

## **Developing survey questions on sexual identity:**

### **Rationale and design of sexual identity questioning on the Integrated Household Survey (IHS)**

Data Collection Methodology for Social Surveys.  
Ethnicity Identity & Inequalities Team.  
Office for National Statistics

December 2008

## Contents

1	Introduction and overview .....	2
2	Question design.....	3
2.1	Face-to-face administration .....	3
2.2	Telephone administration.....	4
3	Meaning of the question.....	5
4	Introduction to the question .....	6
5	Question stem.....	6
6	Response categories .....	7
6.1	Heterosexual/straight .....	7
6.2	Gay/Lesbian .....	7
6.3	Bisexual.....	8
6.4	Other .....	8
6.5	Response category order .....	9
7	Assistance with meaning of response categories .....	9
8	Missing values .....	9
9	Interviewer-administered or self-administered .....	10
10	Proxy data collection.....	11
11	Use of interpreters.....	11
12	Location of the question.....	11
13	Interviewer training.....	12
14	Transgender.....	13
15	Qualifications .....	13
16	The future.....	13

# 1 Introduction and overview

Informed by the research conducted by the Sexual Identity Project of the Office for National Statistics (ONS)<sup>1</sup>, a sexual identity question has been designed for use on the component surveys of the Integrated Household Survey (IHS). These surveys include a number of identity type questions and the sexual identity question has been designed to sit alongside these. The question has not been designed for specific or detailed studies of sexual behaviour and attraction where a series of more detailed questions and answer categories might be more appropriate.

This document specifies the question wording, response categories and the question administration, both face-to-face and over the telephone. It also provides an explanation of the design.

*The research carried out that informs the design is as follows:*

## **Information reviews**

- Review of the literature
- Review of the legal framework
- Reviews of UK and International Surveys

## **Quantitative research**

- Four staged Omnibus survey trials
- Longitudinal General Lifestyle Survey (GLF) pilot<sup>2</sup>
- Analysis of proxy data<sup>3</sup>

## **Qualitative research**

- Telephone interviews with those who preferred not to answer
- Feedback from interviewers and field observations
- Focus groups with members of the public
- Cognitive/in-depth interviews with members of the public<sup>4</sup>

## **Overview**

The research suggests that the question design will be successful in the field when introduced on the component surveys of the IHS in January 2009. These include:

Annual Population Survey

Labour Force Survey

English Housing Survey

Living Costs and Food module (formerly Expenditure and Food Survey)

General Lifestyle module (formerly General Household Survey)

Opinions module (formerly Omnibus Survey)

---

<sup>1</sup> Papers relating to the project can be downloaded from:

<http://www.ons.gov.uk/about-statistics/measuring-equality/sexual-identity-project>

<sup>2</sup> Expected publication date early 2009.

<sup>3</sup> Expected publication date early 2009.

<sup>4</sup> Expected publication date early 2009.

During testing the reaction from the public was generally positive, as was the reaction from survey interviewers. No visible effect on survey response rates was observed. Item-non-response rates were at an acceptable level. Furthermore the question can be administered in a similar way to other core survey questions, with the exception of the mechanism to maintain privacy when interviewing concurrently. This includes the use of concealed show cards in face-to-face interviews and repetition of response categories when administered over the telephone, both of which have proved successful, from the interviewer and respondent perspective.

## 2 Question design

The question is for administration to all people aged 16 or over, except where translators are being used or the questionnaire is being asked by proxy. The question should be treated as an opinion question.

### 2.1 Face-to-face administration

The question shown below is designed for Computer Assisted Personal Interviewing (CAPI). A concealed response show card form of administration should be used.

ASK ALL AGED 16 OR OVER

INTERVIEWER: Allocate all cards, then ask the question to all

[NAME] SHOWCARD 1, [NAME] SHOWCARD 2, [NAME] SHOWCARD 3 etc

**Which of the options on this card best describes how you think of yourself?**

**Please just read out the number next to the description.**

(ONLY IF CONCURRENT INTERVIEW)

**The numbers on each card are different for each person.**

*[Blaise table P1, P2, P3 etc]*

27. Heterosexual / Straight

21. Gay / Lesbian

24. Bisexual

29. Other

*(Spontaneous DK/Refusal )*

Spontaneous 'don't know' answers and refusals can be recorded by the interviewer but are not presented as options to respondents.

Interviewers read only the question stem.

The issue of privacy was a concern to some research participants. When interviewing more than one person concurrently, to maintain privacy as far as is possible in a household environment, each person should be given a different show card. An eight show card system is suggested based on the maximum likely number of concurrent individual interviews. Numbers should be unique across all cards.

Respondents answer using the number on the show card associated with the description. The numbers on the card should not be consecutive, so that no inference can be drawn as to the response category or its associated number.

There is an additional sentence for interviewers to read out when one or more people are being interviewed concurrently. Interviewers thought this important so respondents know why they might read out a different number to other household members. It might also enable respondents with disclosure concerns to answer honestly.

In CAPI the response options should not be visible on the interviewer's computer screen. This is again to help maintain privacy should the interviewer's laptop screen be visible to someone else in the household.

In cases where respondents are blind, have sight difficulties or are unable to read, the interviewer can read out the response options from the card but only if they are interviewing one person on their own and privacy is assured.

'Please specify' is not included with the 'other' category for privacy reasons.

If there are any legitimate circumstances in which the question cannot be administered interviewers should code 'refusal' rather than using the 'don't know' key because 'don't know' could be a valid response to this survey question.

## **2.2 Telephone administration**

The question below is for administration in telephone survey interviewing. The design is intended to maintain privacy if the respondent is being interviewed in the presence of other people (whether or not they are also being interviewed on the survey).

ASK ALL AGED 16 OR OVER

**I will now read out a list of terms people sometimes use to describe how they think of themselves.**

(INTERVIEWER: read list to end without pausing.

Note that 'Heterosexual or Straight' is one option; 'Gay or Lesbian' is one option. )

1. Heterosexual or Straight,
2. Gay or Lesbian,
3. Bisexual,
4. Other

*(Spontaneous DK/Refusal)*

**As I read the list again please say 'yes' when you hear the option that best describes how you think of yourself.**

(INTERVIEWER: Pause briefly after each option during second reading.)

### **3 Meaning of the question**

Sexual orientation is one of the six equality strands covered by a range of legislation, along with gender, ethnicity, religious belief, age and disability. 'Sexual orientation' is the term used in the legislation as an umbrella term, encompassing several dimensions, including sexual identity, attraction and behaviour. The question is intended to ask about sexual identity, which is the dimension most related to certain groups' experience of disadvantage and discrimination.

Sexual identity is a subjective view of oneself that a person can derive in many ways. Indeed, it is up to respondents to decide for themselves how to answer the question. Essentially, it is about what a person is, not what they do. It is about the inner sense of self. The research indicates that feelings of sexual identity are more salient for some groups of people than for others. It is least salient for those who classify themselves as heterosexual.

Conceptions of sexual identity also vary across groups. Heterosexual respondents can be considered 'latent identifiers' as the concept was not necessarily something they had ever thought about before being asked the question. Sexual minority groups were much more likely to have considered the concept before and found it easier to answer the question. Gay or lesbian respondents can be considered 'conscious identifiers'. Bisexual respondents include 'reluctant identifiers' – those who were unhappy with being asked to classify themselves– 'political identifiers', and 'fluid identifiers' whose identity would change depending on time and context.

It is important to recognise that the question is not specifically about sexual behaviour or attraction, although these aspects might relate to the formation of identity. A person might answer questions about sexual identity, sexual behaviour and attraction consistently, but for

some people the answers can differ. A person can have a sexual identity while not being sexually active.

Since sexual identity is subjective, and for some people can change over time or in different contexts, this is an opinion question rather than one of verifiable fact. If a respondent asks what it means, they should be told it is whatever it means to them. If a respondent has problems comprehending particular terms in the response categories, assistance can be provided, but without being prescriptive. See section 7 below.

The question works best if it does not include the term 'sexual identity': people are able to draw appropriate meaning from the wording used in the question and the response categories given.

## **4 Introduction to the question**

As a general principle, the sexual identity question should be treated just as any other survey question. Testing has shown that the question should not be given any special explanation about its meaning, purpose or confidentiality (the last two can be explained if a respondent asks). This is because in doing so the question is highlighted to respondents through the fact that it is being treated differently to other core classificatory questions. A general explanation at the start of the survey or in accompanying literature about the voluntary nature of questions is sufficient.

However, following interviewer debriefings and the cognitive/in-depth interviews, a decision was taken to include a sentence for interviewers to explain that different response option numbers appear on different show cards when interviewing concurrently. It is worth noting that a general introduction to the use of show cards at the start of the survey interview was tested during the cognitive stage of the research. Findings indicated that respondents were unable to absorb and remember what had been said later on in the interview. Therefore this latter approach is not recommended.

## **5 Question stem**

The term 'sexual identity' is not mentioned in the stem of the question. Qualitative research shows participants' familiarity with, interpretation of and preference for the terms 'sexual orientation', 'sexual identity' and 'sexuality' varied. For some participants they were synonymous, while for others their meanings varied. Using any of those terms might lead to some respondents inferring that the question was about sexual attraction or behaviour rather than identity, as defined above. Indeed, some transgender respondents equated the term 'sexual identity' with biological sex. To include the term with a definition would be difficult since it cannot be defined sufficiently briefly in such a way that could be commonly understood. Only a suite of questions on different dimensions of sexual orientation would be able to address the complexity of the concept of orientation, which would be inappropriate for a core classification question.

Furthermore, the word 'sexual' might surprise or inhibit some respondents, particularly in a concurrent interviewing environment where they could be interviewed alongside other family members.

Similarly asking '*...do you consider yourself to be?*' or '*which of these are you?*' were considered too direct and could be interpreted to be asking about attraction or behaviour. Qualitative research participants preferred '*...best describes how you think of yourself*' because it was more 'open', 'softer' or 'friendlier', and reflected better the concept of sexual identity. Furthermore, this form of words demonstrated recognition that answers were qualified, fluid or could be different in different contexts, so appeared not to be trying to label people in the same way as the alternative question stems tested, which also implied mutual exclusivity. In addition, this 'softer' approach gave respondents a little more thinking time because it was less direct and therefore came as less of a surprise during the introductory classification section of the interview. However it is important to note that changes in the question stem wording did not affect the estimates produced when tested quantitatively.

## **6 Response categories**

Discrete response categories are recommended, rather than measuring the concept on a continuum which is more relevant to sexual attraction or behaviour.

### **6.1 Heterosexual/straight**

Evidence from this research and elsewhere indicates that not all heterosexual people understand the term 'heterosexual'. The term 'straight' further defined it for some but not all. Among some young women and some older ethnic minority women, one or other term was not understood. However, no acceptable alternative term was found. Straight is a commonly used, albeit colloquial, term. While it could be considered that the opposite of straight is 'bent', this was mainly of concern to heterosexual participants and not to gay/lesbian/bisexual participants who were familiar with and used the term 'straight'. No other terms used to describe heterosexuals would be acceptable. 'Normal' was a descriptor used by some. This would carry the implication that lesbian, gay and bisexual (LGB) categories were abnormal, and assumes that LGB people do not describe themselves as normal.

It is recommended that the terms heterosexual and straight be separated by the use of a slash (/) rather than the word 'or' when administered face to face using showcards which can imply that the terms have different meanings. The use of inverted commas around the word straight is not recommended as it makes the term stand out.

### **6.2 Gay/Lesbian**

Gay and lesbian form a single category, rather than separate categories, on the assumption that sex or gender is asked as a separate question. This was generally accepted by qualitative research participants.



Again the use of a slash is more appropriate when administered face to face using showcards than the use of the word ‘or’, which can be inferred as making a distinction on gender lines, since some women identify as gay rather than lesbian.

The term ‘homosexual’ should not be included in the response options. Gay/lesbian participants felt that there were historical connotations associated with the term homosexual which was a term used when homosexuality was still illegal and as such could be seen as derogatory. Furthermore, the term ‘homosexual’ was often imposed on them by the medical profession and was considered to be related to behaviour and not identity. It is recognised that the term ‘gay’ was currently the preferred term making it more acceptable and accessible but that this may change over time.

### **6.3 Bisexual**

The term ‘bisexual’ was disliked by some of the bisexual participants because it labelled them on the basis of their attraction to, or relationships with people of both genders. To them gender was irrelevant. They disliked being categorised at all. Categorisation was seen as a need of other people, to find a place for those who did not consider themselves to be heterosexual/straight, or gay/lesbian. From this perspective some people developed a bisexual identity in a political sense in order to be distinct from other groups which tried to ‘claim’ them; and to be counted and be ‘visible’ in society.

No suitable alternative terms for bisexual have been identified in the research. The only terms suggested would imply that attraction/behaviour were being asked about, or refer to concepts other than sexual identity (for example, pansexual, polyamorous, gender apathetic). However, participants did understand the term bisexual and recognised that it was commonly used and understood by others. They said they would be willing to answer ‘bisexual’ to a sexual identity question.

### **6.4 Other**

Sometimes including an ‘other’ response option is used in question design as a catch all and respondents are used to seeing/hearing it offered in surveys. While the three substantive categories might cover all sexual identities some people may feel no sense of sexual identity at all, such as those who would describe themselves as asexual, and may be unhappy about being forced to choose if only three are offered.

Evidence from quantitative trials showed that the proportion who answered ‘other’ was small and comprised mainly those who were heterosexual but did not understand the terminology (particularly before the introduction of the word ‘straight’ in the question design). A follow-up ‘specify’ question is not recommended for this reason, and because it would be difficult to administer orally when other people might be able to overhear.

However, some of the non-heterosexual qualitative research participants said they would choose the ‘other’ category because they were either anti-categorisation (mentioned by some bisexual participants) or rejected the “*simplistic*” male/female “*gender binary*” (mentioned by some transgender participants).

It should be recognised that those who use this category will in effect be excluded from the analysis.

## **6.5 Response category order**

Reliability has been tested by varying the order in which the response options were shown in a quantitative split sample experiment. Differences between the estimates produced were not statistically significant. Some heterosexual respondents were confused by the term bisexual thinking the two terms equated, so it was decided to keep the response order as detailed above, rather than listing gay, lesbian or bisexual before heterosexual.

## **7 Assistance with meaning of response categories**

When the question is interviewer-administered, some assistance may be given if a respondent asks what any of the response categories mean. For example, "*heterosexual or straight might mean, for example, that a person is attracted to people of the opposite sex*". For gay/lesbian, substitute '*the same sex*' for '*the opposite sex*'. For bisexual, substitute '*both sexes*'. It is likely that such explanation would only be required for heterosexual respondents.

If a respondent says "*I'm normal*" or "*I'm ordinary*", the interviewer should not assume that means heterosexual/straight, but probe by asking, "*So which answer category would you choose?*" If the respondent is uncertain, assist as above.

If the interviewer suspects an erroneous answer, it should nevertheless be accepted without question unless the respondent specifically asks for any help with meaning.

## **8 Missing values**

In designing the question, attempts have been made to strike a balance between maximising the substantive data and minimising differences due to data collection mode. No separate response category of 'prefer not to say' (PNTS), or 'object to answer' or similar, is recommended when the question is interviewer-administered (face-to-face or telephone). However, 'don't know' and 'refusal' options should be made available for the interviewer to use if these answers are spontaneously given.

There was little evidence to suggest that respondents objected to the absence of a PNTS option when interviewer administered. There was sensitivity to the questioning among people who identify with a religion where homosexuality is not accepted, although the question was still answered. While they were not against the questioning as such, it was the administration in a concurrent household environment that was the concern. That is, how they would be perceived by other household members when answering the question.

Over the course of the development the proportion of respondents reporting PNTS, don't know or refusal reduced from 5 per cent in the first Omnibus trial to 1 per cent in the last Omnibus trial. It is thought that this is partly due to comprehension problems with the term heterosexual

(the introduction of the term straight aided comprehension in this respect) and interviewer confidence in administering the question. In the first trial respondent characteristics associated with PNTS were: being older, being female (particularly if the interviewer was male); having no qualifications; belonging to a lower socio-economic group; living in London; and coming from a deprived neighbourhood.

The Omnibus survey interviews only one person in the household. However in results from the GLF pilot, which interviews all people in the household, the proportion of missing data rose to 5.8 per cent in the first four months of the trial. Analysis indicates that this may primarily be the result of a few interviewers who appeared reluctant to ask the question<sup>5</sup>. This issue will need to be addressed in interviewer training.

## **9 Interviewer-administered or self-administered**

Testing has shown that the question can be administered face-to-face using show cards. The method for face-to-face interviewing allows interviewers to read the question stem but to use show cards simultaneously with more than one respondent. This method also keeps the interview speed going.

In cases where respondents are not able to read a show card for whatever reason, as mentioned above the interviewer can read out the response options if privacy is assured.

Initially the question was trialled using Computer Assisted Self Interviewing (CASI) as it was thought to be the best way to ensure privacy and maintain confidentiality. Since the sexual identity question was the only identity question administered by CASI, the use of the laptop interrupted the flow of the interview, and highlighted the question to respondents as being potentially sensitive. Furthermore, not all respondents were able to use a laptop computer.

In addition it was felt that in the concurrent interviewing environment, where more than one person in the household is interviewed at the same time but where only one laptop computer would be available for use, it would be even more time consuming to pass the laptop around and would provide household members with the opportunity to view and discuss each others responses. It was also felt that interviewers would have less control over the administration.

Providing interviewers with the option to choose mode is not recommended. They would not be able to reliably judge which mode was most appropriate simply based on their opinion of the respondent's ability to use the laptop or preference for one mode. Furthermore, a mixed-mode methodology would mean the reintroduction of a PNTS option which is appropriate in self-completion but can be replaced by spontaneous DK/REF options in CAPI thereby reducing the proportion of missing data.

---

<sup>5</sup> Preliminary results of the GLF pilot can be found on the ONS website at: [www.ons.gov.uk/about-statistics/measuring-equality/sexual-identity-project/quest-test-and-implem/index.html](http://www.ons.gov.uk/about-statistics/measuring-equality/sexual-identity-project/quest-test-and-implem/index.html)

## 10 Proxy data collection

Proxy data collection for sexual identity is not recommended on grounds of acceptability and accuracy. Qualitative research participants said it was an individual's right to provide this information for themselves. This applied even in respect of spouses or partners. Furthermore, participants said that since it was obvious to them that the accuracy of the data would be affected by asking for proxy information, they said it would affect how they thought about the quality of the organisation collecting the data.

Excluding proxy data collection may impact on the precision of the estimates due to the reduced effective sample size. However work carried out by statistical methodologists suggests that the inclusion of proxy responses would make little difference to the numbers of years of aggregated data required to achieve lower geographical estimates from the IHS.

There is however a potential for bias and this is still under investigation. Analysis completed so far has mapped the demographic characteristics of proxy respondents. Further work is being conducted to assess whether proxy respondents differ systematically from respondents who are interviewed in person to assess the likelihood of the estimates being biased as the question is not asked by proxy<sup>6</sup>.

## 11 Use of interpreters

It is recommended that the question should not be asked if an interview is being conducted using another household member, friend or relative as an interpreter, for reasons of privacy and accuracy of translation. This includes signers for people with hearing impairments.

However, the question could be administered through a professional interpreter not known to the respondent if the respondent is being interviewed alone in private.

The question has not been translated into any other languages.

## 12 Location of the question

The sexual identity question should be located with other identity questions immediately following the question on ethnicity. On the Omnibus survey trials, the question was located in the identity sequence as follows: National Identity, Ethnicity, Sexual Identity, Religion. In development work interviewers and qualitative participants said it made more sense to them and required less explanation if asked at this point, so maintaining the flow of the interview. In fact, some participants said they would be suspicious if the question was asked out of context and appeared to be '*slipped in*' later in the interview.

---

<sup>6</sup> Proxy response report is scheduled for publication in early 2009

The location of the question before or after the religion question was trialled on the GLF pilot, with the order of the sexual identity and the religion question being switched half-way through the trial to test for order effects. In the first three months sexual identity was asked before the question on religion, and in the subsequent three months after religion. The evidence so far supports its inclusion before religion as in the earlier Omnibus trials. Results of the GLF pilot using the full six months data will be published early in 2009. However, this is based on preliminary analysis using only the first four months data. The analysis is therefore incomplete with only one month's data available where the question was placed after religion. Analysis on the full dataset is currently taking place and is scheduled for publication in January 2009.

### **13 Interviewer training**

Interviewers should be aware that all respondents aged 16 or over must be given the opportunity to answer the question. The research indicated that some interviewers were reluctant to ask the question. In particular, some male interviewers were reluctant to ask the question of women over the age of 60.

One of the key elements of interviewer instructions and training should be to give confidence to interviewers about asking the question. Interviewers should be assured, on the basis of testing with respondents and feedback from the majority of interviewers involved in the quantitative trials, that in general they will not have any problems.

It is important that interviewers do not appear nervous or hesitant when administering the question. It is also important that interviewers are not perceived to react in a negative way; a chance remark or a sign that the interviewer finds the question amusing could influence a person's response or whether they give a response at all. This applies to telephone and face-to-face interviewers alike. Even facial expressions and body language that gave an impression of nervousness or awkwardness on the part of the interviewer could have the same effect.

It is important for the interviewer to be perceived to take the issue of privacy seriously.

Furthermore, interviewers should not make any assumptions about the way in which respondents may answer the question based on observation. For example, being married may be a good predictor of sexual identity but interviewers should not assume that because someone is married to someone of the opposite sex that they would automatically describe their sexual identity as heterosexual.

It is important that interviewers understand why the data on sexual identity are being collected in case respondents ask for that information or object to the questioning. A form of words will be provided for them to use in the interview. It is also important for interviewers to understand the concept of sexual identity. Therefore detailed written interviewer instructions will be provided but the administration and conceptualisation of the question should also be fully briefed in survey field training.

## **14 Transgender**

Transgender is not one of the response options for the sexual identity question. Transgender is often mentioned along with lesbian, gay and bisexual, for example there are LGBT support groups and lobby organisations. It has sometimes been listed as a response option at questions on sexual orientation/identity. However, transgender (along with other related terms such as transsexual and intersex) is a different concept, relating to gender identity and biological sex, not to sexual orientation.

The research suggests that transgender people are able to answer a question on sexual identity on the same basis as everyone else involved in the research. However, similar to some of the bisexual participants, there are those who reject the premise of categorisation on the basis of whom people are attracted to and would prefer to use the 'other' category.

As mentioned in section 5, among the transgender community 'sexual identity' has a different meaning to that described above. It means biological sex, as distinct from gender identity.

## **15 Qualifications**

It should be noted that the questioning is, de facto, about sexual identity: i) in the environment in which the question is being asked; and ii) at the time it is administered. A person's public sexual identity can differ according to context, for example a lesbian, gay or bisexual (LGB) person might be 'out' at home but not at work, or with friends but not family. It is also possible that a person's sexual identity could change over time.

The development work indicates that some socially desirable answers can be expected. For example, an LGB respondent might answer heterosexual for a number of reasons. These include perceived disapproval or prejudice on the part of the interviewer, concerns about privacy in the presence of other people during the interview, concerns over confidentiality and who will be able to access their data, and concerns about the uses to which it will be put. The extent of socially desirable answers given is unknown.

The testing has been limited to an interviewer administered household survey environment. The question has not been tested in other environments, such as surveys conducted by post, the internet or in public places.

## **16 The future**

It is likely that over time changes in society with respect to the subject matter may affect the data and how the question is administered. That is, as the questioning becomes more common place and the terminology more familiar, respondents will become more confident in answering. This may mean that the way in which the question is administered can be reviewed. Also, such changes in society may affect how sexual identity is conceptualised, as well as the actual language and terminology used by the public. For example, an additional answer option might be

required for those against categorisation on basis of gender attraction. All of which may need to be reviewed in the future.

As data are accumulated over time investigating cohort changes will be interesting as would utilising the longitudinal element of the IHS at an individual level.