Understanding Society

Working Paper Series

No. 2010-04

November 2010

Understanding Society Innovation Panel Wave 2

Results from Methodological Experiments

Jon Burton, Heather Laurie, S.C. Noah Uhrig

Contributors: Sarah Budd, Annette Jackle, Stephanie McFall, Alita Nandi, Lucinda Platt, Steve Pudney, and Emanuela Sala

Institute for Social and Economic Research, University of Essex

Understanding Society
THE UK HOUSEHOLD LONGITUDINAL STUDY

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

Understanding Society Innovation Panel Wave 2 Results from Methodological Experiments

Jon Burton, Heather Laurie, S.C. Noah Uhrig

Non-technical summary

Understanding Society includes what is known as an 'Innovation Panel' sample. This sample of 1500 households is used to test different methods for conducting the study in order to produce the highest quality data. The results from the Innovation Panel provide evidence about the best way to conduct a longitudinal survey which is of relevance for all survey practitioners as well as influencing decisions made about how to conduct *Understanding Society*. As a longitudinal study where the continued co-operation of sample members is extremely important for the study to be viable over the longer term, testing methods on the main *Understanding Society* study sample would be problematic. The Innovation Panel therefore provides a test-bed where different ways of approaching sample members to gain their co-operation and different ways of asking questions can be tried out in advance of the main study going into the field. This paper reports the results of some of this testing work that was carried out at wave 2 of the Innovation Panel in the spring of 2009 and comments on how the results from the experiments influenced decisions made for the main survey.

Understanding Society Innovation Panel Wave 2 Results from Methodological Experiments

Jon Burton, Heather Laurie, S.C. Noah Uhrig

Abstract

This paper presents some preliminary findings from the Wave 2 Innovation Panel (IP2) of *Understanding Society*: The UK Household Longitudinal Study. U*nderstanding Society* is a major new panel survey for the UK. In April 2009, the second wave of the Innovation Panel was fielded. This paper describes the design of IP2, the experiments carried and the preliminary findings from early analysis of the data. The main design features of *Understanding Society* are outlined and the design and conduct of IP2 described. The results of methodological experiments carried at IP2 are reported and the impact of IP2 on the design of the main survey reviewed.

Key words: longitudinal, survey methodology, experimental designs, data collection

JEL codes: C80, C81, C83

Contact: S.C. Noah Uhrig, scnuhrig@essex.ac.uk

Contributors: Sarah Budd, Annette Jackle, Stephanie McFall, Alita Nandi, Lucinda Platt, Steve Pudney, and Emanuela Sala

1. Introduction

This paper presents some preliminary findings from the Wave 2 Innovation Panel (IP2) of *Understanding Society*: The UK Household Longitudinal Study. *Understanding Society* is a major new panel survey for the UK. Wave 1 of the survey went into the field in January 2009 and will continue through to December 2010. Wave 2 of the survey went into the field in January 2010, with interviewers aiming to re-contact the households who were interviewed during the first wave. Thus, whilst each wave of data collection is spread over two years (comprising of 24 sample months), each individual is interviewed annually. Individuals who move out of the originally sampled households are followed, whilst they are still in the UK, and interviewed at their new address along with anyone else they are living with. New entrants into a household become eligible for interview. In this way, *Understanding Society* is dynamic and continues to represent the changing composition of households and the UK population over time.

One key feature of *Understanding Society* is the Innovation Panel. This sample of some 1500 households was first interviewed in January-April 2008, one year before the main sample of *Understanding Society* was issued. The Innovation Panel is used for methodological testing and is an opportunity for experimentation that would not be feasible on the main sample. The Innovation Panel is used to test new questions, and new ways of asking old questions.

In April 2009, the second wave of the Innovation Panel (IP2) was fielded. This paper describes the design of IP2, the experiments carried and the preliminary findings from early analysis of the data. Section 2 outlines the main design features of *Understanding Society*. Section 3 describes the design and conduct of IP2. Section 4 then reports on the experiments carried at IP2. Section 5 reviews the impact of IP2 on the design of the main survey and Section 6 concludes the paper by outlining plans for IP3.

2. Understanding Society

Understanding Society is an initiative of the Economic and Social Research Council (ESRC) and is regarded as one of the major investments in social science data resources in the UK. The Institute for Social and Economic Research (ISER) at the University of Essex provides the scientific leadership while the fieldwork and delivery of the survey data

is undertaken by the National Centre for Social Research (NatCen). This study aims to be the largest survey of its kind in the world, with a target sample size of 40,000 households across the UK, and will provide high quality, longitudinal survey data for academic and policy research.

The design of *Understanding Society* is similar to that of the British Household Panel Survey (BHPS) and other national panels around the world (for example, the German Socio-Economic Panel, the US Panel Survey of Income Dynamics and the Household Income and Labour Dynamics in Australia panel). At the first wave of data collection, a sample of addresses was issued. Households were then identified at those addresses and up to three households at the address were sampled. These sample households were then contacted by interviewers and the membership of the household enumerated. All adults aged 16 or older were eligible for interview. Children aged 10-15 were also eligible to participate in the survey and asked to fill out a self-completion questionnaire. At the second and subsequent waves of data collection, individuals are re-contacted and, if eligible, re-interviewed. If an individual has moved house within the UK, they are followed and interviewed in their new address along with anyone else living there. If there is a new entrant into the household, they become eligible for interview. Thus, *Understanding Society* is dynamic and captures household composition change over time. From Wave 2, the BHPS sample has been incorporated into the *Understanding Society* sample.

3. Innovation Panel Wave 2 design

The sample for the second wave of the Innovation Panel consisted of all productive households from Wave 1. The details of the first wave of the Innovation Panel, with descriptions of the experiments and preliminary findings, are in the *Understanding Society* Working Paper 2008-03 (Burton, Laurie, and Uhrig 2008). The sample was originally an equal probability sample of addresses in 120 areas (PSUs) across Britain, south of the Caledonian Canal. These were selected from the small user Postcode Address File (PAF). At IP1 the household response rate was 59%, with 1,489 households contacted and productive (where at least one adult gave a full interview). The individuals within these 1,489 households were issued as the IP2 sample. The questionnaire at IP2 followed the standard format used at the first wave of *Understanding Society* and IP1. The interview included:

- Household roster and household questionnaire (15 minutes per household)
- Individual questionnaire: 32.5 minutes for each person aged 16 and over (includes the collection of tracking information and data linkage consents)
- Youth self-completion: 10 minutes for each child aged 10 – 15 years
- Proxy questionnaire – 10 minutes for adults aged 16 and over unable to be interviewed

With the exception of the youth self-completion questionnaire which was paper, the interviews are conducted using the software 'Blaise'. In addition to the questionnaires, respondents were asked for permission to link to administrative data sources including education records and health records.

4. Experimentation in the Innovation Panel Wave 2

IP2 included a number of experiments, some of which related to routing and question wording within the survey instrument whilst others related to fieldwork procedures. This section outlines the experiments carried at IP2; briefly explaining the reasons for carrying them, describing the design of the experiment and giving an indication as to the initial results from early analysis of the experimental data.

4.1. Interview mode

Whilst IP1 was fully face-to-face, one of the main experiments at IP2 was the use of a mixed-mode design (Lynn et al, 2010). Households were randomly allocated to one of three treatment groups. One-third of the sample was allocated to a uni-mode design, and issued directly to face-to-face CAPI interviewers. The other two-thirds were issued to the NatCen Telephone Unit for a sequential mixed-mode of CATI and CAPI fieldwork. Of those issued to CATI, half (one-third of the total sample) were allocated to an "early transfer" treatment. Within this group, when one person in the household refused to participate or was unable or unwilling to participate over the telephone, the whole household was transferred immediately to a CAPI interviewer to carry out a face-to-face interview. The other half of the CATI group (one-third of the total sample) were allocated to a "late transfer" group. In this group, the household was only transferred to a CAPI interviewer when a CATI interviewer had tried to contact and interview all eligible adults in the household. It was expected that the "late transfer" treatment might result in a larger

proportion of CATI interviews (which are less costly than CAPI interviews), while the “early transfer” treatment might result in a higher overall response rate (as more respondents were being approached face-to-face).

Whilst the mixed-mode design was not the only experiment at IP2, it did have the biggest effect on the way the questionnaire and fieldwork procedures were designed. Translating a face-to-face interview protocol for telephone use triggers measurement issues (De Leeuw and Van der Zouwen 1988). *Understanding Society* extensively uses show cards for multi-code and choice questions. Multi-codes were translated into a series of forced-choice items, even though measurement differences between the two forms are well documented (Smyth et al. 2006; Thomas and Klein 2006). Response options were read out to respondents for other choice questions. Response options were not changed in other ways, but question stems and wording required some alteration.

CAPI and CATI operations used different organisational systems. CAPI staff operated a “craft” model where staff were issued households which they managed, interviewing all household members themselves. CATI work, which was centralised, was conducted in shifts where interviewers worked from a pool of available cases. In our experiment, households could have multiple phone numbers and cases would remain in the pool until all eligible individuals had been worked. Both features represented differences from other studies carried out by the NatCen telephone unit, so it was necessary to redesign the telephone work allocation system to accommodate the household design.

Household and individual outcome codes were used to identify cases for transfer from CATI to CAPI staff. Cases were issued to field with a record of previous contact made by the telephone unit along with the reason for the transfer. Cases were transferred as and when they were available for transfer, which meant that face-to-face work was issued piecemeal. Many cases were transferred late in the fieldwork process, so advance mailings would have occurred several weeks before CAPI interviewers called at households.

Looking at the effect of the mode design on response rates, we find that, overall, the household response rate (at least one adult gave a full interview) was 72.7%. However, this single measure hides a significant difference between sample members approached in person compared to those approached by telephone. For the CAPI sample, the response rate was 76.9%. In contrast to our original hypothesis, there was no difference in response between the other two treatments: 69.8% (early transfer) and 71.5% (late transfer). Both

mixed mode treatments resulted in slightly higher refusal rates than the CAPI treatment, as well as more “other non-response” outcomes (Table 1). It is also worth noting that the complete household co-operation rate was significantly higher in the face-to-face CAPI interviews, something which is important in the context of a household panel design where all household members are followed. While not a significant difference, the late transfer cases also performed better in this regard, probably due to the fact that more cases had been completed on the telephone initially than in the early transfer cases.

Table 1. Household response

	CAPI %	CATI (early transfer) %	CATI (late transfer) %	Total %
Complete household	61.4	49.8	53.5	54.9
Partial household	15.5	20.0	18.0	17.8
Productive household	*76.9*	*69.8*	*71.5*	*72.7*
Non-contact	5.9	6.5	4.1	5.4
Refusal	15.5	18.5	16.9	16.9
Other non-response	1.8	5.4	7.6	5.1
n	*513*	*519*	*521*	*1561*

Of individuals in productive households, 84.2% gave a full interview (Table 2). This proportion was higher with CAPI (86.3%) than with either mixed-mode treatment (82.6% for early transfer, $p<0.05$ and 83.6% for late transfer, $p<0.1$). Of all completed individual interviews, 78.3% were completed by phone in the early transfer group and 79.8% in the late transfer group.

Table 2. Individual interview response

	Face-to-face %	CATI % (early transfer)	CATI % (late transfer)	Total %
Complete individual	86.3	82.6	83.6	84.2
Partial individual (break-offs)	0.5	1.7	0.7	1.0
Proxy	5.5	3.9	2.3	3.9
No contact	1.0	1.2	4.0	2.0
Refusal	5.4	6.4	6.7	6.1
Other non-productive	1.3	4.5	2.6	2.7
n	*735*	*665*	*700*	*2100*

The CAPI unimode design obtained a higher response rate than either of the mixed-mode designs. The advantages of CAPI were both in achieving a higher proportion of household interviews and in achieving interviews with a higher proportion of the individuals within those households.

The two mixed-mode designs performed similarly to each other in terms of response rates. We cannot conclude that either mixed mode design is preferable to the other or that either is capable of equalling the response rates achieved by CAPI alone.

Mixed-mode designs are still under consideration for future waves of *Understanding Society* and further experimentation may take place, including with web survey methods.

In terms of lessons learned and trying to understand the effects of the sequential mixed-mode design used, the piecemeal transfer of cases from CATI to CAPI could have de-motivated field staff so transferring all cases within a PSU at once may have been more efficient. A second advance letter to transferred cases before the field interviewer's call may also have been helpful in increasing response at that stage. Additionally, it was necessary to extend the field work period for the two mixed-mode groups in order to fully work cases transferred to CAPI following delays in the transfer.

4.2. Effects of Mode on Measurement

The mode experiment was also used to examine the effects of mode on measurements within the survey. Different types of questions were analysed separately: Continuous variables were analysed using *t*-tests to compare the means of the two groups (CAPI and CATI); Chi-square tests of differences in proportions were used to compare the

distributions of categorical variables including dichotomous outcomes. The findings presented here do not correct for non-response within mode so we cannot rule out sample differences in the findings (mode effects on non-response) as distinct from mode effects on measurement. Here we discuss significant mode effects by question type. Below in Section 4.3, we describe mode effects in the context of measuring ethnic and personal identity specifically.

Continuous measures. Of 37 variables with numeric responses, we found three significant differences at the *P*=0.05 level between modes, or approximately 8% of continuous items. CATI respondents were more likely to report a higher number of miles driven in last 12 months than CAPI respondents; CATI respondents report higher hourly rates of pay (for people paid by the hour) than CAPI respondents; CATI respondents report giving more to charity in the last 12 months than CAPI respondents. The first two may result from sample differences by mode, whereas the third may be reflective of social desirability effects. Though the social distance is greater via CATI and therefore people are believed to be more honest, this might imply CATI respondents would report giving less to charity, not more. It may, on the other hand, have to do with the familiarity of returning the same interviewer to CAPI respondents as at wave 1. If greater familiarity leads to more honest reporting, we might expect this effect as the CATI interviewer will be new to respondents.

Yes/No variables. Of 41 variables, we found six significant differences between modes, or approximately 15% of yes/no items. CATI respondents were more likely to report living with both biological mother and biological father from birth to age 16; CAPI respondents were more likely to report a long-standing illness or disability; CATI respondents were more likely to report doing sport in the last 4 weeks; CATI respondents were more likely to believe that people in the UK will be affected by climate change in next 30 years; CATI respondents were more likely to have given unpaid help to an organisation or charity in last 12 months; CATI respondents were more likely to have donated money to charities or other organisations in last 12 months; CATI respondents were more likely to report saving regularly. As with continuous measures, the first two may represent sample selectivity across modes whereas the latter items may represent social desirability effects in measurement.

Categorical variables. Of 88 variables tested, we found 28 variables with significant mode differences, or approximately 31% of categorical items. Across a range of categorical variables, the CATI respondents appear to report more socially desirable behaviours: CATI respondents report visiting libraries and historic sites more frequently than CAPI

respondents; CATI respondents seem to be less likely to report leaving their TV on standby overnight, less likely to keep their tap running when they brush their teeth, more likely to put on clothes when cold rather than turning up the heating, more likely to not buy something because of its packaging, more likely to buy recycled paper products, more likely to take their own shopping bag when shopping, and more likely to consider the environmental impact of their transportation choices; CATI respondents are more likely to report feeling calm and peaceful in the last 4 weeks, and more likely to report having "a lot of energy"; CATI respondents are more likely to report a "very" or "fairly strong" support for a political party; and CATI respondents seem to take part in leisure activities generally more frequently than CAPI respondents.

A number of other significant effects occurred at categorical variables which do not seem to conform to any pattern, including areas such as specific leisure activities engaged in, the degree of access to various services including shopping, health care, and use of public transport generally. Employed CAPI respondents were more likely to report working for a private firm. Also, the sorts of benefits received varied between modes though the responses did not seem exhibit any discernable pattern although CAPI respondents were more likely to receive pension related benefits and child related benefits rather than CATI respondents which may be indicative of sample selection differences across modes.

In summary, mode effects in measurement do seem to be related to the format of the measure. Only 8% of continuous items where a number was requested seemed to be affected by mode while about one-third of categorical variables were affected. We anticipate further examination of these measures to explore the nature and reasons for these mode effects.

4.3. Measurement of identity

There is a need for good measures of ethnic identity that can be carried in large quantitative surveys, thus the IP tested some prototype questions on identity. The identity questions tests were based on versions carried in the Citizenship Survey in 2007/8. In total, respondents were given 13 identity categories with which they could identify (or not). We find statistically significant differences in responses by the mode of interview only for a few questions (see Table 3). It is not immediately apparent why these mode differences are observed and further work will be conducted to examine the reasons for the mode effects observed in these items.

Table 3 Differences in responses by mode

	Significantly more likely if interviewed CAPI rather than CATI	Significantly more likely if interviewed CATI rather than CAPI
Important to your sense of who you are	* Your racial and ethnic background	* Political beliefs * Level of education
Not very important to your sense of who you are	* Level of education	* Sexual orientation
Not at all important to your sense of who you are	* Political beliefs	* Family
Don't Know	* Occupation/profession * Father's ethnic group	

In addition to examining mode effects in the measurement of identity, a split-ballot question wording experiment was carried to evaluate whether the word *occupation* or *profession* seemed more salient to respondents in relation to their own identity.

There were no statistically significant differences between the responses to the two versions nor were there differences within sub-groups based on whether employed, retired, in paid employment, unemployed, gender, age, ethnicity, whether born in UK and mode of interview.

4.4. Use of showcards and survey measurement

To look at the independent effect of visual aids on response distributions as distinct from mode effects, CAPI respondents were randomly allocated into groups interviewed with showcards and those interviewed without showcards. Of 35 categorical variables, a little more than a quarter (9 variables) showed significant differences between the two groups.

Two items exhibited some evidence of primacy and recency effects. First, respondents were asked which arts related leisure activities they had done in the prior 12 months. Although the item is a 'choose all that apply' question, those with showcards were more likely to chose the first option (17% with showcards vs 11% without showcards, $p < 0.05$), whereas those without showcards were more likely to chose the last option presented to them (61% with showcards vs 72% without showcards, $p < 0.05$). Second, respondents were asked to report the frequency of engaging in their stated leisure activities. For those who indicated that they participated in dance, theatre or acting, a higher proportion of respondents interviewed with showcards said they did this at least once a week (option 1, 35% with showcards vs 13% without showcards), whereas a higher proportion of

respondents interviewed without showcards chose an option lower down the list (option 3, 55% of non-showcards vs 24% those with showcards).

A third item suggests that respondents without visual cues choose the most 'moderate' option over more specific options. The visual cue may be implicated in the response process inducing greater thought in the recall of events, for example. Respondents were asked how frequently they attended different arts events. We found significant differences in the proportion of respondents choosing the frequency "less often than once a month but at least 3 or 4 times a year" in 3 of the 8 arts events queried whereas a higher proportion of respondents with showcards selected the more specific options of "twice in the last 12 months" or "once in the last 12 months" over other categories. Primacy type effects were observed in two of the arts events where the first option was chosen more frequently by showcard respondents.

In the remaining set of showcard/no showcard items examined, significant differences between treatments occurred but no discernable pattern emerged from the marginal distributions. Further work will be conducted to explore the nature and causes of differences in measurement due to visual cues.

4.5. Incentives

As in IP1, IP2 included an incentive experiment. At IP2 we wanted to see the effect of reducing the level of the incentive at a second wave of a longitudinal survey. There was speculation that it might not be the level of incentive that is related to participation at a second or subsequent wave of a panel survey, but whether there is any incentive at all. In addition, there was a conjecture that participants might not remember that they had received an incentive at IP1, or if they did they may not remember the amount.

The allocation to experimental group depended in part on the IP1 incentive group. At IP1 there were three incentive groups; £5 for each participating adult, £10 for each participating adult and £5 for each participating adult increasing to £10 for each adult for a whole-household response (all eligible adults participating). For all those who had received £5 at IP1, all were allocated to a £5 group at IP2. A random half of households who had received £10 at IP1 were allocated to receive £5 at IP2, the remaining half staying at £10. Likewise, the group of households who initially received £5 but where this increased to £10 for a whole-household response were also randomly allocated to two groups, with one getting £5 and the other group retaining the £5-to-£10 treatment.

Table 4 Incentive groups

IP2 Incentive group	IP1 incentive type	IP2 incentive type
1	£5	£5
2	£10	£10
3	£10	£5
4	£5-to-£10	£5-to-£10
5	£5-to-£10	£5

The incentive was pre-paid and sent to the sample member with the advance letter, informing them of the new wave of interviews. The incentive is, therefore, unconditional; even if the sample member did not take part at IP2, they were still able to keep the incentive. An advance mailing (either a card or a letter, see Section 4.5) with the incentive was sent to all eligible household members at IP2 – including those who had not participated at IP1. Everyone in the household was allocated to the same incentive group, and split-off movers from a household were allocated to the same incentive group as the original household.

We found that there was no statistically significant difference between the response rates across the incentive groups. Two of the groups who received the £5 incentive had the lowest household-level response rates; 67.6% for those who received £5 at both IP1 and IP2 and 69.5% for those who had been in the £5-to-£10 group at IP1. Response for the group who had received £10 at IP1, but this had reduced to £5 at IP2 was similar to the group who had received £10 at both waves (72.3% and 72.8%). These two groups had just slightly lower response than the group who had received £5-to-£10 (73.2%), although none of these differences were statistically significant.

This overall finding masks the fact that sample members were contacted using different modes; either a unimode face-to-face design or one of two mixed-mode sequential designs as described in section 4.1 (telephone followed by re-allocation to face-to-face for non-respondents). To look at the effect of the mode and incentives on household-level response rates, we split the sample to enable us to analyse (i) those who were allocated to the face-to-face mode, (ii) those allocated to the telephone “early transfer” group and (iii) those allocated to the telephone “late transfer” group (see section 4.1 above). Table 5, below, shows the household-level response rates (where at least one adult took part in the survey).

Table 5: Household-level response rates, percentage of eligible households with at least one adult interview

	£5 same	**£5** (ex £10)	**£5** (ex £5 to £10)	**£5 to £10** (same)	**£10** same	***n***
	%	%	%	%	%	
Overall	67.6	72.3	69.5	73.2	72.8	*1543*
Face-to-face	72.0	79.1	82.4	73.6	76.1	*510*
Tel – early transfer	61.9	69.8	67.4	67.8	69.2	*510*
Tel – late transfer	68.9	68.2	60.0	78.2*	73.0*	*523*
N	*488*	*260*	*266*	*261*	*268*	

* = significant at p<0.05 compared to the lowest response rate for that mode

Household response rates were generally highest in the face-to-face group, lower in the mixed-modes groups. The only significant relationships between the incentive type and response were for the sample that were in the "late transfer" telephone group, where those receiving £5-to-£10 and those receiving £10 in both IP1 and IP2 were more likely to respond than those receiving £5-to-£10 at IP1 and just £5 at IP2. Within incentive groups, those who received £5 at IP2, having had £5-to-£10 at IP1, were significantly more likely to participate if interviewed face-to-face, compared to being in the 'late transfer' group. There was no significant difference between these incentives groups when the two telephone modes were combined.

To simplify the incentive groups they can be collapsed to form three groups; those receiving £5, £5-to-£10 and £10. Overall there was no significant difference between these incentive groups in terms of obtaining a household-level response. Within the different interview modes, however, those in the 'late transfer' group were more likely to be productive if they had received £5-to-£10 (78.2%) rather than £5 (65.5%), while receiving £10 (72.2%) was not statistically different to receiving £5. There was no significant effect of the incentive value among the face-to-face or the 'early transfer' groups.

Another way of collapsing the incentive groups is to note the change from the previous year; did the participant receive the same as at IP1 or less. Once again it is in the 'late transfer' group where there is some significant difference, with those receiving the same value of incentive as the previous year being more likely to participate (71.9%) than those who received a reduction to £5 (62.9%).

At the individual level, 84.2% of eligible adults (within co-operative households) gave a full interview. Again, there was no significant difference between the incentive groups. Table 6, below, shows the percentage of eligible adults who gave a full interview in each of the incentive and mode treatment groups. Within the face-to-face and the mixed 'early transfer' mode, there was no difference in individual response by incentive group. Within the 'late transfer' group, however, those who received £5 at IP2 after receiving £10 at IP1 and those who received £5-to-£10 had a significantly higher response rate that those who had received £10 at both IP1 and IP2.

Table 6: Individual-level response rates, percentage of eligible adults giving a full interview within co-operating households

	£5 same	**£5** (ex £10)	**£5** (ex £5 to £10)	**£5 to £10** (same)	**£10** same	***n***
	%	%	%	%	%	
Overall	83.7	86.5	83.6	84.8	82.8	*2100*
Face-to-face	85.2	83.0	87.0	88.3	89.1	*735*
Tel – early transfer	81.3	88.1	79.5	79.4	85.0	*665*
Tel – late transfer	84.1	89.1*	84.2	86.0*	75.0	*700*
N	*650*	*363*	*372*	*348*	*367*	

* = significant at p<0.05 compared to the lowest response rate for that mode

Collapsing the incentive groups we find that overall there was no significant difference between those receiving £5, £5-to-£10 or £10 in terms of obtaining a full response. Within the different interview modes, however, those who were contacted face-to-face were significantly more likely to respond if they received £5-to-£10 or £10 (both 75.2%), compared to £5 (66.5%). In the 'late transfer' group, this effect was reversed, with the £5 group being significantly more likely to participate (69.9% compared to 61.2% for the £5-to-£10 and 60.0% for the £10 groups). There was no significant effect of the incentive value among the 'early transfer' group. There was no significant difference in response either overall or within interview mode between those who had received the same and those who received less as the previous year.

To summarise, the level of incentive was not significantly related to either household or individual-level response, a finding which is consistent with the view that it is the fact of having an incentive which is most important rather than the amount of the incentive. It is

also interesting that in the context of a longitudinal study where the incentive level was reduced relative to the previous wave, this did not have a statistically significant effect on response as might have been expected even though response rates were generally lower. However, for the face-to-face sample, a higher incentive was likely to lead to higher response at the individual level but not at the household level. This is important for the household panel design, where all individuals are followed and having complete data for every year of the study is critically important. For the 'early transfer' mode there was no statistically significant relationship with incentive level. Those who were in the 'late transfer' mode, where the telephone interviewer was likely to have needed to make more calls, a higher incentive was associated with higher response at the household level, whilst a lower incentive seemed more successful at the individual level.

4.6. Advance materials

There is interest in whether various types of advance communications can influence household and individual survey response rates. The advance materials experiment compared the effect on household and individual response outcomes of two types of advance materials; (i) a card in a greeting card format and (ii) a formal letter on printed letterhead stationery with the survey logo (see Appendix A). The cards and letters were equal in length, text, incentives, and information related to legitimacy of the study and privacy concerns. Thus, differences are in terms of appearance and format only. The envelopes for both cards and letters were personally addressed. The letter was also personally addressed internally.

Households were randomly assigned to receive cards (n=781) or letters (n=785). The correspondence was sent to each adult in interviewed households at IP1 and to rising 16 year olds who would be eligible for an adult interview at IP2. This assignment was independent of other experimental manipulations.

The outcomes examined were contact at IP2 for those interviewed at IP1 and refusal at IP2 for those who had been contacted. We expected that any effects would be with refusal rather than contact. Among eligible households, 3.8% of the letter group and 3.0% of the card group were not contacted. Among contacted eligible households, 18.6% of the letter group and 16.8% of the card group refused.

For individual-level outcomes, there was no difference in the percentage not contacted by type of advance material (1.2% letter, 1.1% card). There was also no difference in the percentage refusing an interview (5.3% letter, 4.8% card).

Some potential interaction effects on individual refusal were examined. There were no significant or consistent patterns related to a shift in the delivery of incentives. There was a significant interaction ($p = .004$) of advance materials and region. In Scotland, 14.6% of persons receiving letters refused an interview, vs. 3.1% getting the card. However, the sample size within each experimental group in Scotland and Wales are quite small and so this result should be treated with some caution.

Taken together, this provides slight evidence in favour of using cards rather than formal letters as advance mailings. Notwithstanding this finding, a decision was taken for main stage to use formal letters rather than cards. From wave 2 onwards, respondents are sent a tailored advanced mailing depending on their interview outcome at the previous wave and their sample status at the current wave. All communication with respondents is personalised and overprinted with their name and address details. Continuing respondents would receive one type of personalised mailing, non-participating respondents would receive a different mailing, and new 16 year olds eligible for an adult interview at the current wave would get a third tailored letter inviting them to participate. In terms of feasibility on a large sample such as the main stage of *Understanding Society*, this sort of personalisation of cards could not technically be obtained for the available budget. As the cards did not present any major response advantage for the sample as a whole and the costs of the cards was significantly higher, the letters were used from wave 2.

4.7. Measurement of subjective satisfaction questions

It is well established in the survey methods literature that the way you ask a question may have a big influence on the answer that you get. Recent research using the BHPS revealed the significant impact that a change in survey features can have on response behaviour (Conti and Pudney 2010) and IP2 is being used to examine this issue in greater detail, using a randomised experiment approach.

At IP2, scales measuring the participant's satisfaction with their health, family income, leisure, job (if applicable) and their life overall were used to analyse how changes in the way the question is presented affect response. As well as the mode experiment, described above, households were independently assigned to treatment groups formed by varying

question design, delivery and position within the interview. All eligible adults within a household received the same experimental treatment. Table 7, below, sets out the treatment groups at IP2 for this experiment.

Table 7: Experimental treatment groups

Group	Interview Mode	Response Scale	Timing of Question	Sample *n* FTF	Tel
Job satisfaction and Life satisfaction questions					
A	CASI	Full labels: 1-stage	Late	167	0
B	CASI	Polar labels	Late	136	0
C	CAPI+s'card	Full labels: 1-stage	Late	55	0
D	Tel + F2F	Full labels: 2-stage	Late	78	301
E	F2F	Polar labels	Late	69	0
F	Tel + F2F	Polar labels	Late	68	323
G	F2F	Full labels: 1-stage	Early	65	0
H	Tel + F2F	Full labels: 2-stage	Early	67	145
I	F2F	Polar labels	Early	56	0
J	Tel + F2F	Polar labels	Early	69	167

The standard question-set involved an initial question, "How dissatisfied or satisfied are you with the following aspects of your situation: (a) your health; (b) the income of your household; (c) the amount of leisure time you have" (see Appendix B for the standard question-set). This was then followed by the question, "Using the same scale, how dissatisfied or satisfied are you with your life overall?" The responses were measured using a seven-point scale. Groups C, E, G and I had the response categories presented visually on a show-card. Groups D, F, H and J had no visual cue, the question was purely oral. Groups A and B were presented with the computer and asked to complete the question by themselves (CASI). Those participants who were employed or self-employed were asked, "All things considered, which number best describes how dissatisfied or satisfied you are with your job overall?"

As well as the different delivery methods (showcard, oral, CASI), the response scale was presented in three different ways. For groups A, C and G each of the points on the seven-

point scale was labelled; *7 Completely satisfied; 6 Mostly satisfied; 5 Somewhat satisfied; 4 Neither satisfied nor dissatisfied; 3 Somewhat dissatisfied; 2 Mostly dissatisfied; 1 Completely dissatisfied*. This is the "Full labels: 1-stage" response scale. Participants in groups D and H were also able to answer using the fully-labelled scale, but the question was broken into two parts, with the participant first being asked, "How dissatisfied or satisfied are you with your (life/job) overall? Would you say that you are… (1 Dissatisfied; 2 Neither dissatisfied not satisfied; 3 Satisfied)." If the participant indicated that they were either dissatisfied or satisfied they were asked the follow-up question, "Are you somewhat, mostly or completely (dissatisfied/satisfied) with your (life/present job) overall? (1 Somewhat; 2 Mostly; 3 Completely)". This is the "Full labels: 2-stage" treatment. The third treatment, for groups B, E, F, I and J, is the "Polar labels" option. In this treatment group, only the labels for the extreme points on the scale were conveyed; i.e., completely dissatisfied and completely satisfied.

The timing of the job satisfaction question in the questionnaire was constant for all participants who had a job, following a section about employment or self-employment. The placement of the life satisfaction questions could be varied, however, and were either asked early on in the interview (about a quarter of the way through) or late in the interview (very near the end).

Analysis (Pudney, 2010) shows that the placement of the life satisfaction question (early vs. late) was not associated with any significant shifts in the response distribution. In common with (Kalton and Miller 1991), it was found that the labelling of points on the response scale had a more significant impact for female participants than for males. Third, the method of question delivery (CASI v CAPI and telephone v face-to-face) has a significant impact for women in relation to life and job, but for men it is only life satisfaction in the CASI v CAPI where this effect is significant. Finally, there is a significant difference between the response distributions produced by the 1-stage and 2-stage question structures, this was more significant for women than for men. Overall, these results suggest that women's response behaviour may be sensitive to multiple aspects of the interview process. In terms of the form of this impact, we find that polar-point labelling, face-to-face interviewing and 1-stage question design tend to reduce the mean level of reported satisfaction relative to the alternative of full labelling, telephone interviewing and 2-stage question wording. The effect is less clear for CASI v CAPI, where the differences go in different directions. However, these relatively uniform differences in the mean conceal complex differences in the underlying response distributions.

4.8. Measures of change

Measuring how people's social and economic circumstances change over time is a key purpose of household panel surveys. Levels of change are, however, often over-estimated in panel surveys and responses to a question are often not consistent across interviews, even if the respondent's situation has not changed (e.g. Lynn and Sala 2006). Dependent interviewing (DI) has been proposed and used as a tool to reduce spurious change and increase the longitudinal consistency of responses. However simply reminding respondents of a previous response or adding an edit check does not necessarily solve problems with the underlying question (see Jäckle 2008). In addition, researchers frequently voice concern that reminding respondents of responses they have given in previous interviews may encourage acquiescence (i.e., agreement with a statement regardless of its content) which may swap the problem of spurious change for the problem of spurious stability.

Methodological studies on measures of change from panel surveys so far have either documented the extent of spurious change and seam effects (Kalton and Miller 1991; Moore 2004; Moore et al. 2009), or studied the effectiveness of dependent interviewing at reducing both problems (for an overview, see Jäckle 2009). Research on the causes of longitudinal inconsistencies appears to be lacking. This is in stark contrast to cross-sectional surveys, where a large body of literature has studied response effects and factors affecting the reliability of questions (see Sudman, Bradburn, and Schwarz 1995). There is a real and urgent need to understand the causes of longitudinal inconsistencies in order to be able to improve survey questions and hence the reliability of estimates of change.

To address this research, a theoretical framework was developed and a series of split-ballot experiments was designed, to be carried out on IP2 and IP3. The ultimate aim is to understand the processes that lead a respondent, whose situation has not changed from one interview to the next, to give a different response to a survey question in different interviews. This is likely to depend on the probability that the response process is similar in the interviews at time *t* and *t+1* and this occurs under the following conditions:

(i) when the terms and concepts used in the question are less vague,
(ii) when the nature of expected answers are clearly defined,
(iii) when the inferences respondents can draw about the intended meaning of a question are consistent across waves (questionnaire context effects),

(iv) when the stimulus provided by the interviewer in administering a question is consistent across waves.

Additionally (v), it is likely that question wording (use of dependent interviewing) affects the extent to which respondents agree with previous information presented to them.

To test these hypothesis several split-ballot question wording experiments were designed. An overview of the experiments including question wording and treatment groups is reported in the Appendix C. Respondents were randomly allocated to two treatments groups and were administered the wave 1 IP question or the experimental question. The choice of the questions to experiment with was subject to a certain number of constrains including considerations on power of the analysis (e. g. sample size of the eligible respondents to a certain question), questionnaire stability and questionnaire time and costs. Some of these experiments were digitally recorded.

To date only the IP2 data are available. Except for the experiment with implicit vs. explicit requests for dates, all hypotheses underlying the experimental work in IP2 require two waves of data from IP2 and IP3. Only the experiment on the collection of date information is therefore reported here.

Asking respondents explicitly for a date (e.g. 'in which month and year...') indicates that an accurate date is required. Asking implicitly for a date (e.g. 'when...' / 'how long...') may suggest that an approximation is good enough. Respondents may be more likely to use estimation strategies and heuristics, instead of attempting to recall a precise date.

With the experiments designed for IP2, date information collected with implicit requests can be expected to be less accurate. Durations between the event and the current interview are more likely to result in 'heaping' (i.e. disproportionate number of durations of 6 months, 12 months, etc.) and dates are more likely to falsely fall before the previous interview (for questions about events since the previous interview)[1].

To test these hypotheses, experiments were carried out with date questions relating to disability, paid work and regular saving (see appendix C for question wording). Treatment Group A was administered the implicit date questions whereas Treatment Group B was asked the explicit date. Preliminary analyses suggest that the implicit questions yield date information of poorer quality than the explicit questions, as expected as the durations between the event and the current interview are more likely to be heaped. Figure 1 illustrates this for a question asking for the number of years since the respondent joined a

[1] Due to the small number of eligible cases, we are not able to test this hypothesis.

pension scheme (implicit: “How long have you been a member of this pension scheme?” vs explicit: “In which year did you join this pension scheme?”). Compared to the explicit date question, the implicit question produced a disproportionate number of durations that are multiples of 5 years[2].

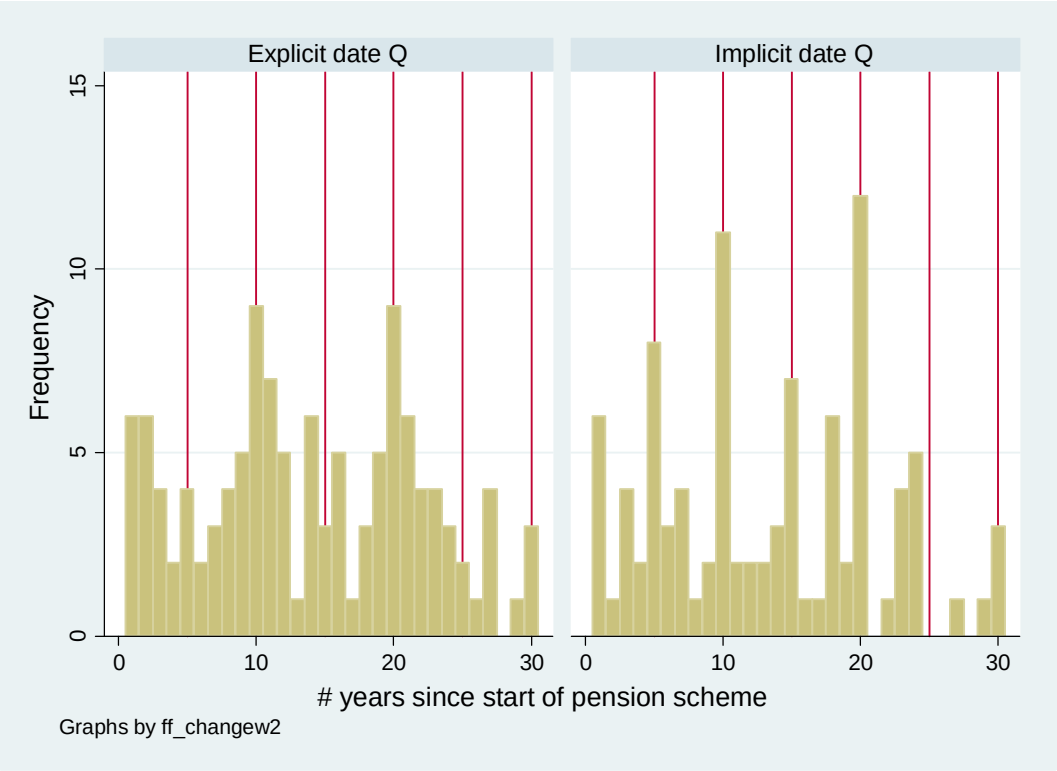

Figure 1. Number of years since joining a pension scheme

4.9. Panel conditioning

“Panel Conditioning”[3] is a term which suggests a systematic effect of panel participation on either response behaviour or on the behaviour targeted by survey measurement (Waterton and Lievesley 1989). If conditioning effects are large, then the utility of panel data is compromised (Cantor 1989; Cantor 2008). A number of studies have demonstrated that survey participation induces such diverse activities as voting, immunization and diagnosis of arthritis (Bartels 1999; Battaglia, Zell, and Ching 1996; Clausen 1968; Kraut and McConahay 1973; Wilson and Howell 2005; Yalch 1976). Other work suggests that panel participation affects opinion formation. For example, Sturgis and colleagues (2009)

[2] We conducted a similar analysis for the other two questions (available upon request to the authors) and, overall, the same pattern emerges.

[3] Other terms used to name the phenomenon include, but are not limited to, ‘Time in Sample Bias’, ‘Rotation Group Bias’, ‘Repeated Interviewing Effect’, ‘Panel Membership Effect’, and ‘Panel Bias’. These terms often provide a clue as to how the author designs their analysis and may also indicate particular assumed causal mechanisms.

observe reduced variances in opinion items over several waves of a panel, an increase in wave-lagged correlations over time, and a decrease in "Don't Know" responses. On-going participation has also been shown to enhance accuracy in survey reports in other ways (Ferber 1953; Sturgis, Allum, and Brunton-Smith 2009; Waterton and Lievesley 1989). An alternative strand of work suggests that conditioning could inhibit accuracy in reporting (Bailar 1975; Cohen and Burt 1985; Ghangurde 1982; Kemsley 1961; Mooney 1962; Neter and Waksberg 1964; Silberstein and Jacobs 1989; Turner 1961). The extent to which on-going participation in a panel survey affects panel responses and subsequent respondent behaviour, therefore, remains an open question.

Disparate findings in the literature could occur because analytic designs typically used are not ideal for disentangling the factors which affect survey data quality (Holt 1989). One strategy is to compare fresh cross-sectional data to concurrent panel data. Such designs are usually deficient because they rarely contain the same sample survey conditions, survey design, sample design and procedures, and identical questions (Cantor 2008; Holt 1989; Menard and Elliott 1993; Sturgis, Allum, and Brunton-Smith 2009). Thus, work on panel conditioning tends to be atheoretical, with little attention to the mechanisms which might generate effects. An experiment carried at IP1 and IP2 of *Understanding Society* focuses on exposure to survey content as a potential mechanism leading to conditioning effects. At IP1 a random half of the sample was asked to self-report their height and weight, and to provide their opinion about the existence and likely effects of climate change. At IP2, the entire sample was asked to self-report height and weight, and to report on the frequency of engaging in various environmentally beneficial behaviours[4]. If exposure to survey content enhances reporting accuracy, then known biases in self-reports of height and weight should be attenuated due to conditioning. And, if prior survey content induces a crystallisation in opinions and behaviour, then conditioned respondents should exhibit less variant environmental behaviour as opposed to unconditioned respondents.

Results: Environmental Behaviour. We formed a single standardised multi-item environmental behaviour scale. Unconditioned respondents seem to have a *higher* average inter-item correlation on the scale ($\mu_{unconditioned} = 0.16$, $\mu_{conditioned} = 0.08$, $p < .0001$). A comparison of variances for the scale computed for each group shows that the

[4] Items include: Leaving TV on stand-by for the night; Switch lights off in rooms not used; Keep tap running while brush teeth; Add cloths rather than turn-up heat when cold; Not buy product with too much packaging; Buy recycled paper products; Use own shopping bag; Use public transport rather than car; Walk/cycle short journeys; Car share; Take fewer flights were possible. With response options: Always, Very Often; Quite Often; Not Very Often; and Never. A scale combining these items is standardised.

conditioned sample has a significantly lower variance (F = 1.22, df = 862, 833, p < 0.01). This means that although unconditioned respondents tend to answer the items unidirectionally, conditioned respondents are more internally consistent in their answers. There are those who ignore climate change and resist engaging in pro-environmental behaviour and others who adhere to climate change ideas and are avidly green in their actions. Though the unconditioned samples should include a mixture of both types of respondents, we used IP1 data to split the conditioned sample into those who frequently consider the carbon implications of their actions and those less often considering them. Not surprisingly, the carbon aware are significantly more "green" than others (p < .001, μ_{carbon} = -0.12, μ_{others} = 0.16), though there is no difference between these groups in terms of average inter-item variance nor variance in the scale itself.

Results: Height and Weight. Conditioning seems to reduce non-response of self-reported weight (p < 0.10) and encourages recent weighing among women (p < .10). Overweight conditioned women report a heavier weight (p < 0.05), while conditioned shorter men report a shorter height (p < 0.10). The effect on self-reports of weight meant that conditioned women were more likely to be categorised as either overweight or obese relative to normal weight(p < 0.10), and obese relative to overweight (p < 0.10). These findings imply that conditioning could enhance accuracy in the reporting of height and weight as a typical finding is that women tend to under-report their weight while men tend to over-report their height (Spencer et al. 2002).

5. Implications for survey design in the future

The Innovation Panel has two main aims. Firstly, to provide a test-bed for data collection procedures and question design which feeds directly into decisions on the design and conduct of the main *Understanding Society* study. Secondly, the IP is intended to add to the methodological literature with regard to longitudinal survey methodology, in particular through the use of randomised experiments aimed at key areas of concern for data quality including response rates and attrition, measurement error, mode effects, bias, and panel conditioning.

In terms of the experiments carried at IP2, some, such as the incentive and card vs letter experiments provided evidence that directly influenced decisions made for wave 2 of the main study. At wave 2 of the main study, respondents receive a reduced incentive relative

to wave 1 (from £10 to £5) and the results of the IP2 experiment provide some reassurance that this will not adversely affect response rates as might have been assumed. Similarly, receiving a letter which is more flexible to personalise and less costly to produce than a gift card was shown to have no significant effects on response rates for the sample as a whole. These experiments therefore enabled difficult decisions to be taken which resulted in significant cost savings at wave 2 while not impacting on the quality of the study. They are also experiments which other survey practitioners should find useful when balancing the cost of strategies for maintaining response rates in a longitudinal survey against other areas of expenditure.

The mixed mode experiment has a longer time frame for *Understanding Society* as the current plan is to have a mixed mode wave at either wave 4 or wave 5 of the study. Testing the variants for the sequential mixed-mode approaches, early vs late transfer of cases from CATI to CAPI, provides us with valuable information on which to base any future sequential mixed-mode design. It also adds to the existing literature in this area where an experiment of this type has never been carried out in a longitudinal context.

The measures of change experiments will not come to full fruition until after IP3 when we have longitudinal measures for the items concerned. However, the experiments with the collection of date information already point to the importance of good question design which is unambiguous, clear, and specific to produce high quality data. Similarly, the satisfaction scale experiments provide clear evidence of the best way to ask and present these types of questions to respondents and on *Understanding Society* we have adopted a fully labelled scale which is self-completed. When we come to carrying out a mixed mode wave, this experiment will allow us to assess which of the measurement options for CATI are the most robust and likely to provide the greatest longitudinal comparability relative to self-completion collection.

Finally panel conditioning is an issue which is a concern for all longitudinal studies but one which it is very difficult to gain firm evidence about. As a result the literature is rather mixed in its findings but conditioning is generally perceived as a potential effect which, if present, will reduce rather than enhance data quality. From the IP2 experiment on just a few measures, we do see some effects of having been asked the same questions previously. In the case of the height and weight questions the effect seems to be a positive rather than a negative one. Respondents who had been asked the questions at the previous wave gave more accurate responses than those being asked the questions for the first time.

The IP is, to date, meeting its aims of generating important and useful methodological findings which are having a direct effect on what is a major study in the UK. Now that *Understanding Society* has bedded down, the IP is being opened up to other researchers outside ISER who can, through an annual call for proposals, make suggestions for their own methodological experiments. The IP is therefore developing into a resource not only for *Understanding Society* but for the research community as a whole.

6. Plans for IP3

IP3 will go into the field in mid-April 2010 and, unlike IP2, will be fully face-to-face. Sample that were not contacted at IP2, or who gave 'soft' refusals will be issued for IP3. A number of the experiments to be carried at IP3 are continuations of experiments carried at IP1 and/or IP2 and longitudinal in nature. The experiments to be carried are:

i. Advance materials. An experiment to gauge whether the content of the between-wave mailing influences responses in the interview. There are two different versions of the "Participants Update" report which is sent to sample members between IP2 and IP3. The report contains five short descriptions of some initial analysis of IP2 data. The only difference between the two reports is that the section about financial wellbeing in one report is replaced by a section on environmental behaviour in the other report. The reports are randomly allocated and analysis will look at the difference between the two groups to a battery of questions about environmental behaviour to be carried at IP3.

ii. Incentives. There will be further experimentation with the incentive groups on the Innovation Panel. Most people will get the same incentive at IP3 as they received at IP2. However, one-sixth of the sample will receive a reduced incentive – similar to the procedure at IP2 (see 4.5, above). This will allow analysts to study whether reducing the incentive at the third wave, rather than the second, has an effect on response.

iii. Measures of change. This is a continuation of the experiment discussed at 4.8, above. The experiment is independent of all other experiments. It examines the effects of explicit and implicit date questions and question ambiguity on observing change between waves of a panel. The showcard experiment carried at IP2 is subsumed within this work as it relates to measures of change in longitudinal studies. There are two parts to this experiment:

a. Question ambiguity/dates: two versions of the questionnaire, each with a mixture of ambiguous and improved questions, one version asks for dates explicitly, the other implicitly.
b. Showcard experiment: one group having the interview conducted using showcards, the other group without showcards. Half of each of the treatment groups at IP3 would have had the opposite treatment at IP2. This means that after IP3 there will be four groups; one who had showcards at IP2 and IP3, one who didn't have showcards at either IP2 or IP3, one that had showcards at IP2 but not IP3 and the last group that did not have showcards at IP2 but do at IP3.

iv. Job satisfaction measures. This is a continuation of the experiment described at section 4.7 above. The treatments for the measurement of job satisfaction differ by whether the response scale is fully labelled, or just the polar categories are labelled, and whether the scale is implemented in CASI, with showcards, without showcards or using an unfolding brackets design.

v. Life satisfaction measures. This is also a continuation of an IP2 experiment and is implemented in the same way as the job satisfaction measures experiment above.

vi. Panel conditioning. This is a continuation of the IP2 experiment described at section 4.9 above. One group will be asked questions about environmental behaviour, height and weight, whilst the other group will not.

vii. Branched versus unbranched questions. This experiment aims to see whether the responses received are different if participants are asked a question in which the response categories are listed out or whether there is an unfolding bracket, or branching, to simplify the question.

viii. Improving measures of wealth. This is a split ballot measurement experiment involving evaluating four design protocols for the collection of debt and asset data for the purposes of measuring wealth. The treatment groups differ in whether the participant is asked to itemise sources of wealth or to aggregate them, and whether they are reporting on their own behalf or that of the household.

A number of these experiments will be audio-recorded for additional analysis. The participant will be asked at the beginning of the interview if they consent to having sections of the interview recorded.

The IP3 questionnaire will be made available on the *Understanding Society* website at:

http://www.understandingsociety.org.uk/design/materials/questionnaires/wave3/

An *Understanding Society* working paper will also be published which will summarise the experimental findings from IP3, to accompany this working paper and the paper on IP1.

Appendix A. Advance materials

(see next page)

Understanding Society

Last year members of your household kindly agreed to be interviewed for *Living in Britain*, a study conducted by researchers at the University of Essex. The study has now been renamed as *Understanding Society* and because it is concerned with how people's lives change over time, we would very much like to interview you and your household again.

Understanding Society covers important subjects such as our health, our opinions, our families and our work. The findings from *Understanding Society* will help us build up a detailed picture about the lives, experiences, behaviours and beliefs of people across the UK in the 21st century and provide an important understanding of diversity within the population. It will help us understand the long term effects of social and economic change in the UK and assist in future decision-making.

Your household's help in the past was very much appreciated and we would like you and your household to take part in the next stage of the study, which is being carried out by NatCen. An interviewer will be in touch with you to arrange a convenient time for an interview that should last around half an hour.

As a token of our thanks, your gift voucher for this year's interview is enclosed.

If you have children aged 10 – 15 we hope you will allow them to complete a short self-completion questionnaire about their hobbies, friends, school-life and hopes for the future. Each child will receive a £3 gift voucher as a thank you for completing the questionnaire.

This year, the questionnaire is available on-line at http://youth.natcen.ac.uk Details for your child(ren) to use to login and complete their questionnaire can be found on the back of this letter. Parents can also access a copy of the questionnaire on-line.

If your household's contact details have changed please let us know by calling Freephone **0800 252 853** or emailing **contact@understandingsociety.org.uk** For more information on the survey visit **www.understandingsociety.org.uk/participants**

Your participation is completely voluntary but we hope you will find the time to help us with this important study and that you enjoy taking part.

Yours,

.....

Nick Buck

Director, *Understanding Society*

This study is being conducted in accordance with the Data Protection Act. This means your personal details will be kept strictly confidential and you and your household will not be identifiable from the data.

Understanding Society

Last year you kindly agreed to be interviewed for *Living in Britain*, a study conducted by researchers at the University of Essex. You may remember we sent you a short report of some of the initial findings a few months ago under the study's new name, *Understanding Society.*
As the study is concerned with how people's lives change over time, we would very much like to interview you again.

Understanding Society covers important subjects such as our health, our opinions, our families and our work. The findings from *Understanding Society* will help us build up a detailed picture about the lives, experiences, behaviours and beliefs of people across the UK in the 21st century and provide an important understanding of diversity within the population. It will help us understand the long term effects of social and economic change in the UK and assist in future decision-making.

Your help in the past was very much appreciated and we would like you and your household to take part in the next stage of the study, which is being carried out by NatCen. An interviewer will be in touch with you to arrange a convenient time for an interview that should last around half an hour.

As a token of our thanks, your £5 gift voucher for this year's interview is enclosed.

If you have children aged 10 – 15 we hope you will allow them to complete a short self-completion questionnaire about their hobbies, friends, school-life and hopes for the future. Each child will receive a £3 gift voucher as a thank you for completing the questionnaire.

This year, the questionnaire is available on-line at http://youth.natcen.ac.uk Details for your child(ren) to use to login and complete their questionnaire can be found on the back of this card. Parents can also access a copy of the questionnaire on-line.

If your household's contact details have changed please let us know by calling Freephone **0800 252 853**
or emailing **contact@understandingsociety.org.uk**
For more information on the survey visit
www.understandingsociety.org.uk/participants

Your participation is completely voluntary but we hope you will find the time to help again with this important study and that you enjoy taking part.

Yours,

Nick Buck
Director,
Understanding Society

This study is being conducted in accordance with the Data Protection Act. This means your personal details will be kept strictly confidential and you and your household will not be identifiable from the data.

Appendix B. Subjective satisfaction questions

Next are some questions about how you feel about your life. Please look at this card and tell me number which you feel best describes how dissatisfied or satisfied you are with the following aspects of your current situation:

a) Your health

b) The income of your household

c) The amount of leisure time you have

SHOWCARD

7 Completely Satisfied

6 Mostly Satisfied

5 Somewhat Satisfied

4 Neither Satisfied nor Dissatisfied

3 Somewhat Dissatisfied

2 Mostly Dissatisfied

1 Completely Dissatisfied

INTERVIEWER: WRITE IN NUMBER CHOSEN.

1 - 7

Using the same scale, how dissatisfied or satisfied are you with your life overall?

INTERVIEWER: WRITE IN NUMBER CHOSEN.

1 - 7

Appendix C. 'Measures of Change' split-ballot question wording experiments

	Version A	Version B
Unclear question instructions: implicit (Version A) vs. explicit (Version B) date questions. Version B is the IP Wave 1 version.	How long have you lived at **this** address?	In what month and year did you move to **this** address?
	When were you first told you had [Health Condition]?	In which year were you first told you had [Health Condition]?
	How long have you been a member of this pension scheme?	In which year did you join this pension scheme?
Vagueness: response categories not mutually exclusive (Version A) Version A is the IP Wave 1 version.	Which of these **best** describes your current employment situation? 1 Self employed 2 In paid employment (full or part-time) 3 Unemployed 4 Retired 5 On maternity leave 6 Looking after family or home 7 Full-time student 8 Long-term sick or disabled 9 On a government training scheme 10 Unpaid worker in family business 97 Doing something else	Which of these describe your current employment situation? 1 Self employed 2 In paid employment (full or part-time) 3 Unemployed 4 Retired 5 On maternity leave 6 Looking after family or home 7 Full-time student 8 Long-term sick or disabled 9 On a government training scheme 10 Unpaid worker in family business 97 Doing something else IF MORE THAN ONE EMPLOYMENT STATUS MENTIONED: And which would you consider your main current employment situation?
Vagueness: unclear terms and concepts (Version A). Version A is the IP Wave 1 version.	Do you have any long-standing physical or mental impairment, illness or disability? By 'long-standing' I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months. 1 Yes 2 No	Have you been, or are you likely to be, troubled for at least 12 months by any physical or mental impairment, illness or disability? 1 Yes 2 No
	Can I just check, did you do any paid work last week - that is in the seven days ending last Sunday - either as an employee or self-employed? 1 Yes 2 No	Thinking back to the seven days ending last Sunday. During this period, did you do any work for which you receive money or a share of profits? 1 Yes 2 No Did you spend 15 hours or more doing unpaid work in a family business? 1 Yes 2 No

	Do you save any amount of your income for example by putting something away now and then in a bank, building society, or Post Office account other than to meet regular bills? Please include share purchase schemes, ISA's and Tessa accounts. 1 Yes 2 No	The next questions are about any money you save from your income. Excluding any money you put away to pay for regular monthly or quarterly bills, do you ever put any money away in a bank, building society or Post Office account? 1 Yes 2 No And do you ever put any money away in share purchase schemes, PEPs, Life Insurance, TESSAs or ISAs? 1 Yes 2 No

References

Bailar, Barbara A. . 1975. The effects of rotation group bias in estimates from panel surveys. *Journal of the American Statistical Association.* 70:23-30.

Bartels, Larry M. 1999. Panel Effects in the American National Election Studies. *Political Analysis* 8 (1):1-20.

Battaglia, Michael P., Elizabeth R. Zell, and Pamela L. Y. H. Ching. 1996. Can participating in a panel sample introduce bias into trend estimates? *1996 Proceedings of Survey Research Methods Section of the American Statistical Association*:1010-1013.

Burton, Jonathan, Heather Laurie, and S. C. Noah Uhrig. 2008. Understanding Society. Some preliminary results from the Wave 1 Innovation Panel. In *Understanding Society working papers, No 2008-03*. Colchester: ISER, University of Essex.

Cantor, David. 1989. Substantive Implications of Longitudinal Design Features: The National Crime Survey as a Case Study. In *Panel Surveys*, edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh. New York John Wiley.

Cantor, David. 2008. A Review and Summary of Studies on Panel Conditioning. In *Handbook of Longitudinal Research: Design, Measurement, and Analysis*, edited by S. Menard. London: Academic Press.

Clausen, Aage. 1968. Response validity: Vote report. *Public Opinion Quarterly* 32:588-606.

Cohen, Steven B., and Vicki L. Burt. 1985. Data collection frequency effects in the National Medical Care Utilization and Expenditure Survey. *Journal of Economic and Social Measurement* 13:125-151.

Conti, Gabriella, and Steve Pudney. 2010. Survey design and the analysis of satisfaction. *Review of Economics and Statistics* forthcoming.

De Leeuw, Edith D., and Johannes Van der Zouwen. 1988. Data Quality in Telephone and Face-to-face Surveys: A Comparative Meta-analysis. In *Telephone Survey Methodology*, edited by R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. Nichols II and J. Waksberg. New York: John Wiley.

Ferber, R. 1953. Obervations on a consumer panel operation. *Journal of Marketing* 17:246-259.

Ghangurde, P.D. 1982. Rotation group bias in the LFS estimates. *Survey Methodology* 8:86-101.

Holt, D. 1989. Panel Conditioning: Discussion. In *Panel Surveys*, edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh. New York: Wiley.

Jäckle, Annette. 2008. Dependent Interviewing: Effects on Respondent Burden and Efficiency of Data Collection. *Journal of Official Statistics* 24 (3):1-21.

Jäckle, Annette. 2009. Dependent Interviewing: A Framework and Application to Current Research. In *Methodology of Longitudinal Surveys*, edited by P. Lynn. New York: Wiley.

Kalton, Graham, and Michael E. Miller. 1991. The Seam Effect with Social Security Income in the Survey of Income and Programe Participation. *Journal of Official Statistics* 7 (2):235-245.

Kemsley, W.F.F. 1961. The household Expenditure Enquiry of the Ministry of Labor. *Applied Statistics* 10 (3):117-135.

Kraut, R.E., and J.G. McConahay. 1973. How being interviewed affects voting: An experiment. *Public Opinion Quarterly* 37 (3):398-406.

Lynn, P., and E. Sala. 2006. Measuring change in employment characteristics: The effects of dependent interviewing. *International Journal of Public Opinion Research* 18 (4):500-509.

Lynn, P., N., Uhrig, and J. Burton 2010 Lessons from a Randomised Experiment with Mixed-Mode Designs for a Household Panel Survey In *Understanding Society working papers, No 2010-03*. Colchester: ISER, University of Essex.

Menard, Scott, and D.S. Elliott. 1993. Data set comparability and short-term trends in crime and delinquency. *Journal of Criminal Justice* 21:433-445.

Mooney, H. 1962. Methodology in two California health surveys. In *US Public Health Service, Public Health Monograph No. 70*. Washington, DC: US Government Printing Office.

Moore, J. C. 2004. *Exploiting Computer Automation to Improve the Interview Process and Increase Survey Cooperation*, *Survey Methodology, No. 2004-01*. Washington DC: Statistical Research Division U.S. Census Bureau.

Moore, Jeffrey, Nancy Bates, Joanne Pascale, and Aniekan Okon. 2009. Tackling Seam Bias Through Questionnaire Design. In *Methodology of Longitudinal Surveys*, edited by P. Lynn. Chichester: John Wiley & Sons.

Neter, J., and J. Waksberg. 1964. Conditioning effects for repeated household interviews. *Journal of Marketing* 28:51-56.

Pudney, S., 2010. An experimental analysis of the impact of survey design on measures and models of subjective wellbeing In *Understanding Society working papers, No 2010-01*. Colchester: ISER, University of Essex.

Silberstein, Adriana R., and Curtis A. Jacobs. 1989. Symptoms of repeated interview effects in the Consumer Expenditure Interview Survey. In *Panel Surveys*, edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh. New York: Wiley.

Smyth, Jolene D., Don A. Dillman, Leah Melani Christian, and Michael J. Stern. 2006. Comparing Check-All and Forced-Choice Question Formats in Web Surveys. *Public Opinion Quarterly* 70 (1):66-77.

Spencer, Elizabeth A., Paul N. Appleby, Gwyneth Davey, and Timothy J. Key. 2002. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutrition* 5:561-565.

Sturgis, Patrick, Nick Allum, and Ian Brunton-Smith. 2009. Attitudes Over Time: The Psychology of Panel Conditioning. In *Methodology of Longitudinal Surveys*, edited by P. Lynn. New York: John Wiley & Sons, Ltd.

Sudman, S., N. M. Bradburn, and N. Schwarz. 1995. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. New York: Wiley.

Thomas, Randall K., and Jonathan D. Klein. 2006. Merely Incidental? Effects of Response Format on Self-reported Behavior. *Journal of Official Statistics* 22:221-244.

Turner, R. 1961. Inter-Week Variations in Expenditure Recorded During a Two-Week Survey of Family Expenditure. *Applied Statistics* 10 (3):136-146.

Waterton, Jennifer, and Denise Lievesley. 1989. Evidence of conditioning effects in the British Social Attitudes Panel. In *Panel Surveys*, edited by D. Kasprzyk, G. Duncan, G. Kalton and M. P. Singh. New York: John Wiley & Sons.

Wilson, Sven E., and Benjamin L. Howell. 2005. Do panel surveys make people sick? US arthritis trends in the Health and Retirement Survey. *Social Science and Medicine* 60 (11):2623-2627.

Yalch, R.F. 1976. Pre-election interview effects on voter turnout. *Public Opinion Quarterly* 40:331-336.