Subjective Well-being:

a qualitative investigation of subjective well-being questions

(Working paper for the Technical Advisory Group on 29 March 2012)

Katherine Ralph
Kim Palmer
Jayne Olney

December 2011

Contents

1   Executive Summary

1.1   Methodology

In October and November 2011, 44 interviews were conducted to examine eight subjective well-being (SWB) questions. The first four questions (satisfaction, worthwhile, happy and anxious) are already included on the Integrated Household Survey and the remaining four (personal relationships, loneliness in daily life, loneliness yesterday and control) are extra questions that ONS has been investigating for possible inclusion following feedback in February 2011 from users.

The four subjective well-being questions are currently administered to respondents by telephone as well as face-to-face interviewing. Therefore the questions in this study were administered both by telephone interviews and through face-to-face interviews. For telephone interviews, respondents were called from the interviewers' mobile phones outside their homes (and asked the first four questions (satisfaction, worthwhile, happy and anxious), they then participated in face-to-face cognitive interviews immediately afterwards.

Purposive sampling techniques were used to obtain the sample. The primary stratifiers within the sample were sex, age and socio-economic group. The sample also included respondents from a range of minority ethnic groups, varying educational attainment, a range of household types, a geographical spread and those with long term limiting health issues. It should be noted that qualitative samples are not systematically random samples from which statistical inference can be drawn. The groups are chosen on the basis of them having an association or relationship to the research subject.

1.2   Results

1.2.1   General reaction and understanding of the purpose of the well-being questions

The general reaction of respondents to the eight questions and understanding of the questions' purpose was investigated. Respondent's reports were characterised as:

- Positive: An enthusiasm for completing questionnaires and/or an interest in the topic area. In terms of the purpose of the questions it was felt that the results could be used to help people who were experiencing problems in their lives.

- Negative: The value of the well-being questions and the project in general was questioned in light of the current economic climate. As people may be reluctant to admit to being anxious or unhappy there was a concern the results would not be accurate. Additionally apprehension and unease over the possible negative impact of the questions on respondents was expressed. There was also concern that the purpose of the questions was for government propaganda.

- Neutral: For some respondents it was felt that *"these things need to be done"*, so answering the questions was not a problem.

1.2.2 The preamble to the well-being questions

The preamble explained the purpose of the questions before continuing to make clear to respondents how to score their response. Previous research had found some misinterpretation of the preamble (Dalton, 2011). Therefore the new introduction to the questions, currently being used in ONS household surveys was tested. However, the purpose was not remembered, probably due to the cognitive demands of applying the scale. This did not have any significant impact on respondents' ability to answer the questions.

1.2.3 Question testing

Each of the questions was considered in terms of:
- understanding of individual questions and terminology
- alternative terminology
- the process for selecting answers and level of difficulty in responding
- potential interviewer effects and societal expectations

Where relevant the time frame used to answer questions was also investigated.

Overall, how satisfied are you with your life nowadays?
The word 'satisfied' was not understood uniformly, although there was evidence that respondents with a positive interpretation had understood the question as intended. When respondents interpreted the word 'satisfied' in a negative or neutral state it was felt that this was something not to aim for or that it meant something that was neither good nor bad.

In wave three of the project the word 'content' was tested as a possible alternative to 'satisfied'. The word 'content' was thought of as comparable to 'satisfied' but less likely to be associated with a negative state. This suggests that the term 'content' could be a viable alternative to 'satisfied' should further concerns be identified with the latter.

When respondents were asked about the term 'nowadays' they described this term as old fashioned and was either not understood or ignored.

Overall, to what extent do you feel that the things you do in your life are worthwhile?
Good comprehension of the worthwhile question was demonstrated by most groups within the sample. However, answers were affected by what was seen as social desirability, which led to inflated scores. Where the question was not understood this tended to be by those with lower educational attainment. This group simply did not understand the term 'worthwhile'.

For the most vulnerable respondents, answering this question was distressing and in some cases respondents became visibly upset. It is recommended that ONS investigate the possibility of creating a flier that interviewers can leave with respondents, which tells them where they can seek help if it is required.

There was also uncertainty about the timeframe within which to consider the question.

The word 'purpose' was tested as a possible alternative in wave three. Further testing is recommended as this is potentially a more useful term that may be better understood and have a less negative impact on respondents (section 4.5.2 for discussion).

Overall, how happy did you feel yesterday?
In relation to this question the terms 'content' and 'joyful' were tested as alternatives to 'happy'. Respondents felt that 'content' was a less transient state and more stable. In contrast 'joyful' was met with negative reactions as it was thought to sound religious.

The original wording of the question using the term 'happy' was widely understood. However there was resistance by respondents to being asked to provide a rating for a single day. Respondents preferred to give *"an accurate"* account of their happiness over a longer time period, rather than how it was on one day.

In wave three of testing respondents were asked to report how happy they felt yesterday, even if it were not a typical day. This appeared to give the question stronger face validity and solve the problems identified in earlier waves of testing.

On a scale where nought is 'not at all anxious' and 10 is 'completely anxious', overall, how anxious did you feel yesterday?
Terminology within the question was understood well, however there was some preference for the term 'stress', which was thought to be a more useful term as it was felt to have fewer stigmas attached.

The scale was a particular problem at this question because a group of respondents reversed it when answering. This was because in the previous questions, scores close to 10 reflected a positive state but at this question, scores close to zero represented a positive state. The use of show cards is recommended for face-to-face interviews.

Overall, how satisfied are you with your personal relationships?
The term 'overall' was understood as meaning in general or on average. In contrast the wording 'personal relationships' led a group of respondents to think of their partner, or lack of a partner rather than to consider wider relationships, such as parents and friends. An instruction was added asking respondents to consider these other relationships; however testing of this was limited. The addition of a guidance note to this question may be useful to ensure that respondents consider relationships beyond that of partners or significant others.

On a scale of 0 - 10 where 0 is 'not at all lonely' and 10 is 'extremely lonely' how lonely do you feel in your daily life?
The scale was problematic at this question as respondents reversed its polarity, meaning incorrect scores were given. In addition, respondents who were not in employment thought the term 'daily life' was not relevant to them as it implied routine, which they no longer had. Thinking about loneliness also resulted in vulnerable respondents becoming distressed.

Overall, how lonely did you feel yesterday?
The term 'overall' was understood well. Lonely' was seen as having two aspects to it: physically being on one's own and the feeling of being lonely even with others present.

On a scale of 0 - 10 where 0 is 'no control at all' and 10 is 'complete control' how much control do you feel you have over important aspects of your life?
Those who had difficulty answering this question wanted clarification as to what 'important aspects' of their life they should consider. Arriving at a response was

difficult for people in relationships who felt they had joint control over important aspects of their life, such as finances.

1.2.4   Applying the scale

Respondents tended to require the scale to be repeated to them in order for them to answer the questions. More importantly however, respondents mistakenly reversed the scale, meaning that high scores were given instead of low scores and vice versa. Show cards were very effective in addressing this but are clearly only useable in face-to-face interviews.

1.2.5   Questions that ask about 'yesterday'

These questions were contentious for a group of respondents who found it challenging to remember the previous day. Respondents did not like scoring one day as it was not seen as being representative of their general state and not a true reflection of themselves. They did not understand the methodology behind asking for the previous day and so thought the Office for National Statistics (ONS) was acting in error, which left them with a poor impression.

In the third wave of testing these questions had an additional instruction: 'Please think about 'yesterday' even if it was not a typical day'. This instruction had very positive results and none of the three points listed above were mentioned by respondents again. This is likely to be due to added clarity.

1.2.6   Question order

There was a desire to place positive questions first as they were seen as being easier to answer and started on a positive note. The anxious question was always placed last as it was seen as the most negative of the four existing questions. The second set of four questions was perceived as being negative and an order was harder to define.

1.2.7   Show cards

Two designs of show cards were tested.   One had the words 'not at all' and 'completely' at its extremities, the other had question specific labels, for example, 'not at all anxious' and 'completely anxious'. The former did not appear to help respondents applying the scale and so the latter was tested.   This proved very effective. However show cards are only useful in face-to-face interviews, therefore implications for comparisons between this data and data from telephone interviews would need consideration.

1.2.8   Answering in front of other household members

When respondents did not object to this, it was because they felt that everyone in the household knew them and nothing was concealed from others. When respondents were not comfortable about this, it was because the topics within the questions were seen as being private and not for open discussion within the household.

1.2.9   Social desirability

There was evidence that respondents wanted to appear in a more positive light than their feelings or circumstances actually were.

1.2.10 Answering by telephone

This was perceived by respondents to require greater cognitive resources than face-to-face interviews. Respondents who actually answered the well-being questions on the telephone felt there was greater risk of errors in this mode due to greater cognitive burden. Telephone surveys were also seen as being less serious than when an interviewer actually calls at a respondent's home.

There was an acknowledgement that answering the well-being questions by telephone would maintain confidentiality as other household members would not hear the questions.

1.2.11 Aspects of well-being

Respondents largely thought that the topics in the questions were relevant to well-being. However, surprise was expressed that health was not included as this was seen as crucial to well-being. Other aspects said to be part of well-being were: job satisfaction, economic security, the environment, crime and religion.

1.2.12 Field issues

Interviewing does not routinely take place on Sundays; therefore the questions that ask about 'yesterday' would always miss out Saturdays. Students noted that their responses could be significantly different in term-time and holiday time.

1.2.13 Data security

High profile breaches in data protection were mentioned when respondents were asked how they felt about providing ONS with well-being data. However, the picture was quite balanced and the topics were not thought to cause great concern regarding data protection and confidentiality.

2 Introduction

2.1   Background

As part of the National Well-being Programme[1] the Office for National Statistics (ONS) has developed four subjective well-being (SWB) questions. The four SWB questions were introduced on the Integrated Household Survey and continue to undergo testing and development. Research to date has indicated that overall respondents feel able to answer the questions with low item non-response rates (Hicks and Tinkler 2011, and Dolan 2011).

The current qualitative research project was undertaken to further understand how respondents answered the SWB questions. Respondent interpretation of words and concepts underlying survey questions is an important consideration of data quality. This project investigated the cognitive processes respondents undertook to understand the SWB questions. Ensuring that the burden placed on respondents is minimised is also a key requirement and further testing was required to see whether any changes to the questions and pre-ambles were important. Another aim of this research was to examine how other potential subjective well-being questions, concerning social relationships and control, were viewed by respondents.

The aims of this project therefore were to;

- assess the functionality of the preamble
- examine the use and suitability of adjectives within the questions and test alternatives
- gain an insight into how respondents understand the questions;
- gain an understanding of how respondents use the scale
- explore privacy and sensitivity issues
- investigate order preferences
- examine the functionality of the well-being questions when administered by telephone

3 Methodology.

3.1   Literature review

A small scale literature review of past well-being research was conducted. Useful papers were available from Stiglitz et al (2009), Helliwell, (2011), Hicks (2011), Waldron (2010), Dolan et al (2010), Pudney (2010) and Rees et al (2010).

Three main conclusions can be drawn from previous findings;
- that subjective well-being is a valid construct that can be reliably measured enabling the researcher to gain insight into a respondent's subject view of aspects of their life;
- that the information acquired should be divided into discrete categories evaluating the concepts of evaluation, experience and eudemonics should be investigated as three separate phenomenon. As an aside, it is also recommended that as well as the above constructs also be evaluated at an

---

[1] http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/index.html

integrated level whereby the researcher can identify any relationships between the three variables which may impact on the data obtained.

- it is also important that order effects be taken into consideration, more specifically the potential impact that question order have on the quality of the data obtained.

ONS has previously conducted a cognitive testing study to investigate the placement of the subjective well-being questions within the Integrated Household Survey (IHS) and ascertain the potential risk of order effects. (Dalton et al, 2011). Conflicting advantages and disadvantages were found with the three tested positions of the questions and further research was recommended.

3.1.1 Background to the current research

The Labour Force Survey (LFS) conducted a small scale (16 interviews) cognitive test of the four primary SWB questions (ONS, 2011) earlier this year. The problems highlighted by this study are as follows:

- the scale had to be repeated at all four questions at the request of respondents
- the 'worthwhile' question was considered for longer than any other question before an answer was given (possibly due to societal desirability)
- when respondents were asked how 'happy they were, they questioned why this only related to 'yesterday'
- respondents mistakenly reversed the scale when asked the question relating to anxiety.

Due to sample size, the LFS research did not include a range of age groups, ethnicity, educational attainment, employed and unemployed, the long term sick/disabled or a geographic spread. Despite this, the LFS research findings are very similar to the current findings, although the depth and scope of the latter are much more detailed.

3.2   Data collection method

3.2.1 Sampling frame

A proportion of respondents were recruited from an existing ONS register of people who expressed a willingness to take part in research. The remainder were recruited through snowballing. Snowballing is a technique whereby researchers identify individuals with the characteristics of interest via other individuals who have the same or similar traits.

Purposive sampling techniques were used in order to obtain the sample. Purposive samples deliberately target groups within the scope of the research to allow a detailed analysis of these groups. The groups are chosen on the basis of them having an association or relationship to the research subject. Qualitative samples are not systematically random samples from which statistical inference can be drawn.

3.2.2 Fieldwork

Fieldwork took place in three waves:

- wave one was comprised of 14 individuals who were interviewed face-to-face for both the administration of the questions and the qualitative interview

- wave two contained 21 individuals. In this wave, the questions were administered by telephone to 13 individuals. Interviewers telephoned respondents using mobile phones outside respondents' homes to administer the questions. The interviewer then conducted the qualitative interview face-to-face with the respondent immediately after the telephone call. This allowed for a more realistic discussion about being asked the questions by telephone, rather than simply asking respondents for their opinion about this. The remaining eight respondents were interviewed face-to-face

- wave three comprised of 10 interviews six of these were face-to-face interviews and four were telephone interviews

A pause and review was held after each wave to inform the next wave of testing. It was unlikely that the geographical position of respondents would be relevant to their understanding of the questions, so geography was not a sampling factor. However, within the sample, there were respondents from England, Scotland and Wales, and from rural and urban areas. The full sample description is illustrated in appendix A.

See Appendix B for further methodological detail concerning conducting qualitative research, sampling frame, calling strategy, incentive payment, recruitment strategy, interview mode and strict quality control measures that were applied.

3.3 Analysis, reporting and interpretation

Qualitative data is disordered and complex therefore data from the interviews went through four key analysis stages; the development of a thematic framework, the ordering and summarising of the data, classification and categorising the data and finally, interpretation. A matrix was used to assist with these stages, which allowed analysis within and between cases.

The report describes the findings from the data; from the themes that developed when ordering and summarising the data. Verbatim comments to illustrate themes are shown in italics. Recommendations are made wherever possible.

Qualitative techniques produce rich, detailed data but cannot be used to inter statistical conclusions. No statistical inference can be drawn from the findings, as this is not the purpose of qualitative research.

4 Main Findings

4.1 General reaction to the primary four well-being questions

Interviewers documented respondents' reactions to the four questions by recording verbatim comments and probing for further clarity when appropriate.

There were positive, negative and neutral reactions to the questions. Positive reactions included:

- a willingness to respond
- an expressed interest in the topic area.

When a willingness to respond was expressed, this was associated with an enthusiasm for surveys. When interest in well-being was expressed, the topic had been considered previously or thought about seriously.

Negative reactions included:

- wasting money
- government manipulation
- causing negative mood

Respondents who had a negative reaction thought that both the four well-being questions and the well-being programme as a whole were a waste of public money, this was due to the current economic climate and publicised cuts to front-line services:

When the programme was seen as a government manipulation it was felt a lot of people would not want to admit they are anxious or unhappy and the questions do not have a context. Therefore it was felt the government can gain easy positive results.

It was felt the questions could lead to a negative mood as they forced respondents to examine their life honestly. Neutral reactions included a feeling that *"these things have to be done"* so taking part was not a problem.

4.2 Purpose of the well-being questions

When asked why these questions would be asked in a survey interview and what the data collected might be used for, the well-being questions were thought about in the following ways:

- to allow help to be provided to people experiencing problems.
- to be used by the government as propaganda for the next election campaign.
- it was not known what the purpose was

When the purpose was seen to be to provide help, it was thought different groups could be compared to assess the equality of society and then targeted help could be given. It was also thought the information might be used to assess the mental health of the population.

When the questions were seen as providing propaganda, distrust of government motivation was expressed. It was felt that as no context for the questions was provided they could be used to make things appear more positive than they are and this could be used as a justification for further cuts to services. ONS was seen as a government agency collecting data for David Cameron and it was not known or believed that ONS was an independent statistical agency.

When the purpose was not known, it was felt it had not been clear that the questions were about subjective well-being and this caused confusion regarding the point of providing answers.

4.3 The preamble to the well-being questions

A preamble was tested previously by ONS (Dalton 2011) during the survey design stage where decisions were made regarding which questions and preamble to include in the ONS Integrated Household Survey and the Opinions Survey. This version was:

> Next I would like to ask you some questions about how you see your well-being. For each of these questions I'd like you to give an answer on the scale nought to 10, where nought is 'not at all' and 10 is 'completely'.

This version was found to be problematic in that the term 'well-being' was thought of in terms of 'health' and so the questions were not what respondents expected. Dalton also recommended removing the instruction about the scale because there was evidence it could cause concern among respondents about their ability to use the scale. Placing the scale instructions within the question was also recommended. This was tested in this research with the 'anxious' question however, and was not successful.

For the present investigation, the preamble currently being used by the IHS and the Opinions Survey (OPN) was used, in which the questions were preceded by wording that explained to respondents what to expect, reassures them and instructs them how to score each question. The term 'well-being' was replaced with 'aspects of your life':

> Next I would like to ask you four questions about your feelings on aspects of your life. There are no right or wrong answers. For each of these questions I'd like you to give an answer on a scale of nought to 10, where nought is 'not at all' and 10 is 'completely'.

During cognitive probing, respondents said they did not remember the two sentences prior to the instruction on how to score the questions. It is likely that this is due to the cognitive burden resulting from the instruction. Respondents must listen to the instruction, understand and remember it, then put it into operation while simultaneously processing the demands of the question.

However, if respondents cannot remember the first two sentences, it is not particularly problematic and is quite reassuring that cognitive resources were devoted to scoring.

Recommendation

The first part of the preamble was not remembered; however no evidence could be found to suggest this had any effect on responding to the questions. There is therefore no concern about the preamble being used as it is.

4.4 Satisfied with life

> Overall, how satisfied are you with your life nowadays?

4.4.1 Understanding of the question

In general being satisfied was interpreted as being in a positive state. However there were instances where satisfied was identified as either a negative or neutral statement. Therefore the term was not consistently understood.

When interpreted as implying a positive state, the question was thought about as the extent of:

- feeling one's needs have been completely met
- peace and contentment felt in life
- happiness felt in life

Presumably the intention is for 'satisfaction' to be understood in a positive way, therefore this group of respondents understood the question as intended.

When interpreted as implying a negative state, 'satisfied' was understood as *"not something to aim for"* as it meant *"just about OK"*. The implication of this could be inflated scores as respondents are scoring a state that is mediocre, rather than a positive state.

When the question was seen as implying a neutral state, it was understood as being *"at a point where things aren't very bad, but neither are they very good"*.

Recommendation

The questions and data are still experimental and therefore further testing is recommended to explore these findings and their impact on the survey information collected.

4.4.2 Alternative terminology

In wave three, the word 'content' was examined. It was understood as being either a positive or neutral state. It was described as *"having a… good time"*, *"no worries"* and *"happy with your lot"*. When neutral, this was seen as being *"OK"*, *"mulling along"* and a *"lack of extremities"*, meaning neither particularly happy nor sad.

When 'life satisfaction' and 'content' were compared to each other they were thought to be similar, this was because it was felt you cannot have one without the other and that they were both *"middle of the road"* measures implying neither particularly happy nor sad.

4.4.3 Understanding the term 'nowadays'

Respondents were asked what the term 'nowadays' meant during cognitive probing. It was interpreted with five varying time frames:.

- the present moment
- the past, as far back as the last five years
- the present and the past with varying time frames
- the past, present and the future with varying time frames
- dependent on when key life events had occurred

When 'nowadays' was thought of as referring to the present moment, this was described as *"now"*, *"as we speak"* and *"this very moment, just as we are"*.

The varying times frames of the present, past and the future included the last year, *"recent months"*, *"the last chapter"* and *"slightly in the past, the present and slightly in the future; six months either way".*

Additionally, if change had taken place in respondents' lives, they thought back to a point when life was different. This could be when they were in employment or before ill health had occurred. This life event change went back several years in some cases. Respondent reports in this study suggested that they tended to consider their own lives rather than make comparisons with other peoples lives.

'Nowadays' was considered to be a dated term by both younger and older respondents and described as *"old fashioned"* and *"something my Granny would have said as she reminisced".* There was also evidence among younger age groups that 'nowadays' was not understood at all; a definition of it could not be provided because it was not known. These respondents had ignored the word and given a response to, 'Overall, how satisfied are you with your life.'

The LFS project concluded that there was not a problem with the term 'nowadays'. However that research was specifically designed to over-sample the 50-69 years age group, and therefore 'nowadays' was tested with fewer young people. This could explain the difference in findings with the current research.

The term was also thought of as *"vague"* and surprise was expressed that statistics would be published from a term that was perceived as very open to interpretation.

In wave three, the term 'these days' was cursorily explored. When it was felt to mean the same thing as 'nowadays', there were varying time frames attached. 'These days' was interpreted as referring to:

- the past few weeks and the near future
- the past week but not including the future
- the current generation

If 'these days' was felt to be different to 'nowadays' this was due to a perceived difference in time frame, with these days referring to a smaller period in time. There were no opinions expressed that 'these days' was considered old fashioned.

Respondents were asked if they had a preference for 'these days' or 'nowadays' as a term. No clear pattern of preference was discernable within the data. However, when 'these days' was preferred, it was felt this was more concise and referred to the present, whereas 'nowadays' was a more general term that did not imply a specific time frame. When 'nowadays' was preferred, this was as it sounded more professional.

Respondents were asked during cognitive probing whether their answers to 'Overall, how satisfied are you with your life?' would be similar or different to 'Overall, how satisfied are you with your life nowadays?' When answers would be similar, this was due to a feeling things would not change much in the future, for example for retired people who had stable lifestyles that did not fluctuate much.

When answers would be different, this was because 'nowadays' was seen to imply a present timeframe. For instance:

"I think I would give a different answer…when you say nowadays it makes you focus on right now and you may have had a bad week, split up with your girlfriend or something, which might drag the score down, whereas your overall life might be quite good."

Recommendation

Consider removing the word 'nowadays' and implement further testing of 'these days' as an alternative term as this was only explored in wave 3.

4.4.4 The reference period

When respondents were asked if they were thinking about the past, present or the future when answering the question, it emerged that they were using the following three combinations to answer:

- the present compared with the past
- the present compared with how the future could be
- the present, past and future considered together

When the present was compared to the past, this was related to a recent significant event that had changed their life course or acted as a marker for progression onto a new stage. For instance, comparing how life had changed since a family member had suffered an accident or since the loss of a job leading to present unemployment:

"Based on a comparison with the recent past and further past. Mostly about the fact I really hated my last job and the one before that. I am quite satisfied with the fact I'm not working and I have got a different life now."

When the present and future were thought about there was a comparison between what was satisfying in life now and what could be obtained in the future to increase satisfaction, or a weighing-up of the likelihood of achieving future goals that might lead to increased satisfaction, but that were not present currently:

"Where I am currently compared to where I want to be. If you think of it as a race, and that's where you're trying to get to in life, I gave an answer for where I think I am now."

When the future was not considered when reaching an answer it was because it was seen as a *"bad place"* and therefore not thought about. Therefore those less optimistic about the future may tend to look backwards to the past.

The period of time considered when responding ranged from six years in the past, to *"a few months"* in the future. Respondents thought back to specific life events and considered how satisfied they were since then, hence the lengthy time period.

4.4.5 Processes in selecting an answer

In order to arrive at an answer, a number of different variables were considered. This included perceived personal satisfaction with: relationships, control over life events, health, finances, work, debts, mental health, social life and home life. Answers were also based on:

- perceived contributions to other people and the community.

- a comparison between one's own situation and that of friends in a difficult situation or others in the world who may be living in third world countries, war zones and under oppressive regimes.

4.4.6   Level of difficulty in responding

When life satisfaction had been thought about before, the question was not difficult to answer and a lack of self-consciousness about answering was present. Also an event occurring recently, which was seen as a goal that was known to increase personal satisfaction, made answering easier, such as obtaining a new job.

When difficult, there was evidence that a lot needed to be thought through in order to answer. There were often delays in responding, which can be indicative of cognitive burden. It was thought a 'don't know' response option should be included for those who find the question too difficult to answer.

Recommendation

As adding a 'don't know' option may increase the number of people using this answer (Krosnick and Fabrigar, 1997) and therefore reduce data quality, Data Collection Methodology advises against the inclusion of a 'don't know' category Additional investigation would be required to ascertain whether a 'don't know' option should be considered in light of some respondents providing 'don't know' as a spontaneous option.

4.4.7   The interviewer effect and societal expectations

Satisficing in surveys (first suggested by Krosnick, 1996) suggests that optimal question answering by a survey respondent involves a great deal of cognitive work and that some people will reduce that burden by choosing socially desirable answers.

In the current study this was evident in two ways:

- expressions of concern regarding what the interviewer would think of scores
- societal expectations being considered when judging a score

Concern was expressed about being thought of as depressed if a low score was given in front of the interviewer and therefore the score was adjusted to be higher.

It was also thought that a high level of satisfaction was present, however a lower score was given as being happily unemployed in society is not socially acceptable:

> *"To be perfectly honest, I am really happy at the moment, even though I am not working and have no money. I love the house where I live and I love the people around me, and I feel really, really, content… but in order to have a better quality of life, I need to work. So I'd better say, I'd better say seven."*

Therefore answers were reached by judging how much societal expectations had been met rather than an individual's authentic feelings.

During cognitive probing exploring what would make a score of seven higher, there was further evidence that being satisfied with being unemployed is not socially

acceptable and an increased satisfaction score should be associated with a show of wanting a job:

> *"I suppose I should say a job would be marvellous but frankly that would be a lie and I could get one I hated as much as the last one and go down to a four or five."*

Due to the primary objectives of the testing, these comments were not explored in detail, as would be necessary to fully investigate how the questions are susceptible to a social desirability effect and/or to investigate if this possible social desirability effect varied according to mode of data collection (face-to-face versus telephone).

Recommendation

The question was interpreted and conceptualised in a variety of ways and with a large variety in time scale, due to each person's subjective view regarding what satisfaction is and what time period should be referred to. Societal expectations also affected scores given. As the questions are experimental, further testing is required to ascertain the affect of these findings on the survey data collected.

4.5   Worthwhile

> Overall, to what extent do you feel that the things you do in your life are worthwhile?

4.5.1 Understanding of the question

Good comprehension of the question was demonstrated by most groups within the sample. When the question was understood, it was interpreted in two main ways:

- to what extent the activities done in one's life are useful.
- how much of an impact one has on other people.

When the usefulness of activities was considered, voluntary work, paid careers of a caring nature and things done to benefit others were mentioned. When other people were considered, relationships with friends and family and the wider community were thought about.

4.5.2 The term 'worthwhile'

During cognitive probing, the term worthwhile was understood well and interpreted in the following variety of ways:

- being in paid employment
- something of non-financial value to other people and society
- something that would be a better option if compared with doing nothing
- things that make people happy
- something of benefit both to oneself and others

However there was also evidence among younger less educated age groups that the question was not understood at all and therefore a score was not given. This was evidenced by an inability to describe what the term meant to them. For example, the following took place between an interviewer and a 19 year old with low educational attainment:

> Interviewer: "*And the word worthwhile. What does that word mean to you.*"

> Respondent: (Long pause, shrugs her shoulders, shuffles, looks embarrassed).

> Interviewer: "*There's no right or wrong answers. It's just whatever it means to you.*"

> Respondent: (Long pause) "*I don't know.*"

> Interviewer: "*OK, let's say, for instance, someone says they do a job that is really worthwhile. What would you think they meant?*"

> Respondent: "*Honestly? I honestly don't know.*"

In wave three, the word 'purpose' was examined as an alternative to 'worthwhile'. It was defined as a goal or reason to undertake an activity:

> "*Having some sort of direction for all your actions, which you can orientate your life towards, like a goal or a person...a direction*".

> "*A reason to do something.*"

Respondents were also asked whether they thought the terms 'purpose' and 'worthwhile' were similar or different. Those stating that they were similar described both terms as:

> "*Similar; 'worthwhile' is doing something for a reason and 'purpose' is also doing something for a reason.*"

They were also seen as intrinsically linked:

> "*For something to be worthwhile it has to have a purpose.*"

When the two terms were seen as different, purpose was viewed as undertaking an action to achieve an end result, whereas worthwhile was seen as undertaking an action to gain general satisfaction.

Recommendation

If the 'worthwhile' question is retained, consider the comprehension needs of lower educated respondents. Provide instructions at interviewer briefings regarding help that can be given with understanding of the questions if it is needed. Also consider further testing of 'purpose' among the sample, in particular with lower educated respondents, to ascertain if this is a better understood term.

4.5.3 Overall

Due to time constraints within the interview, this term was only specifically probed at the worthwhile question. The term 'overall' was described as:

- taking everything into account.
- an average.

Taking everything into account was thought of as including:

"Everything from cleaning the toilet to looking after the family."

"Taking all things into consideration and a lot of acceptance."

"I'd say it means you should take everything into consideration."

When thought of as an average, this was described as *"generally, for the majority of the time, not covering everything."*

There was also evidence during cognitive probing that the term was not noticed when originally answering, yet this did not make a difference to responses.

Recommendation

This term did not cause any issues with this question; therefore it should not be removed. Consider further investigation of the term 'overall' as it is understood for the other three well-being questions.

4.5.4 The reference period

There was some uncertainty about what the reference period should be for this question. A worthwhile activity might have been undertaken in the past but it was ambiguous whether this should be included.

Recommendation

Consider the importance of ambiguity concerning the reference period and the possible inclusion of interviewer instructions.

4.5.5 Comparing the 'worthwhile' to the 'life satisfaction' questions

It was investigated whether or not respondents perceived a difference between the 'life satisfaction' and 'worthwhile' questions.

When a difference was apparent, 'satisfied' was seen as being about feeling whereas 'worthwhile' was seen as making a judgement. Also worthwhile was seen as how much you do for other people, whereas satisfied was seen as how much you do for yourself. For instance:

"Very different…the first one is about how satisfied you are with your life and that might not mean you're doing something that is worthwhile to other people. You could be a burglar and be very satisfied with your life or you could be the head of a blood diamond company and you could be satisfied with your life, but if you had to think about it you might think that it's not worthwhile."

When no difference was apparent and the questions were seen as similar, this is because it was believed the same things had to be thought about when reaching an answer.

4.5.6 What was considered when answering the question

When answering the question, one of the three following variables was considered:

- things done for oneself
- things done for others
- things done for oneself and others

Things done for oneself included passing a driving test, getting a degree, eating healthily and having a paid job that contributed to tax revenues. Things done for others included looking after a disabled relative, doing voluntary work and doing paid employment that helped others:

> *"My job of being a performer helps others. 'Not worthwhile' would be a job not helping anyone but myself or rich and greedy people."*

4.5.7 Level of difficulty in responding

When the question was reported to be easy to answer, this was associated with having considered the topic before. When found difficult, this was related to finding it hard to express a positive opinion about oneself:

> *"Quite difficult to credit yourself for things you do."*

This expressed modesty could potentially lead to deflated scores.

4.6 Happiness yesterday

Overall, how happy did you feel yesterday?

4.6.1 Understanding the question
This question was widely understood and paraphrases such as these were offered:

> *"What sort of emotional state was I in yesterday; was I in a happy mood, a good day."*

> *"How happy are you yourself as a person yesterday, as in the day before today."*

> *"To think about what I did yesterday and think how happy I was over the day."*

> *"I think it's asking whether I felt OK in myself, I mean not sad, it's the opposite, isn't it."*

However, for certain respondents it was important to give *"an accurate"* account of their happiness, rather than how it was on one day. Therefore, although the question was understood by this group, they were not scoring for the day before but giving a more general picture.

Issues taken into consideration when answering the question were:

- activities, or lack thereof, the previous day
- events that occurred the previous day
- how relationships had been the previous day
- what the weather was like the previous day
- how respondents saw themselves in general

The latter of these is clearly not adhering to the wording of the question as it was not specific to the day before.

Recommendation

Further discussion is needed, with consideration to the following three options:

1. Do nothing: the risks involved in leaving the question as it is might be acceptable.
2. Add instructions: the question could have the addition of an instruction to specifically direct respondents to consider only the previous day. Interviewer instructions would also be necessary to help them to direct respondents.
3. Re-word the question so that a general sense of happiness is asked for.

Testing of a preamble to this question instructing respondents to only think about yesterday even if it was atypical was adopted for use in a later wave (see 4.6.5).

3.6.2 Understanding the terminology used in the question

The word 'overall' in the context of this question was taken to mean *"in general yesterday"*, *"on the whole yesterday"* or *"all things considered yesterday"*, which are all acceptable similes.

The word 'happy' was more difficult for respondents to describe compared with the other three main questions and elicited comments such as:

> *"I don't even know what that means, happy."*

> *"Happy's a hard word to describe. It's not being sad, not being lonely."*

Such respondents did give scores to this question however, and once they overcame their difficulties, they described 'happy' with words such as 'content' and 'mood'. Indeed there was some preference expressed for the word 'content' being used instead of 'happy'. The latter being thought of as too positive and transient, whereas 'content' was thought of as a more stable, ongoing, state:

> *"You can be happy with a specific thing, but I see content as a more overall word. Happy is more associated with an action or a timescale. I think content is a more overall thing, but they are similar."*

However, 'joyful' was not thought to be suitable:

> "[Joyful is] *something that made you happy, you enjoyed or made memorable, I wouldn't ever put it in a sentence that I feel joyful, I think it's more specific and a term I wouldn't necessarily use."*

> "[I] *don't think I could ever say I feel joyful but I can say I feel happy. It is not a term I would use, it seems almost a religious term rather than everyday speak. Christmas carols spring to mind."*

4.6.2   Asking about yesterday

There were two issues related to this question asking about the previous day. Firstly, there was scepticism about the value of asking for a score on any subject for just one day. Respondents were surprised that an official organisation should ask a question in such a way as to *"miss the big picture"*. For these respondents, ONS was making an error and this left them with a negative view of the organisation. Added to negative comments about *"the government spending money on unnecessary surveys like this"*, these respondents were not left with a good impression.

Secondly, because the previous day was atypical, it did not sit well with certain respondents:

> *"It goes against the grain to give a wrong impression, especially for something official like this."*

Such respondents wanted to give a more accurate account of their happiness and so adjusted their answer:

> "*I thought it was strange it was just one day and yesterday was very different from the whole month or the last six weeks. I wanted to…lie and put a better day because it was so different from the last six weeks."*

> *"It would not be an accurate picture. So I cut-out* [a distressing event that happened yesterday] *and considered the whole and gave my answer."*

These respondents offered their own solution to this problem; they suggested asking about a longer reference period:

> *"'Overall'* [or] *'over the last few weeks'…might be better. I think just yesterday is not a good measure for how you feel now."*

This is further evidence that respondents do not understand the methodology behind asking about the previous day and many adjust their response.

A positive reaction to being asked about yesterday was the theory that ONS was doing so because:

> *"It is the most recent full day".*

> "*To make sure you get something that's very recent, so yesterday is very recent. It's a quick way of assessing how your mental state is right now".*

They also surmised that 'yesterday' was chosen because it would be easily remembered; although this was not the case in reality.

The report of the LFS research states that respondents found it *"unusual"* to be asked about yesterday. If that research had probed for further information about this, it is likely that the same issues would have been found.

4.6.3   Remembering yesterday

A group of respondents within the sample found it surprisingly difficult to recall the previous day:

> *"Crikey, I've gone blank. I just can't remember. Bear with me."*

> *"I didn't know what day it was yesterday, and I couldn't remember what I did."*

> *"At my age I cannot even remember yesterday...what did I do yesterday?*

These comments were followed by considerable pauses.

Recommendation

Add useful 'memory joggers' to the interviewer instructions, such as telling respondents what day it was: "Yesterday was Monday."

4.6.4   Ease of responding

Two issues arose that accounted for respondents saying the question was difficult to answer. Firstly, as mentioned, remembering 'yesterday' could be challenging. Secondly, if the previous day had been unhappy or stressful, recalling these memories added cognitive burden as respondents tried to assimilate these events into their response:

> *"…you had to think about yesterday and the bad bits you were trying to put to the back of your mind."*

> *"*[It was] *hard. Because… my mood was a bit on the flip side yesterday, I didn't receive good news."*

4.6.5   Wave three changes

An instruction was added and tested in a final wave of interviews. The question and its instruction were:

Overall, how happy did you feel yesterday? Please think about 'yesterday' even if it was not a typical day.

As previously mentioned, there were four issues that arose with the original wording:

- a dislike of being asked about one day, which might not be typical
- changing of responses if the previous day was atypical, so as to give a more general impression
- remembering the day before
- a negative impression about ONS resulting from respondents not understanding the methodology behind asking about one day

None of these issues were found in wave three. Caution must be used here however, because balanced with the strength of feeling in waves one and two, this is a surprising finding. Nonetheless, the addition of the instruction does appear to have had a very positive effect:

"It was good as it made me think about a specific thing, rather than this is just how I am feeling overall."

"When you say yesterday it is easier to answer the question as it brings more specific events into your mind."

"I think you have to take more into account when you ask how happy do you feel. More to think about. But when you ask about yesterday it is easier to think just about that."

"[Yesterday] *was fresh in my mind."*

Ideally this instruction would be tested further. However, it does appear that the addition of the instruction has led respondents to focus on the previous day and more importantly, not mind doing so. By adding clarity about 'yesterday', in the instruction, the negativity was stemmed before it had the opportunity to develop.

Had this research stopped at wave two, the recommendation would have been to re-word the question to remove 'yesterday' and get a more general impression. However, wave three clearly suggests the instruction works beyond its literal remit.

Recommendation

Investigate further the inclusion of the instruction on yesterday even if atypical either as part of the question stem or as pre-amble before the affect questions. Further cognitive research looking at this would be advantageous.

4.7   Anxious

On a scale where nought is 'not at all anxious' and 10 is 'completely anxious', overall, how anxious did you feel yesterday?

4.7.1 Understanding of the question

The overall meaning of the question was understood:

"…how much anxiety was in your life yesterday…"

"Were you worried, concerned about anything."

"…are you feeling nervous…"

"It meant how stressed I was yesterday and how worried."

"Whether I was wound-up yesterday or not."

"It meant how stressed I was yesterday and how worried".

4.7.2 What was considered when answering

Various aspects of life were taken into consideration when answering this question:

- personal life and relationships with partners and children
- issues at work
- unemployment
- financial problems
- health problems

The sample included respondents who told their interviewer they had mental health issues. These respondents said that they compared yesterday to good and bad days they had experienced during their illness to help them arrive at a score.

4.7.3 Interpretation of terminology used in the question

The terms 'anxious', 'not at all anxious' and 'completely anxious' are used within the question. These terms were examined in the cognitive interview with respondents.

'Anxious'

Respondents offered a range of synonyms for 'anxious':

- stress
- nervousness
- wound-up
- panic
- worried
- things on your mind
- unsettled
- tense
- failure

It was described as being more serious than 'worry' and having physical symptoms as well as psychological:

> *"Anxious is what I feel about work, it's the stress I feel…it's more than worry. Anxious seems like when worry starts to have real symptoms like high blood pressure and chest pain, which I have."*

> *"Unsettled, you're worried, you feel sick. There's different types of anxious though. There's people who won't leave the house or people who really worry about little things."*

'Not at all anxious'

This phrase was described as:

- calm
- nothing to worry about
- happy with your life
- content
- not at all nervous
- relaxed
- 'happy go lucky'
- not having a 'care in the world'

There was discussion about whether it was possible to be 'not at all anxious' because *"everyone worries about something."*. For example:

> *"You're not normal. 'Not at all anxious', you must be so laid back you are horizontal."*

> *"I would never say nought. I think everyone's a little bit anxious about something. So 'not at all anxious' is unachievable."*

'Completely anxious'
This was described as might be expected and no comprehension issues were found:

> *"10 would be totally breaking down, not being able to cope."*

> *"Totally stressed to the max."*

> *"Not leaving the house, checking your text messages, locks everywhere, padlock and bolts, being a complete recluse, always questioning everything, double checking and locking, not wanting to take part in anything for fear of what is going to happen or being anxious every time you do something or questioning yourself and your ability and other peoples' ability."*

4.7.4   Alternative terminology
The word 'worried' was compared to 'anxious'. Where difference was discussed, this tended to be about the contrast between 'worry' being of lesser severity and 'anxious' being so severe it has physical symptoms:

> *"Worry is just a general niggle, anxious is full on panic."*

> *"…we all worry don't we, we worry the kids have got to school on time but anxiety is anxious and jittery and you cant settle, and fidgety, and that's what I am like when I am anxious, I can't settle and can't focus on anything."*

> *"It's a more physical version of worry where you get ill, rather than just worrying."*

> *"Worry is in the head and anxiety is in the body, one is mental and one is physical."*

Respondents stating that they were the same or similar tended to be male and under 30 years old. A typical description was:

> *"Anxious, stressed, worried; I think they're all the same. They're just different words for being stressed-out."*

Indeed the word 'stressed' was commonly used in discussions about anxiety, except with the group of respondents who had suffered anxiety, in which cases the language was of a more psychological nature:

- nervous
- panic
- neurotic
- psychosis

Recommendation

Future research should explore a question about 'stress' instead of 'anxious'. 'Stress' has fewer stigmas attached to it and is therefore more likely to elicit robust responses.

4.7.5   Asking about yesterday

Respondents understood that the question was asking about the previous day. However, they applied that understanding in one of two ways. They either:

- gave a true score for the previous day, or
- worked out an average over a longer period of time

The latter approach was adopted when respondents did not like reporting for one day. Looking at a single day was thought to be unrepresentative. Even where a true score was given, respondents commented that *"yesterday was not typical."* Those finding it most unpalatable made comments such as:

> *"I'm surprised at the statistics organisation asking a question like that. It's obvious that non-standard things happen so surely you want an overall feel for how people are feeling?*

4.7.6   Wave three changes

For wave three the reminder about how to use the scale was removed as this appeared to be creating confusion, rather then helping respondents choose a score. Also, an instruction to only think about the previous day was added. The question therefore became:

Overall, how anxious did you feel yesterday? Please think about yesterday even if it was not a typical day.

The use of show cards showing the 0-10 scale was also recommended, the design of which had its extremities labelled, although another version, with no labelling was also tested. See Appendices C and D.

The changes proved very positive. Respondents welcomed the showcards that included labels, and their understanding of the question was not hindered by the removal of the scale. Most importantly however, as described in the wave 3 changes to the 'happy' question, not only did the changes result in the scale being used correctly, but respondents now claimed to not mind answering about the previous day.

Recommendation

Adopt the wave three question and labelled show cards.

4.7.7   Level of difficulty in responding

In waves one and two, there was evidence that only having to consider one day made responding more difficult. Two techniques for remembering the previous day were used. Respondents would either think chronologically through the day or link yesterday to specific events that happened, to jog their memory. Remembering the events of the previous day was said to be more difficult than giving an overall sense of their anxiety.

"It was difficult because it was yesterday. If it had been in the last week it would have been easier."

There was also some resistance and negativity towards giving a score for questions such as this:

"I can't sum-up my life in numbers."

"Do you mean for the whole day? That's a weird question, how can you quantify it."

4.7.8 Stigma
The sample included respondents who had experienced anxiety, which gave a useful insight when comparing these respondents to others. Those who did not report a history of anxiety were *"not bothered"* about being asked this question. However, they did state that they thought people who were suffering from anxiety might feel stigmatised and would worry about giving a high score:

"It didn't bother me but I think it would if people were scoring a nine or 10. As I didn't score high it didn't bother me."

Those who had suffered from anxiety painted a bleaker picture:

"To admit to someone you are anxious is always hard, it is like you have failed."

"That's close to psychosis, completely anxious. Like when there is nothing else at all and if I'm completely anxious I don't have any self-soothing going on. I have no other abilities, it would be terrible".

5   Additional four well-being questions

These questions are additional questions currently being considered by ONS for possible inclusion following feedback last February from users. They have been tested on the Opinions Survey but not the Integrated Household Survey. Preliminary quantitative findings from this survey have indicated they are more invasive, less desirable to answer and the question concerning control over life events is particularly cognitively burdensome.

4.1 Satisfaction with personal relationships

> Overall, how satisfied are you with your personal relationships?

5.1.1 Understanding of the question
Respondents answered this question by thinking about who was the closest to them and therefore, considered:

- partners
- parents
- children
- friends
- neighbours

This suggests the question is largely understood, which is illustrated here:

"How satisfied are you with your friends or family and maybe your partner if you have one."

"The people in my life...my friends, how they treat me and how I treat them, how I feel when I'm with them and how I feel when I'm not with them. Am I important to them, are they important to me."

There were instances where respondents queried what the question was asking:

"Are you talking about friends, family or in general?"

If this happened with a General Field Force (GFF) interviewer, the protocol would be for them to follow strict guidelines set out in their briefing material.

Recommendation

Interviewer instructions would need to be thought through carefully to ensure they address this issue appropriately.

5.1.2 Understanding of the terminology

The term 'overall' was interpreted as suggesting respondents take an average, or wider view, of their satisfaction with personal relationships:

"I would take that as an average of how satisfied I am with different sets of people, my acquaintances."

The word 'satisfied' is understood in different ways:

- negative: *"In this respect it's not a term I would use. I'd be really offended if* [my partner] *said he was satisfied with me or* [my son] *said I was a satisfactory mother…It's just an odd term to use with personal relationships."*
- neutral: *"Neither good nor bad."*
- positive: *"You don't need much more to make things better."*

Their interpretation could have an impact on their score. For example, a respondent who interpreted 'satisfied' to mean, *"really content"*, might be prone to give a low score.

The word 'personal' was said to make respondents think of their partners, as that kind of relationship was viewed as 'personal'.

The complete term, 'satisfied with personal relationships' was described as:

"It is having someone with you to help you make decisions, security and consistency, that's what [it] *means to me."*

"How happy are you, how would you change it, I am not ecstatic but also I am not in floods of tears, generally it is ok."

"I'd be pretty upset if my life and my relationships was just satisfactory."

30

5.1.3 What was taken into consideration when answering

Respondents took one of two approaches when answering this question. They either considered the person closest to them or considered a wider group and decided on an average satisfaction score because, *"I think the clue is in the question. It says 'overall' so it's telling you average it out."*

5.1.4 Who was omitted

It is interesting that work colleagues were not considered when answering this question. It had been surmised that as working respondents potentially spend significant amounts of time with colleagues, and so build up relationships, that respondents would include them. This was examined and it was found that colleagues were simply not important enough to be considered a 'personal relationship'. However, where colleagues had become friends, their status had changed:

> *"I'm not actually friendly per se with people from work. I mean I wouldn't call them friends, if I did call them friends, then they'd be friends, not colleagues so I still wouldn't consider work actually."*

5.1.5 Level of difficulty in responding

Where respondents found this question more challenging, it was because although the question was easy to respond to, they said it was emotionally difficult to think about:

> *"It was easy as I was thinking about my Dad and my boyfriend, but emotionally difficult, all of the questions are emotionally difficult as it makes you stop and think."*

In the first wave of interviewing, respondents commented that it was more difficult to respond because they had no guidance as to who to include:

> *"Fairly difficult as the relationships were not specific."*

After internal discussion, the question was changed for wave two:

> Overall, how satisfied are you with your personal relationships? For example, friends, colleagues, family or partners.

By adding the qualifiers to the question, this gave prompts to respondents to make sure the main relationships were all covered. In practice however, this did not lead respondents to include work colleagues. However, it did encourage those who only considered their partners, or lack thereof, to include their families:

> *"I gave a three because of how I feel about my boyfriend and the split but if I'd thought you meant everybody I'd have been over five definitely coz I love my mam and my family. We're really close."*

This clearly suggests that individuals who are not in a relationship but would like to be, may be prone to giving scores that are lower than necessary as they did not consider all personal relationships.

Recommendation

Keep the instruction at the question to include friends, colleagues, family or partners as it encouraged respondents to think more widely.

5.1.6 Reference period

Although the question does not include a time period, respondents were asked what time frame they were thinking of. This ranged from a year ago to the present. Again, when respondents went back in time it was because they were thinking of a specific incident or important occurrence that had happened. For example, to a time they had a partner.

5.2   Loneliness in daily life

> On a scale of 0 - 10 where 0 is 'not at all lonely' and 10 is 'extremely lonely' how lonely do you feel in your daily life?

5.2.1 Understanding of the question

When respondents were asked during cognitive probing to explain what the question meant, good comprehension was demonstrated. The question was generally understood as the extent to which you have others in your life to help and support you, whom you feel a personal connection with. For example:

> *"If you have somebody there to help and support you."*

> *"Being around people but still feeling lonely…no personal connection felt."*

> *"How many people you have in life who you can lean on if something goes wrong."*

> *"Not having anyone to talk to."*

> *"How isolated and detached you feel day to day."*

Those who had previously experienced mental health issues felt that loneliness was compounded more for them as a feeling of not being able to connect with others. This group also felt the question was *"intrusive"* and *"huge"*. Furthermore, it was commented that it felt insensitive to be asked this question after the previous question, 'How lonely did you feel yesterday'. This was due to a feeling of being probed more deeply because a high answer was recorded for the previous question.

5.2.2 Understanding the term 'daily life'

The term 'daily life' was interpreted in two main ways:

- Monday to Friday, not a weekend
- every day in the week, including weekends

It was thought 'daily life' referred to Monday to Friday as it implied a working week. Those who were unemployed and retired felt that 'daily life' no longer applied to them as they did not have a routine. For instance:

> *"'Daily life' to me is something I would have said when I was working. I don't have a 'daily life' exactly now, since I retired, because one day is just the same as the next."*

Recommendation

Consider the impact of this on the data that would be gathered from this question as non-working respondents' interpretation of daily life could lead to erroneous responses. However misunderstanding of the term was not found and there was no evidence that respondents would alter their score due to the term being used.

5.2.3 'Lonely'

The term lonely was understood well and considered in a variety of ways, including the absence of a partner and the extent of social connections experienced. Lonely was thought to imply both physical and emotional aspects:

> *"Not physically, but emotionally."*

> *"Physical aspect of being alone and the emotional need to talk to someone."*

> *"Being entirely isolated like Tom Hanks in Castaway on a desert island".*

5.2.4 What was considered when answering

When answering the question, three main aspects were considered:

- the extent of being happy with one's own company
- recent changes in circumstance
- the quality of close personal relationships

Respondents considered how happy they were with their own company and compared this with how much they felt they wanted to be with other people but were not able to.

Examples of recent changes in circumstances included:

- mothers who were not working due to childcare responsibilities considered the extent of loneliness felt since their children had started being cared for by childcare providers
- students considered that the time of year meant they were more alone than is usual, as they were asked the questions during summer holidays, therefore the score would be different if they answered during term time
- those who had recently experienced relationships breaking down thought to when they were with their partners

An assessment of the quality of close personal relationships included whether or not respondents had a partner and how much they saw close family members.

It was also thought that loneliness was linked to satisfaction and happiness:

> *"I think the loneliness ...its all part of the parcel of the way you feel psychologically. If you feel lonely you become depressed, you become*

unhappy, which in turn can lead to anxiety. That's how it would affect me, I think."

5.2.5   Level of difficulty in responding

Similar to the 'worthwhile' and the 'life satisfaction' questions, when respondents found the question easy to answer it was associated with already having thought about the topic. Unemployed respondents reported thinking more recently about loneliness due to spending more time alone.

When difficulty was experienced, this was due to the reversal of the scale causing confusion about how to answer. Additionally, it was felt thoughts and feelings are hard to quantify as they can change quickly:

> *"Feelings and thoughts change. I might wake up in the morning feeling a bit lonely then by the afternoon feel I don't want anybody and like everyone is getting on my nerves. So you can't measure a whole day, people do not feel that way all day, I'd feel it for twenty minutes. I might feel it for an hour, for six weeks. How do you measure how long a feeling lasts?"*

5.3   Loneliness yesterday

Overall, how lonely did you feel yesterday?

5.3.1   Understanding the terminology used in the question
'Overall' was understood as in previous questions and so no issues were found. The word 'lonely' was described as:

- not having people to share one's life with
- being alone
- being alone but also having nothing to do

The state of loneliness tended to be considered a negative state. However, for a group of respondents, being alone was a positive experience as it gave them time to relax and enjoy a peaceful period of time:

> *"…it was chance for time out; a bit of R and R."*

5.3.2   Understanding the question
Respondents understood that they had to think of the previous day, even if they were not happy to give a score based solely on that one day. Therefore, the key issue with the question is that for an unknown proportion of the population, their score is not based on the previous day.

Recommendation
Consider the implications of data being provided that are not about the previous day.

Loneliness itself was seen as something physical, not being with others, or emotional, a feeling of loneliness even if with others:

"How long were you on your own for."

"Just whether yesterday I was feeling alone."

Therefore, the concept of 'loneliness' appeared to be understood correctly. However as it was not always seen as an emotion this suggests it's use as an affect question needs to be reconsidered.

5.3.3 What was taken into account when answering

Respondents described considering the following areas when deciding upon their score:

- activities, or the lack thereof, the previous day
- contact with other people, even by phone, text or email the previous day
- always having to put other people first
- whether they are, in general, lonely people

Again, the latter two points are not specific to 'yesterday'. For respondents who self-reported unstable moods, asking about one day was cognitively more difficult to process. This was because they not only had to remember the previous day, but also, they had to work out a score for varying moods across the day. These respondents were irritated by being asked for one day and could not understand the rationale behind it:

"[This question] was asking for a specific day but I don't understand why because it was different from all other days yesterday."

"I can't understand the point of that question ever. What does it mean? I could have lost a family member yesterday; I could have fallen and broken my hip yesterday. I can't understand the reason."

Recommendation

See section 4.6.1 for discussion of yesterday questions and recommendations.

5.3.4 Ease of responding

There was evidence to suggest those with other problems in their life, or mental health issues, found this question more challenging because they were being asked specifically about their loneliness:

"[I] found the concept of loneliness quite tricky…and I wouldn't know how to proportion what out of that general misery was loneliness and which other elements were making it up as well."

5.3.5 Comparisons to the 'lonely in daily life' question

This question was compared with asking about loneliness in daily life:

> On a scale of 0 to 10, where 0 is 'not at all lonely' and 10 is 'extremely lonely', how lonely do you feel in your daily life?

Respondents preferred being asked about loneliness in daily life because they felt it gave a more general picture of how they are, which they thought was more accurate:

> *"[It] gives a better picture or a complete answer, as yesterday had been a good day, but most yesterdays are not."*

> *"It was like comparing it with, not just the day yesterday, but how you feel overall."*

> *"There can be a big difference in how you feel yesterday and how you feel overall…[The] difference between the two is are you feeling sad and depressed all the time or have you just not had a very good day today…[So an] overall view on how you are feeling, or is it just a little blip."*

> *"I think [the 'lonely in daily life'] question is better 'cause yesterday was just about my mam's operation but I'm fine today and I usually am so I think that's better."*

For further discussion of the 'lonely in daily life' question, see Section 5.2.

5.4 Control over important aspects of life

On a scale of 0 - 10 where 0 is 'no control at all' and 10 is 'complete control' how much control do you feel you have over important aspects of your life?

5.4.1 Understanding of the question

When asked during cognitive probing what the question meant, there was evidence for both good comprehension and difficulty. When comprehension was observed, respondents described their understanding of the question in the following ways:

- perceived financial control.
- the extent that life events can be undertaken within the individual's authority

When it was believed the question only related to finances, debt was considered in relation to control:

> *"I took that to be asking about things in my life like bank overdrafts. To me it completely relates to financial issues. Where you have debt you can manage it completely and get a satisfactory outcome."*

When life events were considered, this was conceptualised in the following ways:

> *"If you believe it's in your power to feel better or worse in your life."*

> *"Do I have the power to make decisions, I suppose. Or to decide what I do and don't want to happen."*

Physically disabled people who were not working and not likely to do so in the long term thought about the extent of their dependence on other people to carry out physical activities.

When difficulty with comprehension was experienced, this related to uncertainty about whether the question meant events that are realistically within a person's control or to include events such as death, which are usually not within a person's control. There was also evidence the question was not answered because the meaning of 'important aspects' was not understood at all, particularly within lower educated groups.

5.4.2   'Control'

The term 'control' was understood well and described as:

- how much influence can be exerted over a situation.
- how much power is held to enable change or manipulation.
- being able to manage the outcome of an event

Among women, control was also associated with other people's influence, in particular the experience of being in a controlling relationship:

> "What my ex-husband did to me, I hate that word. Being in control is hard to explain but it automatically took me back to my previous marriage, I was thinking about being controlled. It is not a word I like."

5.4.3   Important aspects

'Important aspects' were considered to be a range of different areas of life including:

- work
- relationships with other people
- pets
- finances
- life goals
- environment
- employment
- the future
- mental state
- physical health

When the future was considered, students thought about university and making decisions such as which modules to study. When mental health was considered, women thought about not losing themselves to what other people want.

When confusion was expressed, it was felt 'important' and 'important aspects' should be defined as it was unclear what they might refer to. There was evidence this confusion led to the question not being answered.

5.4.4   Process in selecting an answer

When answering the question four main aspects were considered:

- what could be personally controlled

- what could not be personally controlled
- the influence of others
- current affairs

When considering what could be personally controlled, this included one's own thoughts and behaviour. The unemployed considered they had much more control over their spare time since being made unemployed/redundant. Additionally the employed felt that income could be personally controlled and it was felt there could be no control over anything without money.

When aspects not under personal control were considered, this included death and sudden accidents:

> "I was thinking about a car accident I had recently and how I had absolutely no control over that, control was completely and utterly taken away from me. You can't control your life, your death".

Birth and marriage were also considered not under control:

> "We're expecting a grandchild in November; I have no control over that. Our daughter's getting married; we have no control over that. So very important aspects of our life we don't have control over".

Older people considered aspects of physical health, such as fatigue and illness, as not being under personal control.

When the influence of others was considered, long-term carers thought about not being able to control much in their lives due to caring responsibilities. Females from minority ethnic groupings thought about the extent their parents controlled their lives. Physically disabled people not working thought about how much they could do for themselves without the help of other people. Women who had experienced controlling relationships considered how much that had influenced their lives.

When current affairs were considered, this included the economic situation of Great Britain and how much control an individual can have within the constraints of the current climate.

The question was not understood among lower educated respondents and there was evidence this led to an answer of 'don't know'.

5.4.5 Level of difficulty in responding

When the question was viewed as easy to answer, this was due to control being thought about previously in relation to a personal situation such as recently being in a car crash or having a physical disability, both of which had led to respondents considering the level of control they have.

When difficult, it was commented that a lot of different memories and feelings needed to be thought about in order to reach an answer and that control was very changeable and complex. Also this question was considered the most difficult to answer of the set of questions as it was unclear what 'control' should encompass:

> "That was the most difficult one to answer so far…because there's control and there's control. Are you asking whether I have control over things I can

control or control in general… well I can't control getting run over by a bus or getting cancer, but nobody can so I don't know whether you want my sense of that, or realistic things."

Recommendations

Consider the addition of interviewer instructions that can be given to respondents experiencing comprehension difficulties with this question. As this question was viewed as the most problematic to answer, the consideration and discussion of the validity of this question is recommended.

6   Question order

The first four questions were asked in this order:

Overall, how satisfied are you with your life nowadays?

Overall, to what extent do you feel that the things you do in your life are worthwhile?

Overall, how happy did you feel yesterday?

On a scale where nought is 'not at all anxious' and 10 is 'completely anxious', overall, how anxious did you feel yesterday?

Respondents were asked to comment on the order. Where respondents said they would not change the order, this was because these respondents did not think question order mattered. The group of respondents who did want to change the order all said they would position the 'happy' question first. This was because respondents thought it was best to start on a positive note:

"I think if you put the happy question first it would put people at ease."

"I think it's good to start with the happiness question because it starts things off nicely."

"It's easy to answer, it's a nice question."

The 'anxious' question was always placed last. This was because it was viewed as potentially being more sensitive and of a negative nature, with possible stigma attached:

"[I'd leave] *the worst one last. Leave the worst to the end."*

"You might put people off if you start with that [anxious] *question. I mean if they're not right in themselves you don't want to start on sadness."*

"Well that would just put a downer on things, if I can use that expression, and people could be more negative about the others."

There was concern that the 'anxious' question would over-shadow the others if it were first. There seemed to be a desire not to make responses unnecessarily negative. When respondents were challenged by interviewers and asked if the 'happiness' question would make responses overly positive if it were first, this was acceptable however.

The order of the 'worthwhile' question and the 'satisfied' question were seen as less important:

> *"…the other ones could go either way."*

> *"…the middle two could have been in any order really."*

However, where a preference was stated, this was for either the 'worthwhile' or the 'satisfied' question to follow the 'happy' question because they were seen as:

> *"'…an extension of the happiness question."*

The second set of questions were:

Overall, how satisfied are you with your personal relationships?

On a scale of 0 to 10, where 0 is 'not at all lonely' and 10 is 'extremely lonely', how lonely do you feel in your daily life?

Overall, how lonely did you feel yesterday?

On a scale of 0 to 10, where 0 is 'no control at all' and 10 is 'complete control', how much control do you feel you have over important aspects of your life?

This set of questions was perceived to be *"potentially quite depressing"* and more difficult to answer:

> *"I think these ones are harder 'cause there's not a happy one."*

Respondents struggled to order these questions; saying either that the existing order was *"fine"* and that the questions[2]:

> *"…blend into one another. Each question leads into another."*

Or that order did not matter because of the questions being more negative:

> *"They aren't the cheeriest of questions really so it doesn't really matter what order they go in. There's not one that stands out as being the first."*

Respondents preferred the first set of questions because they were perceived as being more positive.

7 Usability of the scale

When considered easy to use, the scale was described as *"straightforward"* and an aid to forming thoughts about the question by supplying a framework within which to answer. However there were also several difficulties.

---

[2] It should be noted that ONS was never considering asking all the social relationship questions together but to explore which one of the three (personal relationships, lonely yesterday and lonely in daily life) worked best.

When seen as hard to use, it was thought difficult to quantify an emotion into a number. It was commented that words instead of numbers might make the scale easier to use, for instance 'very satisfied, somewhat satisfied', instead of numbers.

The LFS research (see section 3.1.1 – background to the current research) showed that respondents asked for the scale to be repeated at the satisfaction and worthwhile questions, as did these respondents. This is likely to be due to cognitive burden.

At the control question, it was felt that control changes so much from day-to-day that it was impossible to reach an answer using the scale:

> *"Depends on different days how much control I have. Some days no control and some days complete control. Some days nought and some days ten. Some days seven and some days six, some days five."*

This led to no answer being provided. Other problems included finding the scale for this question difficult as it was felt being in a relationship implied a joint control, which could not be reflected on the scale:

> *"If you live with someone you may be on a mutually agreeing pattern, to actually saying complete control could come over as a control freak, so on reflection the nought-ten is an odd scale but not sure how you would do it. My wife and I have complete control over our lives but sometimes it is a joint control and we compromise and we have give-and-take. One person might make a suggestion but it is a joint control."*

It was also felt it was unclear what 'no control' might mean.

A difficulty with polar point labelled scales is that the meaning of the unlabelled categories is open to respondent's interpretation. Different respondents can interpret the middle categories differently, often increasing measurement error. Few people hold, or express, their opinions in numerical terms so numbers can be removed from fully labelled scales to eliminate the extra processing step of converting opinions to numeric terms. However, they can only be effectively used for scales containing two to five points. Fully labelled scales rate higher on reliability, validity and respondent preference (Dillman, 2009), however the well-being scale has 11 points.

As a compromise, show cards were designed with the extremities labelled only, which was very well-received by respondents in wave 3 (see section 8 for discussion of show cards).

7.1　　　　Reversal of the scale at the anxiety and loneliness questions

As described in the report of the LFS research, the scale at the anxiety question was not universally understood. Where misunderstanding occurred, this was because respondents had reversed the scale, giving low scores instead of high, or vice versa. This became evident at one of the following two points.

- the interviewer being asked to clarify or repeat the question at the point of responding
- during the cognitive interview

When respondents ask for clarification, it is a clear indicator that a question is not intuitive or easily processed. However, it is of equal concern that respondents, who gave what appeared to be straightforward, confident responses, were only found to be incorrect when probing took place during the cognitive interview. For example, a respondent who initially said her score was three for the anxiety question said:

> *"It was at the forefront of my mind so it was easy enough to decide."*

In the cognitive interview however, she discussed some stressful events and the interviewer challenged the score. It then became apparent that she had reversed the scale and said:

> *"It was just me not paying attention…Looking at* [the questions] *again, I think it's because you have to score high if things are OK* [for the preceding questions] *but then score low if things are OK on this one."*

Therefore, it can be surmised that interviewers in the field might not have any idea that an error was made because nothing alluded to this until thorough probing occurred.

As the quotation above indicates, the confusion stemmed from the previous three questions. Here, the closer to 10, the more positive the response. Then for this, the final question, the closer to 10 scored, the more negative the response i.e. a score of 10 is 'completely anxious', which is a negative state.

To further illustrate the problem, a respondent with a history of mental health issues, including anxiety said:

> *"What would completely anxious feel like? Probably a one* [on the scale]."

These findings were repeated for the loneliness questions; interviewers were asked by respondents to repeat the scale at the point of responding as uncertainty was expressed over the polarity. There was also evidence that incorrect answers were given. In discussion during cognitive probing answers were changed when it was realised they had reversed the scale.

Respondents who reversed the scale said they would welcome a visual aid:

> *"If I had the questions in front of me I wouldn't have to second guess. It's good to have a visual aid. I can now see where I've gone wrong."*

Recommendation

The use of the labelled show cards in face-to-face interviews is recommended (version A, shown in Appendix C – see section 8 for rationale) in order to increase usability of the scale. However as the Integrated Household Survey is a mixed mode vehicle where the telephone is also used, further consideration should be given to the effectiveness of the administration of these questions by telephone (see also section 12).

7.2 Ratings

For respondents who could provide an answer using the scale, their ratings for each question had broadly similar ranges. The extreme ends of the scale were often not

chosen and there was evidence that a rating of five was thought of as the average on the scale and answers were placed above or below five based on whether it was thought a person was above or below the average.

7.2.1 Worthwhile

Worthwhile ratings ranged from two to 10. However, for a group of respondents it was felt a rating of 10 could never be reached as improvements can always be made:

> *"I would never give it a 10 as I am a bit of a perfectionist, I always think there is some way or somewhere I can improve things, can make things better".*

Individuals choosing to give a rating of 10 were described as *"stupid and naïve"* as it was thought 10 was unachievable. A score under five was viewed as negative and not what should be answered. This led to a score of six being recorded instead of a score of two:

> *"The starting point for me is five, I think. So it's just over six and now I'm thinking maybe it should have been a two…I just think the numbers below five are almost like negatives to me so it had to be over five."*

Evidence like this suggests that certain respondents do not want to be viewed in a negative light. They inflate their score to appear to be doing worthwhile activities.

7.2.2 Satisfaction

Satisfaction ratings ranged from two to 10, with five perceived as the mid-point:

> *"Straight down the middle, it could be better and it could be worse".*

It was again thought no-one could ever give a rating of 10 as there is always something that could be better. Eight was considered as far enough up the scale to be happy, while seven was seen as a realistic score to aim for.

7.2.3 Happy yesterday

Happiness ratings ranged from three to 10, with this question yielding more scores of 10 than any other. Respondents giving a score of 10 described the previous day in surprisingly uneventful ways. There was no external stimulus to make the day better in any way; this group simply appeared to be willing to score 10 because nothing negative occurred.

However, the contrasting group were much more reserved about scoring 10:

> *"Well I couldn't say 10 because that would come across as being arrogant or something."*

> *"[I hesitated when responding] because I'm very happy, there's not much I'd change. A lottery win would be nice. But I really am a very happy bloke but I couldn't say 10 because that would be showing off."*

7.2.4 Anxious yesterday

Ratings ranged from nought to eight. No respondents gave a nine or 10, which was surprising given that the sample included respondents who were suffering from anxiety. The reason for this group not giving the highest scores was that nine and 10 were considered to represent the most debilitating anxiety:

> *"Ten; you can't get any worse, you'll have a panic attack and lie down and pass out. Eight; you're not quite there yet."*

Those reluctant to give nought seemed incredulous that nought could exist:

> *"… you can always worry about something; [I] don't think there would ever be a day when I would give a nought"*

> *"I was thinking I must have felt a bit anxious about something at some point."*

> *"If I said zero they would think I'm lying."*

7.2.5 Personal relationships

Ratings ranged from two to 10. It was particularly interesting to note that when respondents did not have a partner they gave a lower score, despite being satisfied with other relationships:

> *"In the grand scheme of things a four was a bit low, I have got my mates but haven't really had a relationship since coming to university, and that's what I thought about, having another half. In all honesty I would probably change my answer now, I'd now give it a five, six, probably a six."*

In such cases the question was also viewed as more sensitive:

> *"[It's] a bit rude to ask this question if you are not in any kind of personal relationship."*

7.2.6 Lonely in daily life and lonely yesterday

Loneliness in daily life ratings ranged from zero to eight. Five was seen as the midpoint. Lonely yesterday ratings ranged from zero to eight, with no respondents even considering a nine or 10.

7.2.7 Control

Control ratings ranged from zero to 10. Six and seven were seen as a level of control that is needed to "*keep your head above water."*

Higher scores were seen by one group as undesirable as they were associated with being a *"control freak"* and *"over-powerful"* or *"domineering".*

8 Show cards

Show cards were used to test whether usability of the scale could be improved by using a visual aid. Previous evidence (discussed in earlier sections) had shown that

respondents checked whether they had understood or remembered how the scale worked:

"Is this one to ten again?"

"Which way round is it again?"

Also, the scale was completely reversed, meaning that high scores were low and vice versa. Although qualitative research cannot allude to the magnitude of any finding, it was felt this is a significant issue. Show cards were therefore tested to see if this could be improved.

There were two designs (see appendices C and D). The optimum design, that reduces cognitive burden the most, had labels at each end of the scale; for example, 'not at all anxious' at zero, and 'completely anxious' at 10. Respondents were very positive about this design and said it would mean that other respondents would not have to ask for a reminder about the scale. They also said it would make choosing a score quicker and easier:

"I didn't really need the card to give an answer but it made it faster because it reminded me which end of the scale needed looking at."

"…you don't have to think in your mind where [you are] *on the scale."*

"The wordings on the cards would make it easier to judge which way the scale went and easier to choose your answer."

This design featured one question per show card. There was a concern about adding numerous show cards to the General Field Force interviewers', already numerous, set of show cards. Therefore a second design was tested that was one show card. The drawback with this design however, was that labels at each end of the scale could not be used, which by design, could never help with the problem seen with the 'anxious' question in which the ends of the scale were confused. However it was tested to gauge user-acceptability. Respondents said it was only useful in terms of a visual aid but *"not as useful"* as the labelled version and provided *"little help"*.

8.1   Show card limitations

Clearly show cards can only be reliably used when the questions are administered face-to-face. When questions are administered by phone interviewers are unable to use show cards. Even if they are posted in advance of a telephone interview taking place, there is no guarantee respondents would have them available. They could:

- be lost in the post
- be mislaid within the household
- a household member may receive them then fail to notify other household members

If they are sent electronically there are similar issues.

Respondents who had had the questions administered by phone would also have welcomed show cards however:

"It would have been easier with a show card. When you have got the statements at either side it is easier to position yourself somewhere between them, it is easier to visualise rather than pick a number out of your head."

"[It is] helpful, showing the different ends of the scale."

Recommendation

Consideration should be given to the use of the show cards that feature the scale with labels at the extremities as they prevent the scale being mistakenly reversed (in face-to-face interviews).

9   Distress experienced during the interview

Disabled and unemployed respondents thought 'worthwhile' meant being in paid employment and they became visibly upset when answering this question. Also there was evidence that if loneliness was being felt at the time of answering the loneliness questions then this would have caused personal distress; indeed interviewers observed crying. If an emotion of loneliness was being felt, the interviewer could leave a person feeling distressed with no-one to turn to. It was also commented by respondents who had experienced mental health issues that the eight well-being questions were therapeutic questions that counsellors and mental health practitioners use. Therefore they may produce difficult feelings or cause the emergence of emotions that were not previously thought about. These are issues for serious further consideration.

9.1.1   Ethics

It is important to remember that the opposite of well-being is something much bleaker. The British Psychological Society Code of Ethics and Conduct (2009) states that respondents, especially vulnerable adults, should be given ample opportunity to understand the anticipated consequences of their research participation as part of the informed consent process. Also Graham et al (2007) found that as part of GSR ethical procedures, respondents expect to be "left without negative feelings of participation".

Recommendation

It is important that a strategy is developed showing that ONS has considered the potential impact of asking the well-being questions. As it might not be evident that an individual has been affected, or because it might not affect them until sometime later, a blanket approach in which all respondents are given guidance should be explored.

This could take the form of a flier that interviewers would give to respondents at the time of the interview. As well as including information about where to seek help, it could also reassure respondents about confidentiality and reiterate the purpose of the survey.

10   Disclosing well-being in the household environment

10.1 Answering in front of others in the household

Respondents within the sample were either happy to answer in front of other household members, stating that they knew them so there was no problem in answering in front of them, or this would be problematic.

When it was felt the questions could be answered in front of the household, this was due to a feeling that the household already knew this information and so there would be no surprises:

> *"I'd answer them in front of any family and my good friends because they'd know my state of mind anyway."*

> *"There would only ever be my wife here really, on an odd occasion my children, but they're grown up and I have nothing to hide so that's fine."*

Also honesty was associated with higher scores:

> *"As I say, I'm a very contented man so I wouldn't need to make-up answers."*

When it was felt the questions could not be answered in front of the household, this was due to feeling they were private and it was inappropriate to be discussed openly.

Strategies for dealing with this situation were to give:

- refusals
- inflated scores

Although it could potentially be an awkward situation, respondents who said they would find it too difficult to give scores in front of other household members, said they would refuse to respond:

> *"I wouldn't answer them in front of the household or with my father present, or if I did I would not have answered them at all honestly."*

> *"Impossible. Really not, it has to be confidential."*

> '[I] *would not feel very confident about it at all. Simply because they are questions about thoughts and feelings.* [I would] *feel a bit exposed; bearing a bit of your soul to people who you may not want to, no matter how close you are....deepest feelings really."*

When inflated scores would be given, this was particularly associated with 'negative' scores, which could highlight, or reveal, a problem for the respondent to other household members, of which they may not have been aware. If something was being withheld it would be inappropriate to start talking about it for the first time in an interview situation. There was evidence that answers would then be changed to become more positive if they were asked in front of other household members:

> *"I would probably give a less honest answer, say 'happiness' I would answer a ten."*

> *"I would make sure the response was a positive one and not a true picture of how I felt."*

10.2   Group dynamics

There was evidence that scores would be changed to conform to the group answer. It is perhaps unsurprising that individuals might take the lead from the first person to respond. For example, a parent might follow the lead of the other parent to show a positive united front to their children or so as not to be seen to be dealing with a situation differently. Or, as in this example, students in a shared household might seek to empathise and bring their scores in line with others:

> *"If someone else had said they are really anxious I might say I was anxious too because I wouldn't want them to feel bad."*

10.3   Computer Assisted Self-Interviewing

Respondents were asked if they would be willing to complete their responses on the interviewer's laptop. It was felt that Computer Assisted Self-Interviewing (CASI) would be a better alternative:

> *"I think entering the score on a laptop would be better as no-one in the household would be able to see the score and you could answer honestly."*

> *"Self-completion on a laptop or on a one-to-one basis. Any way of making sure it was confidential in front of other household members."*

However, there are limitations with this approach as elderly respondents or those with visual difficulty are less confident with this mode.

10.4   Answering in front of a Social Survey interviewer

Answering in front of a social survey interviewer was found more acceptable. This was because it was assumed they would be a professional and it would be easier to answer in front of just the interviewer rather than others in the household:

> *"I don't mind. I assume they'd be professional about it."*

> *"Well I assume interviewers are professionals and are confidential so it's no issue for me."*

> *"Probably easier that way. If whoever it is doing it for a reason, then I don't mind."*

There was also a feeling that if a person is not known personally then there is nothing to worry about what they might think of answers:

> *"If you don't know the other person, why would you care?"*

However there was evidence that refusals and inflated answers would still be given.

10.5   Social desirability effects

There was evidence that respondents wanted to appear in a more positive light than their feelings or circumstances actually were. When answering the 'satisfied' question respondents made comments such as:

> *"[I]* *didn't want to give a four in case I sounded depressed."*

At the 'worthwhile' question there was evidence of respondents both not wanting to appear *"too worthy"* and *"like I do nothing for other people."*

Similarly, the 'happy' question was adjusted to conform to what respondents thought was more realistic:

"If I said 10 nobody would believe me so eight is more realistic."

Examples were found at other questions also, which suggests that the well-being questions tested in this research are prone to adjustment by respondents with the aim of presenting a view of themselves that is perceived as more desirable by others. This concurs with other research in the area. The use of CASI would help reduce this issue by making responses more private.

11 Additional issues relating to the additional four questions

The second set of four questions was felt to be more personal and private and therefore less likely to be answered honestly or at all in front of the household than the first set, particularly the questions regarding loneliness:

"I don't think I'd want my son to hear me answer these ones because they are getting more at the nitty-gritty. Say I'd had a really lonely day, I wouldn't want to say that in front of him. A) I'd feel stupid and B) I wouldn't want him to worry."

"I think these are worse in front of other people. It's quite personal. It's self-perception, but it's not just 'are you happy', it's lonely, it's such a judging word… If you had any form of relationship with the people in the room, and you were saying you weren't very happy with that relationship, then it would be really bad."

There was a preference expressed for answering these questions face-to-face rather than on the telephone:

"I wouldn't want to answer these ones by phone; they're more personal…sensitive in some way. It makes me think having someone here would be better for these ones."

Recommendation

Answers to the well-being questions could be seriously impacted by their administration in the household if others are present. Consideration should be given to the administration of the questions. Other methods range from simply instructing interviewers to attempt to interview all household members separately where ever possible, to the use of CASI (see section 10.3).

12 Answering by telephone

When respondents were asked during cognitive probing how they felt about answering the questions on the phone, there were a range of reactions:

- the questions would be refused
- the questions would be easier to answer
- it did not matter which mode

Reasons for refusing to answer on the telephone included not wanting to disclose personal information to someone that could not be seen. This was because it was felt an assessment of trustworthiness could not be made of someone who is not

physically present. Also it was felt the telephone implied a lack of seriousness and is associated with unwanted sales calls and negative experiences with call centres:

> *"I would associate speaking on the phone with somebody with things like calling a call centre and being put on hold while I was waiting around for someone who knows what they're talking about. There would be an underpinning sense of dissatisfaction and unwillingness to engage before I had even started. So I think doing them on the phone would have an impact on the quality of the answer. I would say I am busy...if you can't be bothered to take this seriously then I can't either".*

When the telephone was perceived a preferable mode, it was felt privacy and confidentiality could be maintained more easily. It was also felt answering on the phone would lessen feelings of self-consciousness, anxiety and pressure to answer as no-one would be physically present:

> *"I feel on the phone, because you can't see them* [the interviewer] *you don't mind buying thinking time, because you don't feel pressured as there are no pair of eyes staring at you. Whereas face-to-face you may not be as open and honest actually, because you are under pressure to answer these questions quick fire. It might feel too personal as well."*

When no preference was expressed for either mode, it was felt it would make no difference to how the questions were answered.

The second set of four questions was viewed as less suitable to be answered by telephone than the first set due to the perception they are more sensitive:

> *"I wouldn't want to answer these ones by phone…they're more personal…sensitive in some way. It makes me think having someone here would be better for these ones."*

13 Comparison between telephone responses and face-to-face responses

A sub-set within the sample was asked the well-being questions by telephone. Interviewers telephoned respondents from mobile phones outside respondents' homes and administered the questions, prior to entering the home to conduct the qualitative interview. This allowed the assessment of differences in understanding and use of the questions and the scale, however, the same issues and difficulties were apparent in both modes.

Telephone respondents said that answering the questions by telephone increased the likelihood of wrong answers. It was perceived that more concentration is necessary when answering the questions by telephone than in face-to-face interviews. Those interviewed face-to-face preferred this mode and thought telephone interviewing was not appropriate due to the sensitive nature of the questions.

There is a trade-off between the likelihood of errors caused by the telephone method and the fact that the telephone allows more privacy and confidentiality when answering. However studies have shown substantial differences in responses to scalar questions when asked by telephone versus visual modes, in that more positive responses are given in the telephone mode (Dillman et al, 2009). In addition the

evidence gathered from the present study suggests the telephone is associated with misunderstanding and decreased rapport with the interviewer along with a decreased desire to take part.

Recommendation

Further research is needed. This current qualitative endeavour has highlighted the problems that can occur when using the telephone mode to administer well-being questions, however, an idea of magnitude is now required to show the extent of the problems.

14 Aspects of well-being

Respondents were asked whether they thought the following question topics were representative of well-being:

- satisfaction with life
- things in life being worthwhile
- happiness yesterday
- anxiety yesterday
- satisfaction with personal relationships
- loneliness in daily life
- loneliness yesterday
- control over important aspects of life

Respondents were in one of two camps. In the first of these, they thought the questions gave a brief overview of how people were feeling and were quite happy with this:

> *"I think they do* [reflect well-being] *for me…I just think they tell you a bit about me."*

In the second camp, rather than saying the questions did not reflect well-being, it was difficult to move these respondents on from having their say about the questions:

> *"They do to an extent but I can't believe you really get to know anything about the country by questions like these. You need loads of questions to really understand how the nation is feeling or it's pointless.* [It is] *just a tick box exercise."*

> *"I don't think any questions like these can tell you how people are feeling really. It's all too subjective, how you feel from one minute to the next."*

> *"You'd have to do a proper* [qualitative] *interview like you're doing now to even scratch the surface."*

Respondents thought there were other aspects that fed into well-being that were very important and were surprised that they were not included.

Health

Both physical and mental health was described as being a very important part of one's well-being because without one's health, life can be very challenging. It was said to be the foundation on which other aspects of life sit and therefore strange that it had not been the topic of one of the questions:

"Your health is one of the key fundamental things."

"If you don't have the physical wherewithal to do what you want your well-being will be low so I think health is very, very important. It's only when it's taken away do you realise just how much."

"If you don't have your health you aren't really as happy as you are if you're healthy."

"[Health is] *a very important one for us given how my wife's* [health] *is…it's changed our quality of life, definitely, because we can't do the things we'd like to do."*

Job satisfaction

The desire to have a job or satisfaction while in a job, was also of great importance. Employment was said to play a significant part in an individual's well-being and as the quotes illustrate, job satisfaction is particularly related to emotional/mental health:

"Now that is important because you spend so much of your day at work. If things aren't going right there, it's a complete nightmare."

"That was very important to me when I was working. I'd had jobs over the years where it was terrible and it does affect you."

"I really love my job. I feel happy when I get up in the morning…I think it's really important 'cause my mam had a job she hated and it made her miserable."

"…economic security didn't come into it at all but…job satisfaction did in my case. That is part of the reason I am unhappy that I don't have a job."

Respondents who were unemployed said they would have scored higher at the 'happy' question, had they been in employment.

Economic security

In the minds of respondents, this issue appears to have equal weighting to many of the other issues raised and it was felt money can improve wellbeing:

"Money is the crux of most of this though. I know they say money can't buy you happiness but honestly, if you have money you have the freedom to do so many things that can make you happy; and that's what well-being is really isn't it, just happiness."

"You can do what you want when you have money so it's really important."

"We're very lucky in that we're financially sound. If we weren't, all these health issues would have been harder to bear."

"Money is at the root of it for me because it means I can and can't do things that make me happy."

The environment

Those who mentioned the environment as having an impact on well-being were particularly thinking of the local, rather than the global environment:

''…I feel very relaxed and calm….and I think its this area that does it.''

"I think [the environment] *is important if you are in an awful one, like if you lived in a polluted area, that would get you down…if you were near a factory that produced smells or lots of noise, for example."*

"It would be awful if you lived somewhere where it was really polluted…it would affect your health."

"If you lived in a polluted area or there's loads of graffiti or dog mess around all the time, on some awful council estate for example, that would definitely get you down and I can't see your well-being being great."

Crime

The interviews took place following the riots in London and other major cities. They were mentioned in interviews and may have resulted in respondents being more sensitive to crime as an issue:

"I'd hate to live somewhere where you were frightened to go out at night or where there was gangs of kids. I'd never leave the house…that would become depressing…it would have a negative impact on your well-being."

"If you were frightened to go out 'cause of the crime around you, youths hanging round or drug dealers.''

Religion

Religion was described as offering *"inner peace"* and *"happiness"* and respondents related this to well-being. However there was a strong opposition to this view:

"No, I think religion has a lot to answer for. Stuff that has nothing to do with well-being."

15 Field issues

15.1 Monday to Saturday interviewing

Interviewers in the General Field Force do not routinely work on Sundays. This means that when respondents are asked about yesterday, Saturday would always be missed out. As Saturdays can be quite different, being for many people, a non-working day, this could mean a significant omission in the picture.

15.2 Term-time interviewing

Students noted that their responses would be different if they were interviewed in holiday time or term time.

16 Data security

Respondents were asked how they felt about providing this data to ONS so as to gain an insight into:

- how much confidence people had in disclosing personal well-being information to ONS

- whether they thought it would be kept confidential
- whether this would affect them answering the well-being questions

Respondents held one of three positions:

- they would give any data to ONS
- they would only give selected data to ONS
- they would not give any data to any organisation

The first group were very relaxed about providing data and trusted ONS:

"I'm very happy to give my details. I can't see any great risk really."

The second group were more selective but said that they would answer the well-being questions:

"Well I have to say there has been a few little hiccoughs, hasn't there…well papers going missing in government departments…It makes me feel less comfortable than I was before…no, it wouldn't [stop me answering these questions].*"*

However, they would not give ONS data on financial issues or sexual health/orientation, for example.

The third group were the most guarded and were cautious about providing personal data to any organisation:

"Well, there have been a couple of mistakes made in government departments so I'm not as confident as I was but I wouldn't give anybody my bank details or very personal information, so it's not just you lot."

17 Conclusion

The primary four well-being questions that asked about satisfaction were considered less negative in tone and more acceptable than the additional four well-being questions about satisfaction with personal relationships. The anxious yesterday question was considered to be contentious and respondents would welcome the term 'stress' due to it being more socially acceptable.

Questions that asked respondents to consider 'yesterday' were met with surprise and disbelief because respondents found reporting on one day to be unrepresentative of themselves and therefore an error on the part of ONS. An instruction was added to these questions asking respondents to only think of the previous day. This had a greater positive effect than was expected as it seemed to confirm in the mind of respondents that ONS definitely wanted respondents to only consider the day before the interview took place.

There were significant issues with the scale with respondents finding it difficult to apply. These difficulties ranged from requiring the scale to be explained or repeated to them, to it being completely reversed. This resulted in responses that were the opposite of that which respondents thought they were giving. To overcome this, show cards were designed that were tailored to each question (the extremities were labelled in line with the question topic). This proved very effective in the face-to-face setting. Clearly they cannot be used when the questions are administered by telephone however.

Administration by telephone was seen as advantageous in terms of confidentiality in that other household members would not know which questions were being asked. However, in terms of cognitive burden, respondents thought that the task was potentially more difficult by telephone and could lead to errors.

More vulnerable respondents exhibited clear signs of distress when responding to the questions. Although the questions simply require a numeric response, these respondents talked through their thought processes while trying to arrive at a numeric value. Interviewers noted that respondents were made physically uncomfortable by the questions and often displayed signs of distress.

18 References

The British Psychological Society. (2009) *Code of Ethics and Conduct: Guidance Published by the Ethics Committee of the British Psychological Society.*

Dalton, S., Draper, K., Ralph, K., Waruszynski, P. (2011). *Cognitive interview testing: Subjective well-being questions and potential order effects.* Office for National Statistics.

Dillman, D.A., Smyth, J.D.& Christian, L.M. (2009). *Internet, mail and mixed mode surveys: The tailored design method.* John Wiley and Sons Inc.

Dolan, P. (2011)

Dolan, P., Layard, R. & Metcalfe, R. (2010). *Measuring subjective wellbeing for public policy: recommendations on measures.* Office for National Statistics.

Graham, J., Grewal, I. & Lewis, J. (2007). *Ethics in social research: The views of research participants.* Government Social Research Unit.

Helliwell, J.F. (2011). *How can subjective well-being be improved?* Canadian Institute for advanced research and department of economics, UBC. Second draft of paper for the CSLS symposium in honour of Ian Stewart.

Hicks, S. (2011) *Spotlight on: Subjective well-being.* Office for National Statistics.

Hicks, S. & Tinkler, L. (2011)

Krosnick, J. A., Narayan, S. S., & Smith, W. R. (1996). *Satisficing in surveys: Initial evidence.* In M. T. Braverman & J. K. Slater (Eds.), Advances in survey research (pp. 29–44). San Francisco: Jossey-Bass.

Krosnick, J. A. & Fabrigar, L.R. (2007). *Designing rating scales for effective measurement in surveys.* In L.Lyberg, P.Biemer, M. Collins, L.Decker, E. De Leeuw, C.Dippo, et al (Eds). Survey measurement and process quality. New York: Wiley Interscience.

Pudney, S. (2010) *An experimental analysis of the impact of survey design on measures and models of subjective wellbeing.* Institute for Social and Economic Research.

Office for National Statistics (2011). *Labour Force survey update report: 2012 questionnaire development.*

Rees, G., Goswami, H., Bradshaw, J. (2010). *Developing an index of children's subjective wellbeing in England.* The Children's Society.

Stiglitz, J.E., Sen, A. & Fitoussi, J.P. (2009). *Report by the commission on the measurement of economic performance and social progress.* OECD.
Waldron, S. (2010) Measuring subjective wellbeing in the UK. Office for National Statistics.

White, A. Martin, J. Bennett, N. and Freeth, S. (July 1998) *Improving advance letters for major government surveys: final results.* Survey Methodology Bulletin, No 43.

19 Appendix A: SAMPLE COMPOSITION

Sampling criteria	Number

Sex

Male	21
Female	23
Total	*44*

Age band

16-24	12
25-44	15
45-60	10
61 and over	7
Total	*44*

Socio economic group

Employed	16
Unemployed	11
Economically inactive (including students, long term sick, retired, parents looking after children, long term carers)	17
Total	*44*

Ethnic group

White British	38
Non-White	6
Total	*44*

Location

London	8
Outside London (includes Wales, North England and South England)	36
Total	*44*

Household Type

Lives alone	10
Lives with related adults	19
Lives with unrelated adults	6
Lives with related adults and children	9
Total	*44*

Highest educational qualification

No formal qualifications	4
Below A-level/equivalent	11
A-level/equivalent and above	29
Total	*44*

Mode

Face-to-face	28
Telephone	16
Total	*44*

Mental and physical health Issues

Currently experiencing mental/ physical health issues	10
No health issues	34
Total	*44*

20 Appendix B: Method

Qualitative research

Research that employs qualitative methods must be prefaced with a definition of the term and an understanding of how this data can be used. The term 'qualitative research' encompasses a range of investigative approaches, all of which have at their core, depth, detail, accounts, systems and interpretive epistemology. It can offer great insight into the research subject by exploring, often in minute detail, the associations within the data; it can also be used as an evaluation tool.

Where qualitative techniques are commonly used within ONS however, is in the testing and development of questions; the concepts within questions and the feasibility of operationalising these within surveys.

Qualitative techniques produce rich, detailed data but they cannot be used to infer statistical conclusions. No statistical inference can be drawn from the findings, as this is not the purpose of qualitative research. Hence 'quantitative' language - such as numbers, proportions or percentages of respondents exhibiting particular behaviours or opinions are avoided in the reporting of the findings. It cannot be known with certainty whether, for example, a finding unique to one of our respondents is actually common in the general population, or, conversely, that a common finding is rare. However, qualitative findings are of excellent value when explaining the conditions under which phenomenon occur.

Where it may be possible to generalise the findings to the general population and the implications for the questionnaire, this is discussed. Such judgement is based on our experience of question testing, knowledge of survey research and an assumption that the achievement of our purposive sample design targets provides representation of the broad range of relevant characteristics which may impact upon responses.

Sampling frame

Respondents were found using the respondent register and snowballing. The Respondent Register is a list of individuals who have been previously interviewed for the Opinions Survey and said they agreed to being contacted for further research. This provides a very useful sampling frame, which allows the selection of specific respondent characteristics. Snowballing is a technique whereby researchers identify individuals with the characteristics of interest via other individuals with the same or similar characteristics.

Calling strategy

Calling Strategy
Respondents were contacted at different times of the day and all possible attempts were made to trace the named contact. When the potential respondent was unavailable, times were noted of their availability and attempts were made to contact them again. This continued until contact was made and a decision on whether to take part in the study was given.

The two most frequently stated reasons for not taking part in the research were:

- not enough time

- the predicted interview duration (one hour) was too long

Other reasons were:

- they had given enough time to ONS already

- they could not remember having taken part in the Opinions Survey

- lack of Interest in the study

Incentive payments

Initially, incentive payments were set at £15. This was a token amount to thank respondents for their time and can often be helpful when recruiting. After the first wave of interviews, the amount was increased to £20. This was because the length of the interviews and emotional impact on some respondents had been under-estimated. It was therefore felt that £15 had been too low a sum.

Recruitment strategy

Recruitment took place from October to November 2011. Potential respondents were contacted by telephone. They were reminded that they agreed to take part in further research, this study was explained to them and they were told about the incentive payment. They were then asked if they would allow an interviewer to visit them. If they refused, interviewers tried to persuade them otherwise. Where time allowed, letters were sent as a reminder of the interview time, duration[3] and date. This serves to remind respondents and also verify the authenticity of the telephone contact (White et al 1998).

Interview mode

To gain as much insight as possible into the questions tested, interviewers had to conduct hybrid interviews, in that they were both cognitive and in-depth in nature. This method enabled the interviewers to probe for the detail about well-being, and the understanding of these questions, that the research required.

The interviewers were not bound by a script but instead used a topic guide as an aide memoir. This is an outline of key issues and areas of questioning used to guide a qualitative interview. The topic guide covered the key areas of the study and included some probes. However, the main feature of the topic guide was that it evolved as the study progressed. As new findings emerged they were added to the topic guide so that they could be discussed in later interviews.

To make the most effective use of the interview an audio recording was made. This allowed the interviewer to fully concentrate on the content of the interview without having to make notes, which would not be as comprehensive as a recording. Informed consent was always sought from respondents before recordings were made.

---

[3] The initial estimation of interview duration was 45 to 60 minutes, which some individuals found "*excessive*". However, after 10 interviews were completed actual timings were examined. This showed a range from 35 to 105 minutes

Quality control measures

A project management approach was adopted to provide full quality control. This system engenders close and regular liaison between the research team and those responsible for operational functions. ONS is committed to Total Quality Management (TQM), a continuous process whereby all staff aim to make ongoing improvements in meeting customers' needs.

Throughout the research, quality control measures were implemented. These included the following.

- agreeing the topic guide before use

- signing-off the sample specification

- sending email updates giving progress and raising any issues of concern

- making regular updates to the topic guide to ensure all emergent findings were included; sending updated topic guides to interviewers

- giving interviewers feedback on their performance

- protecting respondents' confidentiality

- cross-checking the ordering and summarising of the data in the analysis chart

21 Appendix C: SHOWCARDS VERSION A

Not at all satisfied   0   1   2   3   4   5   6   7   8   9   10   Completely satisfied

Not at all worthwhile 0   1   2   3   4   5   6   7   8   9   10   Completely worthwhile

Not at all happy   0   1   2   3   4   5   6   7   8   9   10   Completely happy

Not at all anxious   0   1   2   3   4   5   6   7   8   9   10   Completely anxious

22 Appendix D: SHOWCARDS VERSION B

Not at all   Completely

```
  0       1       2       3       4       5       6       7       8       9       10
  |_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|
```