



Understanding Society
Working Paper Series

No. 2011 – 02

March 2011

Effect of interview modes on measurement of identity

Alita Nandi

(Institute for Social and Economic Research University of Essex)

and Lucinda Platt

(Centre for Longitudinal Studies, Institute of Education)

Effect of interview modes on measurement of identity

Alita Nandi

(Institute for Social and Economic Research University of Essex)

and **Lucinda Platt**

(Centre for Longitudinal Studies, Institute of Education)

Abstract

In this paper, we investigate the effect of interview mode on responses to attitudinal and autobiographical questions: a 13-item multidimensional identity module. We find small but significant mode effects of three to seven percentage points on the response pattern for the domains of national identity, religion, racial or ethnic background, political beliefs and sexual orientation. We also find very small but significant mode effects of one to two percentage points on item non-response for occupational identity and identification with father's ethnic background. We conclude that mode effects on these questions can convincingly be interpreted as stemming from social desirability bias.

Keywords: interview mode, measurement, identity

JEL classifications: C83

Author contact details: Alita Nandi, ISER, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, United Kingdom. Email: anandi@essex.ac.uk

Acknowledgements: This research was supported by ESRC funding for the UK Longitudinal Studies Centre.

Effect of interview modes on measurement of identity

Alita Nandi

(Institute for Social and Economic Research University of Essex)

and **Lucinda Platt**

(Centre for Longitudinal Studies, Institute of Education)

Non-Technical Summary

In surveys we ask people a series of questions to get some specific information. The quality of this information (or response) depends on a number of factors: how the question was worded, how it was asked (e.g. face to face, over the phone or by paper self-completion), how demanding it was to come up with the answer, whether the question asked about sensitive issues that a person would be uncomfortable answering and the cognitive abilities of the person asked. If we do not ask people the same questions in exactly the same way, there may be systematic differences in the answers. For example, we ask a person to choose her three favourite cars from a very long list of cars that is read out, she is not likely to remember all the cars when answering the question and end up choosing from among the ones she does remember. On the other hand if we had given her a paper with the list written on it, she may have taken her time, looked over the entire list and given a different answer. When we find such systematic differences in response due to differences in the way the interviews are administered we call that mode effect.

In this paper we estimate the mode effect (the interview modes being telephone and face-to-face) on the responses to a module of 13 questions measuring identification in different domains, such as occupation, religion, age or life stage. The intention was to ask these questions in the second wave of the *Understanding Society*: the UK Household Longitudinal Study (UKHLS) and at regular intervals thereafter, providing researchers with a rare opportunity to investigate questions relating to determinants of identity change as well as formation over time. While it was decided that the second wave of the UKHLS would be conducted face-to-face, part of the sample may be interviewed by telephone in future waves. If there are mode effects then we will not be able to say if the observed change in the identity measure is due to real change in identification or due to a change in the interview mode. It was therefore important to identify their extent and impact prior to inclusion in the main survey. The Innovation Panel, a UK representative household sample of 1500 households, fielded prior to the main UKHLS survey provides a critical resource for testing methodological issues of importance to the main survey. In the second wave part of the Innovation Panel sample was interviewed face-to-face and the rest by telephone, thus providing a unique opportunity to estimate mode effects for identity questions. We examined the effect of interview mode on overall response patterns as well as item non-response. Since the questions were asked in the same way across both modes, the main impact on response is likely to be the physical presence or absence of the interviewer. This can affect the willingness of respondents to provide less socially acceptable answers, with face to face response typically demonstrating more of such 'social desirability bias'.

We find small but significant mode effects of three to seven percentage points on the response pattern for the domains of national identity, religion, racial or ethnic background, political beliefs and sexual orientation. We also find very small but significant mode effects of one to two percentage points on item non-response for occupational identity and identification with father's ethnic background. We conclude that mode effects on autobiographical attitudinal questions appear to be small, and can convincingly be interpreted as stemming from social desirability bias.

Effect of interview modes on measurement of identity

Alita Nandi

(Institute for Social and Economic Research University of Essex)

and Lucinda Platt

(Centre for Longitudinal Studies, Institute of Education)

Introduction

In this paper, we investigate the effect of interview mode on responses to questions (including item non-response) on a multidimensional identity module asked in the second wave of the Innovation Panel that was conducted in March-June 2009. These questions were autobiographical and attitudinal in nature with response options that were closed, three point fully-labelled rating scales. The intention was to ask these and other similar questions (across additional ethnic identity dimensions) in the second wave of *Understanding Society: the UK Household Longitudinal Study* (UKHLS) and at regular intervals thereafter, providing researchers with a rare opportunity to investigate questions relating to determinants of identity change as well as formation over time. While it was decided that the second wave of the UKHLS would be conducted face-to-face, part of the sample may be interviewed by telephone in future waves. Any mode effect would confound estimates of measures of identity change.

Similar identity questions had been asked in cross-sectional surveys where the interview mode was only face-to-face and so estimates of mode effects was not available. However, there has been extensive research into mode effects, especially for sensitive questions, which we could draw upon (Tourangeau and Smith 1996, Tourangeau 2004). It was not possible directly to translate their findings to these identity questions since we did not have much information about the sensitivity and social norms regarding the responses to these questions. Therefore, as part of an ongoing project to design ethnic identity questions for the UKHLS, we aimed to test the mode effects by fielding a set of identity questions in the second wave of the Innovation Panel. The Innovation Panel is a longitudinal survey of 1500 households, which takes place around nine months prior to the main UKHLS survey with the aim of testing methodological issues of crucial importance to the UKHLS and other longitudinal surveys. At sweep 2, as part of a wider mode experiment, part of the Innovation Panel sample was interviewed face-to-face and part by telephone. This provided a unique opportunity to estimate mode effects for the set of identity questions and explore how they accord with existing knowledge about response to sensitive and attitudinal questions.

We estimated multivariate multinomial ordered logit models to identify the effect of interview mode (face-to-face vs telephone) on response patterns and univariate logit models to identify the effect of interview mode (face-to-face vs telephone) on item non-response ('don't knows' and refusals). We found that there were small but significant mode effects (3-7 percentage points) on response pattern for five of these questions and very small but significant mode effects (1-2 percentage points) on item non-response for two of these questions. We argue that these effects can be understood with reference to social desirability bias.

Background: Mode effects

Mode of data collection in surveys, which can range from face to face interviews to telephone to paper (self-completion) and increasingly web, is recognised to have a potential impact on the ways in which respondents answer questions. It therefore has implications for question validity and for equivalisation of responses across modes. At the same time, with challenges of response

rates and rising costs to achieve samples facing surveys in many countries, using a combination of modes can offer a way in which to maximise response (de Leeuw 2005). In such circumstances the issue of potential mode effects becomes acute (Jäckle et al. 2010).

The model of survey response popularized by Tourangeau, Rips and Rasinski (2004)¹ breaks down the response process into four sub-processes which may be sequential or simultaneous: “comprehension of the item, retrieval of relevant information, use of that information to make required judgements, and selection and reporting of an answer” (Tourangeau, Rips and Rasinski 2004). These processes are affected by characteristics of the interviewer, question wording and survey instrument and, of course, the characteristics of the respondent themselves. The interview mode enters into this equation via the physical presence or absence of interviewer and the visual or aural stimuli of the survey instrument. For example, it is hypothesised that the presence of the interviewer affects the response process at the comprehension stage by helping respondent with difficult questions and at the final reporting stage by assuring the respondent of the confidentiality of her responses. Alternatively, the presence of an interviewer may invoke social norms and make the respondent tailor her responses to conform to those norms (social desirability).

Having described the overall issues, we discuss how face-to-face and telephone interview modes may affect the response process for attitudinal questions specifically. Attitude towards an issue is conceived of as a repertoire of beliefs, feelings and opinions about that issue (Hippler *et al.* 1987). Compared to many factual questions attitudinal questions are both tend to be more sensitive, and are typically more cognitively burdensome, though this will vary with the particular questions being asked. With attitudinal questions as with other types of questions respondents may have a ready-made answer in which case these processes have already been completed prior to the interview and so they don't need to retrieve the information during the interview. In such cases, the response is less dependent on the interview mode (except for the judgement stage). But if it is not well formed then it is the context within which a question is asked that determines which one of the repertoires of beliefs, feelings and information a respondent will retrieve and how she will form a judgement. The impact of context is stronger if the question is ambiguous and the respondent tries to decipher what she is being asked by the context i.e., at the comprehension stage. Some of the facets of the interview that constitute the context are the questions prior to this, the general purpose of the survey, the characteristics of the interviewer, or even the response options. Where mode is telephone rather than face to face the characteristics of the interviewer are likely to be less salient as they can only be heard rather than seen.

Often during face-to-face interviews showcards are used to show response options. This reduces the cognitive burden on the respondent in trying to remember the response options, especially when there are many response options. It is practically very difficult to provide showcards for telephone interviews and almost never done. Related to this issue is that of recency effects where respondents are more likely to respond to the category that they have heard recently and primacy effects where they are likely to respond to the ones they have heard at first. These effects tend to be different between modes when one can see all the categories in one (face-to-face with showcards or self-completion) but needs to remember them in the other (telephone) (Krosnick 1999, Holbrook, Krosnick, Moore and Tourangeau 2005). Reducing the number of categories available to respondents can assist with this difficulty. A narrow range of categories may also limit the extent of mode effects at the extremes of a scale, as found by Jäckle et al. 2010.

¹ This was adapted from Cannell, Miller, and Oksenberg's (1981) (see Tourangeau, Rips and Rasinski 2004)

Telephone interviews often increase the pace of the interview giving the impression that there is less time to respond. This may result in more off the cuff responses (*satisficing*) in a telephone interview (Holbrook, Green and Krosnick 2003), though Jäckle et al. 2010 found no evidence of satisficing in their study of modes in the European Social Survey. In addition, a person may be multi-tasking during a telephone interview but is less likely to do that during a face-to-face interview. Questionnaire length is likely to be an issue here, and the questions themselves have to be sufficiently clear and straightforward that speed of response does not impact on the quality, and sufficiently varied that the respondent is not encouraged simply to repeat options across long strings of questions.

At the final stage of response construction, respondents are likely to edit their responses so as to conform to a socially acceptable answer, known as social desirability bias (De Maio 1994 in Roberts 2007). This is higher the more sensitive (to social pressure) an issue is. The presence of an interviewer is likely to exert more social pressure than if she is just heard over the telephone. On the other hand, if the respondent has concerns for the confidentiality of her response then seeing an interviewer who can reassure her, may reduce item non-response ('don't knows' and refusals) as well as averting socially desirable responses (Holbrook, Green and Krosnick 2003).

While it is usually assumed that face to face interviews invoke social desirability more than phone interviews because of the norms around social interaction, in Holbrook, Green and Krosnick's study, they found *more* socially desirable responses in telephone interviews. Their explanation was that respondents felt more comfortable and so were more truthful in face-to-face interviews. In their study as in most studies external sources are used to decide what is a socially desirable response. An alternative, that we use here, is to allow the data to guide us: on the assumption that social desirability bias is the main factor in our (attitudinal) questions liable to account for mode effects we can scrutinise the direction of the effect for different sorts of question and whether respondents more likely to be subject to social desirability bias show greater mode effects. As we show, our results are consistent with social desirability bias being more prominent in face to face interviews.

Implicit in this discussion is the effect of the respondents' own characteristics on the response quality. A respondent with higher cognitive ability or conscientiousness is more likely to put as much effort as is necessary to arrive at a "correct" response. In other words, satisficing is likely to be low no matter what the interview mode. These last two points imply that we need to pay attention both to the context of the interview and the characteristics of the respondent if we are to identify – and potentially adjust for – mode effects.

Data, Design and Questions

Understanding Society: The UK Household Longitudinal Study (UKHLS) is a large-scale multi-purpose household panel survey of a sample of individuals in approximately 40,000 private households in UK, covering four separate samples. The main sample of around 28,000 households and an ethnic minority boost sample started in 2009, while the former BHPS responded joined the main sample for the beginning of the second wave in 2010. A further methodological panel of around 1,500 households (the Innovation Panel) is a stand alone sample that is used to inform survey practice in the main sample, and other longitudinal studies. It enters the field around 9 months prior to each wave of the main sample. Across all samples, adults household member (ages 16 and over) are interviewed and are interviewed every year and followed over time to anywhere in the UK. Children aged 10-15 complete a separate self-completion instrument and are allocated the main questionnaire when they reach 16. About half of the questionnaire content in the Innovation Panel overlaps with that in the main samples.

However, it is specifically set up to enable experiments to be carried, which can involve different forms of questions and different question modes, as well as other aspects of the survey process such as experiments relating to incentives, and so on. For more information see: <http://www.understandingsociety.co.uk/>.

The second wave of the Innovation Panel (IP2) was in the field during March-April 2009. The individual interview lasted around 30 minutes. A total of 1,785 adults were interviewed either by telephone or face to face. IP2 incorporated a range of experiments on: Mixed Modes and Response, Incentives and Response, Advance Materials and Response, Use of Show Cards and Survey Measurement, Measurement of Job Satisfaction, Measurement of Life Satisfaction, Measurement of Identity, Measures of Change. All were independent of each other except for the measures of job and life satisfaction which were nested within the showcards experiment.

We are concerned here with the Mixed Modes and Measurement of Identity experiments. The autobiographical attitudinal questions we explored in IP2 are intended to capture salient aspects of identity. Aspects of identity had been identified as an important element to measure in *Understanding Society*, which offers the ability to evaluate variation across subpopulations, including minority ethnic groups, and its development or metamorphosis over time. It was therefore important to test whether questions would be subject to mode effects, since the mode of interview could be expected to change (at least for some respondents) between incidences of the question being asked, as the survey moved towards the anticipated mixed mode design. In order for researchers to be able to answer the questions of interest around determinants of identity change, it would be important that any evaluation of change would not be spurious by deriving in fact from mode effects.

The identity questions consisted of thirteen questions which asked respondents the importance 'to their sense of who they are' of a suite of characteristics often associated with identity. The characteristics were: profession, ethnic or racial background, religion, national identity, political beliefs, family, father's ethnic group, mother's ethnic group (if different from father's), marital or partnership status, gender, age and lifestage, level of education, sexual orientation (see Appendix).

The first of these questions which measures occupational identity also incorporated a wording experiment, allocated independently of mode, which constituted the Measurement of Identity Experiment. We wanted to explore whether higher quality response as exemplified by lower non-response could be found when respondents were asked about their 'profession' compared to their 'occupation'. Those who were asked either of these response options were randomly allocated across the three mode assignments (see below). Using data from the Citizenship Survey 2007 we had found a high degree of item non-response for this question and that it was highest among those not employed. We speculated that the word 'occupation' is interpreted as something to do with the current job and that perhaps using a word such as 'profession' may be interpreted as reflecting long term occupation.

Our aims in designing the form of the questions (and adapting them from the form they took in the Citizenship Survey) was first to minimise and second to isolate mode effects through our design. We therefore aimed to adopt a 'unimode' approach (Dillman 2000). To minimise differences in cognitive burden we avoided the use of showcards regardless of mode; and to reduce cognitive burden and recency and primacy effects we retained only three response categories: Important, Not very important, Not at all important. We also attempted to avoid complexity in the questions. The questions were largely straightforward and were modelled on a suite which had been regularly used in face to face mode in a cross-sectional survey and so had

been well tested in the field as well (our module differed on the identity dimensions measured and the reduction in response options from four to three).

To avoid excessive repetition on the one hand and cognitive burden on the other, the 13 identity measures were split into smaller groups. The entire question was read out only before each such sub-section. If the whole question were read out with each of the 13 items then that would have made the questions extremely tedious and repetitive. On the other hand just reading out the entire question once, right at the beginning, would have put enormous cognitive burden on the respondents as it would have required them to remember the question for all the 13 items. The grouped repetition was thus designed to mitigate potential tedium and cognitive burden which could also lead to differences between modes. The first set of measures comprised occupation/profession, racial and ethnic background, religion and national identity; the second set were political beliefs, family, father's and mother's racial or ethnic background; the third group consisted of marital or partnership status, age or life stage, education and sexual orientation.

To isolate mode effects we ensured that the question order, general purpose of the survey as understood by the respondent, the response options and question wording were kept identical in the two modes. Our aim was to ensure comparability of respondent characteristics across modes through random assignment but that was not wholly achieved (as we discuss below). The only other difference was whether the interview took place over the telephone or face to face. Having isolated actual mode as the only means through which differences in responses could occur, the channels for any mode effects remained as social desirability and/or confidentiality, via the observable physical presence /absence of the interviewer or their role in providing reassurance, and satisficing via faster interview pace and multi-tasking in telephone interviews or lack of clarification from the face to face interviewer.

The suite of 13 questions varied in terms of sensitivity. For example, it is recognised that expressing the importance of politics can seem socially undesirable, particularly when the other person's political orientation is unknown. At the same time, the degree of sensitivity may depend on the characteristics of the person. For example, occupational identity may be a sensitive issue for unemployed or retired people. The physical presence (or absence) of the interviewer may serve as a context for the question, and what is inferred about them. For example, if the ethnic background of the interviewer is recognisably different, this may influence responses to questions on own ethnic background. There is no reason to believe that interviewers differed systematically in terms of their characteristics across the two modes.

In summary, we expected to be able to identify any social desirability mode bias in terms of respondents choosing a particular response option in one mode rather than another, with the more 'desirable' response being more common in the face-to-face interview. Satisficing would result in higher item non-response in telephone interviews.

Delivering the mode experiment

Random assignment of respondents across the two modes ensures that differences in respondent characteristics do not confound mode effects. While the assignment was random the achieved interview mode was not totally random, as the mode experiment also involved exploring final achieved response rates. To be specific, households were initially randomly assigned to three groups. The first group was to be interviewed by CAPI. But if they refused they were given the option to be interviewed by telephone. The second and third groups were initially assigned to CATI. For group 2 if one household member could not do the telephone interview then all household members would be transferred to CAPI. For group 3, attempts would be made to interview all household members by telephone, and only those who refused or

were not contacted would be transferred to CAPI. So, those interviewed face-to-face include respondents who refused to be interviewed by telephone as well as respondents who agreed to be interviewed face-to-face. If these telephone ‘refusers’ are systematically different from telephone respondents in terms of how they respond to the identity questions then any observed mode effect will partly be a consequence of this bias.

Table 1 illustrates the initial assignment of individuals and their ultimate interview mode. As it shows, only 6 per cent of individuals assigned to face to face were subsequently interviewed by telephone, but 20 per cent of those in households assigned to telephone interviewing were subsequently interviewed face to face.

Table 1: Assigned and achieved interview mode

Final wave 2 interview mode	Initial assignment			All
	Group 1: Face-to-face	Group 2: Telephone Transfer entire household to Face-to-face if one person refuses	Group 3: Transfer to Face-to-face only those who refuse	
Face-to-face	600 (94%)	120 (21%)	118 (20%)	838 (47%)
Telephone	36 (6%)	439 (79%)	472 (80%)	947 (53%)
Total	636	559	590	1785

Comparing respondents who remained in their assigned mode with those who switched from telephone to face-to-face we find that there are some differences in characteristics between the switchers and non-switchers. We controlled for a number of socio-demographic variables – age, sex, educational attainment, indicators of income and wealth (whether own a house, household or self has access to cars), marital status, number of children, employment or activity status (retired, in paid employment, other), region of current residence (England, Scotland or Wales), general health, religious beliefs (not religious, Christian or non-Christian), race/ethnicity (whether white or not), country of birth (born in UK or not) to estimate the difference between switchers and non-switchers. In case of initial assignment to group 2, we find that switchers and non-switchers differed in terms of education and access or ownership of car. With those assigned to group 3, the difference between switchers and non-switchers was in terms of region of residence, birth cohort and education level. We are not concerned with interpreting these differences here, only with being able to adjust for observed differences between the groups. We therefore decided to control for these variables when estimating mode effect.

Table 2: Distribution of individual non-response rate in the 2nd wave of the Innovation Panel across different (household level) assignment and final interview modes

Final interview mode	Initial assignment			All
	Group 1: Face-to-face	Group 2: Telephone Transfer entire household to Face-to-face if one person refuses	Group 3: Transfer to Face-to-face only those who refuse	
Face-to-face	14%	33%	36%	21%
Telephone	5%	10%	9%	9%
Total	13%	16%	16%	15%

We also found that the individual response rate among responding households who were to be interviewed face-to-face was much higher than those who were to be interviewed by telephone. If non-respondents are systematically different from respondents in terms of how they respond to the identity questions then part of the mode effect that we estimate could be explained by non-response bias. If the difference in response behaviour is completely explained by differences in socio-demographic characteristics of respondents and non-respondents then controlling for these characteristics in the models would take care of this problem.

Results

We measured item non-response by the proportion of don't knows and refusals (most of which are don't knows). Using Pearson's chi-squared test as well as Fisher's exact test we found that the interview mode matters only for occupation and father's ethnic origin questions, the probability of item non-response being higher in face-to-face interview mode (see Table 3). Looking deeper into the characteristics of the non-responders we find that 35 out of 37 persons who report don't know to the occupational identity question are not employed and 18 of the 21 persons who report don't know to the father's ethnic group identity question are white. For these particular groups the corresponding identity domain is less relevant and so less likely to be a formed concept but rather something that needs to be retrieved and response constructed during the interview. However, if this were the case then we would expect to see a higher degree of non-response in telephone interviews. On the other hand, if social desirability was at play, and person who is not employed is uncomfortable responding to an occupational identity question then we would expect to see the above result. Similarly ethnic heritage (as measured by father's ethnic group) may be less relevant for someone from the majority group, but they may feel uncomfortable in the presence of an interview supplying a 'not important' response. Thus, these results suggest that satisficing is not a concern with these identity questions but some may be subject to social desirability bias and the impact of that is greater in face to face interviews, consistent with other studies.

Table 3: Tests of differences in item non-response by interview mode

	Proportion of item non-response		Pearson's Chi-square test	Fisher's Exact test
	Face-to-face	Telephone		
Your occupation?	3.34%	1.16%	0.00	0.00
Your ethnic or racial background?	1.07%	0.74%	0.45	0.46
Your religion?	0.72%	0.21%	0.11	0.16
Your national identity?	0.95%	0.74%	0.62	0.80
Your political beliefs?	0.60%	0.32%	0.38	0.49
Your family?	0.12%	0.21%	0.64	1.00
Your father's ethnic group?	1.79%	0.63%	0.02	0.03
Your marital or partnership status?	0.84%	0.32%	0.14	0.20
Your gender?	1.19%	1.80%	0.30	0.34
Your age and life stage?	0.95%	0.63%	0.44	0.59
Your level of education?	0.48%	0.11%	0.14	0.19
Your sexual orientation?	1.67%	1.58%	0.89	1.00

Note: Item non-response is “don't know” in almost all cases. There is only a handful “refusals”.

“Your mother's ethnic group” was asked of only those whose mother's race/ethnicity was different from that of her father and so there were only a handful of cases who were asked this question. So, we have omitted this question from the analysis.

We also found that item non-response for both the variants of the occupational identity question was around 2 percent, with higher rates for face-to-face interviews (3% compared to 1% for telephone). Thus while alternative wording was not able to reduce level of non response it was relatively low in each case.

We used multivariate ordered logit models to evaluate mode effects in response patterns for each of the 13 questions. This allowed us to take account of the hierarchical/ordinal nature of the response categories. As noted, although the initial assignment into the three different mode categories was random, modes were switched if one or more household member refused an interview. As there were some differences in characteristics between those who switched and those who did not, any observed mode effect may be partly due to differences in characteristics between those interviewed by telephone or face-to-face. So, we controlled for these characteristics in the model.

Table 4: Mode Effect as measured by the predicted probability of response if everyone were interviewed face-to-face compared with if they were interviewed by telephone (only questions for which the mode effect was significant are shown)

	Face-to-face			Telephone		
	Important	Not important	Not at all Important	Important	Not important	Not at all Important
Racial and Ethnic Background	0.56	0.29	0.14	0.51	0.32	0.17
Political beliefs	0.33	0.46	0.21	0.42	0.42	0.16
Education	0.69	0.25	0.05	0.74	0.21	0.04

Note: Response pattern model estimated by ordered logit with only interview mode as covariate

In general, mode appears to have little effect. In the unadjusted analysis, those questions where significant mode differences were found were racial and ethnic background, political beliefs and educational level (see Table 4).

Table 5: Mode Effect as measured by the predicted probability of response if everyone were interviewed face-to-face compared with if they were interviewed by telephone, estimated at the mean value of all covariates (only questions for which the mode effect was significant are shown)

	Face-to-face			Telephone		
	Important	Not important	Not at all Important	Important	Not important	Not at all Important
Racial and Ethnic Background	0.58	0.29	0.13	0.52	0.32	0.16
Religion	0.34	0.49	0.18	0.29	0.50	0.21
National Identity	0.77	0.19	0.04	0.74	0.21	0.05
Political beliefs	0.34	0.47	0.19	0.41	0.44	0.15
Sexual orientation	0.68	0.23	0.09	0.64	0.26	0.11

Note: Response pattern model with interview mode and other covariates estimated by multivariate ordered logit

Once we controlled for socio-demographic characteristics we found that mode effects for education question was not significant while those for national identity, religion and sexual orientation were (see Table 5). There were, then, only 5 questions which exhibited significant mode effects once controlling for relevant differences between the two experimental groups, and the difference in the probability of choosing 'Important' varied between 3 to 7 percentage

points. In all these domains except for political beliefs, those interviewed by telephone were less likely to choose that domain as important to their sense of who they are.

We have argued that social desirability bias influenced by the presence of the interviewer is likely to be at the heart of any mode effects in responses, given the way we designed the experiment. Moreover we would expect, if the direction of effects is consistent with that expectation, that if a social norm dictates a particular domain be chosen as important (or unimportant) then it will more likely be chosen in a face-to-face interview than by telephone. That is, if a person is inclined to select the socially undesirable category she will have a higher likelihood of choosing the socially desirable option in a face-to-face interview than when interviewed by telephone. For example, if a person does not belong to any religion but it is socially undesirable to say that religion is not an important part of her identity, she will be more likely to choose “important” in a face-to-face than a telephone interview.

The predicted higher probabilities of choosing “important” in face-to-face interviews for racial and ethnic background, religion, national identity and sexual orientation identities might suggest that “important” is the socially desirable categories for these identity domains if interviewers do indeed exert a social desirability effect. By similar reasoning we could say that “important” was the socially undesirable category for political identity. If this was the case, we would expect to find higher mode effects in terms of selecting ‘important’ in the face-to-face interview where the category was apparently less salient in fact, that is for white respondents in the case of racial and ethnic background identity, for those not belonging to any religion in case of religious identity, and for those not born in UK in case of national identity.

In the case of political identity, if social desirability in the face-to-face context was stimulating a greater tendency to answer ‘unimportant’ then we would expect higher mode effects where political identity was *prima facie* more salient. We therefore estimated whether there were differences in mode effects for those who reported having done some voluntary work, especially in activist organisation in the last 12 months.

It was more difficult to evaluate the operation of social desirability in the face of the sexual orientation question. The direction of a social desirability effect and whom it would impact more were *a priori* hard to predict. However, we considered that the ‘true’ salience of this aspect of identity might well differ by gender, with women being expected therefore to be more subject to mode effects.

To explore whether these inferences held, we therefore interacted interview mode with white/non-white ethnicity variable for the racial and ethnic identity question, with religious beliefs for the religious identity question, with country of birth for the national identity question, with voluntary activity & voluntary activity in activist organisations for political identity, and with gender for the sexual orientation identity question.

As Table 5 shows, our inferences of social desirability and its impact in face-to-face interviews were substantially confirmed. We found that the if the interview mode were changed from telephone to face-to-face then the probability of choosing “important” would increase by 0.05 for white but not for non-white respondents, by 0.04 for those who have reported not to have any religion but not for those who reported having a religion. We found that the if the interview mode were changed from telephone to face-to-face then the probability of choosing “important” would decrease by 0.11 for those who did some voluntary work in the past 12 months and by .06 for those did not (see Table 6). These findings were highly consistent with our expectations and demonstrate that, to the extent that we have mode effects in responses,

they can be interpreted as social desirability causing some respondents to adjust their evaluation of the importance of an element of their identity.

Table 6: Average change in predicted probability of choosing “important” if interview mode were changed from telephone to face-to-face

Your racial and ethnic background?	Ethnic group belonging to		
	White	Non-white	
	.05*	.15	
Your religion?	Religion belong to		
	None	Christian	Non-Christian
	.04***	.01	-.11
Your national identity?	Born in UK?		
	Yes	No	
	.04	.04	
Your political beliefs?	Any voluntary activity in last 12 months?		
	Yes	No	
	-0.11**	-0.06**	
Your sexual orientation?	Gender		
	Women	Men	
	.07**	.01	

Note: Response pattern model with interview mode and other covariates estimated by multivariate ordered logit

* p<.10, ** p<.05, *** p<.01

#political, conservation, the environment and heritage, animal welfare, justice and Human Rights, local community, neighbourhood or Citizen’s groups, Trade Unions

Furthermore, in the case of sexual orientation question we found that if the interview mode were changed from telephone to face-to-face then the probability of choosing “important” would increase by 0.07 for women but not for men. This is also consistent with our hypothesis that sexual orientation is less salient for women but is subject to social norms in the interview context.

Conclusions

We can conclude then that very few of these autobiographical identity questions on identity are subject to any mode effects and the magnitude of the mode effects we did find are small, varying from 3 to 7 percentage points. This alone is encouraging for our future analysis. More generally, we can infer that, once measures have been taken to minimise mode effects through eliminating showcards in face-to-face interviews, ensuring clear and simple questions with a small number of response categories these (attitudinal and autobiographical) questions do not raise serious concerns for comparability across modes. We have also demonstrated that there is a strong case for interpreting the mode effects we find as located in social desirability bias. This is supported by the finding that mode effects are stronger for some groups where there is likely to be greater susceptibility to the pressure of social norms represented by the presence of an interviewer. To the extent that we can identify factors implicit in the greater susceptibility of particular responses and respondents to mode effects, the analysis also offers indications of appropriate adjustments that may be needed adequately to take account of them in the face of a move from one mode to another.

References

- DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena Volume 2*. pp.257-282 New York: Russell Sage Foundation.
- Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2nd ed.). New York: John Wiley Co.
- Edith de Leeuw. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys" *Journal of Official Statistics*, Vol. 21. No. 2. Pp. 235-255
- Hippler, H., N. Schwarz and S. Sudman. 1987. *Social Information Processing and Survey Methodology*, Springer-Verlag, New York.
- Holbrook, A. L., M. C. Green and J. A. Krosnick, (2003) "Telephone versus face-to-face interviewing of National Probability Samples with Long Questionnaire" *Public Opinion Quarterly*, 67: 79-125
- Holbrook, A. L., J. A. Krosnick, D. Moore and R. Tourangeau (2005) "Response Order Effects In Dichotomous Categorical Questions Presented Orally: The Impact of Question and Respondent Attributes" Available at <http://comm.stanford.edu/faculty/krosnick/Gallup%20Paper.doc>
- Jackle, A., C. Roberts and P. Lynn, "Assessing the Effect of Data Collection Mode on Measurement. 2010, *International Statistical Review*, 78 (1): 3-20s
- Krosnick, Jon A. 1999. "Survey Research." *Annual Review of Psychology* 50:537-67.
- Roberts, C. 2007. "Mixing modes of data collection in surveys: A methodological review" *ESRC National Centre for Research Methods Briefing Paper, NCRM Methods Review Papers* NCRM/008
- Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55: 775–801.
- Tourangeau, R., L. Rips and K. Rasinski *The Psychology of Survey Response*, Cambridge University Press.
- Tourangeau, R., and T. Smith. 1996. "Asking Sensitive Questions: The Impact Of Data Collection Mode, Question Format, And Question Context." *Public Opinion Quarterly*, 60: 275-304

Appendix

Identity Questions:

We'd like to know how important various things are to your sense of who you are. Please think about each thing I mention, and tell me whether you think it is important, not very important or not at all important to your sense of who you are?

READ OUT EACH AND CODE

- 1 Important
- 2 Not very important
- 3 Not at all important

- (a) Your occupation (alt. version: profession)? INTERVIEWER: IF DK PROBE: Is that because you are retired?
- (b) Your ethnic or racial background?
- (c) Your religion?
- (d) Your national identity?
- (e) Your political beliefs?
- (f) Your family?
- (g) Your father's ethnic group?
- (h) Your mother's ethnic group? *{Asked only if father's ethnic group is different from mother's}*
- (i) Your marital or partnership status?
- (j) Your gender?
- (k) Your age and life stage?
- (l) Your level of education?
- (m) Your sexual orientation?