be offered one (or more) alternative modes via which to provide their data. If the cheapest mode is used first, this is potentially a method of reducing survey costs.

Our main interest in initiating this review was with the collection of longitudinal data. Type MM-3 is, therefore, a main focus of interest. The idea lying behind this type of mixed mode design is that the early waves of a longitudinal study might need to use expensive methods to ensure representative coverage of a large population, and to build a relationship with the respondents. Having gained some commitment to the study, and time to collect details needed for other types of data collection modes (eg. Telephone numbers; email addresses), a longitudinal study can then use less expensive methods to continue to collect data from respondents. It is the case that the specific UK longitudinal surveys that are considering new uses of mixed-mode designs in future are those where other modes have been used to collect data at earlier waves. Type MM-3 is definitely of relevance to this review, therefore, both in terms of the experience of using such methods and in the evidence about effects on data quality that result. There is currently, however, a relative

scarcity of examples of the use of mixed modes (including web) across longitudinal data collection waves.

Table 1. Types of mixed mode designs at the Response Phase , their main rationales and effects on survey quality

Mixed mode design	Type of data/ Rationale for implementation	Effects on survey quality
Type MM-1.		
One sample, one time period One questionnaire	Cross-sectional survey	
Different modes applied to different sub-samples *Options* Sequential Concurrent	*Reduce cost *Improve coverage *Improve response	*Reduce coverage and non-response error *Mode effects on measurement /confounded with subgroups
Type MM-2.		
One sample, one time period	Cross-sectional survey	
Different modes applied to different parts of the questionnaire (for the same person)	*Improve privacy of measurement *Reduce social desirability/ acquiescence bias	*Improve data quality on sensitive questions
Type MM-3.		
One sample, multiple time periods.	Longitudinal survey	
Different modes used to measure same person at different points in time *Options* Sequential Concurrent One mode per wave only	*Reduce cost	*Measurement differences which confound time effects and mode effects.
Type MM-4.		
Different samples.	Cross-national comparisons Mainly cross-sectional	
Different modes used on different samples (plus sometimes different times and different questionnaires.)	*Different research traditions *Different coverage *Different facilities and access *Different cost structure	*Coverage error *Non-response error *Measurement error *Incomparability

Longitudinal studies that have adopted mixed mode data collection from the outset are currently also uncommon. They are not distinguished specifically as part of De Leeuw,'s (2005) set of categories. One might think that such studies do not necessarily exhibit the risks to survey quality that come from switching to mixed mode following the use of single mode data collection at an earlier wave. However, the use of more than one mode at any one data collection wave means that individuals in a longitudinal study will always be able

to switch modes between data collection waves. Even mixed mode longitudinal studies planned from the outset, therefore, will have similar issues to face as the ones that start to use mixed mode data collection later in their trajectories if sequential or concurrent mixed mode choices are made at some waves.

However, it is also the case that longitudinal surveys consist of a series of cross-sectional waves of data collection, and so each new wave has elements in common with Type MM-1. For this reason, examples and evidence about type MM-1 are also of interest. The relatively scarcity of examples and evidence about Type MM-3 uses of mixed mode (including web) for longitudinal data made it even more important to include learning from Type MM-1. But in addition, a longitudinal study can engage in even more complex mode changes since it is possible to have both between wave mode changes, as described above, at the same time as within wave mode changes, for example by adopting a sequential mixed mode approach at one or more waves, when this was not the approach used earlier at the start of the survey. Added complications are likely to result to the measurement errors from such complicated use of multiple modes, putting them more on a par with the complexity of measurement errors implied in mixed mode type MM-4. This complexity is evident in studies such as the LSYPE and YCS (see Appendix 6).

This review largely ignores cross-national studies fitting into Type MM-4 as a highly complex and specialist area. We also failed to find any substantial examples of Type MM-2 using web as one of the modes.

There is a further distinction between single frame and dual frame designs according to whether the all potential respondents can be approached via a single mode because the information is available about their contact details on a single sampling frame; or whether more than one sampling frame, dual frame for example, is required to cover the contact details of all members of the sample, or more than one sample is needed to gain coverage of a whole population.

What do we mean by 'the web'?
The ways of accessing 'the web' as a mode of data collection has more variations than tends to characterise other single mode studies. This is because the routes to using the web have been proliferating via different hardware options (for example, fast and slow dial up broadband connections in the home or in public spaces; mobile phones connected to the internet; other hand-held devices with internet connections). Different manufacturers of the hardware also load or impose different software options, operating systems, browsers and screen settings onto their products, some of which interface with other software and some do not. Respondents can therefore face constraints on their ability to fill in web questionnaires, even when they are motivated to do this and they potentially have at least one access route to the internet. This is part of the fast changing environment in which survey data collection is having to navigate. This review is able to provide and note some statistics about access routes to the web. However, the literature considering the web as a mode in data collection, has not currently ventured into making distinctions between different web access routes in its discussions of the implications for data quality of using the web via these different alternative access routes.

The rest of the Review
The rest of this Review proceeds as follows: Section 2 provides a general overview of the types of evidence and experience of surveys and survey research on which we draw in this review showing the range of datasets by types of mixed mode designs that we found. Section 3 considers the evidence on effects on survey quality of collecting data using mixed mode designs of varying types. This is prefaced by a review of the range of quality issues that concern us. In Section 4 we present the final conclusions.

2. Overview of studies - evidence and experience

Sources of evidence and experience
The sources that make up this review offer a range of experience and evidence about the use of mixed mode data collection that involves the web as one of the modes. Our approach to assembling this material is described in Appendix 1.The sources consist of the following.

(a) The experiences of longitudinal surveys which have used mixed modes of data collection at one or more waves. These experiences come from written documents and from conversations with fieldwork agencies and data custodians. Surveys are included as 'longitudinal' in the list by satisfying a minimal requirement that they have at least one main data collection stage *and* one follow up at a different point in time, either of the whole initial sample or of a sub-sample.
(b) The experiences of cross-sectional surveys that have used mixed modes of data collection – focussing more heavily on large-scale rather than small-scale studies. These are included on the basis that they have things in common with the collection of a single wave of a longitudinal data set.
(c) Summaries of response in single mode studies where the single mode is the web. Such studies have the potential to tell us something about likely response from sub-populations who are approached to do a web interview as the first part of a sequential mixed mode design. There have also been innovative designs attempting to compare the quality of data obtained by administering the same survey instruments in more than one mode in parallel in separate and distinct surveys, so-called Dual Frame studies (eg. Link and Mokdad, (2006) on a health surveillance survey; Schonlau et al , 2003 on a large-scale survey on health care; Revilla and Saris, (2010) for the European Social Survey- Netherlands; and Duffy et al (2005), comparing a web panel with a F2F survey – for some further details see Appendix 5).
(d) The experience and learning from internet 'panels'. Volunteers are recruited who agree, in principle, to respond to regular sets of survey questions put to them by the fieldwork agent that recruited them, in exchange for payment. One could argue that the prior 'commitment' of someone on an internet panel and their agreement to be re-contacted regularly has elements in common with a cohort member of a longitudinal survey.
(e) Experimental studies specifically designed and controlled to reveal knowledge about mode effects. These are the most rigorous in their design, but many are small scale and applied to sub populations which leave caveats about their applicability to more general or whole populations. However, there are some obvious exceptions. National

statistics agencies, particularly in the Netherlands (van der Laan and Nunspeet, 2009; Cuppen et al, 2010), Sweden and Norway, have been pioneering the use of mixed mode data collection in large-scale population surveys. They have taken the opportunities to embed experiments into the design and alongside data collection of the survey data. Another innovative approach adopted in a recent research initiative used experiments along with quantitative data analysis and cognitive interviewing across three modes of data collection (F2F, Telephone and web) to test out hypotheses about mode effects under carefully controlled conditions (SDMI, Appendix 8).

It is also important to highlight that this is a fast changing field. Generalisations that were common in the 1990s, differed from those in the early twenty-first century, and things are still moving on apace. Clearly the gradual increase in access to and use of the web by the general public is making things possible that previously were not. This also means that some of the limitations of the web mode of data collection are gradually being reduced as time goes on. This review tries to give brief overviews of the direction of movement of the trends in our knowledge in this field, and tries to give more emphasis to the recent than to the past studies on this topic.

Mixed mode example studies

A set of example surveys were identified where mixed mode designs had been used to collect the data. They cannot claim to be exhaustive, but it is likely they represent the range and frequency of applications under each of the mixed mode headings. In order to be included in the table, the survey had to be reasonably large scale. In most cases samples sizes are above, often substantially, 2,500 cases.[1] The survey also needed to have collected its data for a purpose other than merely for an experiment. Some experimental studies have been included where they were embedded as part of the data collection, a practice that has been common in a number of Scandinavian countries. An attempt is made to draw a division between surveys of general population groups, and those of sub populations. However, clearly, this distinction is not always clear cut.[2] For some of these surveys further details were available, either from written publications, or from conversations and information supplied by fieldwork agencies. These are indicated by '**Details**' in the final column and are presented in Appendix 6. Datasets that were considered and contacted but found not to meet our specifications are also listed in Appendix 1.

Longitudinal surveys.

Longitudinal surveys that are using (or have used) mixed mode data collection, including the web (Type MM-3), are relatively infrequent. Under the sequential mixed mode heading, 8 cases of general population studies are listed in Table 2 as well as a further 6 longitudinal

[1] However, sub groups who were followed up from mainstage surveys, for example refusers, may not reach the 1000 cases threshold even though the main survey contains a much larger sample size. Three surveys included in the lists had only 1000-1500 cases these are Converse et al (2008), Meckel et al (2005) and Dannetun et al (2007).

[2] Some of the surveys of sub-populations, for example of people over 50, can be large and arguably general surveys. Large-scale surveys of such groups have more in common with whole population samples based on national random address databases, than they do, for example, with students or academics that are known users of email and the web.

surveys of sub-populations. They are all cases where an existing longitudinal study has turned to mixed mode data collection after starting out with (mainly) single mode collection, but their nature does not provide the learning opportunities one might have hoped for. Of the 8 surveys under the general population heading are two cases following up refusers to the main longitudinal survey using web based data collection (SHP, Dutch LFS). These were trials to see if this was a route to helping minimise attrition. A further two cases (ALLBUS, US consumer attitudes) were attempts at using the longitudinal data about respondents and their access to the internet, from earlier waves, to identify potential cases for an internet panel. These latter examples only served to demonstrate that this was a highly inefficient way of trying to recruit for an internet panel. One more case (Dutch LFS) was using the LFS to recruit for another survey as part of the new Dutch approach to integrating their data collection across surveys. Things were learnt from all of these exercises, but they are not things which will particularly help existing longitudinal datasets of general populations to adopt, avoiding obvious pitfalls, sequential mixed mode methods. The longitudinal studies of sub-population that have used sequential methods have more learning points to offer (along with the cross-sectional sequential mixed mode approaches), except that we need to remember they focussed on particular sub groups who were mainly experienced web users.

Of the six cases of longitudinal sub-group surveys using sequential mixed modes, five are on populations of young people or students, and four are UK/England datasets. Again these datasets turned to mixed mode and web data collection at some point in their history, rather than from the outset. Examination of the details of some of these surveys (eg. LSYPE, ELS, YCS) revealed that they can be considered to have been highly successful in delivering good quality survey data using a sequential mixed mode approach – certainly as measured by response rates.[3] It is also worth noting that the instruments used in these studies were all reasonably substantial (see details Appendix 6) lasting 25 minutes or more, unlike many of the instruments from the cross-sectional surveys using sequential approaches.

Across the general and sub-population studies there are 7 examples of starting off the sequence with a (cheaper) web instrument and moving through to more expensive options for those not responding at the earlier stages. It became clear in talking to UK fieldwork agencies that they have accumulated much experience in carrying out these sequences, albeit more especially on populations of young people. They have procedures and algorithms for cleaning up emails, judging when to initiate the next stage, how many reminders are needed and which routes to send them, and so on. It was also felt to be important not to allow respondents to refuse in the middle stages of the sequence, especially if a more persuasive (human) interviewer was the last phase and likely to have the best chance to succeed in securing a response. Response was found to be higher among those where they had a valid email, than those where they did not. However, there was also the potential for very high rates of invalid email addresses and ‘phone data for young people, after the 3.5 year gap between waves of the DLHE survey, with a high loss of contact with these cases resulting (Appendix 6).

[3] The representation and coverage quality issues for surveys can be argued to be slightly less important for a longitudinal survey than for a cross-sectional survey since the strength of longitudinal data is its potential for uncovering relationships between past experiences and present outcomes.)

It is also worth noting, that in nearly all of the cases using sequential mixed mode data collection for longitudinal data, the data documentation does not highlight or draw attention to potential measurement error issues to consider when combining the data across modes. The NSS survey is an exception here (see Surridge 2006, 2008). Nor did most of the datasets refer to question wording having considered the potential for mode effects. LSYPE and YCS were exceptions in this respect where it was clear that the change to mixed mode data collection (eg. Wave 5 in LSYPE) had led to revisions to question wording using the unimode approach (see below for more details about this approach).

In communications with several longitudinal data providers in the EU and the USA caution was expressed about the prospect of moving towards using web based data collection. It was considered too soon to be thinking of this, given the state of knowledge of this mode and its consequences. A number of risks and barriers were identified. One large risk is that of losing contact with valuable panel/cohort members; in particular whether this risk varies by the mode of data collection in an earlier wave. There is currently little, if any, information about mode effects on attrition that include web as a mode. There was felt to be insufficient information about whether the complex instruments and recall instruments, often used in the major panel/cohort studies, would transpose successfully to web instruments; and about how to adapt the use of a Household Grid, on which other interviews depend and flow from, to a web environment. The fact that funding for longitudinal data has to be planned and allocated for much longer time periods into the future may also give such datasets a small cushion from having to rush into changing to cheaper data collection modes, than would be the case for one-off cross-sectional surveys. Even where pressures to move to cheaper forms of data collection were being experienced by longitudinal studies, there was recognition of the many unknown factors. Lynn (2011-Appendix 9 presentation; Lynn, 2011) draws attention to the issues raised for household surveys in which all adult household members are to be interviewed. It has been traditional to first get a Household Grid filled in to be able to see the list of eligible household members for interview, and have individual interviews all lead off from this one pivotal instrument. Opening up the Household Grid to be filled in as a web questionnaire, also opens up the possibility that more than one person might try to fill it in, adding confusion and scuppering the clarity of the subsequent fieldwork tasks (Lynn, 2011-Appendix 9). Adopting the web as a new mode of data collection during the course of a longitudinal study does open the way to use information about the households and individuals in them, collected at earlier waves, to ease the transition, or to offer new within-household forms of communications. This would be a new form and use of dependent interviewing. However, here too, there is little knowledge on which to build.

Table 2. Examples of longitudinal surveys with sequential mixed mode designs where web is one mode.

'Details' =further information about this survey contained in Appendix 6**.**
Key: → indicates the order of the sequence.

Type MM-3 Longitudinal* SEQUENTIAL ***General populations***	**References**
Stockholm County Council Public Health Survey, 2007 Follow up of 2002 survey of 18-84 year olds. 2002 mail survey only. 2007 mail/paper → mail/paper+web; and, web→web+mail/paper	Holmberg et, al (2009) Lorenc (2010) Statistics Sweden.
Swiss Household Panel. (**SHP)**. To replace usual Telephone mode, offered Wave 12 (2010) refusers possibility of web questionnaire.	Voorpostel and Ryser (2011) **Details.**
Labour Force Survey, (**LFS)** UK. Online Pilot 2010/11[4] Follow up of sample of those who had completed 5 quarterly waves of LFS by CAPI (W1)+ CATI (W2 to W5) with request to do web survey	Dawe and Wilson (2011- Appendix 9) **Details.**
Dutch Labour Force Survey. Follow up of 2005 refusers to LFS panel element (age 15+). Households who refused at Wave 2 CATI were either: 1) Assigned CAPI if no telephone number, CATI otherwise.(RR=76.6% for these options) 2) Ask only a subset of Basic questions. CATI if phone number; Web or mail if no phone.(RR=45.4% for these options).	Banning and Schouten (2009)
German ALLBUS (**GSS**), 2006. Respondents with internet access invited to do a web follow up; 46% were internet users; 37% or those agreed; and 64% of those responded after much effort.	Couper (2007)
US Survey of Consumer Attitudes Those interviewed earlier by phone and were willing were followed up to join an internet panel. Found to be very inefficient way of recruiting a web panel.	Couper (2007)
Health Interview Survey, 2008, Netherlands Embedded in Labour Force Survey. LFS respondents asked for permission and details to be recontacted with another questionnaire.	Cuppen et al, 2010.

[4] This pilot work aimed at moving the UK's Labour Force Survey from being mixed mode CAPI-CATI to include a web mode by dates from 2012 onwards in order to reduce costs is typical of work that is taking place, at differing stages of development across many EU countries who are obliged to carry out a Labour Force Survey (eg. Denmark, The Netherlands, Germany) and among other countries who also have similar surveys (eg. Canada and New Zealand).The papers presented at the past two annual conferences of the *European Workshop on Labour Force Survey Methodology* show the plans of various countries; for the 6th conference in Wiesbaden on 12-13 May 2011, see http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/Content/Events/LFS/Papers,templateId=renderPrint.psml
For 5th conference in Paris, 15-16th April 2010 see http://www.insee.fr/en/insee-statistique-publique/default.asp?page=colloques/wlfsm/wlfsm-papers-presentations.htm.

Approached by email address with web questionnaire (if email)→CATI. and CATI if only had Phone number.	
*Generations and Gender Survey, (***GGS)** Dutch. (ages18-79) Wave 1 (2004) and Wave 2 (2007) were both CAPI Wave 3 (2010) sequential web→CATI (if phone number)→CAPI	Liefbroer, 2011 Appendix 9.**Details.**
Sub-populations	
Longitudinal Study of Young People in England, (**LSYPE)**, Wave 5, 6 and 7 use Web →Phone→F2F	Department for Education (2010a). **Details**
Youth Cohort Study (**YCS**) Cohort 13, Wave 2, 3 use Web →Phone→F2F	Department for Education (2010b). Carpenter (2010). **Details**
Education Longitudinal Study of 2002 (**ELS**:2002) USA. Base year (2002) with 2-year follow ups of high school leavers in 2004 and 2006. Early contacts with students done in school with tests+paper questionnaire. Second follow up in 2006 was web instrument. Non-responders contacted for CATI or CAPI interview.	Ingels et al (2007) http://nces.ed.gov **Details**
Beginning Teacher Longitudinal Study, USA Study of teachers starting to teach in 2007-08 followed up for 3 waves. Wave 1. Mailed Paper →Telephone or F2F follow up Total RR=86% Wave 2 Sent Web details→Telephone+ Mail Paper. 24% reply by web before follow up. Total RR=86% Wave 3. Sent web details→Telephone. Total RR=86%	Tourkin et al (2010) Kaiser and Cross (2010). Keigher (2010) http://nces.ed.gov
*Destinations of Leavers of Higher Education (**DLHE -Longitudinal**).* Web→Postal/paper→Phone. Followed up at 6 months and 3.5 years after leaving higher education. Email sent to those with email to do Online. Paper/mail sent 2 weeks later to those not done online or with no email.	HESA (2009) **Details.**
Effective Provision of Primary and Secondary Education. (**EPPSE 3-14, and 16+.)** *Children's Questionnaire*, England. Cohorts of children, N=3000, followed up 2003 – 2013 (ages 3 to 16+). In 2003 children aged 6-7 given own Paper Questionnaire to fill in. At age 11, Paper Questionnaire mailed → mailed choice of another Paper copy; Telephone or F2F for non-responders. At age 14 Paper Questionnaire mailed→ another mailed copy plus choice of web/emailed electronic copy/ Phone/ F2F for non-responders. Plus Text reminders. Plus incentives for hard-to-reach. 5% of total sample opted for web return.	http://eppe.ioe.ac.uk/eppe/eppeintro.htm

* Surveys are included as 'longitudinal' in the list by satisfying a minimal requirement that they have at least one main data collection and one follow up at a different point in time, either of the whole initial sample or of a sub-sample.

There were three cases of longitudinal datasets were the mixed mode came from having separate waves of longitudinal data collected by different modes (Table 3). However, the largest case, the German NEPS, has yet to complete its data collection plans. The regular web surveys, from the US Health and Retirement Survey, sandwiched in off-years, between the 2-yearly core data collections (using CAPI/CATI) have been found to be useful for methodological investigation, for collecting supplementary data on particular topics and for keeping in touch with respondents (see Appendix 6 item for HRS).

We were unable to find examples, fitting the specification, of concurrent mixed mode designs used on longitudinal data (Type MM-3 Concurrent), either in general or sub populations. Perhaps this is not surprising, since it would add complexity in the consideration of measurement error, without necessarily adding to the response and even risking losing cases. Some of the experimental work reported below even suggests that offering mode options concurrently to respondents can even reduce response, so concurrent mixed mode use may have been seen as too risky to try, and possibly not offering substantial cost savings.

Table 3. Examples of longitudinal studies using mixed mode by one mode per wave.[5]

Type MM-3 Longitudinal* ONE MODE PER WAVE ONLY	
Sub populations	
German National Education Panel Study (**NEPS**). Cohorts of students in education. Wave1. Pap&Pen questionnaire (recruiting) Autumn 2010 Wave2. Phone interview. Winter 2010/2011 Wave3. Grp administer competence tests (Pap&Pen) Spring 2011 Wave4. Web survey (Autumn 2011) Wave5. Phone interview (Spring 2012) Wave6. Web survey (Autumn 2012) Wave7. Phone interview + competence test (Spring 2013) Wave8. Web survey (Autumn 2013).	Von Maurice et al (2011). Aschinger et al (2011)
Danish National Birth Cohort. (**DNBC**) The 11-year follow up of initial sample of 91,661 pregnant women in 1996-2002 doing web data collection with mother and cohort child. Mothers have previously mainly done CAPI-CASI interviews.	Meder and Larsen (2011)
Health and Retirement Survey (**HRS**) USA. Aged over 50. Two-yearly cycle for core data collection waves using mix of phone and F2F. Off-year sub-sample web surveys in 2003, 2006, 2007, 2009 and 2011.	http://hrsonline.isr.Umich.edu/index.php?p=avail **Details**.

'Details' =further information about this survey contained in Appendix 6.

Cross-sectional sequential mixed mode surveys

Cross-sectional surveys using sequential mixed mode methods (Type MM-1) were slightly more frequent in number than longitudinal examples, and they were substantial survey uses of mixed mode data collection (see Appendix Table A2.1). Twelve examples of sequential mixed mode studies on general populations and six on sub populations were identified.

[5] A smaller study of this type not reaching the 1000 case sample size can be found in D'Agostino et al (2006)

What is notable about the general population studies is that the majority are from The Netherlands. Their policy of moving towards use of mixed mode in all major survey work, has increased the examples of using (or trialing) mixed mode (with web) for a range of general population surveys. Interestingly, the Netherlands also has a methodological programme developing inference for mixed mode surveys, for example, for sequential use of mixed modes (see Beulens and van den Brakel, 2011). Across these studies, the general view is that they have been instructive and positive in their outcomes. Certainly the Netherland experience has not discouraged or caused a rethink on their strategy, to the extent that Eurostat is now advocating a similar approach be adopted for the collection of all countries' contributions to official EU social statistics measures (Bohatá, 2011). Statistics Netherlands now has general overviews of what can be achieved using sequential mixed mode data collection in general population surveys (summarised under Section 3 on quality below). The sequential approach has also even been productive in surveying hard-to-reach immigrant populations in The Netherlands (Kappelhof, 2011). On the other hand, Norwegian experience is less enthusiastic. They had considerable success initially using mixed mode data collection which included the web for business surveys, but these surveys were mandatory, and had financial penalties for failing to return. Subsequent experience with the Rent Market Surveys of 2006 and 2007 (Lagerstrom, 2007) found low web response rates, dominated by men and younger age groups. If web was used as part of a sequential mixed mode approach, the same total response rates could be obtained as were available by other data collection options, but the costs were then also parallel to these other options and not any cheaper (see Lagerstrom, 2011-Appendix 9).

The sub-populations studies identified as using sequential data collection in cross-sectional data collection show a predominance (5 out of 6) of young people, students' or education-related surveys. It is also worth noting that some of these cross-sectional cases (across general and sub populations) employed very short instruments (eg. 7 questions in the *Employer Views* survey; 22 in the *National Students Survey*).[6] This might suggest that sequential mixed mode is more highly suited to high volume but quick and short instrument collection. Fortunately, there is experience, as cited above, with longer instruments being successful using this approach. However, commentators generally consider that some reduction in instrument length is advisable when making an instrument suitable for a sequential mixed mode survey that will need to use two or more of web, phone, paper and F2F modes. There is general recognition that splitting instruments into more modules and running them at different times should be considered. This practice has been seen as beneficial by at least one longitudinal survey as a way of keeping in touch with panel/cohort members between major data collection sweeps.[7]

We found 12 cross-sectional surveys (Type MM-1) applied to general populations using the concurrent mixed mode approach that included a web option (see Appendix Table A2.2). The largest users of concurrent mixed mode for cross-sectional data are national statistics agencies engaged in Census data collection. Canada stands out as having had an extensive programme of research on the use of mixed modes which included web. Census

[6] Details about instruments were not always clear from the documentation in other cases. Since some of the other studies are also student evaluation surveys, it is likely that they too were short instruments.
[7] Communication from Prof Mary Beth Ofstedal (University of Michigan) about the US Health and Retirement Survey's practice of having off- year sub sample surveys, usually web or mail, between the 2-year cycle of mainstage core data collection waves.

data collectors in these countries have also first trialled and many then adopted web as a concurrent option. There are other surveys in this concurrent list, (British Gambling Prevalence, and American Community Survey) who carried out trials using the web for data collection, but who subsequently decided not to continue with it into later surveys. The additional complexity added by the mixed mode, given the low take up of web options, was probably discouraging. Of the four sub populations studies using concurrent mixed mode for cross-sectional data collection, the three that are completed did not have high response rates. Since Census data collection is done on such a large scale, even modest response rates, 15-20 per cent of the population filling in a web questionnaire, may make it worth considering for the financial savings.

A further set of 4 more complex combinations of sequential and concurrent mixed mode approaches, all undertaken through embedded experiments and all using Scandinavian data sets are listed in Appendix Table A2.3.

Conclusions from the examples

Given our interests in longitudinal data and more general populations rather than sub populations, especially if limited to young people or students, the findings were disappointing. This highlights the importance of forthcoming research of the sort planned for the Understanding Society's Innovation Panel Wave 5 (see Appendix 7). On the other hand, we need to weigh carefully the fact that some colleagues and custodians of longitudinal dataset around the world see mixed mode data collection including the web as premature, given the state of knowledge and the risks. As a positive element, the major UK fieldwork agencies all now have familiarity and experience with sequential mixed mode data collection. However, their experiences are heavily weighted to a focus on young people and students, as known web users. Among this sub group, email addresses and mobile phone numbers change at alarmingly high rates. Unfortunately we do not have equivalent data for older age groups but would suspect it to be less of a problem. For experience and evidence about the use of web in data collection for general populations, we need to look to The Netherlands, Scandinavia and Canada. We might consider why Scandinavian countries have been leading the field on the use of web in general population surveys. Their levels of internet coverage may be an important element.

Internet access and awareness

The cross-country figures on the extent of internet and broadband in EU country households are displayed in Table 4. Rates of internet and broadband access in Scandinavian households, the Netherlands included, have been well ahead of many other EU countries. This has probably encouraged them to think they could make a major switch into using web based data collection by the end of the first decade of the 21st century. Their plans to make this shift, even for surveying general populations, were noted in 2009 documents (van der Laan and van Nunspeet, 2009) although some of the preparations and feasibility assessments were being made earlier. Such decisions would have been less viable in the majority of other EU countries, including the UK, at the same point in time, given the web access figures (a conclusion reached earlier by Blyth, 2008a).[8] However, the

[8] It is also worth noting that country-specific sources of data give a different picture than the Eurobarometer figures for the UK; For example IDATE/industry data from OFCOM suggested that household broadband was available in 28% of UK households in 2008 (27% in France, 30% in Sweden; 27% in German; 25% in the

Netherlands, also made early decisions to achieve general population coverage (with random probability samples) using web-based technology by installing it in homes where they did not have the facilities. Even so, it is still necessary to consider access to web facilities by age, as a caveat to these generalisations, since older people can have much lower levels of web access and familiarity, even in countries were access is generally high.[9]

Table 4. Internet and Broadband access by % of households across selected EU countries.

	Internet %				Broadband%			
Country	**2006**	**2007**	**2008**	**2009**	**2006**	**2007**	**2008**	**2009**
UK	63	67	71	77	44	57	62	69
EU27	49	54	60	65	30	42	48	56
Denmark	79	78	82	83	63	70	74	76
Germany	67	71	75	79	34	50	55	65
LuxemB	70	75	80	87	44	58	61	71
Netherlands	80	83	86	90	66	74	74	77
Finland	65	69	72	78	53	63	66	74
Sweden	77	79	84	86	51	67	71	80
Norway	69	78	84	90	57	67	73	87
Iceland	83	84	88	90	72	76	83	87

Source: Eurobarometer: http://europa.eu/rrapid/pressreleaseaction.do?reference=STAT/09/176&gformat=HT

In 2010, the ONS Opinions survey figures for the UK found that 19.2 million households had an internet connection, constituting 73% of UK households, an increase of 0.9 million since 2009 and an increase of 5 million households since 2006. There were 38.3 million individual internet users in the UK in the three months prior to being interviewed, constituting 77 per cent of the UK adult population. Of those, 30.1 million accessed the Internet every day or almost every day, representing 60 per cent of all UK adults. Over 9 million adults had never used the Internet, at 18 per cent, down from 10.2 million in 2009 (ONS, Statistical Bulletin: Internet Access 2010). The number of people who used a mobile device (eg. laptop, mobile phone or palmtop) to access the internet wirelessly, away from the home or the workplace, increased in 2010 with 45 per cent of internet users connecting this way, up from 40 per cent in 2009. The mobile phone was the most popular device used to access the internet wirelessly, away from the home or workplace; just under a third (31%) of internet users connected this way. The adoption of mobile phone technology was found to be highest in the UK among 16 to 24-year-old internet users, with 44 per cent of this age group using a mobile phone to access the internet.

Figures for internet use in the UK also show very large differences by age group. ONS UK figures for 2010 suggest that the majority (60%) of people aged 65 and over had never used the internet compared with only 22 per cent who had not used it in the 55-64 year age group and only one per cent of 16-24 year olds. There were also regional differences, with

USA). Broadband installation per 100 households in 2008 was 67.1% in the UK, 63.7% in the USA and 63.9% in France).

[9] Studies such as Bech and Kristensen (2009) of older Danish adults aged 50-70 years pointed out that access to the internet was over 86% for 40-59 year olds, but only 57% for 60-74 year olds at the time .

internet use being lower in the North East, among those living in rented accommodation, and among owner occupiers, but higher among those with mortgages.

However, while access to and familiarity with the many versions of this technology are increasing apace, these UK figures suggest there will be challenges to face in the UK in trying to secure web survey responses from general adult populations. Reaching older age groups, especially the over 60s and achieving coverage in all regions will pose challenges unless, like the Dutch, a strategy is adopted of providing people with an internet facility and the necessary training when they do not already have these. However, Marcel Das argued that the profile of internet usage in the Netherlands looked much more like that of the UK (and similar countries) when the Netherlands decided to go down the route of widespread use of the web for data collection. This suggests it is more a decision of intent and planning rather than one about whether the facilities are ready.[10]

3. Survey quality

What are the quality issues?
The most common ways quality is now assessed in surveys is via the components of total survey error; coverage error; sampling error, nonresponse error and measurement error.

Web surveys suffer in comparison with other older modes (mail and telephone) in terms of their *coverage error*. Although improving, the population with access (and familiarity) to web based data collection is much lower than coverage for mail and telephone since there are no population sampling frames for web users. The lack of sampling frames also means that it is difficult to draw random probability samples for web users. However, survey agencies and researchers have been trying to move towards developing such samples for web surveys and for membership of web panels.[11] *Sampling errors* related to web surveys are elements of total survey error that are not examined in any great detail in this review. Coverage has been examined in some studies as web, like respondents opting for other modes, have had their characteristics compared with population characteristics. Both coverage and sampling errors are less relevant to continuing to collect data for existing longitudinal datasets where the sample is already well defined.

Non response error is an area where there is a greater and growing amount of information for web in comparison with respondents from mainly mail and telephone modes. This element of unit response is covered in more detail in this review and is one type of potential mode effect; that is, observed differences in survey outcomes if the survey is administered using different data collection modes. It is important to remember that response rates are only one indicator of non response error. Equally important is the representation coverage offered by the survey. In principle lower response rates can go along side unbiased

[10] This point was made at the NCRM seminar (Appendix 9).

[11] See for example Scherpenzeel and Toepoel (2011) who carried out an experimental study to test a number of recruitment methods which were to inform the desire of the LISS (internet) panel to be based on a probability sample. One solution adopted by Knowledge Networks, some commercial internet panels and the Dutch CentERdata, to gain representative coverage for a web panel is to select a probability sample using an address or RDD database and then install the equipment in households/individuals that have not got it already.

nonresponse characteristics; and higher response rates can sit alongside very biased demographic characteristics coverage. These non response errors are sometimes called selection effects.

Lastly *measurement error* and its variation by mode are also covered in this review. The literature examining mode effects in web measurements compared with other modes has been increasing, often through experimental studies. The 2011 ESRA conference had a large number of new studies under this heading, not all completed. We are able to report research on social desirability, elements of satisficing, and about particular types of questions and their formats.

Given that various designs of mixed mode study were outlined in Section 2 above, we can also note the main areas of quality that are affected by such designs (see Table 1 Final column). Mixing modes can also cause measurement differences between the subgroups of people responding by different modes, or for people responding by different modes over time, with implications for the comparability of data. If the mode changes from one wave of a longitudinal study to the next, questions staying constant, this puts at risk the measurement of social change, which is one of the main aims of a longitudinal survey.[12] Once a mode change had been made in a longitudinal data set, were it then to be kept constant, measurement error due to mode of data collection would cease to be a problem for analysing transitions using the post change data.

In additional there may be interactions between response and measurement error according to the mode and mixed modes of data collection used. In the literature, particular attention has been paid to a number of measurement error topics, namely those of social desirability, satisficing and acquiescence in answering survey questions and sensitive questions. Some attention has been paid in this review to these types of measurement errors in the case of some of the mixed mode combinations used in data collection.

Ideally we wish to know about the 'mode effects'. However, identifying so-called 'mode effects' poses technical challenges, and is rarely possible by examining the data collected ex post. The identification of mode effects needs carefully designed experimental frameworks to be adopted from the outset, where a counterfactual group can then be compared with mode treatment group. There is also the issue that there can be confounding of different types of errors. Research designs and research methods have both been advancing to try to identify and separate the different error components (eg. Jäckle et al, 2008; Lugtig, 2011; Vannieuwenhuyze et al, 2010; Cartenssen, 2011) to separate selection effects from measurement error effects due to mode (for example, as part of the European Social Survey program).

Often in surveys, quality can be traded against cost. Lowering the cost, by using cheaper modes of data collection used to be directly reflected in the lower response rates of the survey for most topics. The potential value of mixed mode data collection is that it potentially reverses this relationship, as far as response rates and their representativeness are concerned, as argued by de Leeuw (2005). In a world where it has got harder to get

[12] The problems of changing mode in this way also apply to repeat cross-section surveys where the purpose is to measure time series trends over time.

people to cooperate and respond to surveys, the representativeness of the achieved sample, particularly if based on landline telephones, has been declining – with hard-to-recruit groups declining compared with others. The use of mixed modes to collect the data has been seen as a way not only of boosting the achieved sample size and response rates, especially among the hard-to-reach, but also potentially improving the representative character of the achieved sample. The representation is improved by each mode bringing a biased sub sample of the population who are more likely to use one particular mode. When combined together, these sub samples can give reasonable coverage of the whole population. It is better coverage in comparison with that obtained from the respondents using each mode considered alone.

De Leeuw (2005) argues that the 'optimal' data collection method is that which is "best affordable" to optimise procedures and reduce total survey error within the available budget and time (see Table 5). While mixing modes can compensate for the weaknesses of each mode, there will be a trade off between cost and non-sampling errors.

Table 5. Potential benefits and disadvantages of some alternative mixed mode designs

	Sequential use of mixed modes	**Concurrent use of mixed-modes**	**Use different mode at two separate waves of longitudinal survey**
Response rates	improve	improve	
Coverage error	Can be reduced	Can be reduced	
Non-response error			
Costs	reduce	reduce	reduce
Design costs	increase	increase	increase
Measurement error from mode differences		May be confounded with subgroup differences	Impacts on ability to measure change over time

Source: De Leeuw (2005).

This Review gives consideration to how far the achieved samples have become more representative and some brief consideration to how far the cost savings from adopting certain mixed modes have materialised.

Survey Quality (1). Response, coverage and representativeness.

Roberts (2007) earlier review noted the generalisations about response rates for separate modes, F2F mode producing the highest response rates but at greatest cost, followed by telephone and then post and web, at the lowest end response and cost end. Cook et al (2000) carried out an early meta analysis of the response rates from 68 web surveys and found that they varied by (a) the number of follow-up contacts/reminders; (b) whether there were personalized contacts; (c) whether there were precontacts (eg advanced letters posted); and (d) the salience of the survey. But web survey response rates did not appear to be influenced by the length of the survey or by whether an online password was required to complete the survey. That response rates increased from the number of contacts/reminders for web surveys was a similar finding to reviews of mail surveys, with the proviso that reminders did not go beyond saturation point, or else they became

counterproductive and lowered response rates. Also similar to mail surveys, the finding was that higher response in web surveys was associated with more personalised correspondence being sent to individuals. Precontact was found to be associated with slightly higher response and slightly less variable response in web surveys. Surveys of less salience similarly had only slightly lower response rates. However, both precontact and salience appeared much less important in web than in mail/post surveys. Lastly incentives in web surveys were associated, in Cook et al's analysis, with lower web response rates possibly confounded with and compensating for the respondent burden.

Since these 2000 findings, many studies have advanced our knowledge about web surveys. In part this has reflected the gradually improving state of access and familiarity with web technology. Clearly, over time, higher proportions of the population have become familiar with the web and practiced users (see for US Horrigan and Smith, 2007; Smith, 2010; for the UK and Europe Blyth, 2008; Table 4 above). Some of the differences in the extent to which countries' populations have web access may explain their reticence, or enthusiasm for pressing ahead to use the web in survey data collection. The research base has been building, therefore, upon a changing context and environment over time, and across countries. So it is not surprising if we also find the research findings changing over time.

However, researchers also pointed out that huge efforts were made in the past to understand how to get the best out of the older modes of data collection. Even up to 2004, there had not been time to match this earlier effort with a similar exploration for web modes of data collection (Kaplowitz et al, 2004). So, it was suggested, the early disappointment of low web response rates should not have been surprising. Since then growing numbers of carefully devised experiments have emerged to try to pin down how to optimise response and recruitment to surveys using the web mode. In terms of the numbers of studies, more have emerged on comparisons of the mail/Paper mode with the web than compare telephone and web modes, but there are findings under each of these headings to summarise.

Response rates in Mail/Paper compared with Web

Earlier studies, largely from concurrent mixed mode designs, found that offering, usually by mail, an option to do a web survey which was usually done by mail/paper could lower response rates relative to those obtained from a mail/paper-only survey (Schneider et al, 2005; Brennan, 2005; Brogger et al, 2007; Griffin et al, 2001; Tourkin et al, 2005; Werner, 2005; see Appendix 3). This finding applied to general population groups, but could also apply to particular sub-groups regarded as internet-savvy (see Hayslett and Wildemuth, 2004 on librarians; Kaplowitz et al, 2004 and Miller and Dillman on US university undergraduates). A study of older adults in Denmark (age 50-70 years) had a web response rate of 17% compared to a mail response rate of 42% (Bech and Kristensen, 2009). The web rate is low compared with the range of web response rates in Appendix 3, but the mail response rate is on a par with those other rates, even though it comes from an elderly population.

There has been some exploration of the reasons why people are attracted by paper but not by web questionnaires (see Millar and Dillman, 2011 for a review). Suggestions include the idea that paper questionnaires can be very enticing to start to look at, and handle. They

also enable potential respondents to get an understanding of the task they are being asked to do – in order to decide whether to go ahead or not. It is not possible to make the same judgements about a web instrument that usually unfolds one question at a time on the screen.

Efforts to experiment with contact strategies, offering varying degrees of information about alternatives, providing incentives and the use of sequential methods have identified some common elements in how to boost web based response rates in general population studies from the early low levels. However, on the whole, they can only be boosted to approach, and not usually reach, response rates from mail-only surveys. However, that still implies the potential for cost savings.

A meta analysis of standardised response rates from 43 cross sectional surveys carrying out mail and web comparisons dating from 1996 to 2006 (Shih and Fan, 2007) explored the factors associated with levels of response rates along six dimensions; namely, the option design, the method of notification and delivery; the mode of option delivery; the population type; whether incentives were used; and the number of follow-up reminders. Some of the findings from this meta analysis obtained from a multivariate analysis of the studies are integrated into the summary of findings below. Studies using the web for surveying general populations, other things equal, had significantly lower web response rates than those surveying particular sub-populations.

Smyth et al's (2010) use of a sequential approach of web first, followed by mail for non-responders put up the share of web responses to 75% but the overall response rate of 55% was lower than the sequential mail leading to web option which gave a 71% response rate, although mostly collected by mail returns. Werner (2005), Holmberg and Lorenc (2008) and Holmberg et al (2009) in a set of studies offering concurrent mode options in surveys of general Swedish populations found that a more web-intensive approach, holding back details of the mail option boosted web response rates. The meta analysis of Shih and Fan (2007) support these conclusions; more web-intensive approaches (less or no mention of alternative modes when invitations first sent out to do the survey) helped to boost web responses; sequential methods led to higher web responses; there was a positive effect of reminders, responses increasing with their number.

Incentives

The use of prepaid cash incentives largely helped to increase response rates in the above experimental studies, but to varying degrees. Goritz's review (2006) found the effects were relatively modest, 2-4 per cent increases attributable to incentives. In the case of Messer and Dillman (2011) the additional effect on the response rate could be as much as a 17 percentage point increase and it was also high in an experiment reported in Laurie and Lynn (2009). The meta analysis of Shih and Fan (2007) found that, other things equal, an incentive was not associated with a significantly higher response rate. However, the measure of 'incentive' used by Shih and Fan was rather crude; some incentive versus no incentive. Other studies have found positive effects of prepaid and cash incentives which Shih and Fan did not distinguish. There is also the caveat to note that contacting respondents by email makes it more difficult to be convincing about the offer of a pre-paid unconditional incentive, as compared with a contact by post (where it can clearly be included).

Routes of contact

The issue of how to contact respondents for the best response has been examined. Hayslett and Wildemuth (2004) reviewed the studies up to 2004 on this topic. They found for studies where mail or email were possible for contacting respondents, evidence that email contact produced better response in some cases, but, in others, email contacts produced lower response rates, but only slightly lower. The experience of the fieldwork agent collecting LSYPE data for young people found response was higher for those with valid emails (see Appendix 6). For general populations, the approach via an advance mailed letter was shown to improve response – even web response (Smyth et al, 2010), compared with an email sent when the subpopulation's emails were known (Messer and Dillman, 2011; Kaplowitz et al, 2004). The meta analysis of Shih and Fan (2007) found, to the contrary, that, other things equal, the sending of email notifications was associated with higher response rates. Miller and Dillman's (2011) experimental work on US college students found that there was better response if contact was made by multiple postal and email contacts. Of course, use of invitations to participate and /or reminders sent by mail all add to the cost of surveys and then offset some of the cost savings of using the web mode. These postal contacts or reminders also tend to increase the overall timing of fieldwork, again detracting from the beneficial elements of the web mode.

Response time has also been examined in a few studies. Hayslett and Wildemuth (2004) reviewed the earlier studies which concluded, as did their own research, that those who opted to do a web survey were generally quick to respond and took a shorter time than mail/paper responders (Akl et al, 2005). The majority of web respondents who used the web option did so within 2 weeks of receiving the request (Hayslett and Wildemuth, 2004). This has been a continuing finding.

Coverage and representativeness mail/web comparisons

The characteristics of mail and web survey respondents have been noted in a wide range of studies. Large-scale population comparisons report differences in terms of age, education, marital status, and labour force activity status. For example, the Canadian Census, in its various methodological testing work, has found that web respondents had lower response rates than mail/paper respondents in all sub groups defined by demographic characteristics (Grondin and Sun, 2006). Roy and Laroche (2006) in the 2004 Canadian Census test found web respondents compared with mail/post were higher in the youngest age group, likely to be in jobs, and have a university degree. Gender has often been found not to be a differentiating characteristic of web or mail respondents in the above or other studies across a range of countries. However, some studies have found more men responding in web surveys than in other modes and this has helped the gender representation in mixed mode surveys (eg. Bech and Kristensen, 2009; Couper, 2000).

Studies have found that even for national panels of households, both web and mail respondents were unrepresentative of the US population (compared with Census figures) on at least some characteristics (Rookey et al, 2008). Mixed mode surveys of people living in small rural towns or communities in the US found those who completed web questionnaires were of younger mean age, more highly educated, more likely married, with children, employed full time and with higher incomes, and more mobile and likely to own a cell phone than mail respondents. A New Zealand general population sample found that

web respondents had higher incomes, were more likely to own a cell phone but had lower mean age than mail respondents (Brennan, 2005). Where age breakdowns are given web respondents are usually in much lower frequencies than mail respondents in the oldest population age groups (eg. 60 and over). However, some studies (eg. some Canadian Census tests) reported web respondents as 'older' than mail respondents when the expectation is that younger people are more proficient in web use. Ashton and Dunn's (2009) web pilot work in the UK found that the youngest age group (16-24) appeared not to be as interested as older age groups in doing web questionnaires.. However, the web pilot respondents were, as expected, slightly younger overall than face-to-face respondents. Possibly the advanced letter to the household may have (unintentionally) excluded the youngest age group, or the (lack of) salience of the survey topic made a difference to the youngest age group.

Response rates in telephone compared with web

One of the initial motivations to develop web-based data collection was the increasing difficulty that telephone interviewing was facing in keeping up high response rates in this particular mode. The proliferation of unwanted cold calling, caller ID barriers on household landline telephones and increasing numbers of surveys were all felt to be increasing the number of calls needed to gain an interview and putting up the cost per interview significantly. Similarly telephone coverage problems have been increasing partly through people switching to mobile telephones (without national sampling frames) and giving up landlines (which had near complete national coverage). Mail surveys appear not to have been affected by declining response and coverage rates. When one looks across a selection of unit response rates in studies using telephone, mostly from experimental studies comparing web and telephone modes, the telephone response rates range widely from in the high 90s per cents to middle range 45 per cents (see Appendix 4) in most cases. Certainly, like mail/Paper response rates, they are usually higher, some considerably higher, than web response rates.

Coverage and representativeness telephone/web comparisons

National telephone sampling lists have in the recent past been providers of high quality representative samples for surveys and coverage of general populations. Statistics on demographic characteristics emerging from smaller scale sample surveys, or experimental studies, may not be as closely representative of the general population. However, it is common for studies to assess themselves on their coverage, against a large-scale national sample survey. Web studies most often have two stages to undergo in this assessment for mail-web comparisons; first they have to establish that people are able to fill in a web questionnaire, through having the necessary facilities and knowledge; and second that they are willing to cooperate. Non-response will feature at both stages. The resulting profiles of demographic characteristics which studies report all suggest non-response is not random. It is perhaps not surprising when studies, such as Fricker et al (2005), found that neither the RDD telephone sample nor the web panel sample closely represented the US adult population. Lugtig et al (2011) found that their CATI sample were older, less often employed, more often single, female, living in urban areas, and on lower incomes than web users. Braunsberger et al (2007) found that the characteristics of their web panel were more representative than the RDD telephone sample whereas Chang (2001) and Chang and Krosnick (2003) found that their RDD and Knowledge Network samples deviated only slightly from the national US demographic profile; their Harris Interactive (online panel)

sample, however, deviated a lot more. Both web samples had higher incomes, were more educated and slightly older than the general (US) population.

Sequential mixed modes, response rates and representation

Some more recent experimental studies have started to examine the effects on response rates and coverage or representation of different types of sequential mode strategies (which include web as one mode), in comparison data collected by concurrent mixed mode or single mode approaches; (for example Dillman et al, 2009 surveying registered consumers; Messer and Dillman, 2011, surveying households; Millar and Dillman, 2011, for highly literate internet users; Manfreda et al, 2001 surveying primary and secondary schools). They have generally found that mixing modes has improved response rates compared with using the web as the sole mode. But, at best, response rates only reaches the best mail response rates. The sequential mixed mode approach has given better representation for general populations, but that the measurement error issues are also real and cannot be ignored. Martin and Lynn (2011) report the results of an experiment comparing the European Social Survey (ESS) F2F survey with two mixed mode approaches, one sequential web, telephone and F2F, and one concurrent using the same three modes. Both mixed mode designs achieved significantly lower response rates than the single F2F mode.

However, the change of policy of the Statistics Netherlands to use mixed mode data collection systematically (van der Laan and Nunspeet, 2009) and embed experiments for learning has produced considerable amounts of information about results from sequential methods. Their many surveys and experiments on general population samples have led them to a number of initial conclusions: An approach to a survey of individuals, adopting CAWI as a first stage, possibly including the option for respondents to send off for a PAPI mode questionnaire, can achieve a 25-40 per cent response depending on the type of survey. Re-approaching non-respondents to do a CATI mode can add another 25-30 per cent. Adding CAPI as a final option can add a further 10-15 per cent of additional response. This 3-stage sequence can, therefore, achieve a response rate range from 60 to 85 per cent. These are very encouraging findings for this approach.

Some of the studies reported in Table 2 (and Appendix 2 Tables with some further details in Appendix 6) also show that it is possible, using sequential mixed modes, that start with the cheapest web mode, to maintain high response rates. The surveys where we were able to obtain sufficient data to inspect were those based on students or young people as respondents. In the cases of LSYPE and NSS, the fieldwork agencies were set expected target response rates by funders. Clearly, given a sequential design it is possible to put effort in at the last (and most likely for success) stage to meet the target, although at greater cost per marginal unit response. However, the fieldwork agencies we spoke to also reported experiencing a learning process in the management of sequential surveys. The trend increase over time in (first stage) web response rates reported in Appendix 6 supports their claims. Cleaning of emails was found to be necessary up front, even where provided by higher education institutions. They also indicated they had learnt how to manage sequential surveys more effectively and efficiently over time, knowing when, how many, and via what routes to send reminders that would be most likely to produce positive responses.

It is of interest to know whether respondents in sequential mixed mode surveys always respond in the same mode. We found two longitudinal data sets that had some data on this

topic (LSYPE and a Norwegian Rent Market Survey Panel and Rent Price survey (Lagerstrom, 2011-Appendix 9). The 2011 Norwegian survey, based on the general population, found that mode was very stable across two panel waves; 90 per cent of telephone respondents at one wave responded by telephone again at the next wave; 70 per cent of web responders were likely to respond using the web at the subsequent wave. LSPYE longitudinal survey of young people displayed less stable mode choices across 3 annual waves of data. Only 4% of the sample who completed the 3 mixed mode waves used all three modes but only half (49%) stuck to the same choice of mode through all 3 waves; 26% of the sample used the web and 21% used the telephone at all 3 waves.

Response rates in web panels compared with other modes

Web panels have gained agreement from their participants to engage in regular surveys, in exchange for payment. One could liken this commitment to attachments some individuals have to a longitudinal survey. However, the response rates to web panels are not totally reassuring in this respect. Survey completion rates for panel members, as cited in the literature range from 5-70 per cent in one case with an average of 15-20% (see Chang and Krosnick, 2009) to an average of 85% in another. The gap between agreeing to participate and actually delivering a completed survey also allows for coverage and representation to move away from national indicators as others have noted (Sparrow and Curtice, 2004; Smith, 2003, Krosnick and Chang, 2001; Chang and Krosnick, 2009). Web panels have been found to be biased towards younger, higher household incomes and those with degrees in the USA (Rookey et al, 2008) and in the UK to owning one's own house, higher social class, car ownership and being employed (Sparrow and Curtice, 2004). These demographic differences have been found to be problematic for the representativeness of results on specific types of questions, namely attitudes, political views and political behaviour (Smith and Dennis, 2008; Sparrow and Curtice, 2004; Rookey et al, 2008). A web plus mail survey from a web panel list gave better coverage and representation than web alone in a study (Rookey et al, 2008).

Web panels have also allowed for some investigation of panel conditioning, another element of concern for longitudinal studies[13]. In principle, conditioning may make the panel less and less like the public they are intended to represent as they accumulate experience in doing surveys. Chang and Krosnick's (2009) brief review of evidence on this topic suggests that there are only small or no effects of panel conditioning. However, there were slightly more effects noted for particular measurements, as indicated below. The NES internet panel built in design features to minimise panel conditioning (DeBell et al, 2010).

Longitudinal attrition rates.

Since we are very interested in uses of mixed modes of data collection for longitudinal data, we tried to find examples of the attrition effects of using different modes in an earlier wave of data collection. There were limited examples that were possible to follow up, but in no cases were we able to find any examples of analyses which have already looked at this issue (given our specification to include a web mode).[14] Data are available to analyse this

[13] For a brief review of panel conditioning effects see Burton et al (2010).

[14] One communication from the HRS team said they had given some examination, albeit without any "clean control groups", to the effect on their core survey wave responses of the off-year surveys on sub samples which used either mail or web modes. Mary Beth Ofstedal noted that there was some evidence that being asked to participate in the CAMS sub-sample survey (a detailed survey about household expenditures by mail

question for some surveys, but we were not able to start to do this within the limits of this review paper. Clearly, this is an issue that is of interest to collectors of longitudinal data and is a gap waiting to be given attention.

Drop out has been investigated at the point of recruitment to a survey, according to the method used to recruit respondents to take part. Although this research has not focussed on longitudinal surveys it may have implications for them. Sakshaug et al (2010) review these studies and point out the relatively high loss of cases when an interviewer is used at the recruitment stage, but respondents are then expected to carry on to complete a self-administered instrument (typically 20% of those who agree to carry on do not make the transition to do the survey). The switch leads to increases in nonresponse error which can outweigh any improvements in measurement error from using self-administered instruments to collect sensitive questions.

Summary Survey Quality (1)

We can note the following conclusions about the mode effects of web surveys on unit response. Much of this work has been devoted to comparing web and mail/paper surveys. Shin et al's (2011) review, focussing on references since approximately 2000, suggested that the findings of studies about mode effects on unit response rates were sometimes inconsistent and varied. They were found to range from zero to substantial mode effects, according to the target population, survey samples and administrative designs. Web surveys generally produced lower unit response rates compared to mail surveys. The few exceptions were when studies of academics or professionals were the target group, when higher unit response rates were then reported for web than for mail, although not in all cases. However, students did not necessarily produce higher web than mail unit response rates.

Shin et al (2011) also noted that the face-to-face mode consistently produced the highest response rates and also the most equal cooperation across demographic groups. Telephone samples were noted as having fewer respondents with low education, low income, who were older or were from minority groups. Postal/mail and web modes produced better response rates from educated, more literate and computer literate individuals but were lower than those obtained from face-to-face or telephone modes.

On the socio-demographic profile of unit response rates, web surveys have been found to produce lower response rates than the mail survey mode in general population surveys across most demographic sub groups- namely those of gender, race, education level, and income groups (Bech and Kristensen, 2009; Couper et al, 1999; Kaplowitz et al, 2004; Kwak and Radler, 2002; Link and Mokdad, 2006; Miller et al, 2002; Shin et al, 2011).

Should we worry about the non-representativeness of web samples in particular? Are these demographic differences in the samples likely to lead to mode effects in the findings? Adjustments of weighting, by using regression models, or by matching using modelled

---

in 2007) had a slight negative impact on participation in the subsequent core wave. On the other hand, the diabetes (by mail in 2003) and early off-year sub-sample internet surveys appeared to have a slight beneficial effect on participation in the next core interview. However, all these effects were all quite small. One benefit of doing off-year surveys the team noted was that it helped to reduce the amount of tracking they needed to do for the core data collection waves.

propensity scores mostly reduce or eliminate the mode effects from modes having their own demographic biases in the characteristics of respondents (Dever, 2008; Schonlau et al, 2003; Shin et al, 2011). However, on some particular questions, usually linked to the motivation of the survey itself, mode effects from response biases arising from mode differences remain.[15]

From the point of view of longitudinal surveys, mail/Paper surveys, despite advantages of lower cost than interviewer-based methods, have a further limitation. They do not offer the facility to use information given by respondents either at an earlier contact wave, or earlier in the same questionnaire, such as feedforward or dependent interviewing. These are all possible in a web data collection environment but not in self-administered paper/mail instruments. They allow for complex routing to be used to avoid unnecessary data collection and reduce respondent burden.

Survey Quality (2). Measurement differences across survey modes

One of the well-established and fundamental mode differences in findings from mode experiments is that significant differences arise in answers given across the divide between questions that have aural (telephone, IVR) compared with visual components (Post/Paper; Web, and CAPI Show cards). De Leeuw (2005) pointed this out. Studies have found that aural/ telephone respondents give more positive extreme answers to opinions than do respondents to mail surveys (Tarnai and Dillman, 1992; Krysan et al, 1994; Christian et al, 2007) and more positive ratings in general (Christian et al, 2007). Web surveys have been shown to be like mail surveys in these respects (Christian et al, 2007).

Studies have looked in detail at the use of scales of various kinds (for example, 11 point scales; 7-point scales; 5-point scales; with varying extents of numbering - fully numbered scales and polar point only labelled scales). Mode effects have been found in the labelling of scales such that respondents answer fully labelled scales (common in visually presented questions) differently from scales labelled at polar endpoints only (more common in aurally presented survey questions). Telephone respondents have been found to give significantly more positive answers than web respondents on scales (Christian et al, 2007) and hence appear to be expressing greater amounts of satisfaction. Surridge's (2006, 2008) multivariate analyses of the sequential mixed mode annual National Students' Surveys (NSS) in England found students who responded by telephone (compared with web) appeared more satisfied with their course/university, other things equal, in every year of data analysed. The mode and format of scales independently influences responses to scale questions, although Christian et al (2007) did not find evidence of interaction effects from the mode and format. This suggests that combining survey responses across modes where the scales have been presented in different modes, visual/aural, is not advisable.

[15] Seeming mode effects remaining after weighting or other adjustments have been found for questions on the following topics: political activism, vote choice and voting intentions; attitudes towards the level of support for the euro, health, increasing National Insurance contributions and environmental issues and values; knowledge-based questions about cholesterol; health behaviours of binge drinking, current smoking and sexually transmitted diseases; decisions about health care needs; and rating one's own health. In some cases researchers have commented that the apparent effects could be confounded with social desirability effects and unable to be unambiguously identified as mode effects.

Social desirability.
Survey respondents with an interviewer have long been known to be more likely to give socially desirable responses than those which are self-administered (De Leeuw, 2005). The expectation for mail surveys and web surveys has been, therefore, that they will produce lower pressure to give socially desirable responses than telephone or face-to-face modes. Little, if any attention has been given in the literature to examining any potential social desirability differences between mail and web surveys therefore. Studies comparing web and telephone modes have largely confirmed this expectation that the telephone mode will display greater social desirability effects, across a range of topics (Lugtig et al, 2011; Chang, 2001; Chang and Krosnick, 2003; Kreuter et al, 2009; Hope et al, 2011-Appendix 8 for sensitive questions).

In a comparison of web and face-to-face surveys (Heerwegh, 2009) the F2F questions showed signs of social desirability, but only in one half of the 36 items. This study, based on students, suggested social desirability may not necessarily be as large or pervasive as is sometimes thought. Duffy et al's (2005) comparison of a F2F with an online panel also found evidence of social desirability in the F2F survey.

Satisficing
Satisficing can be displayed in a number of ways; through tending to agree with assertions because this involves less effort than disagreeing; providing no opinion or Don't know answers; rating a series of items to the same scale point known as non-differentiation; selecting response categories at random; choosing extreme points on a scale, especially attitude scales; or allowing the order response categories are presented to affect the choice of answer. Studies, which include the web mode have been increasing on many of these satisficing topics. Betts and Lound (2010) have a review of satisficing. The starting point was that one would expect satisficing to be greater (1) in self-administered (including the web) than in interviewer administered modes. (2) satisficing is also expected to be greater in telephone than in face-to-face modes and (3) greater in web than in postal modes. As in other self-administered modes, doing surveys by the web allows respondents to multi-task, and the lack of interviewer increases the cognitive burden for respondents. However, the evidence as reviewed by Betts and Lound only weakly supported the expectation that acquiescence was higher in web compared with other modes. Recent SDMI research examined the relationship between mode and satisficing (Hope et al, 2011- Appendix 8). Rather than less acquiescence in modes with interviewers, compared with self-administered modes, they found more acquiescence. Sakshaug et al (2010) show that there can also be interactions and trade offs between satisficing (in so far as it affects an individual's likelihood of agreeing to do the survey and answering its questions) and non response and measurement errors. Putting effort into refusal conversion for reluctant respondents can reduce nonresponse error, for example, but at the expense of increasing measurement error as these respondents put less effort into answering the questions correctly. A brief review of evidence on some other particular satisficing sub-topics is presented below.

Don't Knows (DKs), unsure and accuracy
The web mode generally has been found to be less likely than the telephone mode to produce DK answers, or leave items blank (Fricker et al, 2005) and also less likely than mail surveys to obtain DK answers (Bech and Kristensen, 2009). Much of the quality

difference is likely to be related to many web instruments insisting on having its items filled in with potentially valid answers before being allowed to complete the survey. Fricker et al (2005) also found that answers on the web were better quality than those on the telephone when the task was cognitively more demanding. This type of interaction between mode type and question type was also noted in Schonlau et al (2003). However, one caveat is that forcing respondents to place valid answers in web questionnaires, or refusing to allow no answer, can produce higher reporting of incorrect answers; Grondin and Sun (2006) found evidence of this in the tests for the 2006 Canadian Census as did Best and Krueger (2004) earlier.

Comparisons between the same questions in a face-to-face and web survey found 9 per cent DKs for the F2F and 6 per cent for the web survey given to a set of university students in Belgium (Heerwegh, 2009). Another comparison of F2F with a web panel found a higher number of DKs and 'Not sure' in the web panel than in the F2F survey (Duffy et al, 2005). The authors raised the possibility that the web results were due to the lack of acquiescence when there was no interviewer, in comparison with F2F.

Evidence from web panels on the accuracy of reporting and the possibility of panel conditioning, from having regular experience of doing surveys has reported some increased reporting accuracy among web panel members (Shin et al, 2011; Chang and Krosnick, 2009) when using web compared with another mode; for example, lower DKs in the web than in the mail survey, as well as lower non-response to closed ended and open-ended questions in the web survey. In addition, deciding to do only surveys on topics of personal interest, as a member of a web panel, led to increases in reporting accuracy in a non-probability web panel sample (Chang and Krosnick, 2009). In contrast, an RDD sample displayed more random measurement error, more survey satisficing and more social desirability bias than the non-probability web panel sample data; the probability web panel sample displayed more random error and more satisficing than the non-probability web panel sample (Chang and Krosnick, 2009).

Several studies have been able to check the accuracy of data provided by different modes, from sources that validate their data. Self-reports were more accurate when provided via the web than by telephone (Chang and Krosnick, 2008) by the web than CATI or IVR (Kreuter et al, 2008), and by the web than mail (Laroche and Grondin, 2008).

Item non-response

The 2006 Canadian Census found that use of the web led to lower item non-response and lower invalid response rates compared with mail/paper (Grondin and Sun, 2009). Similar positive effects on web survey item response quality were found in studies by Bech and Kristensen (2009), Denscombe (2006, 2008, 2009), Kwak and Radler (2002), Schaefer and Dillman (1998), and Shin et al (2011).

However, it should be noted that conclusions drawn from Shin et al's (2011) review of mainly web/mail comparisons of mode effects on item non-response focussing on references since approximately 2000, were that the findings of studies were inconsistent and varied in *item* response rates. This was the same conclusion they reached about their review of unit response. Mode effects on item response were found to vary between a zero and a substantial effect, depending upon the target population, survey samples and

administrative designs. Some studies reported higher item response rates to web than mail surveys and some more extended open ended responses in the web mode. Shin et al noted that most of the testing of web mode surveys had been done on computer savvy populations with only a few (eg. Link and Mokdad (2005, 2006; Lorenc, 2010; Bandillar et al, 2003; Bech and Kristensen, 2009; Shin et al, 2011) examining general populations.

Shin et al's study of the national Gallup (internet) panel survey respondents found that the web survey, compared to a mail survey on this group, produced lower item non-response to both closed and open-ended questions and fewer 'Don't Know' to closed-ended and minimal difference in open-ended questions. Any differences disappeared with adjustment for demographic differences in respondents' characteristics.

Comparisons between the same questions in a face-to-face and web survey found equal use of the response option by the two modes, but 8.5% non-response to items in the web survey compared to 0.2% in the F2F for a set of university students in Belgium (Heerwegh, 2009). Revilla and Saris (2010) found little difference between the single item and composite scores in a web panel compared with a F2F European Social Survey.

Recent SDMI research examined hypotheses about the amount of non response in ranking questions and end labelled scales using different modes which imply varying degrees of help being given to the respondents (Hope et al, 2011, Appendix 8). Hypotheses that more help (for example from an interviewer compared with a self-administered web mode) would produce lower amounts of item nonresponse were not supported.

Non-differentiation.
Recent SDMI research (Hope et al, 2011, Appendix 8) examined one hypothesis about non-differentiation; non-differentiation was expected to be less in F2F than in telephone or web modes. The research supported this hypothesis.

Primacy and recency effects
Primacy effects occur if respondents, given a list of items to choose from, pick those occurring first in the list. This is a mode effect most often associated with visual survey modes. If respondents are more likely to choose items listed last in a list, this is the recency effect and is most often associated with aural survey modes. Dillman et al's (2009) summary review of the literature on these two effects concludes that despite at least 82 experiments placed in different surveys to explore these effects, no consistent patterns of effects have been noted. Recent SDMI research (Hope et al, 20011-Appendix 8) tested the hypotheses that there would be less primacy in F2F with Showcards than in using the web mode, and less recency in F2F, without showcards, than in telephone and web modes. Both of these hypotheses were not supported.

Visual-aural differences for satisficing
SDMI research (Lynn et al 2011-Appendix 8) tested out the hypothesis that visual rather than aural presentation will tend to decrease the propensity for the respondent to take short cuts in response due to satisficing. Support was found for this hypothesis. The additional hypotheses that this effect was mediated by the cognitive ability of the respondent and by the characteristics of the question also gained some support from the analyses but not complete support.

Other mode effects.
Some attention has also been given to other potential mode effects, although the literature is relatively sparse on these topics; layout by mode has been examined by Dillman et al (2009) and Dillman (2009), order effects by Dillman et al (2009) and format and design by Couper et al, (2001).

Data collected have also been analysed to test for mode effects in particular question topics and constructs; for example, the psychometric properties of a 20-item scale identifying qualities of leadership, where no web/mail mode differences were found (Cole et al, 2006); psychosocial measures (loneliness, control and social support) in the off-year web collection of the US Health and Retirement Study (HRS) where mode effects were found (Weir, 2004); psychometric and psychosocial constructs related to smoking where no differences were found between data collected by web or telephone (Graham et al, 2006); the collection of qualitative data to measure 'corporate image', where negligible differences between web, telephone or postal surveys were found (Coderre et al, 2004); and the structural properties of 18 scales (39 items) used in Hong Kong university student course evaluations, where no web/paper mode effects were found (Leung and Kember, 2005).

It is also important to note that differences between modes on particular items can be wrongly attributed to mode effects when questionnaires were altered in order to carry out the collection in a different mode. Such apparent mode effects have been noted in data on household ownership of financial accounts and assets (Van Soest and Kapteyn, 2009) and considered most likely to be due to context and wording effects.

Which types of questions are best?
The initial response to learning about mode effects in particular questions was to adopt a unimode strategy to question construction in mixed mode surveys attributed to Dillman with 9 principles as set out in Appendix 10). Alongside removing any mode-specific elements in questions, the time allocated to completing survey instruments was also often reduced. For example, participants to a telephone survey were thought to have limits (eg 20 mins) to their attention span for answering questions by 'phone. This was then regarded as the maximum for all modes if telephone was included as one of the mixed modes.

Dillman (2009) and Dillman and Christian (2005) have since argued for a more nuanced mode-sensitive approach to question construction, taking into account the stability or instability of questions and the potential benefits of particular modes for collecting specific data. Stable information is the sort individuals carry around as obvious answers and are willing to reveal to others, for example, their age, gender, or housing tenure. Unimode questions will do well on these indicators. On the other hand, instability in questions should be expected where there are issues of social desirability, acquiescence, recency or primacy effects, and some order/context, layout or category/code effects. Devising unimode questions here will be more challenging. Particular challenges to the unimode approach come from use of scales in surveys and, to a lesser extent, handling DKs. Dillman (2009) has also noted that accuracy can be improved for questions about dates by use of mode-specific variations. These are aimed at allowing questions to vary across modes, but in such a way that they remain cognitively equivalent to each other. Recent research by De Leeuw et al (2011) has shown that it is possible to devise probes that help reduce the

number of non-informative answers. These can be used, not only in web surveys, but also in telephone surveys, to make them more equivalent, albeit with differences remaining.

*Forced-choice or Check all formats***.**
Developments in web surveys have produced formats that are particularly suited to this mode. One such format is the 'check all that apply' question. Respondents are asked to mark all that apply to them from among a list of options and multiple items can be selected. Telephone use of 'check-all-that-apply' is considered awkward and is often replaced by 'forced-choice' format questions. Here respondents provide a Yes/No answer for each item on the list. The forced-choice format was thought likely to lead to errors if used in a web interview, with respondents continuing to tick the same 'Yes' (or 'No') down the whole list of options. Studies have investigated whether these mode-specific questions are equivalent and the results cast doubt on this assumption. Forced-choice formats have been shown to produce more options being chosen than the check-all-that-apply format and more options being marked positively. Paradata also showed that respondents spent longer in answering the forced-choice format question (Smyth et al, 2006). Respondents who fill in check-all-that-apply questions are argued to employ a weak satisficing strategy making the forced choice format preferred to the check-all format. Recent work by the SDMI project found support for the hypotheses; that there would be a higher percentage of items chosen in the telephone mode with the Yes/No format than in the F2F in web modes using 'Mark all that apply'; that the difference between modes will be greater for more difficult questions; and that a deeper level of processing is carried out by the respondent when using the 'Yes/No' than the 'Mark All' format (Nicolaas et al, 2011-Appendix 8). The SDMI team concluded, therefore, that it is not possible to consider the Yes/No format as functionally equivalent across the three modes of CAPI, CATI and CAWI (Nicolaas et al. (5), 2011-Appendix 8).

Unimode or another?
There are now, in principle, at least 4 approaches to devising questions for mixed mode surveys: (1) the unimode approach; (2) the mode-specific construction (modifying the question structure, wording or presentation to use the particular capabilities of each mode); (3) mode enhancement (use the features not available in all modes to improve quality in one particular mode); and (4) generalised design (purposely designing questions differently in different modes aimed at being cognitively equivalent).

Attempts have been made to provide principles for making the choice between the unimode and another approach, on particular survey items (see US Census Bureau web site, Martin et al, 2007). Our investigations with UK fieldwork agencies, as part of this Paper, found that unimode questions for surveys that involve sequential mixed mode is still the main approach practised. Very detailed research, question by question, has also been taking place by a UK SDMI research team on these issues (Gray et al, 2011; and Nicolaas et al (6), 2011-Appendix 8). However, the principles of when and where to depart from the unimode route are not always entirely clear and this is still an area under development. The SDMI team concluded that the three approaches, other than unimode, were not sufficiently supported with evidence to be likely to offer practical alternatives to 'unimode' in the near future (Nicolaas et al (6) 2011-Appendix 8). The SDMI team also concluded that the following formats are not portable across modes; these are the Yes/No format, branching; agree/disagree; end-labelled scales; ranking questions; and rating questions. It would seem that there is a long hard road ahead, checking items types one by one, before a more

complete picture can be built up about mode effects or construct equivalent mode-specific constructions for particular question types.

Coding.
The literature has recently begun to explore the issue of coding for information provided via self-administered modes compared with interviewer surveys. Items that usually involve post fieldwork coding are the occupation and industry classifications of a person's job. Tijdens (2011) discusses the 3 different options when in web mode for collecting a person's job details (open response format with office coding; closed response format with a search tree; or open plus closed response format with text string matching). The latter option of text string matching is only starting to be considered.[16] Tijdens (2011) presents results from offering web respondents a search tree linked to an occupational database, using similar data for 3 countries (UK, Belgium and Netherlands). This study found fairly high drop-out rates during the self-coding process using the search tree. Drop-out rates at the various stages were linked to whether the respondents had a job title (adequacy of the question), the survey burden involved in finding the occupation in a search tree, and respondents' cognitive capacities in finding the right search path.

Dawe and Wilson (2011, Appendix 9) reported on respondent self-coding of their job's industry and occupation classification in a pilot of using the web for collecting the UK's Labour Force Survey. Web respondents were asked to enter free text descriptions for the industry and occupation of their employment. These details were subsequently coded using automated procedures and by experienced occupation coders and compared with codes allocated through interviewer-collected data for these individuals. Discrepancies were found between the interviewer allocated codes and the post web allocated codes of up to 50% difference for industry (SIC) and 60% for occupation (SOC) codes. On the basis of these findings, the pilot was not sure whether to advocate collecting, via the web, industry/occupation using free text and coding in-house, or including frames for respondents to use for themselves (eg search trees as used by Tijdens).

Survey costs and operations
Quality is to some extent dependent on survey costs. It is clear, however, that one of the main motivations for using newer mixed mode designs, particularly the sequential mixed modes that start out with the cheaper web mode, is to reduce survey costs. The review, above, of what happens to unit response using sequential mixed modes is reasonably reassuring. It seems possible to maintain respectable, even very good response rates using sequential mixed mode designs that start with the web mode. Experience in the UK is largely restricted to surveys of young people, but does include some longitudinal data sets. But surveys using sequential mixed modes in the Netherlands have more experience with general population samples, but have been largely cross-sectional. The coverage biases need to be taken into consideration, namely that older people and the less educated are less likely to use the cheaper web mode. Also, measurement error issues and problems may be difficult to eliminate. The adoption of a sequential mixed mode approach where CAPI was one of the final stages means that the location of potential participants for such

[16] There is currently a proposal by Eurostat and Tijdens under consideration to develop a tool for an EU-wide measurement /classification of occupations in web surveys using string matching to provide ISCO codes.

interviews is not predicable at the outset of the survey. Interviewers' travel and time costs of reaching these respondents are also uncertain, therefore.[17]

The actual costs and savings of field work are mentioned in some papers which all confirm lower costs when using web and sequential modes that include web as the first stage (Nunspeet et al, 2011; INSEE, 2011; Beukenhorst and Wetzels, 2009; Liefbroer, 2011-Appendix 9). Cost savings are not necessarily reduced if the web is used as a later option in the sequence (Janssen, 2006). Fieldwork agencies we spoke to in the UK about the cost issue all confirmed that the average cost per unit respondent was considerably lower using sequential mixed mode that starts off with the web mode, compared with many traditional single mode (CAPI or CATI) surveys. One caveat noted, however, was that the cost savings that could be made applied to surveys of individuals, but could not necessarily be carried over to household surveys. Costings for a household survey, where more than one member of the household was meant to have an interview were argued to be more difficult to draw up under sequential mixed mode. It might be that one person in the household used the web element, but another household member did not, so that a telephone call or even the most expensive visit to the household was required in the end. This potential for a mixture of data collection modes within any one household was expected to eliminate a substantial element of the potential savings attached to sequential mixed mode designs that start out using the web, when applied to household surveys. The same issues would apply if other measurements had to be done on the survey participant eg. physical measurements requiring a visit. The principles of cost calculations in household panel studies are discussed in Lynn (2011) and Lynn (2011-Appendix 9 presentation).

Commentators have also noted that the fieldwork period is likely to be extended when using a sequential mixed mode approach compared with a concurrent or single mode approach. It takes more time, and better systems to be tracking responses, sending out multiple reminders, and ultimately reallocating non-responders to an alternative mode as the survey is in the field. This time element is seen as a disadvantage by national statistical offices where timetable deadlines are often crucial for producing statistical series.

Betts and Lound (2010b) also noted that there may be ethical and legal issues to consider in carrying out mixed mode data collections. However, on the whole these are the same for any survey; for example, that those who are unable to respond in one mode be offered and alternative mode by which they can respond, according to any disability, language difficulty or access to facilities they may face. They note that surveys which involve the web and or sequential mixed modes, especially if they are household surveys, make a number of elements more complicated: these include seeking consent to do the survey; obtaining permission to link to administrative records; fulfilling statutory requirements about data protection and maintaining within-household confidentiality.

[17] Commentators also raise the issue of whether (CAPI) interviewers will be demotivated by being given the most reluctant respondents to chase up as the last stage of a sequential mixed mode survey. However, this is based on an assumption about which there is little evidence. These potential respondents may not be reluctant to do F2F interviews, only to give their replies via other modes of data collection.

4. Conclusions

Mixed mode surveys which include the web as one of the modes have been increasing in the 21st century, as has our knowledge of the outcomes of using the web for data collection. Knowledge has been advancing in the population about how to access and use computers and the web effectively to fill in web surveys. On the research side, we know more about the consequences for surveys and their quality of opting to use the web, either alone or in combination with other modes of data collection.

The early attempts to use the web to do surveys had to be restricted to survey samples of web-savvy (sub) population groups, known to have the necessary equipment and experience to complete a web questionnaire. Nonetheless, response rates to early web surveys were low. Access to the web and familiarity with using it are still far from complete among the UK's population as well as in the majority of other industrialised countries. The profiles of web users show clear demographic biases by age, education and income. To collect a representative population survey in the UK using the web mode alone would need to factor in installing the necessary equipment (and training) for those who did not have it already.

However, using the web in combination with other modes has been shown to offer fruitful possibilities and outcomes. By including the web in a sequence of modes by which data are collected, research, experiments and survey practice have shown that it is possible to obtain good quality total unit response rates from the mixture of modes, to gradually increase the web share of the total response, and gain benefits to the coverage and representativeness of the survey. The literature suggests that the highest web shares of overall response from sequential studies come from an approach which is web-intensive and involves an unconditional pre-paid cash incentive. Web-intensive means first approaching respondents to use the web, and following this up with several reminders, without offering any other option at the outset. Getting respondents to use the web first is clearly the main driver of reducing costs in a sequential mixed mode approach where the cheapest mode is used first in the sequence.

Despite the demographic biases in data collected via the web, analyses of such data have shown web data to have a number of advantages. They have been shown to be cheaper and often faster to collect and process than data collected by other modes, and can reach population groups that were hard to reach by other modes (eg. those rarely in the house; given up landline phones; living in rural communities; immigrant groups). Web data have also been found to have the advantages of greater accuracy, fewer gaps, fewer DKs and item non-responses, and the lack of social desirability or acquiescence biases that emerge in the presence of interviewers. However, the cost to survey quality of measurement error may be increased from collecting data from the web as part of mixed mode survey.

For longitudinal surveys, data collection using the web has advantages over the other main low cost self-administered alternative of mail/paper. Survey instruments on the web, like CATI and CAPI, can make use of complex routing, feedforward data, and dependent interviewing. They can also include data repair for missing data. These are all

enhancements to the survey data that can also help to reduce respondent burden and improve the overall experience of the survey for respondents.

UK datasets, identified during this review to have used sequential mixed mode data collection, have shown the benefits of this approach when surveying sub population groups, heavily dominated by young people or students, the so-called web-savvy populations. The major UK survey organisations all have successful experience of sequential mixed mode data collection, although not necessarily for longitudinal data. However, these methods have been used successfully in both longitudinal and cross-sectional UK surveys, but there are relatively few cases overall, and very few cases of longitudinal data using such mixed mode approaches to data collection. Many of the examples found had large volume sample sizes and very small instruments. However, among them were also studies which had carried out sequential mixed mode data collection successfully, including and starting with the web, using fairly substantial instruments.

For experience of the use of web and sequential mixed mode in data collection for general populations, we need to turn to Scandinavian countries. Scandinavian countries appeared to have the widest overall coverage for web access in the period 2000-2010, although even here there is lower access among the older population. Their high levels of web use may have encouraged them to go faster down the web route to data collection. The Netherlands has contributed most to these developments, as their approach to collecting national statistics has made decisions to integrate their data collection and adopt mixed mode data collection wherever possible. In general, cross-sectional surveys are more evident in these experiences of web and sequential mixed mode approaches to surveys.

Another major source of experience of using mixed modes including web data collection for general populations was found to be the collection of Census data across the world. These applications have mainly been trials and then uses of concurrent options for web or mail, and more recently, sequential mixed modes (including the web). Again, these are cross-sectional surveys.

The major longitudinal resources being collected in Europe and the USA appear cautious about making use of web-based data collection, or sequential mixed mode data collection which includes the web. Relatively few have made any steps in the direction of using mixed modes, including web-based data collection. It is probably not surprising that cautious approaches are most evident. There is an absence of any research on the consequences of adopting these approaches for the quality of subsequent waves of longitudinal data collection that would follow. The potential for loss of members from a panel or cohort is a very serious issue of data quality. There is the older age group to consider who are known to use the web much less frequently than young and middle-aged adults. There are often complex measurement instruments using recall to the last survey contact to build in which may be challenging for inclusion in web instruments. There are also the complex measurement errors to consider, from changing the mode, potentially both between and within data collection waves. In addition, funding is often on a longer timetable for longitudinal studies, such that pressures to reduce survey fieldwork costs may not always be felt immediately. Household panel studies face another set of issues: of mapping out, in a web context, the Household Grid of all household members and their relationships to each other. This instrument is the central pivot around which the rest of the data collection

for the individual household members fans out. There is little knowledge or experience of how Household Grids can be adapted successfully to a web environment. All of the above issues would work in the direction of adopting a more cautious approach.

There are gaps in our understanding which need to be addressed to progress the effective use of mixed mode data collection using the web, as follows:

1. Research into the consequences of using mixed mode data collection (including web) in a longitudinal survey – on the next wave's response rates by last wave's mode of data collection.
2. More research and experience on how household surveys can adapt to use the web and face the issues related to using Household Grids, dealing with consent, data protection and within-household confidentiality and gaining other permissions (eg to access and merge administrative records with the survey data).
3. Continuing the programme of research to extend our understanding about the measurement errors involve in mixed modes of data collection.
4. Particular attention being given to the way question wording can have mode effects.
5. Continuing to research, develop and refine guidance about how and when to use mode-specific questions in mixed mode surveys and when to default to the unimode style of question construction in order to avoid mode effects on particular survey questions. This requires greater understanding by mode of stimulus, cognitive equivalence of tasks, burden, and benefits.
6. Flags to be routinely entered into the documentation of mixed mode surveys drawing attention to this form of data collection, with a variable included indicating the mode by which each case was collected and pointers to where advice can be obtained on how to proceed in analyses.
7. Guidance on how to combine and analyse data from different modes that has complex and confounded measurement errors that would be likely to result from a mixed mode data set.
8. Guidance on how to combine and analyse mixed mode data collected in a longitudinal data set, both after the initiation of the change and across the cross-over point of change in data collection modes along the survey's lifetime.
9. Carrying out more analyses to find out if there are likely effects on the findings from changing the mode of data collection during the process of collecting a longitudinal survey.

References

Akl, E.A., Maroun, N., Klocke, R.A., Montori, V. and Schunemann, H.J. (2005). 'Electronic mail was not better than postal mail for surveying residents and faculty.' *Journal of Clinical Epidemiology*, 58. 425-29.

Ariel, A., Giesen, D., Kerssemakers, F. and Vis-Visschers, R. (2008). 'Literature Review on Mixed mode Studies', Statistics Netherlands internal paper.

Aschinger, F., Epstein, H., Muller, S., Schaeper, H., Vottiner, A. and Weiβ, T. (2011). 'Higher education and the transition to work' in *ZfE*, 14: 267-82.

Ashton, K. and Dunn, E. (2009). 'Online Opinions - A Pilot to Extend Social data Collection Capabilities within the Office for National Statistics', *Survey Methodology Bulletin* 65, pp. 28-38.

Bandillar, W., Bosnjak, M. and Altdorfer, P. (2003). 'Survey administration effects? A comparison of web-based and traditional written self-administered surveys using the ISSP environment module'. *Social Science Computer Review*, 21, 235-43.

Banning, R. and Schouten, B. (2009). 'A mixed mode follow up of the panel refusers in the Dutch LFS'. Statistics Netherlands Discussion Paper 09035.The Hague: Statistics Netherlands.

Bates, N. (2009). 'Cell Phone-Only Households: A Good Target for Internet Surveys?' *Survey Practice*, November 2009: www.surveypractice.org.

Bates, S. C., and Cox, J.M. (2008). 'The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors'. *Computers in Human Behavior,* 24(3), 903-916.

Bech, M. and Kristensen, M.B. (2009). 'Differential response rates in postal and web-based surveys among older respondents', *Survey Research Methods*, 3, 1-6.

Best, S.J. and Krueger, B. (2004). *Internet data collection*. Thousand Oaks CA: Sage Publications.

Bethlehem, J. (2009). 'Use of Web Surveys in Official Statistics'. Paper presented at the Fourth Internet Survey Methodology Workshop, Bergamo, 17-19 September 2009.

Betts, P. (2010). 'The application of alternative modes of data collection in UK Government social surveys: Report on a workshop held at the Office for National Statistics, 8-9 December 2009', Office for National Statistics. Available at: http://www.ons.gov.uk/about/who-we-are/our-services/data-collection-methodology/reports-and-publications/alternative-modes-of-data-collection/index.html

Betts, P. and Lound, C. (2010a). *The Application of Alternative Modes of Data Collection in UK Government Social Surveys. Literature Review and Consultation with National Statistical Institutes*. United Kingdom: Government Statistical Service.

Betts, P. and Lound, C (2010b). *The Application of alternative modes of data collection in UK Government social surveys: A report for the Government Statistical Service*, Newport: Office for National Statistics. Available at: http://www.ons.gov.uk/about/who-we-are/our-services/data-collection-methodology/reports-and-publications/alternative-modes-of-data-collection/index.html

Beukenhorst, D. (2008). 'Mixed mode: Update on Past Research and Plans for the Near Future'. Statistics Netherlands Internal paper.

Beukenhorst, D. and Wetzels, W. (2009). 'A Comparison of Two Mixed mode Designs of the Dutch Safety Monitor: Mode Effects, Costs, Logistics'. Paper presented at the European Survey Research Association Conference, Warsaw, 29 June – 3 July 2009.

Biemer, P. and Lyberg, L. (2003). *Introduction to Survey Quality*, New Jersey: John Wiley & Sons.

Blyth, B. (2008). 'Mixed mode: the only 'fitness' regime?'. *International Journal of Market Research* 50 (2), pp. 241-266.

Blyth, B. (2008a). 'The implications of variation in national data collection mode access and rats of access change: a European over view'. Paper in Proceedings of the 3mc2008 conference.

Bohatá, M. (2011). 'Paper Eurostat: The Modernisation of European Social Statistics'. Paper presented at 97th Conference of the Directors-General of the National Statistical Institutes (DGINS) Wiesbaden, Germany, 27 Sept 2011. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/Content/Events/DGINS/Document__Session__II__0,property=file.pdf

Brady, S., Stapleton, C. and Bouffard, J. (2004). '2003 National Census Test: Response Mode Analysis', US Census Bureau Internal Memorandum.

Braunsberger, K., Wybenga, H., and Gates, R. (2007). 'A comparison of reliability between telephone and web-based surveys'. *Journal of Business Research*, 60, 758-764.

Brennan, M. (2005). 'The Effect of a Simultaneous Mixed-Mode (Mail and Web) Survey on Respondent Characteristics and Survey Responses'. Proceedings of ANZMAC 2005 Conference, Marketing Research and Research Methodologies (quantitative).

Buelens, B. and van den Brakel, J. (2011). *Inference in surveys with sequential mixed-mode data collection.* Statistics Netherlands Discussion Paper 201121. The Hague: Statistics Netherlands.

Burton, J., Laurie, H. and Noah, Uhrig S. (2008). 'Understanding Society: Some preliminary results from the Wave 1 Innovation Panel', Understanding Society Working Paper Series No. 2008-03, available at: http://research.understandingsociety.org.uk/innovation-panel-wave-1

Burton, J., Laurie, H., and Noah Uhrig, S.C. (2010). Understanding Society Innovation Panel Wave 2: Results from Methodological Experiments, Understanding Society Working Paper Series No. 2010-04, available at: http://research.understandingsociety.org.uk/innovation-panel-wave-2

Carpenter, H. (2010) *Youth Cohort Study – Cohort 13 Sweep 3 technical report* http://www.esds.ac.uk/doc/6024/mrdoc/pdf/6024ycs_sweep_3_tech_report.pdf.
Chang, L. (2001). A comparison of samples and response quality obtained from RDD Telephone survey methodology and internet survey methodology. Doctoral dissertation, Ohio State University.

Cartenssen, J. (2011). 'Separating selection from mode effects when switching from single (CATI) to mixed mode design (CATI/Web) in a health survey'. Presentation give at ESRA 2011 conference, Lausanne. http://surveymethodology.eu/conferences/lausanne-2011/presentation/472.

Chang, L. and Krosnick, J. (2003, 2008). 'National Surveys via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality', available at: http://comm.stanford.edu/faculty/krosnick/

Chang, L., and Krosnick, J.A. (2009). 'National Surveys via RDD Telephone Interviewing Versus the Internet'. *Public Opinion Quarterly,* 73(4), 641-678.

Christian, L. M., Dillman, D. A., and Smyth, J. D. (2007). *The effects of mode and format on answers to scalar questions in telephone and web surveys. Advances in telephone survey methodology* (pp. 250-275), John Wiley & Sons, Inc.

Coderre, F., Mathieu, A., and St-Laurent, N. (2004). 'Comparison of the quality of qualitative data obtained through telephone, postal and email surveys'. *International Journal of Market Research,* 46, 347-357.

Cole, M. S., Bedeian A. G. and Feild H.S. (2006). 'The Measurement Equivalence of Web-Based and Paper-and-Pencil Measures of Transformational Leadership'. *Organizational Research Methods*, vol. 9, pp. 339-363.

Converse, P.D., Wolfe, E. W., Huang, X. and Oswald, F.L. (2008). 'Response rates for mixed-mode surveys using mail and email/web'. *American Journal of Evaluation*, 29 (1), pp-107.

Cook, C., Heath, F. and Thompson, R. L. (2000). 'A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys'. *Educational and Psychological Measurement,* 60 (6), 821 -836.

Couper, M. P. (2000). 'Web Surveys. A Review of Issues and Approaches'. *Public Opinion Quarterly* 64(4), 464-494.

Couper, M. (2005). 'Technology Trends in Survey Data Collection', *Social Science Computer Review* 23, pp. 486-501.

Couper, M. (2007). Web Surveys in a Mixed Mode World. Paper presented at the ESRC Conference 'Survey Research in the 21st Century: Challenges and Opportunities', London, 24th October 2007.

Couper, M. (2008). *Designing Effective Web Surveys*, New York: Cambridge University Press.

Couper, M. Traugott, M. W. and Lamias, M. J. (2001). 'Web survey design and administration', in *The Public Opinion Quarterly*, 65 (2), pp.230-53.

Couper, M. P., Kapteyn, A., Schonlau, M. and Winter, J. (2007). 'Noncoverage and nonresponse in an Internet survey'. *Social Science Research* 36 (1), 131-148.

Couper, M. P., and Miller, P.V. (2008). 'Web Survey Methods'. Public Opinion Quarterly 72(5),831 -835.

Cuppen, M., van der Laan, P. and Nunspeet, W. (2010) 'Challenges of redesigning household surveys and maintaining output quality'. Paper at Quality in Official Statistics Q2010 Conference, Finlandia Hall, Helsinki, 4-6 May 2010.

Curtice, J. and Sparrow, N. (2004). 'Measuring attitudes of the general public via internet polls: an evaluation'. *International Journal of Market Research*, 46 (1), 23-44.

D'Agostino, A., Quintano, C. and Castellano, R. (2006). 'Telephone and Web surveys: a study of the marginal effect of mode'. SMABS-EAM Conference 2006.

Dannetun, E., Tegnell, A., and Giesecke, J., (2007). 'Parents' attitudes towards hepatitis B vaccination for their children. A survey comparing paper and web questionnaires', Sweden 2005: *BMC Public Health* 2007, 7:86doi:10.1186/1471-2458-7-86.

The electronic version of this article is the complete one and can be found online at: http://www.biomedcentral.com/1471-2458/7/86

DeBell, M., Krosnick, J.A. and Lupia, A. (2010). Methodology Report and User's Guide for the 2008–2009 ANES Panel Study, American National Election Studies. http://electionstudies.org/studypages/2008_2009panel/anes2008_2009panel_MethodologyRpt.pdf

De Leeuw, E. (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. and Hox, J. J. (2011). 'Internet Surveys as Part of a Mixed-Mode Design', in Das, M., Ester, P. and Kaczmirek, L. (eds.) *Social and Behavioral Research and the Internet*, New York: Routledge.

De Leeuw, E., Hox, J. and Dillman, D. (2008). *The International Handbook of Survey Methodology*, European Association of Methodology, London: Lawrence Erlbaum Associates.

De Leeuw, E.D., Hox, J. J. and Scherpenzeel, A. C. (2011). Emulating interviewers in an online survey: Experimental manipulation of _do-not-know' over the phone and on the Web. JSM Proceedings, Survey Research Methods Section, 6305-6314. Alexandria, VA: American Statistical Association.

Dennis, J. M. and Li, R.J. (2007). Results of a Within-Panel Web Survey Experiment of Mode Effects: Using the General Social Survey's National Priority and Environmental Policy Attitude Items: The American Association for Public Opinion Research (AAPOR) 62th Annual Conference, 2007.

Denscombe, M. (2006). 'Web-based questionnaires and the mode effect: An evaluation based on completion rates and data contents of near identical questionnaires delivered in different modes'. *Social Science Computer Review*, 24, 246-54.

Denscombe, M. (2008). 'The length of responses to open-ended questions: A comparison on online and paper questionnaires in terms of mode effects'. *Social Science Computer Review*, 26, 359-68

Denscombe, M. (2009). 'Item non-response rates: A comparison of online and paper questionnaires.' *International Journal of Social Research Methodology*, 12, 281-91.

Department of Education (2010a). *LSYPE User guide to the datasets: Wave One to Wave Six*. Prepared by National Centre for Social Research, London. Available online at: http://www.surveynet.ac.uk/sqb/surveys/lsype.asp.

Department for Education (2010b). *Youth Cohort Study (YCS), Cohort 13, Sweeps 1 to 3 Quick Start User Guide*.
http://www.esds.ac.uk/doc/6024/mrdoc/pdf/6024userguide.pdf

Dever, J., Rafferty, A. and Valliant, R. (2008). 'Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias?', *Survey Research Methods* 2 (2), pp. 47-60.

Dillman, D. A. (2000). *Mail and Internet Surveys - The Tailored Design Method (2nd Edition*), New York: John Wiley and Sons.

Dillman, D. A. (2009). 'Some consequences of survey mode changes in longitudinal surveys', in Lynn, P. (ed.) Methodology of Longitudinal Surveys, pp.127-40, Chichester: Wiley.

Dillman, D. A. and Christian, L. (2005). 'Survey Mode as a Source of Instability in Responses across Surveys', *Field Methods* 17 (1), pp. 30-52.

Dillman, D., Smyth, J. and Christian, L. (2009). *Internet, Mail and Mixed-Mode Surveys - The Tailored Design Method (3rd Edition*), New Jersey: John Wiley and Sons.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. and Messer, B.L. (2009). 'Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the internet. *Social Science Research* 38 (1), 1-18.

Duffy, B., Smith, K., Terhanian, G. and Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research,* 47(6), 615-639.

Dunn, E. (2010). 'Online Opinions – A Pilot to Extend ONS Social Data Collection Capabilities'. Paper presented at the 5th conference of the European Workshop on Labour Force Survey Methodology in Paris, 15-16th April 2010. http://www.insee.fr/en/insee-statistique-publique/default.asp?page=colloques/wlfsm/wlfsm-papers-presentations.htm.

Flatley, J. (2001). 'The Internet as a mode of data collection in government social surveys: issues and investigations', *National Statistics Social Survey Methodology Bulletin* 49, pp. 1-10.

Fricker, S., Galesic, M., Tourangeau, R. and Yan, T. (2005). 'An experimental comparison of web and telephone surveys'. *Public Opinion Quarterly,* 69, 370-392.

Frosch, M. and Lauritsen, S. (2010). 'Evaluation of CAWI pilot-the use of CAWI in the collection of household data in the Danish LFS'. Paper presented at The fifth European Workshop on Labour Force Survey Methodology will be held in Paris on 15-16th April 2010. http://www.insee.fr/en/insee-statistique-publique/default.asp?page=colloques/wlfsm/wlfsm-papers-presentations.htm

Gambling Commission (2008). *British Gambling Prevalence Survey 2007*, Birmingham: Gambling Commission.

Gibson, R. and McAllister, I. (2009). 'Online election surveys: Keeping the voters honest?'. *Journal of Political Marketing*, 8 (2), 105-129.

Goritz, A. (2006). 'Incentives in web studies: Methodological issues and a review', *International Journal of Internet Science*, 1: 58-70.

Graham, A. L., Papandonatos, G. D., Bock, B.C., Cobb, N.K., Baskin-Sommers, A., Niaura, R. and Abrams, D.B. (2006). 'Internet- vs. telephone-administered questionnaires in a randomized trial of smoking cessation'. *Nicotine & Tobacco Research,* 8, S49-S57.

Greene, J., Speizer, H. and Wiitala, W. (2008). 'Telephone and web: Mixed-mode challenge'. *Health Services Research*, 43, 230-248.

Griffin, D., Fischer, D. and Morgan, M. (2001). 'Testing an Internet Response Option for the American Community Survey'. Paper presented at the Annual Conference of the American Association for Public Opinion Research Montreal, 17-20 May 2001.

Grondin, C. and Sun, L. (2009). '2006 Census internet mode effect study'. Statistics Canada Internal Paper.

Groves, R. (1989). *Survey Errors and Survey Costs*, New York: John Wiley and Sons.

Groves, R. and Lepowski, J. (1985). 'Dual Frame, Mixed Mode Survey Designs'*, Journal of Official Statistics,* 1 (3), pp. 263-286.

Hayslett, M. M. and Wildemuth, B. M. (2004). 'Pixels or pencils? The relative effectiveness of web-based versus paper surveys'. *Library and Information Science Research*, 26, 73-93.

Heerwegh, D. (2009). 'Mode differences between face-to-face and web surveys: an experimental investigation of data quality and social desirability effects', *International Journal of Public Opinion Research* 21(1), pp. 111-121.

Heerwegh, D. and Loosveldt, G. (2008). 'Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality', *Public Opinion Quarterly* - Special Issue on Web Survey Methods, 72(5), pp. 836-846.

HEFCE (2011.) *National Student Survey: Findings and trends 2006 to 2010*, Higher Education Funding Council for England Issues Paper 2011/11. http://www.hefce.ac.uk/pubs/hefce/2011/11_11/

HESA (2009). *Destinations of Leavers from Higher Education Institutions Longitudinal Survey of the 2004/05 cohort Key Findings Report*. Higher Education Statistics Agency.

Holmberg, A. and Lorenc, B. (2008). 'Understanding the decision to participate in a survey and the choice of the response mode. Paper presented at the European Conference on Quality in Official Statistics (Q2008), Rome, 8-11 July 2008.

Holmberg, A., Lorenc, B. and Werner, P. (2009). 'Optimal contact strategy in a mail and web mixed mode survey'. Paper presented at the NTTS 2009 (New Techniques and Technologies for Statistics) Seminar, Brussels, 18-20 February 2009.

Hope, S. and Nicolaas, G. (2008). 'Mixed modes and measurement error: Study design and literature review'. Paper presented at the 7th International Conference on Social Science Methodology (RC33) Naples, 1-5 September 2008.

Horrigan, J. B. and Smith, A. (2007). Home Broadband adoption, 2007. Pew Internet and American Life project. http://www.pewinternet.org/pds/PIP_Broadband%202007.pdf

Ingels, S.J., Pratt, D.J., Wilson, D., Burns, L.J., Currivan, J., Rogers, J.E. and Hubbard-Bednasz, S. (2007). *Education Longitudinal Study of 2002 (ELS:2002) Base-Year to Second Follow-up Data File Documentation*, Washington: US Department of Education.

INSEE (2011). 'Improving efficiency in data collection by using adequate collection modes'. Paper presented at the DGINS Conference, New conceptual design for household and social statistics, Statistisches Bundesamt Deutschland, Wiesbaden, 27 Sept. 2011. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/Content/Events/DGINS/Document__PaperINSEE,property=file.pdf

Jäckle, A., Roberts, C. and Lynn, P. (2008). 'Assessing the effect of data collection mode on measurement', ISER Working Paper 2008-08. Colchester: Institute of Social and Economic Research, University of Essex.

Janssen, B. (2006). 'Web data collection in a mixed mode approach: an experiment. Statistics Netherlands' Paper presented at the European Conference on Quality in Official Statistics (Q2006) Cardiff, 24-26 April 2006.

Joshipura, M. (2008). '2005 ACS Respondent Characteristics Evaluation' US Census Bureau, American Community Survey Research and Evaluation Memorandum Series.

Kaiser, A. and Cross, F. (2011). *Beginning Teacher Attrition and Mobility: Results From the First Through Third Waves of the 2007–08,* Report NCES 2011-318. Washington: National Center for Education Statistics, US Department of Education.

Kaplowitz, M.D., Hadlock, T.D. and Levine, R. (2004). 'A Comparison of Web and Mail Survey Response Rates'. *Public Opinion Quarterly,* 68 (1), 94 -101.

Kappelhof, J. (2011). 'Mixed-mode surveys among non-western minorities in the Netherlands: does it work? Presentation of Paper presented at ESRA 2011 conference, Lausanne.

Keigher, A. (2010). *Teacher Attrition and Mobility: Results From the 2008–09 Teacher Follow-up Survey* (NCES 2010-353). U.S. Department of Education. Washington, DC: National Center for Education Statistics. http://nces.ed.gov/pubsearch.

Kreuter, F., Presser, S. and Tourangeau, R. (2008). 'Social Desirability Bias in CATI, IVR and Web Surveys – The Effect of Mode and Question Sensitivity', *Public Opinion Quarterly* - Special Issue on Web Survey Methods, 72 (5), pp. 847-865.

Kröhne, U. and Martens, T. (2011). 'Computer-based competence tests in the national educational panel study: The challenge of mode effects'. *Zeitschrift für Erziehungswissenschaft*, 14, pp. 169-186

Krysan, M., Schuman, H., Scott, L.J. and Beatty, P. (1994). 'Response rates and response content in mail versus face-to-face surveys', *Public Opinion Quarterly*, 58, 381-99.

Kwak, N. and Radler, B. (2002). 'A comparison between mail and web surveys: Response pattern, respondent profile and data quality'. *Journal of Official Statistics*, 8, 257-273.

Labrasseur, D., Morin, J-P., Rodrigue, J-F. and Talyor, J. (2010). 'Evaluation of the innovations implemented in the 2009 Canadian Census Test'. Proceedings of the Survey Research Methods Section of the Joint Statistical Meetings, American Statistical Society, pp.4089-97.

Lagerstrom, B. (2008). 'Cost Efficiency in a Mixed-Mode Survey - Evidence from the Norwegian Rent Market Survey'. Paper presented at the 19th Workshop on Household Nonresponse, Ljubljana, 15 September 2008.

Laroche, D. and Grondin, C. (2008). 'Impact of online edits and internet features in the 2006 Canadian Census'. Paper presented at Conference of European Statisticians, United National Statistical Commission and European Commission for Europe, Vienna, 21-23 April, 2008.

Laurie, H. and Lynn, P. (2009). 'The use of respondent incentives on longitudinal surveys'. In Lynn, P. (ed.) *Methodology of Longitudinal Surveys*, Chichester: Wiley.

Leung, D. and Kember, D. (2005). 'Comparability of data gathered from evaluation questionnaires on paper and through the internet'. *Research in Higher Education*, 46, 571-591.

Lindhjem, H. and Navrud, S. (2010). 'Can cheap panel-based internet surveys substitute costly in-person interviews in CV surveys?' Norway: Norwegian University of Life Sciences, Norwegian Institute for Nature Research (NINA).

Link. M. and Mokdad, A. (2005). 'Effects of survey mode on self-reports of adult alcohol consumption: A comparison of mail, web and telephone approaches', *Journal of Studies on Alcohol,* 66, pp. 239-245.

Link, M.W. and Mokdad, A.H. (2006). 'Alternative modes for health surveillance surveys: An experiment with web, mail, and telephone'. *Epidemiology*, 16 (5), pp. 701-704.

Lugtig, P. (2011). 'Using propensity score matching to separate mode and selection effects'. Presentation given at ESRA 2011 conference, Lausanne. http://surveymethodology.eu/conferences/lausanne-2011/presentations/513/.

Lugtig, P., Lensvelt-Mulders, G.J.L.M., Frerichs, R. and Greven, A. (2011). 'Estimating nonresponse bias and mode effects in a mixed mode survey'. in *International Journal of Market Research*. Vol. 53, No. 5, 2011, pp. 669-686.

Lynn, P. (2011). 'The Impact of alternative mixed-mode strategies on non-response in a complex household survey: Constraints and experiments', Paper presented at the International Workshop on Household Survey Non-response 2011, Bilbao.

McCabe, S.E., Diez, A., Boyd, C.J., Nelson, T.F. and Weitzman, E. R. (2006). 'Comparing web and mail responses in a mixed mode survey in college alcohol use research', in *Addictive Behaviors*, 31: 1619-27.

Meder, I. K. and Larsen, P. S. (2011) 'The 11-year follow up of the Danish National Birth Cohort'. Presentation given at EUCCONET Workshop: Surveying Children in Longitudinal Studies, 5-6 May, 2011, ESRI, Dublin. http://www.growingup.ie/fileadmin/user_upload/documents/EUCCONET/Inger_Kristine_Danish_National_Birth_Cohort.pdf

Manfreda, K. L., Vehovar, V. and Batagelj, Z. (2001). 'Web versus mail questionnaire for an institutional survey'. In Westllake, A. (ed.) *The Challenge of the Internet*, London: Association for Survey Computing.

Manfreda, K., Berzelak, J., Vehovar, V., Bosniak, M. and Hass, I. (2008). 'Web surveys versus other survey modes: a meta-analysis comparing response rates', *International Journal of Market Research,* 50(1).

Marcus, B., Bosnjak, M., Lindner, S., Pilischenko, S. and Schutz, A. (2007). 'Compensating for low topic interest and long surveys – A field experiment on nonresponse in web surveys', *Social Science Computer Review,* 25, pp. 372-383.

Martin, E., Hunter Childs, J., DeMaio, T., Hill, J., Reiser, C., Gerber, E., Styles, K. and Dillman, D. (2007). *Guidelines for designing questionnaires for administration in different modes*, Washington DC: US Census Bureau.

Martin, P. (2011). 'What Makes a Good Mix? Chances and Challenges of Mixed Mode Data Collection in the ESS'. Working Paper No. 02- Feb 2011, London: Centre for Comparative Social Surveys, City University.

Martin, P. and Lynn, P. (2011). *The effects of mixed mode survey designs on simple and complex analyses*. ISER Working Paper Series No. 2011-28. November 2011: Colchester: Institute of Social and Economic Research, University of Essex. http://surveymethodology.eu/conference/lausanne-2011/presentation/238/

Meckel, M., Walters, D., and Baugh, P. (2005). 'Mixed-mode surveys using mail and web questionnaires'. *The Electronic Journal of Business Research Methodology*, 3, 69-80.

Messer, B. L. and Dillman, D.A. (2011). 'Surveying the general public over the internet using address-based sampling and mail contact procedures'. In *Public Opinion Quarterly*, 75 (3), pp.429-57.

Millar, M. M., and Dillman, D. A. (2011). 'Improving response to web and mixed-mode surveys'. *Public Opinion Quarterly*, 75 (2), 249-269.

Miller, E.T., Neal, D.J., Roberts, L.J., Baer, J.S., Dressler, S.O. and Metric, J. (2002). 'Test retest reliability of alcohol measures. Is there a difference between internet-based assessment and traditional measures?' *Psychology of Addictive Behaviors*, 16: 56-63.

Nicolaas, G. and Lynn, P. (2002). 'Random digit dialling in the UK: viability revisited', *Journal of the Royal Statistical Society, A* 165 (2), pp. 297-316.

ONS (2010). *2011 Census: Evaluation of the 2009 Rehearsal*, Newport: Office of National Statistics.
http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/2009-census-rehearsal/index.html

ONS Statistical Bulletin (2009). 'Internet access - households and individuals', Office for National Statistics, available at http://nswebcopy/StatBase/Product.asp?vlnk=5672 .

Oosterveld, P. and Willems, P. (2003). 'Two modalities, one answer? Combining Internet and CATI surveys effectively in market research'. In Fellows, D. S. (Ed.) *Technovate* (pp. 141–150). Amsterdam: ESOMAR.

Porter, S. and Whitcomb, M. (2007). 'Mixed-mode contacts in web surveys: Paper is not necessarily better', *Public Opinion Quarterly,* 71 (4), pp. 635-648.

Revilla, Melanie. (2010). 'Quality in Unimode and Mixed-Mode Designs: A Multitrait-Multimethod Approach'. *Survey Research Method,s* 4 (3), 151-164.

Revilla, M.A. and Saris, W.E. (2010). 'Comparison of surveys using different modes of data collection: European Social Survey versus LISS Panel'. Working paper, New Developments in Survey Methodology – seminar series. Universitat Pompeu Fabra, Spain.

Riek, S., Rietz, C. and Kruger, T. (2004). 'Using web surveys in mixed-mode approaches: an experimental comparison with traditional survey modes'. German Online Research Conference (GOR), 2004.

Roberts, C. (2005). 'Expert workshop on mixed mode data collection in comparative social surveys'. Economic and Social Research Council - National Centre for Research Methods report on the 'Expert workshop on mixed mode data collection comparative social surveys', 15-16th September 2005, City University, London, available at:
http://eprints.ncrm.ac.uk/406/

Roberts, C. (2007). 'Mixing modes of data collection in surveys: A methodological review', Economic and Social Research Council - National Centre for Research Methods, NCRM Methods Review Papers NCRM/008, available at: http://eprints.ncrm.ac.uk/418/

Rookey, B. D., Hanway, S. and Dillman, D. A. (2008). 'Does a probability-based household panel benefit from assignment to postal response as an alternative to internet-only?' *Public Opinion Quarterly*, 2 (5), 962–984.

Roy, L. and Laroche, D. (2006). 'The internet response method: Impact on the Canadian Census of Population data'. Proceedings of the Survey Research Methods Section, American Statistical Association, 2006.

Russell, H., Watson, D. and Banks, J. (2011). *Pregnancy at Work: A National Survey*. Dublin: HSE Crisis Pregnancy Programme and the Equality Authority.

Safir, A. and Goldenberg, K. (2008a). 'Mode Effects in a Survey of Consumer Expenditures', US Bureau of Labor Statistics Office of Survey Methods Research paper, available at: http://www.bls.gov/osmr/abstract/st/st080200.htm

Sakshaug, J. W., Yan, T. and Tourangeau, R. (2010). 'Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items'. *Public Opinion Quarterly*, 74 (5), 907-933.

Schaefer, D. and Dillman, D.A. (1998). 'Development of a standard e-mail methodology: Results of an experiment'. *Public Opinion Quarterly*, 74: 68-84.

Scherpenzeel, A. and Toepoel, V. (2011). 'Recruiting a probability sample for an online panel: Effects of contact mode, incentives and information.'. Paper presented at ESRA 2011 conference, Lausanne.

Schonlau, M., van Soest, A., Kapteyn, A. and Couper, M. (2009). 'Selection bias and the use of propensity scores', *Sociological Methods and Research* Special Edition on Web Surveys 37(3), pp. 291-318.

Schonlau, M., Zapert, K., Simon, L.P., Sanstad, K. H., Marcus, S.M., Adams, J., Spranca, M., Kan, H., Turner, R. and Berry, S. H. (2003). 'A comparison between responses from a propensity-weighted web survey and an identical RDD survey'. *Social Science Computer Review*, 22 (1), pp. 128-138.

Shih, T-H. and Fan, X. (2008). 'Comparing response rates from web and mail surveys: A meta-analysis', *Field Methods,* 20 (3), pp. 249–271.

Shih, T. and Xitao, F. (2007). 'Response rates and mode preferences in web-mail mixed-mode surveys: A meta-analysis'. *International Journal of Internet Science,* 2 (1), 59-82.

Shin, E., Johnson, T.P. and Rao, K. (2011). 'Survey mode effects on data quality: Comparison of web and mail modes in a U.S. National Panel Survey'. *Social Science Computer Review*, DOI 10.1177/0894439311404508.

Shury, J., Vivian, D. and Oldfield, K. (2009). *Technical Report. Destinations of Leavers from Higher Education (DLHE) Longitudinal Survey 2004/05*, Prepared for HESA, London: IFF Research.

Shury, J., Vivian, D. and Riley, C. (2011). *DLHE Longitudinal Survey 2006/07 – Technical Report*, Prepared for HESA, London: IFF Research.

Skills Funding Agency (2011). Framework for Excellence, Employer Views. http://ffepublication.skillsfundingagency.bis.gov.uk/EmployerData+Explained#intro.

Smith, A. (2010). 'Home Broadband 2010'. Pew Internet and American Life Project, Pew Research Centre: http://perinternet.org/Reports/2010/home-Broadband-2010.aspx

Smith, T.W. and Dennis, J.M. (2008). 'Mode effects on in-person and internet surveys: A comparison of the General Social Survey and Knowledge Network Surveys'. Paper presented at JSM conference, 2008 Survey Research Methods Section.

Smyth, J., Christian, L. and Dillman, D. (2008). 'Does "Yes" or "No" on the telephone mean the same as "check-all-that-apply" on the web?', *Public Opinion Quarterly,* 72(1) , pp. 103-113.

Smyth, J., Dillman, D., Christian, L. and Stern, M. (2006). 'Comparing check-all and forced choice question formats in web surveys', *Public Opinion Quarterly,* 70 (1), pp. 66-77.

Smyth, J. D., Dillman, D. A., Christian, L. M. and O'Neill, A. (2010). 'Using the internet to survey small towns and communities: Limitations and possibilities in the early 21st century' *American Behavioural Scientist*, 53 (9), 1423-48.

Speize, H., Baker, R. P., Wiitala, W. and Greene, J. (2006). 'Telephone and Web: The Mixed-mode Challenge' Paper presented at the 2nd International Conference on Telephone, 2006.

Surridge, P. (2006). *The National Student Survey 2005: Findings*. A report to HEFCE, Bristol: Higher Education Funding Council for England. http://www.hefce.ac.uk/pubs/rdreports/2006/rd22_06/

Surridge, P. (2008). *The National Student Survey 2005-2007: Findings and Trends*. A report to HEFCE, Bristol: Higher Education Funding Council for England. http://www.hefce.ac.uk/pubs/rdreports/2008/rd12_08/

Tarnai, J. and Dillman, D. A. (1992). 'Questionnaire context as a source of response difference in mail versus telephone surveys', in Schwarz, N. and Sudman, S. (eds.) *Context effects in social and psychological research*, New York: Springer-Verlag.

Taylor, J. (2010). 'The 2009 Census test supplementary study of the new Census collection methodology'. Proceedings of Statistics Canada Symposium 2010, Social Statistics.

Tijdens, K. (2011). 'Drop-out rates during completion of an occupational search tree in web-surveys'. Paper presented at the Internet Survey Methodology Workshop 2011, Aug 30. 2011, The Hague: The Netherlands.

Tourkin, S. C., Pugh, K. W., Fondelier, S. E., Parmer, R. J., Cole, C., Jackson, B., Warner, T., Weant, G. and Walter, E. (2004). *1999– 2000 Schools and Staffing Survey (SASS) Data File User's Manual* (NCES 2004–303). Washington, DC: U.S. Department of Education.

Tourkin, S., Graham, S., Chambers, L., Strizek, G., Parmer, R., Jackson, B., Dial, S., Wang, Y., and Kaiser, A. (2011*). Documentation for the first through third waves of the*

2007–08 Beginning Teacher Longitudinal Study (NCES 2011-319). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

US Census Bureau (2009). *Design and Methodology: American Community Survey*, Washington DC: US Government Printing Office:

Van den Brakel, J., Smith, P. and Compton, S. (2008). 'Quality procedures for survey transitions – experiments, time series and discontinuities', *Survey Research Methods,* 2 (2), pp.123-141.

Van der Laan, P. and van Nunspeet, W. (2009). 'Modernisation of household surveys in the Netherlands: State of play', Statistics Netherlands paper prepared for the special European meeting of Directors of Social Statistics on the European Programme of Social Surveys, Luxembourg, 13 January 2009. Also Discussion Paper 09044, Statistics Netherlands, The Hague: Statistics Netherlands.

Vannieuwenhuyze, J., Loosevedlt, G., and Molenberghs, G. (2010). 'A method for evaluating mode effects in mixed mode surveys'. *Public Opinion Quarterly,* 74 (5), 1027-1045.

Van Nunspeet, W., Cuppen, M. and Van der Laan, P. (2011).' Redesigning Dutch social surveys: Efficiency and mixed mode'. Paper presented at the DGINS Conference, New conceptual design for household and social statistics, Statistisches Bundesamt Deutschland, Wiesbaden, 27 Sept. 2011.
http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/Content/Events/DGINS/Document__StatisticsNetherlands,property=file.pdf

van Soest, A. H. O. and Kapteyn, A. (2010). 'Mode and context effects in measuring household assets' in Ester, J. W. M., Das, P. and Kaczmirek, L. (2010) (Eds.), *Social and Behavioral Research and the Internet: Advances in Applied Methods and New Research Strategies* (pp. 291-323). Oxford: Taylor and Francis.

Von Maurice, J., Sixt, M. and Blossfeld, H-P. (2011). *The National Educational Panel Study: Surveying a cohort of 9th graders in Germany*. NEPS Working Paper No. 3, Bamberg, Otto-Friedrich-Universität, Nationales Bildungspanel.

Voorpostel, M. and Ryser, V-A. (2011). 'Mixed mode data collection as a strategy to decrease panel attrition in the Swiss Household Panel. Working Paper 2_11, FORS, Swiss Foundation for Research in Social Sciences.

Watson, J. (2009). 'Literature review identifying potential sources of mode effect between internet questionnaires and paper questionnaires', Newport: Office for National Statistics internal paper.

Weir, D. (2004). 'Modes, trends and content: A comparison of the 2003 HRS internet survey with HRS 2002 and 2004 core survey responses for evaluative, subjective and objective content'. Internal Paper, Survey Research Centre, Ann Arbor: University of Michigan.

Werner, P and Forsman, G. (2005) ‘Mixed mode data collection using paper and web questionnaires’. Paper in Proceedings of the AAPOR - ASA Section on Survey Research Methods, 2005.

Wilmot, A. and Dewar, A. (2006). ‘Developing harmonised questions for use in a mixed-mode data collection environment’ *National Statistics Survey Methodology Bulletin* 58, pp. 75-87.

Zajac, K., Allmang, K. and Barth, J. (2007). ‘2005 National Census test: Response mode analysis’, Internal Census Bureau Memorandum. Washington DC: US Census Bureau.

Appendix 1. Literature search strategy

The main aim of the literature search was to identify surveys and research on mixed mode data collection which included the web to collect the data. We were particularly interested to find examples of mixed mode methods (including web) used to collect longitudinal surveys. However, studies illustrating the usage of mixed methods on cross-sectional surveys were also considered. Priority was given to studies published after the year 2005 since we quickly discovered the existence of a number of earlier literature reviews or searches that went up, at least, to 2005. The existence of the other literature reviews informed our strategy, once they were identified. Further, we were very fortunate in being able to gain access to a list of over 700 references produced up to 2009 by the SDMI project, (see Appendix 8 for further details). Their list of references covered all types of mixed-mode data collection. Only a relatively small subset were relevant to our search, but very useful none the less, and for which we were very grateful. In the light of these other reviews and searches, our search became focussed on topping up the references from the earlier materials, and had a number of strands.

In Phase 1 we examined the obvious search engines known to be sources of research on survey data collection and/or web modes of data collection, as listed in Table A1.1.

Table A1.1 Phase 1: Literature available through internet search engines.

Name of Engine	Keyword Used	Website
Survey Resources Network		http://www.surveynet.ac.uk/sdmi/outputs.asp#nicolaas
European Social Survey Website		http://www.europeansocialsurvey.org/index.php?option=com_content&view=article&id=67&Itemid=235
ISER, Research Resources Unit, University of Essex	Mixed mode, web internet	http://www.iser.essex.ac.uk/publications/search
WebSM.org	Mixed mode, web internet	http://www.websm.org/
Google	Mixed mode, web internet	http://google.com
JSTOR	Mixed mode, web internet	http://www.jstor.org
ISI Web of Knowledge	Mixed mode, web internet	http://apps.webofknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=X11fcdKHE@62fG4k8eD&preferencesSaved=
webdatanet		http://www.research-projects.uzh.ch/p15383.htm

In addition we examined in detail the past 3 years of issues of the obvious journals that publish papers on methodological issues in data collection. We also engaged in contacting to ask for information or further leads:

- known academic experts in the field of mixed mode data;

- the main UK fieldwork agencies, to ask about their knowledge of UK datasets using mixed mode (including web);
- the custodians and other representatives of the well-known longitudinal household panel and individual longitudinal data sets in Europe and the USA, as well as making a full investigation of the data source's web site documentation and details of data collection modes used (Table A1.2).

We were very grateful to receive many responses to these requests from obviously busy people, although not all our emails were answered. The Acknowledgements show the extent of the community assistance we have been given in assembling this material.

Table A1.2. Contacts with longitudinal data providers

Dataset	Use of mixed modes including web
NEP, National Education Panel, Germany	Yes
LISS internet panel, Netherlands	Not mixed mode, only web panel
ANES web panel	Not mixed mode, only web panel
PSID, Panel Study of Income Dynamics, USA	Not used mixed mode with web
American Community Survey	Not used of mixed mode with web, with exception of trial mentioned in paper.
PAIRFAM, Panel Analysis of Intimate Relationships and Family Dynamics, Germany	Not use mixed mode with web
PASS, Panel Study Labour Market and Social Security, Germany	Not used mixed mode with web
SHP, Swiss Household Panel	Used mixed mode as follow up
CEPS/INSTEAD surveys, Luxembourg	Not used this kind of mixed mode
GSOEP, German Socio-Economic Household Panel Study	Not used mixed mode with web
NLS, National Longitudinal Surveys – USA Cohort Studies.	Not used mixed mode with web
SIPP Survey of Income and Program Participation, USA	Not used mixed mode with web.
SLID, Survey of Labor and Income Dynamics, Statistics Canada	Not used mixed mode with web.
HILDA, Australia	Not used mixed mode with web
GUS, Growing up in Scotland	Not used mixed mode with web
Growing up in Ireland	Not used mixed mode with web for baby cohort or Children age 9 cohort
ELSA, English Longitudinal Study of Ageing	Not used mixed mode with web

A number of recent conferences were also helpful in providing very up to date material; The European Survey Research Association (ESRA) conference July 2011, Lausanne; a Royal Statistical Society day seminar (11-10- 2011, Appendix 8); and the National Centre for Research Methods day seminar, (3-11-2011, Appendix 9).

Appendix 2. Further examples of surveys using mixed modes by type of approach.

Table A2.1 Examples of cross-sectional surveys using sequential mixed mode approaches.

'Details' =further information about this survey contained in Appendix 6.
Key: → indicates the order of the sequence.

Type MM-1 Cross-sectional SEQUENTIAL	
General populations	
ICT survey. Netherlands. 1)CATI survey. Non respondents asked to do web survey. 8% of refusers responded adding 1% to overall response. 24% of non-contacts responded, adding 2.5% to overall response. 2)All sent letter with URL and login info.→Reminder post card→Phone if known to persuade to do online – or by mail/paper. Overall response=37.5%; CATI alone=51%	Janssen (2006) Statistics Netherlands
Labour Force Survey Netherlands– embedded expt. First given choice web or mail/paper→Phone→CAPI	Van der Laan and van Nunspeet (2009) Statistics Netherlands
Labour Force Survey Pilot. Denmark 2009. All sample 15-74 year olds sent Advanced letter with web URL+Psswd (a) Half sample no further contact. RR=20.6% (b) Half sample. Letter said they would be phoned if no response. Sent phone reminder after 4 days→CATI. RR=45.9% Response share web=63.9%; CATI=36.1 % Reminders worked better on older and less well on younger people..	Frosch et al (2010)
2008 *Crime and Victimisation Survey,* Netherlands Web→Mail/paper→CATI→CAPI	Van der Laan and van Nunspeet (2009) Statistics Netherlands
2006-7 *Crime and Victimisation Survey,* Netherlands Web→CATI/CAPI	Van der Laan and van Nunspeet (2009) Statistics Netherlands
2008 *Dutch Integrated Safety Monitor* (**ISM**) Crime and victimization in population of 15+ Advanced letter to do web survey. Can request Paper.→Two reminders→CATI (if number)→ CAPI if no tel. number, or no response.	Buelens and van den Brakel (2011) **Details**
Health Interview Survey, Netherlands. Aimed to replace earlier CAPI survey by providing core data from administrative registers, supplemented with shorter survey instruments and sequential modes. web→Phone→CAPI. Old CAPI total RR=61%. New sequential total RR=64%. Share of sequential response: Web=50%; CATI=33%; CAPI=17%.	Van Nunspeet et al (2011) Statistics Netherlands
Behavioural Risk Factor Surveillance System. Non- institutionalised adults 18 years or more in 4 US states.	Link and Mokdad (2006)

web→mail/paper→CATI	**Details**
National Postsecondary Student Aid Study. (**NPSAS)** USA Surveys of students' financial aid in 2004, 2008, 2012. Web→CATI.	http://nces.ed.gov/surveys/npsas/about.asp
Employers' Views Survey, England. Views of further education provision. Skills Funding Agency, BIS. Further education providers give SFA's survey agent their employers' contact details. 1) those with email sent ULR to do web survey→CATI 2) those without email sent mail/Paper→CATI.	Learning and Skills Council Data Service. Skills Funding Agency (2011) **Details.**
Canadian Census 2009 test. Letter sent with details for web completion→Reminders→Paper Questionnaire mailed.	Taylor (2010)
*National Survey of Sexual Attitudes and Lifestyles III. (***NATSAL3)** UK. This 3rd cross-sectional survey 2010 (Natsal1-1990, Natsal2-2000) on national address frame; 16-74 year olds, uses CAPI-CASI. Some of refusers to main survey are being written to and asked if they will complete a shorter web questionnaire. Interim results: 8% response of those refusers asked respond by web.	On-going. Fieldwork period 2010-2012. Communication from Bob Erens
Sub-populations[18]	
National Students Survey (**NSS**), England. Web→Postal/mail→Phone.	HEFCE (2011) Fieldwork agent. **Details**
Evaluation of Children's Centres collecting data from staff and managers. Web→CATI	Fieldwork agent.
US College students use of alcohol. Emailed a URL with request to do a web survey + 2 reminders→sent a Paper copy to fill in + reminders →sent further Paper copy. Web share= 80.7%; Paper share=19.3%	McCabe et al (2006)
Course evaluation questionnaires from Hong Kong university students. Paper version sent to all students→Email sent giving URL for web version.→both Paper and email with URL sent out again. Share web=50.5%; Share paper=49.5%	Leung and Kember (2005)
Professional NBPTS Teachers in Ohio and South Carolina. (a) Sent postal/Paper survey→email with Web URL. Or (b) sent email URL →follow by postal/Paper. Initial Mail/Paper RR=80.7%. Initial web RR=41.8% Final RR by mail (a)=82.2%; email/web (b)=70.4%.	Converse et al (2008)
Non-western minorities in Netherlands.	Kappelhof, ESRA

[18] Other studies stated to be using sequential mixed mode data collection, but with insufficient documentation, at the time of writing, to be included above are: I. Andreas (2011) Greek Candidate Survey in 2007 web→CASI (ESRA 2011 presentation): Mars, G. and Van der Houwen, K. (ESRA 2011 presentation 537). 'Comparing (non) response between a sequential mixed mode design and a probability-based internet panel (LISS): C. Bayart (ESRA 2011 presentation 459, Lausanne) Lyon Household Travel Survey; refusal to do F2F→web survey.

1) By F2F survey 2) By sequential Web→CATI→CAPI. Response rates by sequential mixed mode slightly better or the same as F2F for each of 5 hard to recruit minority groups.	2011 conference paper.

Table A2.2. Examples of cross-sectional surveys using concurrent mixed mode approaches.[19]

'Details' =further information about this survey contained in Appendix 6.

Type MM-1 **Cross-sectional CONCURRENT**	
General populations[20]	
British Gambling Prevalence Survey, 2007. Household address based survey of all adults 16+ in household. Option to do Self completion paper-Booklet or online. 7% chose online. (Mixed mode option not repeated in 2010 survey.)	Gambling Commission (2008)
UK 2011 Census Rehearsal - Voluntary Choice web or paper. 8% of responses chose web; Total RR=41%.	ONS (2010)
UK 2011 Census. Compulsory Choice of web and paper/mail return. 16/17% chose web.	www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census
Canadian Census 2006, and pretest of method Paper sent to all. Returns by mail or web possible. Share web=18.3%.	Laroche and Grondin (2008); Grondin and Sun (2009); Roy and Laroche (2006)
Canadian Census – 2009 test of changes for 2011 Census. 1)Sub-sample posted letter requesting fill in web Census test 2)Subsample posted Paper Questionnaire by post(with option to post back or do web return) 3)Subsample got Paper questionnaire delivered to house with options to post back or fill in on line	Labrasseur et al, (2010) **Details.**
US National Census tests 2003 Choice of web, paper, IVR or Paper only	Brady et al, (2004)
US National Census tests 2005 Choice of web, paper, IVR or Paper only	Zajac et al, (2007)
American Community Survey Choice of web or Mail/paper	Griffin et al, (2011)
ONS Opinion Survey internet pilot – over 2 months Choice of web or F2F. Month1 web RR =8%; Month2=17%	Ashton and Dunn, (2009)

[19] A smaller study of this type not reaching the 1000 case sample size can be found in Hayslett and Wildemuth (2004).

[20] Other studies stated to be using concurrent mixed mode cross sectional data collection, but with insufficient documentation to be included are Simpanen, M. and Pertti (ESRA 2011 presentation, Lausanne) on a web pilot on Statistics Finland's Finish Consumer Barometer Survey with ransom assignment to (1) web questionnaire; (2) telephone questionnaire; or (3) web and telephone.

Other Censuses use of web and RR% Singapore 2000 (15%) Norway 2001 (9.9%) Canada 2006 (18.5%)	Couper (2007)
Gallup Panel 2006. Voting intentions/World Affairs Comparisons of frequent web user respondents with mail/paper respondents	Rookey et al, (2008)
Parents' survey on attitudes to their children having HepB vaccination. (Sweden) 2 cross sectional samples drawn. 1)First sample sent Paper Questionnaire to fill in. (Initial RR=49%, N=489/995; Final response=55%) 2)Second sample sent URL login and password to fill in on web (Initial RR=18%, N=179/1001; Final response=15%) Both N responses were increased by one reminder.	Dannetun et al (2007)
Sub-populations	
Head Teachers views about new English Baccalaureate Web and CATI	Fieldwork agent. On-going
Pregnant women and employment. Ireland Nationally representative survey of pregnant women (baby born July 2007 to June 2009). Paper version posted to 5000 mothers in 2009 also giving details of how to respond by web, and offering telephone interview for non-English speakers. Overall RR=46% (n=2300 replies) Share response returned by post=87%; by web =12%; by Tel=1%	Russell et al (2011)
Small and medium-sized enterprises in NW England. Mail/paper and Web	Meckel et al, (2005)
University graduates on sensitive topics. One of CATI, IVR or web	Kreuter et al (2008)

Table A2.3. Examples of cross-sectional surveys using mixed concurrent and sequential mixed mode approaches.

Type MM-1. Mixed Cross-sectional CONCURRENT/ SEQUENTIAL comparisons	
General populations	
Dutch Safety Monitor survey EXPT. 1)Web/mail→CAPI/CATI or 2)CATI/CAPI	Beukenhorst and Wetzels, (2009)
Norwegian Rent Market Survey 2007. EXPT This was second of two rent market surveys (also 2006) using mixed mode and web. 1)Letter sent with web URL→URL+Paper→CATI. Total RR=47.6% Web stage RR=11.6%. Web+ Paper RR=24.7% 2) Letter giving Choice of web or mail/paper enclosing mail/paper QU. CATI offered at stage 3 as added choice. Total RR= 48.6% Stage 1 Web option RR=5%; Paper option=13%	Lagerstrom (2008) Statistics Norway. Lagerstrom (2011-Appendix 9)

Stage 2 Web option RR=6%; Paper option=20% Stage 3 Web option RR=7%; Paper option=21%. CATI=20%	
Swedish Health Survey,18-64 year olds. EXPT. 5 Sequential and concurrent designs compared. Best response when web and paper both offered. Web response highest when only web mentioned at the outset (WebRR=47%). Lower if web give as option and paper option mentioned (WebRR=32%). Web option lowest if only mentioned in reminder stage WebRR=2%). Total RR. Web only followed by sequential alternative=73% Total RR. Web and paper as options followed by sequential=76%.	Holmberg et al, (2009) Statistics Sweden
*European Social Survey (***ESS***)* Netherlands mixed mode feasibility testing. EXPT.	Martin and Lynn (2011) Vannieuwenhuyze et al (2010) **Details.**

EXPT. Indicated this involves an embedded experiment.
'Details' =further information about this survey contained in Appendix 6.
Key: → indicates the order of the sequence.

Appendix 3. Further details of post/mail and web comparison surveys.

Reference	Survey details	Response rate findings
Brennan (2005)	Sample from New Zealand electoral register. Cross sectional	Mail only – 40% Web option – 25%
Brogger et al (2007)	Norway, adults 20-40 Cross-sectional	Mail only – 47% Web option – 45%
Griffin et al (2001)	American Community Survey. General US population. Trial	Mail only – 44% Web option – 38%
Werner and Forsman (2005)	Swedish Students housing survey in Linkoping (web access rate supposed to 100%)	Total RR% Mailed Paper +mention web option 84% Web details+Paper to be sent later– 85% Web details – 85% Paper QUEST sent – 84% Email sent with URL for web-84%.
Kaplowitz et al (2004)	Students at US university	Mail/Pap only. 31.5% Prenotice postcard+Email 29.7% Prenotice postcard+Email+postcard 28.6% Email+Post card 25.4% Email only. 20.7%
Brennan (2005)	New Zealand general public	Paper sent by post. (40%) Letter offers web/paper. Opts web (17%) Letter offers web/paper. Opts paper (9%)
Bates and Cox (2007)	Students. Self reports on sensitive questions	RR% in Group setting / Office setting / Individual setting Paper 93 / 70 / 80 Web 73 / 73 / 67
Hayslett and Wildemuth(2004)	Academic reference librarians	Paper invitation, paper survey (43.3%) Paper invitation. Web survey (22.9%) Email invitation, Web survey (33.1%)
Millar and Dillman (2011)	Undergrad students at US university. Internet literate population	1)choice of mail or web (Paper included) 2) Sent paper Quest. only 3)web only details given 4) web→Paper
Messer and Dillman (2011)	General household address sample sent Washington Community Survey	1)mail→web (40.1) 2)adding prepaid incentive (53.6%) 3)offering instruction card for web (46.3%) 4)send follow request by mail 5)add additional incentive with follow up request (52%) 6) Paper only +incentive (56.7%) (7)Paper only + 2xincentive (68.4%)
Smyth et al (2010)	Population living in small towns and communities in USA	Mail only: 71% Web only: 41% Web/mail choice 55%.

Appendix 4. Further details of telephone/web comparison surveys.

Reference	Survey details	Response rate findings
Fricker et al (2005)	Public attitudes to science	RDD Phone sample 97.5% RDD Web sample 51.6%
Kreuter et al (2009)	Sample of recent university graduates in one US university. Sensitive behaviour and questions	Tot CATI Web Initial random assign 1501 338 639 Agreed participate web 617 Started Questionnaire 1107 329 368 Completes 1003 320 363 % completes 66.8 94.7 56.8
Schonlau et al (2003)	California residents about health and healthcare	RDD – 43% Harris Interactive Web Panel –12%
Braunsberger et al (2007)	Random sample of consumer panel (web) and random sample of Telephone list. Health insurance coverage	RDD. -29.3% Web option –38.5%
Dillman et al (2009)	List of names and addresses of purchasers of consumer warranties. Consumer satisfaction survey.	SEQUENTIAL. 1)Prenotice→mail/Paper+cash incent→Tel 2)Tel→Mail/Paper+cash incent 3)Use Tel to recruit IVR→Tel. 4)Use Tel to recruit Web. Post URL+cash incentive→Email reminder→Tel./CATI Phase 1. Phase 2. Total 1) 75.0% 31.7% 82.8% 2) 43/44% 66.3% 80.4% 3) 28.5% 35.9% 50.4% 4) 12.7% 44.9% 47.7%
INSEE (2011)	ICT survey. Done by CATI until 2010. In 2010 two surveys in parallel (a) CATI and (b) web	RR% CATI = 75% Web = 28%

Appendix 5. Comparisons of single mode studies carried out in parallel using the same instrument but different samples/datasets, Dual Frame.

Study subject and modes	Findings	Reference
Socio-political participation and attitude questions. Web panel compared with F2F CAPI omnibus survey	5 questions carried in both surveys. 1)Online purchasing 2)View about amount of information people receive 3)Personal attitudes towards risk. 4)Social pressure.5)rules Attitude Qus results were similar without weighting; voting similar after weighing; results differ on political activism and knowledge of cholesterol.	Duffy et al (2005)
Health surveillance surveys, USA in 4 states. 1) RDD sample drawn and asked to do web survey 2) RDD sample drawn and sent paper/mail survey 3) RDD sample does CATI interview	RR% N=1143 (15%) N=836 (44%) N=2072 (40%) Web and Paper respondents were more likely than CATI to be female, white. Paper most likely to be 65+ Web most likely highly educated and live in larger households. CATI least likely highly educated Variations found in prevalence of some health conditions by mode.	Link and Mokdad, (2005)
European Social Survey (F2F)-Netherlands, compared to LISS (web) Panel, in 2008-09	ESS RR=52-68% LISS Panel RR=31%	Revilla and Saris, (2010)
Large-scale health care survey in California. Survey carried out as (1) RDD phone survey and (2) web survey	RDD CATI RR=43% Harris Interactive (web) Panel RR= 12%	Schonlau et al., (2003)
Voters' polling intentions. UK, US and Germany all have comparisons of online and offline polls in the lead up to elections.- studies are reviewed in this reference	Gibson and McAllister note that mode differences emerge responses in different countries. In Australia, authors found survey mode itself does not have any impact on how respondents answered questions about vote and party choice. No evidence that web responses are more 'honest'. Those not using the web respond differently to questions about political orientation than do those without access.	Gibson and McAllister (2009)
Labour Force Survey Online Pilot UK, 2010-11	Compare Omnibus F2F survey with 2 monthly sample from same address frame asked to do a web survey. RR for web Month (1) =8% RR for web Month (2) = 17% Month (2) added threat of sending interviewer to collect and Advanced letter addressed concerns expressed by month 1 sample.	Dawe and Wilson (2011-Appendix 9)

Appendix 6. Further details of datasets using of mixed modes at contact phase, Web plus Other within one contact.

(Mainly Sequential mixed-mode – alphabetical order by survey name). Key: → indicates the order of the sequence.

Dataset Acronym	Description	Mixed mode and when used	Other fieldwork information	Outcomes	Sources of information
Behavioural Risk Factor Surveillance System.	Non- institutionalised adults 18 years or more in 4 US states. Questions about health conditions	**Concurrent** cross-sectional Web: mail/paper: CATI 1)Random sample from RDD mailed invite to participate via web. 2) different RDD group mailed Paper QUEST. to return by post 3) Regular RDD CATI survey sample asked same questions.		(1) RR% web share=15% (2) RR% Mail/Paper share=44% **(3)** RR% share CATI=40% Found characteristics of respondents varied between mode groups by age, gender and education level. Web: 15% in 65+ age group Mail/Paper: 32% in 65+ age CATI: 23% in 65+ age Web: 15% in 18-34s Mail/Paper: 10% in 18-34s CATI: 20% in 18-34s. CATI: 58% female; web or mail 64/54%female	Link and Mokdad (2006)
Canadian Census – 2009	Test of changes for 2011 Census.	1)**Sequential**: Sub-sample posted letter requesting fill in web Census test with URL and access code→Reminder→sent Paper copy and web details→FIELD Collector will call. 2)Subsample posted Paper Questionnaire by post (with option to post back or do web return) 3)Subsample got Paper questionnaire delivered to house with options to post back or fill in on	Sub sample for (1) were selected as areas of high internet use in earlier Census. Approach (2) non-responders contained further division into 2 groups; one received Tel. Voice message.	Response rates % by modes to: By/ (1) (2) (3) Total Web 42.8 13.3 11.6 32.2 Field 15.6 16.2 15.7 15.8 Mail 18.7 47.8 51.0 29.3 Total 77.1 77.3 78.3 77.3 N 70k 30k 10k 110k	Labrasseur et al (2010)

		line→Reminder.			
DLHE Destination of leavers from Higher Education	Longitudinal follow up of cohorts of graduate leavers from higher education in England in 2002-03, 2004-05 and 2006-07. First contact at 6 months after leaving is Census of all leavers. Second Contact, wave2 at 3.5 years is sample of respondents at 6 months. Individual HEIs have discretion about which modes they decide to use at the Wave 1 (Census) data collection for each cohort, and these can and have varied by cohort. A fieldwork agency handles the data collection at Wave 2 where a more systematic sequential approach is taken.	**Sequential** mixed mode. ***In 2002-03 and 2004-05*** cohorts used at Wave 2: Web→Post/Paper→ Telephone. Email sent to those with email to do Web→Reminder after one week and another reminder a few days later. Paper/mail sent 2 weeks later to those not done online or with no email where postal address available→Reminder sent after one month of mailing →telephone where number known. ***Cohort 2006-07 W2 changes to order of sequential*** **(1)Sequential** for sample. web (by email)→ 3 text reminders →Phone →Post (2)Email to do web survey for rest.	**Instrument at W1**, 25-30 questions. Given up to 7 calls to reply in last Tel. phase. **Instrument at W2** (3.5 years) contained 6 main sections approx. 50 questions for all +more for Research degrees Large loss of data at outset. Figs for yr 06-07; 12682 of tel. numbers were unobtainable. (23% of sample) 25878 emails failed to reach destination.	**Cohort of 2004-05 leavers**: 319,260 of 430,290 leavers in 2004-05 took part (74%) at 6 months. *Wave 2: After 3.5 years.* 1)*sequential sample* 71389 -25875 sent email to do web survey, 4892 response=19% -63043 mailed, 7032 response=11% -48671 tel. 14320 response=29%. Total response 26244 (39%) 2)Of 247870 left, 89604 with emails sent email to do web survey; 15153 responses=17%. **Cohort of 2006-07 – wave 2.** 1)*Sequential sample*, 70958 -41875 emailed to do web survey, 5432 responses = 13% -54074 tele, 20254 responses =37% -37545 mailed, 3661 responses=10% Total response=29338 (44%) 2)Of 261152 left, 153631 with emails sent email to do web survey; 19725 responses=13% Responses: women higher than men; Older grads higher than younger; white grads higher than non-white.	W2 follow up of 2004-05 cohort, Shury et al (2009). HESA (2009). 2006-07 cohort- W2 Shury et al (2011)

ELS:2002 Education Longitudinal Study of 2002 USA.	Longitudinal follow up of high school leavers from 2002 in 2004 and 2006. Early contacts with students done in school with tests+paper questionnaire.	Second follow up in 2006 was web instrument. Non-responders contacted for CATI or CAPI interview **Sequential.** Web→CATI→CAPI.	**Instruments.** Main student survey wave 2 contains approx. 650 items. Further Dropout survey for some contains added 20 items.	2006 wave 88% response rate achieved in total (=14,200 interviews). In following order 1)Offer web before CATI call (20% response N=3270) → 2)Offer web after CATI call (10% N=1635) → 3)CATI call (53% N=8365)→ 4)CAPI call (9% N= 1500)	Ingels et al (2007)
Employer views survey	Cross-sectional. This annual survey of employers using Further Education training provision, first reported for year 2009-10, collects data for a performance measure of each employer's views of the quality of that provision and its responsiveness to employers' needs.	**Sequential.** Providers give SFA's survey agent employers' contact details . Employers with emails are sent URL to fill in web Questionnaire. If no email, employer posted Paper Questionnaire to fill in. Reminders sent by email. Phone/CATI follow up for employers of subset of providers who are close to reach response targets set in advance.	**Instrument.** The questionnaire consists of two classification and nine rating questions which are scored from 1 to 10, eg. from 1 'poor' to 10 'excellent':	For 2009/10 survey. 160,075 invitations to participate were sent out and 154,584 were known to have reached employers. Valid Responses received = 33033 RR = 21.4% Share of response by mode Web =23% Post=76% Automated phone=1%	Skills Funding Agency (2011). RR details from email communica-tion with LSC support desk.
ESS European Social Survey, Netherlands	Cross sectional. Feasibility testing of concurrent plus sequential mixes versus combined sequential mixed modes.	1)**Concurrent** (35% of sample who have known phone number). Tel. contact. Offer CATI, web or F2F 2)**Sequential** (on 35% with known Tel.) Offer, web only if access, or CATI→F2F if not. 3)**Sequential** (on 30% no Tel. number). F2F contact. Offer CAPI→Web→CATI.		Final Response rates 1)+3) =44% covers all population 2)+3)= 43% covers all population 2)+3) was cheaper with higher % by web and lower % by F2F	Martin and Lynn (2011)

GSS Generations and Gender Survey, Dutch Part of GGP, Generations and Gender Programme, a multi-country generation of comparable data sets	Longitudinal. Wave1 (2004) CAPI N=8,200 Wave2 (2007) CAPI N=6,200 Wave3 (2010) mixed mode N=4,300	Wave 3. **Sequential** . Send Advanced letter asking to do web questionnaire. Phone to encourage doing web questionnaire. →CATI for non responders (if known phone) →CAPI.	**Instrument.** Expected to take approx. 55 mins. to complete. Also carrying out further pilot work using mixed modes on other countries in the wider GGP programme Generations and Gender Programme, Currently on Slovenia.	Total Response rate at W3 = 72% of W2 responders (N=4,300). Share of response: Web = 55% CATI=27% CAPI=18%. Web response higher for men than women; for younger than 65+ age group; for more highly educated than lower educated. CATI response higher among 65+ age group. Web respondents had -higher mean scores on loneliness scale -lower % giving permission to contact partner -lower % giving permission to link to official register data -less willing to give firm agreement for future involvement in the survey, although did not refuse.	Liefbroer (2011 Appendix 9) www.nidi.nl
HRS Health the Retirement Study (HRS)	Longitudinal panel with members over 50 interviewed every 2 years using mix of Phone and F2F with web survey supplements some of off-years in between. Longitudinal one mode per wave for off year surveys on sub samples in 2003, 2006, 2007, 2009 (2011	**2003**. (subset of 2002 core survey) Of those with internet and saying willing to do survey, **2006**. (subset of 2004 core survey with internet + respondents to 2003 off-	**2003 Instrument.** 15 min. web survey. Topics: internet, health problem, disability, numeracy, psychosocial, housing, finances. Prepaid incentive $20 sent. **2006 Instrument**. Topics: internet, health and emotional	**2003**. 29.7% reported web use, correlated with better health & socio-demographic indicators 73.3% of Web-users said willing to do the study, more likely among older, minority and low educated Web-users. A sub set of this group were mailed an invite to do the web survey; 78% of them completed the survey. Of those with internet and saying willing to do survey, RR=81.5% Achieved N= 2197. **2006** RR=70.4% Achieved N=1352.	Couper et al (2007) for 2003 follow up of sub sample from 2002. For all off year surveys see: http://hrsonline.isr.umich.edu/

	forthcoming) Off year surveys usually put in the field in March-May.	year) **2007**. (subset of 2006 core survey with internet and not in another sub study) **2009**. (subset of 2008 core survey with internet plus earlier internet samples)	problems, social security, numeracy, expectations, prescription drug use, housing, finance and insurance. Incentive of $20. **2007 Instrument.** Topics: internet, health and emotional problems, social security, numeracy, expectations, prescription drug use, housing, finance and insurance. Incentive of $20. **2009 Instrument**. Topics: covering physical and mental health, health behaviors, psycho - social, economics, retirement. Incentive of $25.	**2007**. RR=71.6%. Achieved N=2665. **2009**. RR=77.2% Achieved N=4433.	
ISM 2008 Dutch Integrated Safety Monitor	Annual cross-sectional crime and victimization in population aged 15+. Consists of national sample and regional booster samples.	From outset of new survey. Advanced letter to do web survey. Can request Paper.→Two reminders →CATI (if number)→ CAPI if no tel. number, or no response.		National sample RR % shares 2008 2009 2010 Web 40.1 47.3 49.5 PAPI 15.4 16.1 13.2 CATI 34.0 25.3 23.8 CAPI 10.5 11.3 13.6 Web responses lower for elderly and non-western immigrants.	Buelens and van den Brakel (2011)

LFS Labour Force Survey, UK	LFS involves collection of data about all household members. Full survey collects 500 Questions from 50,000 households, 120,000 individuals per quarter. RR has been declining, RR 1998=73.5% RR 2009=56%. 2010-11 Online Pilot Survey. Approached sample that had completed 5 panel waves of LFS (W1 CAPI + W2-W5 CATI) to do a W6 online.	Asked those completing W5 if agree to do online survey and provide email.	Av intv length=28 mins To those giving prior agreement, sent Email (if have email) and Advanced letter if not to do online survey repeating many LFS W5 questions. **Instrument.** Estimate duration= 20 mins.	2008 Pilot using Omnibus survey –and address based sample asking people to do web survey RR complete=18% RR partial=2% 2010-11 Pilot using Omnibus survey and address based sample asking people to do web survey. RR Month (1)=8% RR Month (2) = 17% after addressing some concerns from month (1) Total sample response rate= 25% RR for those with email = 40% RR for non-email sample=13%	Dunn (2010) Dawe and Wilson (2011-Appendix 9). Dawe and Wilson (2011-Appendix 9).
LSYPE Longitudinal Study of Young People in England	W1 started in 2004 with cohort of 13-14 year olds. Annual contacts	W5, W6, W7. **Sequential** use of web-online, telephone, F2F **Instrument** had 14 core sections for Young Person at Wave 5 aimed to take 25 mins. Similar length for subsequent waves.	Sent to all: 1)Advance letters and emails used; 2) Unconditional cash incentive 3) Postal Reminders. Main concern to avoid refusal at Telephone phase. **Incentives:** W1-W4 = £5. W5, W6 = £8. W7= £10.	Response rates, %: *OL T F2F All* *W5* 32 44 12 88 *W6* 39 48 13 87 *W7* 45 42 13 90 *web response associated with higher education NEET associated with use of F2F. RR higher for those with email address	User Guide to W1 to W6 data at ESDS. Fieldwork agency conversation for W7. Department for Education. (2010a)

NSS National Students Survey	Annual cross-section survey. Each Higher Education Institution in England provides fieldwork agent annually with students' contact details (email and postal address) .	Since 2005, uses **sequential** approach. Students are sent Web→Email reminders →Post/mail→Telephone. **Instrument**= 22 core questions in 2011 for all students.	Requires cleaning of emails in advance.	Response rate shares by year 2005 2006 2007 2008 2009 2010 2011 Total RR% 60 56 58 63 62 63 65 N 290k 280k 340k 352k 367k 406k 405k Web share % 31 44 56 59 57 61 67 Post share % 23 21 16 14 10 9 6 Tel. share % 45 35 28 26 34 30 27 Differences by institution in these shares and by whether student is FT/PT, but not by gender.	Fieldwork agent. HEFCE (2011) Analyses of mode effects on student satisfaction scores in Surridge (2006, 2008)
SHP Swiss Household Panel Study	Annual contact with nationwide random sample of Swiss households, starting SHP1 in 1999 (SHPII in 2004). Concern to devise strategies to help reduce attrition.	Usual contact mode=CATI. Wave 12 (2010) tried offering individual 'refusers' at W12 a web questionnaire. **Sequential** CATI→web.	Household Grid. Individual interviews. W12 Refusers = (1) refusal at current wave offered CAWI after refusing CATI. (2) past refusers were recontacted and offered CAWI immediately.	W12. 4672 households and 7500 individuals completed interviews by CATI. Of the 1962 individuals not participating in CATI, 87 chose CAWI, but only 43 filled in web questionnaire. CAWI group contained more men, younger age, never married, higher educated than non-participants. CAWI group had irregular past response pattern.	Voorpostel and Ryser (2011)
YCS Youth Cohort Study	Longitudinal Successive cohorts of age 16 school leavers followed for 3 or 4 annual waves. Cohort 13 for 4 waves 2007-2010, England changed to use sequential mixed mode survey.	First use of mixed mode for Cohort 13 at W2 when **sequential** use of web-online, telephone (20 mins.), F2F (24 mins.) Used same instrument as LSYPE from YCS Cohort 13 wave 2. LSYPE and YCS surveys were run as one	Sent to all: 1)Advance letters and emails used; 2) Unconditional cash incentive 3) Postal Reminders	Cohort 13 Response rates, share% : OL Tel F2F All* W1. 69 69 W2 29 40 15 84 W3 31 47 8 86 *Issued & elig W1=10846 W2=7490 W3= 6282	Fieldwork agency conversation. Department for Education (2010b) Carpenter (2010)

Appendix 7. Details of plans for Understanding Society Innovation Panel Research project.

Innovation Panel 5 (IP5) Plans.
IP5 plans to be mixed mode (face-to-face and web), with the fieldwork due to start in March 2012. The aim will be to test the implications of using web in combination with f2f in the UKHLS Innovation Panel – with a view to implementing a similar design for the main Understanding Society survey. The use of a combination of both modes aims to maximise the number of web respondents (to achieve maximum cost savings), and will be looking at cost, non-response and measurement outcomes. A number of tests have been designed to be carried out in this IP5 panel.

For further details see presentation by Lynn

Peter Lynn, ISER, University of Essex
"Opportunities and challenges in introducing web interviewing to Understanding Society: the UK Household Longitudinal Survey"
http://eprints.ncrm.ac.uk/1917/

Innovation Panel 2 (IP2) Completed.
The second wave of the UKHLS Innovation Panel (IP2) was mixed mode, using a sequential telephone/face-to-face design (2 methods) alongside a sub-sample that were face-to-face initially.

For further information about the Innovation Panel (and mixed mode experiments in IP2 – not using the web) see:
http://research.understandingsociety.org.uk/publications/working-paper/2010-04

Appendix 8. Details of presentations of findings from ESRC's Survey Design and Measurement Initiative (SDMI) at Royal Statistical Society, Seminar, 11 Oct 2011 on *Mode effects on survey measurement*.

Mixed Modes and Measurement Error: Principal applicant: Ms Gerry Nicolaas (NatCen)
Co - applicants: Dr Annette Jäckle & Professor Peter Lynn (ISER)
Grant reference: ESRC: RES-175-25-0007

Papers are currently in the process of being written up. Details of other outputs can be found at: http://www.esrc.ac.uk/my-esrc/grants/RES-175-25-0007/read

Results from the project were recently presented at the Royal Statistical Society (RSS) on the 11th of October 2011, including a handout which summarised existing evidence including our results from our study and provided recommendations for designing different types of questions for mixed mode surveys. Presentation slides and the handout can be found at: http://www.surveynet.ac.uk/sdmi/events.asp

Individual presentations from the RSS conference are as follows and are referred to in the text of this Review Paper.
(0)http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres0.pdf

(1)Jäckle, A., Campanelli, P., Lynn, P., Nicolaas, G., Hope, S. and Nandi, A.
How and when does the mode of data collection affect survey measurement?
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres1.pdf

(2) Gray, M., Blake, M. and Campanelli, P.
The use of cognitive interviewing methods to evaluate mode effects in survey questions.
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres2.pdf

(3)Lynn, P., Hope, S., Jäckle, A., Campanelli, P. and Nicolaas, G.
The Role of Visual and Aural Communication in Producing Mode Effects on Answers to Survey Questions.
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres3.pdf

(4)Hope, S., Campanelli, P., Nicolaas, G., Lynn, P., Jäckle, A. and Nandi, A.
The role of the interviewer in producing mode effects: Results from a mixed modes experiment.
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres4.pdf

(5)Nicolaas, G., Campanelli, P., Hope, S., Jäckle, A. and Lynn, P.
Is it a good idea to optimise question format for mode of data collection?
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres5.pdf

(6) Nicolaas, G., Campanelli, P., Jäckle, A., Lynn, P., Hope, S., Blake, M. and Gray, M.
Designing Questions for Mixed Mode Surveys.
http://www.surveynet.ac.uk/SDMI/mixedmodes2011/pres6.pdf

Appendix 9. Presentations from Seminar, General Population Surveys on the web: possibilities and barriers, organised by ESRC's National Centre for Research Methods, University of Southampton, held at the British Academy, London, Nov 3rd 2011.

The presentations can be seen at: http://eprints.ncrm.ac.uk/1917/

Marcel Das, Centerdata, Tilburg University
'Innovation in data collection for scientific research: the Dutch MESS project'

Bengt Oscar Lagerstrom, Statistics Norway
'mixed mode surveys of the general population: the Norwegian experience'

Aat Liefbroer, Netherlands Interdisciplinary Demography Institute
"Mixed-mode approaches in the Generations and Gender Survey: Past experience and future expectations"

Fiona Dawe and Laura Wilson, UK Office for National Statistics
"Building Web Collection Capability on ONS Social Surveys"

Peter Lynn, ISER, University of Essex
"Opportunities and challenges in introducing web interviewing to Understanding Society: the UK Household Longitudinal Survey"

Appendix 10. Principles of Unimode approach to mixed mode question construction.

- Make all response options the same across modes and incorporate them into the stem of the survey question.
- Avoid inadvertently changing the basic question structure across modes in ways that change the stimulus.
- Reduce the number of categories to achieve mode similarity.
- Use the same descriptive labels for response categories instead of depending upon people's vision to convey the nature of a scale concept.
- If several items must be ranked, precede the ranking question with a rating question.
- Develop equivalent instructions for skip patterns that are determined by answers to several widely separated items.
- Avoid question structures that unfold.
- Reverse the order in which categories are listed in half the questionnaires.
- Evaluate interviewer instructions carefully for unintended response effects and consider their use for other modes.