

NatCen

Social Research that works for society

European Social Survey Round 10 self-completion paper questionnaire

User testing report

Authors: Dr Emma Berteen and Natasha Phillips

Date: March 2021

Contents

1	Methods	3
1.1	Background	3
1.2	What is user testing?	3
1.3	Sampling and recruitment	3
1.4	Fieldwork	4
1.5	Analysis	5
1.6	Caveats to user testing	5
2	Findings and recommendations	6
2.1	Overall impressions	6
2.1.1	Main findings	6
2.1.2	Recommendations	9
2.2	Clarity of instructions, numbering and routing	10
2.2.1	Instructions	10
2.2.2	Numbering	11
2.2.3	Routing	11
2.3	Visual design of different question types	13
2.3.1	11-point scale questions	13
2.3.2	Tick-box questions	15
2.3.3	Numerical entry questions	15
2.3.4	Open questions	16
2.4	Occupation questions & education questions	17
2.4.1	C10 & C11: Highest qualification	17
2.4.2	C42, C51 & C54: Partner's and parents' highest level of education	19
2.4.3	C12 & C13: Years of education completed	20
2.4.4	Occupation and income questions	21
2.5	Other question-specific issues	22
2.5.1	Household grid (C4)	22
2.5.2	Human Values Scale (D63-83)	22
2.5.3	Ancestry question (C58)	23
2.5.4	Other questions	24

1 Methods

1.1 Background

Due to the disruption to face-to-face fieldwork caused by the COVID-19 pandemic, the European Social Survey European Research Infrastructure Consortium (ESS ERIC) wanted to explore the feasibility of using a paper self-completion version of the European Social Survey (ESS) Round 10 questionnaire. As part of the development work for this, once a paper self-completion questionnaire had been produced, NatCen were commissioned to conduct a user testing exercise in order to explore the extent to which the paper questionnaire is fit for purpose. This report outlines the key findings from this user testing exercise.

1.2 What is user testing?

User testing is a technique that involves evaluating a product (in this case, a paper questionnaire) to ensure that the interaction between the user and the interface is as smooth as possible. It is a useful tool to ensure that the users are involved at this stage of the questionnaire development and can suggest ways to improve their interaction with the tool. Usability is often assessed by looking at three different measures: accuracy, efficiency and satisfaction.

The user testing conducted for the ESS self-completion paper questionnaire was retrospective in nature, meaning that participants were invited to take part in an interview to discuss their feedback after having completed the questionnaire. During the interviews, participants were asked to reflect on their experiences of filling out the questionnaire and to discuss any difficulties they had. Participants of this exercise included those who do not have/use the internet at home, those aged 60-85, and those who normally opt to fill out forms and questionnaires on paper, since these would be the potential users that would likely complete the real-life survey on paper, rather than online.

1.3 Sampling and recruitment

In total, 20 interviews were conducted with people recruited specifically for this project. The participants varied in terms of age, gender, level of highest qualification, job status, whether they had internet access at home, and whether they normally opt to complete forms and questionnaires on paper or online.

A professional recruitment agency, Propeller Research, was used to assist with the recruitment of participants. We provided the recruiters with a screening questionnaire which they followed when inviting people to take part. Details collected at the screening stage were double-checked at the start of each interview. A confirmation letter outlining details of their appointment and key information about the study was sent to each participant recruited.

Table 1: User testing sample composition

Characteristics		No. of participants
Gender	Male	10
	Female	10
Age	18-39	7
	40-59	6
	60-85	7
Highest qualification	GCSE or equivalent or below	10
	A-Level or equivalent or above	10
Job status	In paid work (or away temporarily)	13
	Self-employed	2
	Retired	4
	Unemployed	1
Normal method of completing forms and questionnaires	On the internet	8
	On paper	12
Internet access at home	Yes	16
	Mobile internet only	2
	No	2

1.4 Fieldwork

Participants were sent a paper questionnaire, accompanied by a cover letter with instructions, around a week before the start of the fieldwork period. They were asked to complete the questionnaire in their own time, ideally in the 24 hours before their interview so that the questionnaire was fresh in their minds. Participants were asked to annotate the document with any issues they experienced as they completed it, and also to note down how long it took them to complete.

Participants then attended an interview with a researcher at NatCen who is trained in user testing methods. During the interviews, participants were asked to give their feedback on various aspects of the questionnaire. Participants were asked probing questions in order to gather the required information. Interviewers used a semi-structured topic guide to ask relevant probes. Probes looked at:

- Overall impressions of the paper questionnaire;
- Clarity of numbering, routing and instructions;
- Visual design of different question types;
- The format and user-friendliness of the occupation and education questions;
- Other question-specific issues, including the layout and instructions for the questions on income, ancestry, the Human Values Scale and the Household Grid.

Participants were interviewed remotely due to the COVID-19-related restrictions in place at the time of interviewing (Spring 2021). Interviews were conducted either via telephone or via the video-conferencing software Zoom, depending on participants' internet access and digital literacy levels. Participants were based across England, specifically in the Midlands, North-West and South-East.

Interviews lasted approximately an hour and interviews were recorded with the participants' consent. Participants were given a £30 high street voucher as a thank you for their time and help.

After the interview, participants were asked to post their questionnaire back to NatCen in the pre-paid envelope provided. Researchers then reviewed the completed questionnaires to identify any further issues that had not been discussed during the interviews. Where necessary, and where consent to re-contact had been given, interviewers made brief follow-up calls to participants to discuss these additional issues.

Procedures for this user testing exercise were approved by the NatCen Research Ethics Committee and the ESS Research Ethics Board prior to fieldwork commencing.

1.5 Analysis

The interviews were recorded with the participants' consent. Interviews were summarised and charted by interviewers who reviewed the recording of each interview. All interview summaries were written into a matrix in Excel. Key findings from each of the scripted probes were recorded in the charts. Data could thus be read horizontally as a complete case record for an individual, or vertically by topic, looking across all cases. Once the matrix was complete, the data in the matrix were reviewed thematically. A debriefing session was organised for NatCen and the ESS ERIC team in order for the interviewers to present the initial findings.

This report presents results from this thematic analysis of the data, and recommendations for any alterations, which were discussed at the debriefing session. The recommendations for changes are presented at the end of each subsection.

1.6 Caveats to user testing

The user testing conducted for this project is qualitative user testing. This can be used to explore the types of issues that people encounter when completing a survey, and potential reasons for the occurrence of these issues. However, as small purposive samples are used, qualitative user testing methods cannot be used to quantify how often the problems reported will occur in practice. Where problems are detected, it is not possible to quantify the impact these problems will have on survey data collected. Although this report provides information on the number of test participants who encountered an issue, these numbers should not be extrapolated as being representative of the wider survey population. Throughout this report, we make recommendations that aim to address the issues raised by the user testing participants. However, there is no guarantee that all potential issues with the questionnaire have been detected or that the solutions that we propose are themselves problem free.

2 Findings and recommendations

2.1 Overall impressions

2.1.1 Main findings

Overall look and initial reactions

- The initial reaction of most participants was that the questionnaire was longer than they had expected it to be. However, all participants proceeded to complete the questionnaire and none of the participants decided not to take part after having seen the length of the document.
- Many of the participants expressed positive views about the overall look and design of the questionnaire, stating that they thought it looked professional and official. Participants reported finding it easy to read and that there were no surprises in terms of the look of the document. One participant remarked: *“I thought it was very well done I have to say”* (EB05).
- Participants liked the use of colour in the questionnaire; some commented that they liked that the colours were simple and *“not too busy”* (FS05). Others found these *“nice and bright”* (FS03). In particular, participants found the red colouring of the question numbers useful, since it made the numbers stand out, which assisted with navigation through the questionnaire.
- The cover page was generally well liked by participants. They found the map and title on the front cover engaging and interesting. One participant (EB03) commented that the text on the front-page states that it is a ‘brief survey’, yet thought that this was slightly misleading, given that they found the questionnaire to be relatively long. Some participants said that they would have liked a little more information on the front cover about the topics covered in the survey, to give them a clearer idea of what the questionnaire would be about.
- The size of the text was considered suitable for all participants. Participants in our test found all text to be easily readable and clear.

Ease of filling in

- In general, participants managed to complete the whole questionnaire without any significant difficulties. Participants reported skipping questions largely only if they had not been routed to them. A few participants did not want to answer certain questions as they found them too sensitive and/or intrusive (e.g. the income question at C40). One participant (NP05) reported finding most of the questions on employment, as well as C6 and A65, quite sensitive and intrusive, but answered these nonetheless.
- Occasionally participants reported leaving questions blank if they did not know the answer (e.g. C51 and C54 about parents’ qualifications). There were also a few cases in which participants left questions blank if they had doubts about how to answer (e.g. EB01 did not answer the Human Values Scale due to finding it unclear). Such question-specific issues will be discussed later on in this report.
- Participants generally did not make many mistakes when completing the questionnaire. Where this did occur, participants mostly indicated these through filling in incorrect boxes/circles or crossing out incorrect textual answers. It was not

always apparent to participants that there were instructions for correcting mistakes on page 2, but most participants felt they intuitively knew how to indicate a mistake. Upon reviewing the questionnaire documents, it was visually clear when participants had indicated mistakes and done this in line with the instructions.

- In terms of the time taken to complete the questionnaire, this varied quite a lot between participants, with some reporting a completion time of 20 minutes and others up to 90 minutes. Most participants did find the questionnaire quite long, but many said they found the process fine once they had got going. The table below outlines self-reported completion length among participants, as well as number of breaks reported. The self-reported estimates of completion length reflect time spent on actual questionnaire completion and do not include the duration of any breaks.

Serial number	Self-reported length of time taken to complete questionnaire (minutes)	No. of breaks reported
FS04	20	0
NM01	20-25	0
FS03	20-30	0
FS02	25	0
NP01	25-30	0
NM04	30	0
NM05	30	0
NP04	30	1
EB04	30-45	0
NP03	35-45	1
NM02	40	0
NM03	40-45	0
FS01	40-45	1
NP02	45	0
EB01	45	1
NP05	60	0
EB05	60	1
FS05	60	1
EB03	60	3
EB02	90	1
	Mean: 42 minutes¹	
	Median: 40 minutes	

¹ Where participants gave an estimated range, the mid-point of the range has been used to calculate the mean length of time.

- Some participants reported completing the questionnaire in one sitting, and others took one or more breaks. For those who took breaks, this was mostly due to finding the questionnaire quite long and getting slightly fatigued (e.g. EB03 who took 3 breaks). These participants reported finding the sections useful in dividing the questionnaire up in order to take breaks.
- There is some, limited evidence of fatigue in the questionnaire documents themselves. There are three cases of missing data appearing towards the end of the questionnaire: NP02 skipped C51-55 and D63-83; NM05 skipped pages 27-31; FS04 skipped D11-80 (but did answer D32-33).

Layout

- Participants found the spacing between questions adequate and generally thought it was all well laid out.
- It was clear to participants where one question finished and the next one started. Having the numbering in colour helped with this. Participants generally did not find the pages too cramped; the spacing and amount of material per page for most participants was “*about right*” (NM04). Some participants did, however, comment that there was quite a lot of text per page, but this did not present any major issues in the testing.
- Participants found having the question text in bold useful.
- The two-column format was generally well liked by participants. During probing, most participants said that they preferred this to a one-column format as it made the booklet shorter. Some added that being able to see multiple questions per page made the routing easier. Participants found it easy to follow the order of questions in the two-column format.
- Some participants commented that the page numbers on the even-numbered pages were not entirely clear, as these are currently right-aligned.

Flow and topics

- Participants generally thought the questions flowed well throughout the questionnaire, commenting that it was “*logical*” (NM01) and “*fluid*” (NM03). However, a few participants disliked that not all topics were always asked together, e.g. NP05 did not like that the questionnaire asked about their job in more than one section and felt that these questions should have been together.
- Participants commented that they liked the variety of questions and question formats, as this kept the questionnaire interesting. Some participants really liked the questions about politics and related issues, whilst others found these less interesting. Some participants commented that some of the questions were unexpected, e.g. the questions about politics, religion, social issues, and their parents.
- Participants mostly found the questions relating to the pandemic at the start of the questionnaire interesting: “*topical*” (FS02), “*you get engaged*” (FS01). Some participants reported finding these very relevant, and thought it was relevant for these to be followed by more general questions about government. Some participants commented that this set the tone for the rest of the questionnaire. However, some participants found this to be quite a heavy start and said they would have preferred simpler questions to begin with (e.g. NP01). Nevertheless, starting

the questionnaire with questions about the pandemic did not put any participants off completing it.

- Many participants stated that it was not always clear whether to answer the questions in terms of today, or instead thinking about pre-pandemic life. This was particularly an issue for some of the questions relating to politics, employment and income.

2.1.2 Recommendations

- Retain the general format and design of the questionnaire. This worked well overall and participants felt that it looked professional and well designed.
- Retain colour in the questionnaire and ensure that the document is printed in colour in all countries. This is particularly important for the red numbering throughout.
- Retain the two-column design. This worked well in general and keeping the length of the document short proved important.
- Make very clear in the questionnaire whether participants ought to answer in terms of today or whether to consider their pre-pandemic situation. We suggest that this is communicated clearly at the outset, and perhaps also at start of each section as a reminder. Consider also communicating this with specific questions that are likely to be particularly affected, e.g. questions about socialising (A61), income (C40) and employment (C19-39), or with any “We would now like to ask you...”/“The following questions are about...” statements preceding these questions.
- Move the even page numbers to the left-hand side of the page to ensure that these are clear and easily readable.

Suggestions to consider

- Consider including a little more information on the front cover about question topics that will appear in the questionnaire so that participants are aware of these before completing it. If there is limited space on the front cover, consider adding this to the advance letter instead.
- Continue to ask questions about the pandemic at the start of the questionnaire as these were generally seen as relevant and engaging. However, consider using any less “heavy” questions on this topic to start with to ease participants in.
- Given that some participants found certain questions quite sensitive and had privacy concerns (e.g. questions on employment, income, C6 and A65), include a statement of reassurance of confidentiality on the front cover if there is space to do so. If there is no space, consider adding this to the advance letter instead.
- It would be helpful to communicate roughly how long it might take participants to complete on the front cover or in the advance letter. Our testing suggested 40-45 minutes on average. Remove the word “brief” (“a *brief survey*”) from the front cover as this was felt to be misleading.
- If questionnaires are to be scanned in any of the countries, and the scanning organisations require mistakes to be formatted very consistently, consider adding a small visual illustration of how mistakes are to be indicated at the top

of page 2. This could help ensure all participants indicate mistakes in the same way.

2.2 Clarity of instructions, numbering and routing

2.2.1 Instructions

2.2.1.1 Main findings

- Participants generally found the instructions relating to how to answer questions (e.g. how to fill out the 11-point scales, how many boxes to tick for tick-box questions etc.) clear throughout the questionnaire.
- Some participants felt that the instructions were not necessary, as it was already clear to them how to answer the questions. This was true for the scale questions, for instance, where some participants felt this was self-explanatory. Consequently, these participants reported not reading the instructions very often if at all. However, other participants reported finding the instructions very helpful and appreciated the clear guidance. This was particularly true towards the beginning of the questionnaire. Towards the end, participants reported not paying as much attention to the instructions as they felt more confident in how to answer questions.
- Having reviewed the completed questionnaire documents, there were not many questions whose instructions were not commonly followed. The main exceptions to this are the education questions (C10 and C11), which are discussed later on in this report. Additionally, there were a few, limited cases of ‘tick one box only’ vs. ‘tick all that apply’ not being adhered to, with some participants ticking more than one box at single-coded questions, e.g. NM01 ticked three boxes at C39 so had missed the “tick one box only” instruction. One participant (NP03) suggested that it would be clearer if these instructions appeared on a new line to visually separate them more from the bold question text.
- Two participants reported not being entirely sure whether they were allowed to select numbers 1-9 on the 11-point scales. This is discussed further in the section on 11-point scales below.

2.2.1.2 Recommendations

- We do not recommend any significant changes to the questionnaire with regard to the instructions relating to how to answer different question types. Instructions relating to 11-point scales are discussed further below.

Suggestions to consider

- Consider making any “Tick all that apply” or “Tick one box only” instructions visually clearer. Suggest that – if possible in terms of spacing and formatting – these are put on a new line to slightly separate these out from the question stem. Consider this in particular for C39. It may also be worth considering placing these instructions in italics.

2.2.2 Numbering

2.2.2.1 Main findings

- Participants reported finding the red question numbering clear and easy to follow. Having the numbers in red was well liked by participants. They remarked that this made them stand out, and helped with routing. Some participants mentioned liking the position of the numbers on the same line as the question stem.
- In general, participants did not notice the small grey numbers next to tick boxes (office-use text required for keying). Where these were noticed, participants did not find these distracting at all, and did not pay attention to them when answering.
- Participants largely liked that the questionnaire was divided into four sections. They stated that this helped to break the questionnaire up into chunks, and those who took breaks found these divisions useful. During probing, participants said they would not have preferred continuous numbering (i.e. no sections), or more, smaller sections.
- Some participants said they would have liked each section to have a title, so it was clear what the questions would be about. Others mentioned that it would be useful if a new section started on a new page.
- Generally, participants felt that questions were grouped together well, but a few participants stated that they disliked that they were asked about the same thing, e.g. their job, in more than one section. Others found this variation useful in terms of staying engaged.

2.2.2.2 Recommendations

- We do not recommend any significant changes to the questionnaire with regard to numbering.

Suggestions to consider

- Retain sections, but – where possible – consider starting a new section on a new page, provided that this would not heavily disrupt the formatting and would not mean that the overall length of the questionnaire is considerably extended.
- Consider titles for each of the sections where appropriate and/or possible.

2.2.3 Routing

2.2.3.1 Main findings

- Overall, participants managed to navigate their way through the questionnaire without significant difficulties, and generally answered the questions that they were supposed to. Some participants commented that following the routing required them to pay attention, and occasionally had to double-check the routing instructions, but that it was manageable overall. No participants reported getting completely lost at any point.
- Some participants mentioned that having the routing text in bold was clear and helpful.

- Some participants reported reading all of the questions anyway, regardless of whether they were routed to them or not, in order to check that the questions did not apply to them.
- Participants found routing instructions in the form of brackets and arrows next to questions (e.g. in A25 on page 4) clear. They stated it was clear which instructions related to each response option. Participants felt that the instructions 'continue to X below' and 'skip to X across' were clear and could follow these. Some participants felt that the word 'across' wasn't necessary, but had no issues with it in terms of comprehension.
- Participants found skipping to questions on different pages fine, but some commented that the page numbers on even pages was not very clear due to not being left-aligned.
- Participants did not always read routing instructions that appeared before a question. During probing, participants commented that if they had been directed to a question, they would generally only read from the question number onwards, rather than reading any preceding instructions. For example, if participants were instructed to skip to C16 (on page 17), many did not read the instructions above this question, and simply read from the number C16 onwards. This resulted in some participants answering this question when they were not supposed to.
- Similarly, some participants missed the routing at the top of page 15, i.e. the pre-question routing instructions before C5 and C8. FS05 stated that they "*mainly ignored that top bit*". Moreover, some participants did not answer any questions on this page. Some said this was because they did not have anyone else in their household so assumed they did not need to fill in. The excess spacing on the page also caused confusion, e.g. NM01 had to "*read and read again to make sure I was doing it right*".
- Routing instructions that appeared after questions were also missed by many participants, especially where these were at the bottom of a page, e.g. the routing after C14 and C43, which many participants said they did not read.
- Participants generally did not have significant issues with longer sets of instructions which appeared at the start of a set of questions (e.g. those at the top of page 18). Most reported being able to understand and follow these. However, others found this a little long and were not able to follow it with ease. Some of the confusion around these instructions related to participants being unsure whether they were supposed to answer the questionnaire in terms of their current situation or in terms of their pre-pandemic situation.

2.2.3.2 Recommendations

- Move any pre-question routing in line with (or below if on the same line is not possible) the question number. This is to help make sure that the routing instruction is read, particularly if participants are asked to skip to a question from a previous one. Make sure that this is checked for each case, as some of the wording of the routing instructions may need slightly editing.
- In some cases, it might make more sense to simplify the routing instructions and place these either above the question stem (in line with the number) or below the question stem. We suggest that this is decided on a case-by-case basis for each question with pre-question routing. For instance, at question C42, this could read:

C42 What is the highest level of education your husband / wife / partner has successfully completed?

➔ **If you do not live with a husband / wife / partner, skip to C51 on page 22.**

or

C42 If you live with a husband / wife / partner, please answer this question. Otherwise, **skip to C51 on page 22.**

What is the highest level of education your husband / wife / partner has successfully completed?

- Move any routing which appears after a question (particularly where this is at the bottom of a page) to the next question, either in line with or below the question number. This is to make sure that this is read. For example, move the routing after C43 onto C44.
- As stated above, clearly communicate whether participants should answer about their current situation or pre-pandemic situation, particularly around the employment questions and the instructions at the top of page 18.
- Left align even page numbers (as stated above).

Suggestions to consider

- Suggest including a clear visual demarcation before C5 to show that this is not part of the Household Grid, and also adding a sentence at the top of the page, e.g. 'The next questions ask about your family and the area you live in'. Alternatively, consider a title at the top of this page. This is to avoid participants ignoring the questions on page 15.
- If there is space to do so, move C8 and C9 into the left-hand column, or try and avoid the extra spacing on page 15 in another way, as this did cause some (limited) confusion.

2.3 Visual design of different question types

2.3.1 11-point scale questions

2.3.1.1 Main findings

- In general, participants found the 11-point scale questions to be engaging and visually appealing. Participants largely found the format to be clear and easy to fill in: *"whoever has done it has done a really good job"* (EB02).
- Participants reported no issues with reading any parts of the scale questions. Specifically, participants felt that the grey-coloured numbers in the circles were sufficiently visible, and the size of all of the text was suitable. Participants found the size of the circles adequate.
- Participants generally felt that the labelling of the scales was clear, and understood what each number on the scale represented. It was clear to participants that the labels corresponded to the end points of the scales. One participant (NM02) commented that it was helpful that the end-labels were slightly smaller in size, which helped differentiate them from the question stem text. Another (NP01)

commented that some of the end-labels were quite long and would prefer if these were shorter, but this participant had no issues understanding that these labels referred to the end points of the scale only.

- Some participants stated that the direction of the scales felt relatively consistent, which made completing these questions easier. They did not feel that each item flipped between positive and negative each time (e.g. NM01). Another (EB04) commented that they felt that the scale questions were free of bias, and felt free to tick any circle they wanted to.
- It was not always clear to some participants that they were able to give answers between 1 and 9 at these questions. One participant (EB02) felt unsure of this at first but then decided this was possible and proceeded to give answers that were not 0 or 10. Another participant (NP02) only answered 0 or 10 throughout. Other participants did not have any issues with this.
- Scales that had previously been converted from grids (e.g. A17-24) worked well. Participants had no issues with the question stem and instructions appearing just once at the top. They did not feel that it was necessary to repeat this before each item. However, some participants (e.g. NM02, NM04) did comment that the end-labels were necessary here so should not be removed. They also liked that the question stem was repeated in brackets on page 4 to reiterate that A20-24 were part of the same battery.
- Regarding the scales in section B, specifically the questions about democracy in general vs. democracy in Britain, one participant (EB03) did not like that B25 asked about democracy in general after a whole set of items asking about democracy in Britain. They reported having to concentrate harder for B25-B30, as these oscillated between democracy in general vs. in Britain (in contrast to B1-12 and B13-24 which are grouped together).

2.3.1.2 Recommendations

- Retain the 11-point scale question formatting.
- Retain the formatting of batteries of scales which have been converted from grids, e.g. A17-24, B1-12. The results from the testing do not suggest a need for the question stem to be repeated before each item.
- Retain the instructions at A1, but consider amending the wording to clarify that 1-9 would be valid answers. Possible alternatives include:
 - Please tick any number from 0 to 10
 - Please tick any circle from 0 to 10
 - Please tick one circle from 0 to 10
 - Please tick a number from 0 to 10

Suggestions for consideration

- Consider re-ordering B25-B30 so that questions about democracy in general vs. democracy in Britain are grouped together. However, this is not essential if this ought to be retained for other reasons. Retain the underlining of “in general” and “in Britain today” in these questions to help with clarity.

2.3.2 Tick-box questions

2.3.2.1 Main findings

- Questions which required participants to tick one or more boxes from a list seemed to work well in general.
- Participants reported finding the size of the boxes adequate.
- Some participants queried the suitability of their ticks or size of ticks. For instance, one participant questioned whether it mattered that their ticks had gone over the lines of the boxes (FS02). There were a few, limited instances of one participant (FS04) not ticking the boxes but putting a line/scribble instead, or ticking next to the text rather than in the box. This was visually clear but may cause problems if scanned by a computer.
- It was generally clear to participants when to tick one box vs. when to tick all that applied. Some participants felt that this was intuitive and did not need explicit instructions. Participants generally reported having seen and paid attention to the “tick all that apply” instructions where stated, e.g. at D1. However, very occasionally these instructions were missed, e.g. NM01 ticked three boxes at C39 so had missed the “tick one box only” instruction. NM04 felt that the “tick all that apply” instructions did not stand out sufficiently. The main issues arose with the education questions (C10 and C11) which are discussed separately below.

2.3.2.2 Recommendations

- If questionnaires are going to be scanned, and keeping neat ticks within the box edges is important to the scanning process, consider an illustration at the start of questionnaire (on page 2) showing participants how boxes ought to be ticked.
- Suggest making any “Tick all that apply” or “Tick one box only” instructions visually clearer. Suggest that – if possible in terms of spacing and formatting – these are put on a new line to slightly separate these out from the question stem. Consider this in particular for C39.

2.3.3 Numerical entry questions

2.3.3.1 Main findings

- The numerical entry boxes generally worked well. Participants felt that the size of the boxes was adequate.
- Some confusion arose where boxes were empty (rather than having shaded letters/numbers in them), e.g. A6 and A8, since some participants were not sure why two boxes were given for hours, i.e. that the format required one digit per box. There was also some uncertainty relating to whether to put single-digit answers into the left-hand or right-hand box. There was variation among participants in their decisions. These answers were all clear upon reviewing but could potentially cause issues if scanned. Additionally, there were some formatting differences for empty boxes; some participants put dashes in these, e.g. 4 5 -, and some entered a zero, e.g. 0 4 5.
- At questions such as C24, one participant (NM05) stated that they felt it would be fine to enter a single number answer into the leftmost or rightmost box. However, this was not a question that the participant was routed to.

- At question C27, one participant (NM01) was confused about having three boxes for their answer. They had to re-read the question to make sure they were answering correctly, and commented "*surely they don't want a three-figure number there?*". Nevertheless, this participant was able to give an appropriate answer (040) at this question.
- Some participants commented that some of the numeric entry questions were difficult to answer, e.g. A6 was difficult to estimate. One participant (EB03) said they would have preferred bands at this question.

2.3.3.2 Recommendations

- Put greyed-out letters (e.g. H H or M M for time questions) or numbers (e.g. 0 0) in numeric entry boxes to help keep formatting consistent.
- For time questions, consider giving an explicit instruction, e.g. "Please enter the time in hours then minutes, e.g. if you did this for one hour and twenty minutes you would enter 01 20".
- At question C27 and any similar questions, consider a single long (empty) box for participants' answers, rather than three separate boxes. Individual boxes seem more appropriate and less confusing for questions where we know in advance the correct number of digits (e.g. dates and times). Consider a single, longer, empty box for all questions where we do not know in advance the correct number of digits.

2.3.4 Open questions

2.3.4.1 Main findings

- In general, participants found open text formats fine. Participants did not report any significant issues with these questions in terms of the formatting.
- Participants mostly felt that there was enough space to write in an answer for these questions.
- One participant (EB04) commented that space felt a little tight given their large handwriting. This was particularly the case for C29 and C32, but this participant wanted to give a lot of detail about more than one job (despite being aware that the instructions at the top of page 18 did not require this). However, the participant managed to answer these questions very clearly.
- NM01 and NP01 stated that they were not sure if they needed to write in capital letters or lower case at the open text questions. Most participants, including NP01, did not use upper case. NM01 wrote in upper case at A81 and in lower case for longer open text questions (e.g. employment questions).
- In some cases, participants' handwriting was not always easy to read at these questions.

2.3.4.2 Recommendations

- If this is important for scanning, consider asking for text to be in block capitals to help with readability.
- Similarly, if required for scanning, consider a 'one box per letter' format for shorter open text questions such as A81 about languages spoken at home.

2.4 Occupation questions & education questions

2.4.1 C10 & C11: Highest qualification

2.4.1.1 Main findings

- A number of participants reported difficulties in selecting their answers at these questions, due to the long list of response options, and found it quite time consuming to do so. Others felt they were able to select their answers easily.
- The instruction to only select the first box from each list where they had passed the examination was broadly missed or misinterpreted. Findings from a review of the returned questionnaires showed that:
 - Twelve participants ticked more than one box at C10 or C11
 - Two participants left both C10 and C11 blank
 - One participant ticked only one option at C10, leaving C11 blank
- One participant (FS02) thought they were being asked to tick the first box that applied to them and then move onto the next option within the list, which suggests that the instruction 'moving down the list' may have been misleading. Participants generally agreed that it should be made clearer that the questions were referring to their highest qualification.
- One participant (NP05) only selected one box at C10 and left C11 blank as she thought they were one question, due to the similar look of the lists. She did not look at C11 after finding the option that applied to her at C10.
- Some participants said that they had missed the option that applied to them due to the number of different qualifications listed under each option. Both EB05 and FS05 said they could not find Higher National Diploma (HND) in the lists, even though this qualification was listed under option 5 at C11. This may have been because the awarding bodies are shown before it, making it more difficult to find when scanning the list.
- A number of participants expressed a preference for writing in their highest qualifications in an open text format.
- Other specific issues with C10 and C11 included:
 - NP05 interpreted the question as asking about the first exam they had passed rather than first option in the list that applied to them.
 - NM02 did not know what level the BTEC he had completed was and was unsure whether to select option 3 at C10 or option 7 at C11.

- NM05 was not sure whether a qualification completed through work would count under option 10 at C11. NP02 was unsure of whether to count an apprenticeship completed through work. EB02 was unsure whether to include professional training so included this as free text underneath C10.
- NP02 and FS04 could not find a qualification that applied to them and did not notice the ‘none of these’ options at the bottom of the lists and skipped the questions instead.
- Some participants suggested adding an explicit instruction to tick only one box. NM04 suggested splitting the lists into two, where the most common qualifications were listed first (i.e. GCSEs and A Levels), and then the others are asked separately. NP04 suggested moving C11 to a different page as there was too much information on one page.

2.4.1.2 Recommendations

- If the grid format is retained, add explicit instructions to select only one option and only the *highest* option at C10 and C11. Since many participants ticked more than one option, we recommend prioritising these changes. Consider amending the question to read something along the following lines:
 - From the list below, what is the highest level of education you successfully completed?
 - or
 - What is your highest qualification?

Please tick one box only in the list below.

If you do not have any of these qualifications, please tick “None of these” at the bottom of the list.
- Alternatively, the instructions could be changed to “tick all that apply”, and a variable could be derived on which qualification is the highest. This could make the process simpler for respondents and the same information would still be collected.
- If length/formatting considerations allow, we suggest moving C11 to a different page to make the distinction between C10 and C11 clearer, and to avoid participants only selecting an option from one list.
- Alternatively, consider replacing C10 and C11 with an open text question asking participants to enter their highest qualification. This would eliminate the need for participants to read through all of the qualifications, however it is possible written responses will not provide the desired level of detail. It is also possible that participants will state what the qualification is in rather than the level. We recommend assessing the answers given for partners and parents in Table 2.4.2 (where this information was known/applicable) to determine whether the level of detail in these answers is suitable, to help inform the decision.

2.4.2 C42, C51 & C54: Partner's and parents' highest level of education

2.4.2.1 Main findings

- There was a considerable amount of variation in answers at these questions. For those who were routed to C42, most participants were able to give their husband/wife/partner's highest level of qualification. However, for parents, they were generally less able to recall this information and either indicated this or left the questions blank.
- NP04 participant wrote the occupation of their partner and mother rather than their highest qualifications.
- Table 2.4.2 below details participants' responses to the open questions at C42, C52 and C54 which ask about partner's and parents' highest level of education.

Table 2.4.2: Written responses to questions C42, C51 and C54 about partners' and parents' highest level of education

Serial number	Response at C42	Response at C51	Response at C54
EB01	<i>(blank)</i>	<i>(blank)</i>	<i>(blank)</i>
EB02	<i>n/a – widowed</i>	<i>Secondary modern</i>	<i>?</i>
EB03	<i>Master's degree</i>	<i>Master's degree</i>	<i>Degree</i>
EB04	<i>(blank)</i>	<i>High school education</i>	<i>High school education</i>
EB05	<i>(blank)</i>	<i>No qualifications</i>	<i>No qualifications</i>
FS01	<i>(blank)</i>	<i>HNC</i>	<i>NVQ3</i>
FS02	<i>GCSEs</i>	<i>Degree</i>	<i>Degree in teaching</i>
FS03	<i>None</i>	<i>None</i>	<i>None</i>
FS04	<i>(blank)</i>	<i>(blank)</i>	<i>(blank)</i>
FS05	<i>Accountant ACCA</i>	<i>?</i>	<i>?</i>
NM01	<i>A Levels</i>	<i>Degree</i>	<i>Degree</i>
NM02	<i>A Level</i>	<i>No idea</i>	<i>No idea</i>
NM03	<i>(blank)</i>	<i>No knowledge</i>	<i>No knowledge</i>
NM04	<i>Degree</i>	<i>NVQ</i>	<i>None</i>
NM05	<i>(blank)</i>	<i>No idea</i>	<i>No idea</i>
NP01	<i>GNVQ</i>	<i>(blank)</i>	<i>(blank)</i>
NP02	<i>(blank)</i>	<i>(blank)</i>	<i>(blank)</i>
NP03	<i>(blank)</i>	<i>(blank)</i>	<i>(blank)</i>
NP04	<i>Plumber</i>	<i>Don't know</i>	<i>Typist</i>

2.4.2.2 Recommendations

- We do not recommend any significant changes to these questions based on the findings of the user testing exercise. However, consider the suitability of the level of detail in the above data. If this is insufficient, an instruction could be added which specifies that the exact level of the qualification is needed (and refer back to C10 and C11 if necessary, or give some examples). For instance: 'If you do not know their highest qualification, please look at the list given at C10 and give your best guess. If they had/have no qualifications, please write "None"'.

2.4.3 C12 & C13: Years of education completed

2.4.3.1 Main findings

- C12 was generally answered well, however some participants were unsure whether to count any education they had completed after a break, e.g. a gap year.
- One participant (NP01) was not sure if he should include university at C12, despite completing a full-time university course.
- Another participant (NM02) entered his age when he had finished school at C12, and at C13 only counted the years of education he had completed after school. The questionnaire review showed that NP04 had written '2' for C13, suggesting she had also not counted her years of primary or secondary education.
- Not all participants fully understood 'full-time equivalents' at C13, however most had not completed any courses part-time so this instruction was not relevant to their answers. NP05 gave the example of going to college rather than sixth form, suggesting further clarity is needed around this term.
- Some participants were unsure of how to count time spent on part-time courses and others were not sure of the earliest level of education that they should count. It was also not clear to some whether they should include professional training.
- A few participants said that the question could include a statement about the age that full-time education typically starts.
- Participants who had completed education whilst working had difficulties with C13. NP02 also skipped this question as he did not know whether to include his apprenticeship. NM05 also left it blank as he was not sure of the aggregate length of time he had spent on the course he had completed at work. FS04 did not think C13 was relevant to him as he had completed an apprenticeship after leaving school so left the question blank.

2.4.3.2 Recommendations

- Consider including an example/explanation of 'full-time equivalents' at C13. For example, "If you completed a part-time course, please count the number of years it would have taken you to complete the same course full-time". This could be included as a help screen in the web survey.
- Consider specifying whether professional training, including apprenticeships, are in or out of scope at these questions.
- Consider adding clarification that full-time education in the UK typically starts at age 4, or is typically 11-13 years. This could also be included as a help screen in the web survey.

2.4.4 Occupation and income questions

2.4.4.1 Main findings

- Most participants were easily able to write in their answers at C29-33 and generally preferred to do so rather than having to read through a list of response options. A small number commented that they thought more space was needed. Most participants had sufficient space.
- There were no notable issues with the questions about partners' and parents' work and occupation, although some participants commented that they were surprised to be asked these questions.
- The instructions at the top of page 18 were generally followed. One participant (FS04) was unsure how to answer questions C19-34 as they had missed the instruction to answer about their most recent job and wrote the names of two jobs at C29. There were some queries from participants here about whether they should give answers about their pre-pandemic situation or current situation.
- Despite the instruction to tick only one box, NM02 selected two options at C30.
- EB01 did not answer questions C25-33 as she is self-employed and did not think the questions were very relevant to her work.
- The income grid at C40 generally worked well, with most able to select their answer from the list. Most participants used the annual income column to select their answer.
- There were no issues with the look of the grid. Some participants did not think the weekly and monthly columns were needed but others thought these were helpful. Participants generally found the grey shading of alternate lines helpful.
- One participant (NM05) felt that the grid was not representative of people receiving benefits, as the lowest income bracket was too high. Another (EB04) was not sure whether to answer about their income before the pandemic or now as their financial situation had changed.
- Five participants chose not to disclose their household income as they felt this was too sensitive.

2.4.4.2 Recommendations

- We do not suggest any significant changes to the occupation and income questions based on the findings of the user testing.
- As stated above, it would be helpful to state whether participants should answer about their current or pre-pandemic situation.
- As stated above, consider putting "tick one box only" instructions on a new line to help these stand out.
- As stated above, consider a statement of confidentiality on the front cover and/or in the advance letter.

2.5 Other question-specific issues

2.5.1 Household grid (C4)

2.5.1.1 Main findings

- Participants generally thought the household grid was clear and easy to fill in. There were no issues with the two-column format of the grid. One participant (NM01) said this section initially looked intimidating, but after reading the questions thought it was easy to complete.
- As intended, no participants entered their own details in the household grid, although one participant (FS04) mentioned that it was not initially clear whether she should do so or not. All participants completed the grid for all other members of their household.
- A few participants appeared not to include themselves when asked to write in the number of people living in their household at C1.
- One participant (EB04) initially entered the details of the second other person under Person 3 rather than Person 2, but then corrected this. Another participant (NP02) entered the details of his wife under Person 6, leaving the other sections blank.
- Although not applicable to him, one participant (NP01) said it was unclear how to answer if there were more than eight other people in a household.
- When exploring routing instructions, it became clear that a few participants had mistakenly skipped questions C5-9. Two of these were people who lived alone and may have seen this page as a continuation of the household grid. Another participant (NM04) who skipped these questions said he thought page 15 was only for households with 8 people. He had not read the instructions at the top of page 15.

2.5.1.2 Recommendations

- Consider making the instruction for people who live alone to skip to C6 more salient, e.g. by underlining 'only person in your household'.
- Consider adding a visual demarcation at the top of page 15 to make it clearer that these questions do not directly relate to the household grid. Consider also adding a sentence at the top of the page, e.g. 'The next questions ask about your family and the area you live in', or a title, e.g. 'Your relationships and where you live'. Also consider rearranging page 15 to remove the white space at the bottom of the first column, if formatting allows.

2.5.2 Human Values Scale (D63-83)

2.5.2.1 Main findings

- Some participants particularly liked answering this section as it allowed them to reflect on their own values. Conversely, others thought the questions were odd and a couple commented that they had to read the question twice to understand what it

was asking. Some suggested that the questions should be asked of them directly, e.g. 'being creative is important to you'.

- A couple of participants were reluctant to answer or skipped the section as they felt it was too personal. NP02 said he did not answer as he was not interested in comparing himself to others. NP05 answered all of the questions but was reluctant to as it felt too much like "a *personality test*".
- One participant (EB01) thought these questions were asking her to reflect on how much she was like her colleagues as the previous section asked about work and colleagues. She skipped all of the human values scale questions as she did not have colleagues so did not think the section was relevant to her. She thought the distinction between the two sections could be made clearer.
- No issues with the gender-neutral language were raised spontaneously. When asked to reflect on the language, the majority of participants thought this was fine, but one (FS01) suggested changing reference to 'their wellbeing' to 'other people's wellbeing' at D74 to make this slightly clearer.

2.5.2.2 Recommendations

- We do not recommend any significant changes based on the findings from the user testing exercise.
- Consider a more prominent demarcation at the top of page 32 to make clear that D63-83 are completely separate from the previous questions about the workplace. Consider adding a title above the Human Values Scale, e.g. 'Things that are important to you' or 'Your values'.

2.5.3 Ancestry question (C58)

2.5.3.1 Main findings

- Some participants said that it was not clear that selecting a second ancestry was optional. One participant (NP03) suggested changing the instruction to "you can choose up to two ancestries" to make this more obvious.
- Some participants seemed to confuse ancestry with nationality at this question:
 - NP01 said that he had understood the question to be asking for the nationality of each of his parents, so he selected 'Bangladeshi' in both columns.
 - Another participant's (NP02) parents were both from India, but he had selected 'Indian' and 'British', reflecting his own nationality. Similarly, NM05 selected 'Jamaican' and 'British', despite both parents being Jamaican.
 - EB05 was unsure whether to answer based on his parents' ancestry, where he grew up, or where he lives now, and ended up selecting one option based on his parents' ancestry and one based on where he lives now.
- Some participants were unsure if they should tick 'English', 'British' or both.
- NM03 ticked one box only, but said that if two were applicable, he would have ticked "what I really think I am" in the first column, suggesting he thought the question was asking people to answer in a certain order.

- There were also inconsistencies in the way participants selected their answers from the two columns: two participants ticked the same option in both column 1 and 2 and seven participants ticked two ancestries in the same column.

2.5.3.2 Recommendations

- Consider changing the instruction to “you may choose up to two ancestries” or “you can choose up to two ancestries” to make it clearer that selecting a second ancestry is optional.
- We recommend removing the second column and allowing participants to tick up to two ancestries in the same column. This may reduce the number of participants who think they have to give a second ancestry and ensure they answer in a consistent way.

2.5.4 Other questions

2.5.4.1 Main findings

- A number of question-specific issues were raised at the end of the interviews. Many of these were minor issues or small comments on particular questions.
- A number of participants said they did not know whether to answer based on their usual circumstances, i.e. before the COVID-19 outbreak, or now, and this distinction would have had a significant bearing on their answers. In particular, a few participants said question A61 (‘How often do you meet socially with friends, relatives or work colleagues?’) was very difficult to answer and should specify the period of time they should be thinking about.
- A review of the returned questionnaires showed that one participant (NP02) had answered about both their adult children at D11-22, and in wrote both of their ages at D12. They then skipped D18-21, possibly because their answers were not the same for both children. Similarly, another participant (EB01) answered about both their parents at D23-39 but was able to complete the remaining questions. These errors likely resulted from missing the instructions above the questions above D11 and D23 respectively.
- FS04 gave the answer ‘I was made redundant / lost my job’ at D96, despite writing that they had last had a job in 2002 at C18. This may have been a result of fatigue, where they had not read the question fully at D96.
- FS01 commented on the amount of text at D91, as she found this “*overwhelming*” to look at. In particular, she thought the length of the end labels of the scales should be condensed.
- One participant (EB03) thought the amount of text at A7 could be simplified to ‘how often do you use the internet’, as the type of device used was unnecessary.
- One of the participants (NP04) works as a private carer for her mother, and for questions D45-D4 about her line manager, she answered about her interactions with her mother. This was felt to be repetitive as she had already answered similar questions about her mother at questions D23-33. There was possible confusion around the definition of line manager: ‘the person you receive work instructions from on a regular basis’.

2.5.4.2 Recommendations

- As stated above, make explicit whether participants should answer about their current or pre-pandemic lives.
- Follow the recommendations for pre-question and post-question routing instructions outlined in Section 2.2.3, to ensure that routing instructions are read for the questions about adult children and parents at D11-22 and D23-39.
- Consider amending the definition of 'line manager' above D45 to make it clearer who should and should not answer this section.
- Consider shortening the end labels at D91: "They placed too much importance on the economy" vs. "They placed too much importance on people's health". Alternatively, consider rephrasing the labels *and* question stem to make both more concise:
 - Please indicate how well you think the UK government balanced protecting the economy and protecting people's health when responding to the coronavirus pandemic.

'Far too much importance on protecting the economy' vs. 'Far too much importance on protecting people's health'.