

2011 CENSUS QUESTION TESTING – THE LANGUAGE QUESTION (2009)

**Jane Eunson
Lorraine Murray
Ipsos MORI Scotland**

On behalf of
General Register Office for Scotland
2009

CONTENTS PAGE

Background, Aims And Methodology	4
Findings	7
Q16 – The language matrix	7
Q17 – How well English is spoken	11
Q18 – language spoken at home	11
Conclusion	12
ANNEXES	13
Annex A: The topic guide	13
Annex B: The questionnaire	17

1. BACKGROUND, AIMS & METHODOLOGY

1.1 Background and aims

As part of the final preparations for the 2011 Census, the General Register Office for Scotland (GROS) commissioned Ipsos MORI Scotland to test draft questions for the language section of the 2011 Scottish Census. More specifically, we were commissioned to test the impact of including Scots (in addition to English and Scottish Gaelic) in the matrix language question. The aim was to test whether respondents' answers to whether they can understand, speak, read or write Scots provide meaningful data, to explore the impact of the inclusion of Scots in the matrix on answers to other language questions and to further explore how respondents interpret Scots.

1.2 Methods

As described in more detail below, the questions were tested among Gaelic speakers and potential Scots speakers, specifically recruited to take part in this project, as well as among those with English as a first language and those with English as a second language, recruited as part of other projects to test the health questions and to test the whole census. This section covers the methods used in testing the questions with those specifically recruited to test the language questions. The methods described are very similar to the methods used in the other projects, however there were some small differences (for example, in the way respondents were recruited). Full details of the methodologies used in the other projects can be obtained in the respective reports.

1.2.1 Cognitive Interviewing

Cognitive interviewing is a widely used approach to critically evaluate survey questionnaires. It allows an investigation into the way target audiences understand, mentally process and respond to survey materials. For example, when a questionnaire is designed it is possible that the author may intend one interpretation of a question but find that respondents presented with the question adopt an alternate understanding. If cognitive interviewing is used successfully in testing questions, survey materials can then be modified to enhance clarity¹.

There are several different techniques that can be used in a cognitive interview. A technique called 'retrospective probing' was deemed to be most appropriate for this study. Retrospective probing involves the interviewer presenting a question to be answered, the respondent answering it and the interviewer following up by probing for specific information relevant to the question or to the specific answer given (e.g. What does this question mean in your own words?). This probing can be done immediately after an individual question is asked or after the respondent has completed all of the questions. Given the short length of the census section used, retrospective probing was done after respondents had completed all questions.

1.2.2 The sample

The sample was not intended to be statistically representative of the Scottish population, but was designed to include people who were more likely to speak Scots

¹ Willis, G.B. (2005). Cognitive Interviewing: A Tool for Improving Questionnaire Design

and people who were Gaelic speakers. Accordingly, we specifically recruited nine potential Scots speakers and four Scottish Gaelic speakers. We also tested the question with 89 other respondents, both those with English as a first language and those with English as a second language, as part of other work to test the health questions and as part of the testing of the whole Census. An additional interview with a Scottish Gaelic speaker was also achieved as part of the testing of the health questions.

The potential Scots speakers were recruited from areas of Scotland (Aberdeenshire, Angus, Ayrshire and Glasgow) where there are dialects of Scots (including 'urban demotic' Scots among young people in Glasgow). Respondents had been born and brought up in the local area and were of C2DE socio-economic status². Gaelic respondents were recruited in Edinburgh and other respondents were recruited in Edinburgh, Glasgow, Aberdeen, Fife, Clackmannanshire, East Ayrshire, and East Dunbartonshire.

1.2.3 Respondent Selection

Recruitment of potential Scots speakers was done using contacts of the research team and through on-street recruitment. The recruitment of Gaelic speakers was done using contacts of the research team and through a Gaelic-medium primary school.

Potential Scots speakers were not informed of the reason that they had been invited to participate. If they knew that they had been recruited on the basis of being 'Scots speakers', this may have changed the way in which they answered the language questions.

Due to the recruitment methods used for Gaelic speakers, it was not possible to recruit them without them knowing that their participation was connected to the fact that they spoke Gaelic.

1.2.4 Fieldwork

In addition to using retrospective verbal probing, respondents were observed while they completed the questions. Points at which they looked puzzled or confused, where they hesitated, where they seemed to be taking care or where they seemed to skim over or ignore instructions or response categories were noted, and respondents were probed accordingly.

To provide context for the questions, respondents also completed other parts of the form. They completed the first page of questions (H1-H5) and all of the 'Individual questions'. It was agreed that H6-H14 would not be completed in order to keep the interviews shorter and because these sections were being tested as part of the whole questionnaire testing. Respondents only completed the 'individual questions' for themselves.

² Socio-economic status is a measure commonly used in social research. It categorises individuals into one of six social grades on the basis of the current or previous occupation of the chief income earner in the household. Groups C2DE correspond to skilled manual and unskilled manual occupations and the economically inactive.

Interviews took place between 31 July and 14 October 2009 and were conducted using a topic guide³ designed by Ipsos MORI in partnership with GROS (attached at Annex A).

All interviews were conducted in respondents' homes or in the Ipsos MORI office in central Edinburgh. On average the interviews lasted around 40 minutes. The discussions were digitally recorded and transcribed. Respondents were given £25 to cover any expenses, to acknowledge that they had given up time to take part, and to encourage participation from a wider range of people.

All interviews were conducted in English. This was on the basis that those with little or no English language ability would tend to delegate the completion of the Census form to someone else to provide substantial help with the interpretation and comprehension of the form in English.

1.2.5 The questionnaire developed for testing

A PDF version of the questionnaire was provided by GROS to Ipsos MORI. The relevant section is included in Annex B.

³ A topic guide is a document that outlines the topics that should be covered in an interview. It often gives examples of the types of questions that might be asked, however, the precise wording used will vary as a result of the interviewer's exchange with the respondent.

2. FINDINGS

2.1 Q16 – the language matrix

2.1.1 The impact on other answers of including Scots in the language matrix

The impact on respondents with English as a second language and non-Scottish people with English as a first language

Although respondents with English as a second language have little understanding of the term Scots (they tend to assume it simply means English with a Scottish accent and perhaps a few Scottish words), this is not impacting on their other answers.

Similarly, it does not affect the answers of non-Scottish people with English as a first language (who also tend to assume Scots simply means English with a Scottish accent and perhaps a few Scottish words).

The impact on Scottish respondents

Including Scots in the language matrix does not generally affect the answers of Scottish respondents. There were, however, two issues that arose.

Firstly, in three cases, respondents who understood Scots to mean simply the version of English spoken by Scottish people, ticked all the Scots options but did not tick any of the English options on this basis (though they then indicated at Q17 that they spoke English 'well' or 'very well').

Therefore, some of the Scottish people who think that Scots simply means the version of English spoken by Scottish people will fail to tick the options under English.

One possible solution might be to assume in data cleaning that anyone who ticks Scots must also understand/speak/read/write (as appropriate) English. However, this assumption that anyone who understands/speaks/reads/writes Scots must also be able to understand/speak/read/write English which may be controversial and may be seen by some as suggesting that Scots is not a distinct language.

The second issue that arose was that someone who interpreted Scots as meaning 'bad/improper English', and felt this applied to them, ticked that they spoke English 'not well' at Q17 on this basis. Others thought that people who speak Scots might tick 'not well' (though this did not apply to them). We do not think this would have happened if Scots had not been included at Q16.

Therefore some Scottish respondents who think Scots means 'bad/improper English', and who think this applies to themselves, may tick 'not well' at Q17 when they might otherwise tick 'very well' or (more likely) 'well'.

In terms of finding a solution, this is perhaps harder to correct through data cleaning and would involve considering a combination of answers to distinguish these respondents from those for whom English is a second language and genuinely do not speak English well. (Perhaps those who tick Scots at Q16, and do not claim to

speak any other language other than English or Scots at Q18, and who answer Scottish and/or British only at Q14 and Q15).

The responses of the five Gaelic speakers were unaffected by the inclusion of Scots at Q16.

2.1.2 The impact on data on Scots

Understanding of the term Scots

Firstly, and as found in previous testing, understanding of what is meant by ‘Scots’ is very varied and there is considerable confusion about the meaning of the term. There is a range of interpretations from those who think Scots is simply the version of English spoken by Scottish people (essentially English with a Scottish accent), to those who think it is a dialect spoken in particular parts of Scotland, to those who feel it is an old language that hardly anyone speaks any more. A further interpretation which emerged in this round of testing was that Scots means ‘bad English’ or ‘not proper English’ which people should be encouraged not to speak.

“Scots is only what I would consider to be English with a Scottish accent.

Who do you think would tick Scots then?

Everybody from Scotland.” (Scottish man)

“I wouldn’t know where to put the difference between Scottish dialects of English and Scots.” (Scottish woman)

“Well a kind of old Scots as in Burns and poems and old Scots words [...] Who do I consider to speak Scots? Well probably somebody from about 250 years ago.” (Scottish man)

“I wouldn’t have said you had to put Scottish there [...] Scots I would say would be English or slang.” (Scottish woman)

“What did you think was meant by the term Scots?

I don’t know, dialect?, but surely that’s just English or people who are just kind of ignorant to the English language, to make up their own words. I don’t have any idea what Scots could possibly mean.” (Scottish man)

As we found in previous testing, the confusion about the meaning of the term and the range of interpretations which are applied will lead to inconsistencies in response (e.g. people who speak in very similar ways will respond differently). The number of Scots speakers will either be overestimated or underestimated, depending on which interpretation data users apply. Thus, the question will not yield any meaningful data on Scots.

We cannot see any way to solve this problem in the context of the Census. Data users should be very clear about the severe limitations of the data collected on Scots.

Secondly, many respondents spontaneously asked ‘what does Scots mean?’ (much more than in the previous version because they were being forced to consider whether they ‘understood’, ‘spoke’ it etc.) and were therefore unsure whether they had answered the question correctly. However, this did not concern them too much and most said they would just guess at the meaning or leave Scots blank. While most said they would not phone the helpline, there could be an increase in the number of calls.

“I wasn’t sure if I should [tick the Scots options] or not so I left blank.”
(New Zealand man)

“I was a little bit confused about what it means by Scots.”

“What do you think it might mean?”

Well originally I thought it was Scottish Gaelic but obviously that’s an option there, so I’m not sure what it means, I imagine slang. But even though I thought that, I wasn’t sure if that was correct, so I just left it blank.” (Scottish man)

Thirdly, some English respondents and English as a second language respondents, who thought Scots meant simply the version of English spoken by Scottish people (a Scottish accent and possibly a few dialect words), said that they could ‘understand’ and ‘read’ Scots, on the basis that they could understand people with a Scottish accent or could understand some Scottish words (‘piece’ for sandwich, ‘tattie’ for potato, and ‘stay’ for live were given as examples).

In testing the previous version (Scots just included at Q18), we found no instances of people who *clearly* did not speak Scots indicating that they did. However, our testing of this version suggests that there will be more cases of people who (by most definitions) do not understand or read Scots indicating that they can.

“When I came to live here, I came here one year or two year ago, I found it very difficult to understand the Scottish accent, but now I understand it very well.” (Sudanese woman)

2.1.3 The impact on the acceptability of the questionnaire

As noted above, many people spontaneously asked ‘what does Scots mean?’ when they reached Q16. In both the testing for the language and health questions, and the testing of the whole questionnaire, there were more spontaneous questions about the term Scots than about any other terms or questions in the Census. This is likely to lead to an increase in calls to the Census Helpline (and Helpline staff should be given guidance on how to respond to such queries). However, reactions in testing suggest that most people would not bother to contact the Helpline on this issue but would either guess at the meaning of the term (and tick the boxes or not depending on what they guessed the term meant) or leave the Scots boxes blank because they were not sure what it meant (effectively answering ‘no’ from a data collection point of

view). Although many people queried the term and were unsure whether they had answered correctly, it was not something that caused them much concern.

This lack of concern was linked, in part, to a view among many respondents that the inclusion of Scots was a bit 'silly'. In addition to there being more queries about the meaning of the term than any other term, there were more comments that it was a 'stupid question' than any other question. However, it should be noted that these respondents did not feel *strongly* that it should not be included: it was a minor issue which had no significant effect on their overall perceptions of the Census.

“What was that chuckle about Scots for?”

Well it could mean anything, it's just a nonsense question.” (Scottish woman)

“I thought that [Scots] was not very clear, so that's why I didn't know how to answer it [...] It didn't particularly bother me in that it's not a quantitative question that I'm misleading anyone, in the same way as I would be if I hadn't put the correct address, or in counting people I shouldn't be counting or something like that, but it bothered me because I didn't know what they were trying to get at [...] If I'm being perfectly honest, I don't think even at that I would have bothered to [phone the Helpline]. I think part of me would have thought, if they can't explain clearer what they are wanting to know, if they can't make it clear and I can't understand it, I wouldn't have gone to those lengths...” (English woman)

Although these problems emerged in previous testing (when Scots was included only at Q18), they are exacerbated by the inclusion of Scots in the Q16 matrix because the term is given more prominence and the format of the question forces all respondents to consider their abilities in Scots.

2.1.4 The 'none of these' box in the language matrix

There is a 'none of these' box in the language matrix which could apply to, for example, babies, people with learning disabilities or people from another country who do not speak/read/write/understand English (or Scottish Gaelic or Scots).

We found in previous testing that many people did not notice the box or did not know what it was for and so did not tick it. While we are aware of the general principle that there should be something to tick at every question to provide information that people have answered it rather than missed it out, it is not serving that purpose effectively. Therefore, as a potential way of saving space in the questionnaire, although not ideal, the 'none of these' box in the language matrix could be removed.

2.1.5 The question format

In previous rounds of testing the language matrix contained two columns (English and Scottish Gaelic). There were concerns that changing the layout of the question to include three columns could cause respondents problems in understanding how to answer the question. However, our testing showed that this was not the case and that the question was straightforward for respondents to complete.

2.2 Q17 – How well English is spoken

This question is working as before. On the basis that the distinction between ‘very well’ and ‘well’ does not matter, it is working well. As mentioned above, there is a possibility that the inclusion of Scots in the language matrix will mean that some of those who think Scots is ‘bad/improper English’, and who think this applies to themselves, may tick ‘not well’ at Q17 when they might otherwise tick ‘very well’ or (more likely) ‘well’.

2.3 Q18 – Languages spoken at home

This question is working well. As in previous testing, respondents generally interpreted using a language ‘at home’ as meaning the language(s) they would speak with other people in their household. Some thought that this type of question was primarily aimed at those who had moved to Scotland from elsewhere and might therefore speak a different language in the household.

There are no other languages that people felt strongly about including as response options. This includes the five Gaelic speakers interviewed who did not feel strongly that Gaelic should be included. A few people mentioned that it was inconsistent to have Gaelic and Scots included at Q16 but not at Q18 but, again, they did not feel strongly that these languages should be included at Q16.

All respondents, who indicated at Q16 that they speak Scots, said at Q18 that they speak English only. No-one wrote in Scots or suggested that Scots should be included as an option. This appeared to be because they did not see it as another language so it did not cross their minds to write it in.

“Was there a reason you didn’t write it in?”

No I just it says here ‘no, English only’. Others please write in, I wouldn’t have bothered.

Why would you not have bothered putting Scots in there?

To put in, I would assume to be a language like Arabic, Polish or whatever.” (Scottish man)

“Did you think about ticking, yes Scots?”

No, as I said here, Scots is only English with a different accent.” (Scottish man)

2.4 Conclusion

Regardless of whether Scots is included in Q16 or Q18, confusion about the meaning of the term and the range of interpretations which are applied will lead to inconsistencies in response (e.g. people who speak in very similar ways will respond differently). The number of Scots speakers will either be overestimated or underestimated, depending on which interpretation data users apply. The question will not yield any meaningful data on Scots and potential data users should be made aware of this. We cannot see any way to solve this problem in the context of the Census.

While the inclusion of Scots in the Q16 matrix does not appear to affect the *interpretation* of the term, its prominence will lead to more people ticking Scots than if it is included at Q18. There will also be more cases of people who (by most definitions) do NOT understand Scots indicating that they do. However, as noted above, in neither case (inclusion at Q16 or Q18) will the data be meaningful.

In relation to the impact on other data and on the acceptability of the Census questionnaire, the inclusion of Scots at Q16 is more problematic than the inclusion of Scots at Q18. To summarise, the problems with its inclusion at Q16 are:

- some people will tick Scots and not English at Q16. This could be relatively easily addressed by data cleaning but only if it is assumed that anyone who understands/speaks/reads/writes Scots MUST be able to do the same in English – which may be controversial and may be seen by some as suggesting that Scots is not a distinct language
- some of the Scottish respondents who think Scots means ‘bad/improper English’, and who think this applies to themselves, may tick ‘not well’ at Q17 when they might otherwise tick ‘very well’ or (more likely) ‘well’. This could be not be completely solved by data cleaning
- there will be more calls to the Helpline from people unsure of the meaning of the term Scots
- some people feel that the inclusion of Scots is ‘silly’ and that it is a ‘stupid question’. This is exacerbated by its inclusion at Q16. However, it should be noted that these respondents do not feel *strongly* that it should not be included: it is a minor issue which has no significant effect on their overall perceptions of the Census.

On balance therefore, we would suggest that Scots should be included at Q18 rather than at Q16 because it is less problematic. However, if Scots were to be included in Q16, while problems with the responses to Q17 and Q18 may result, these would be unlikely to cause significant difficulties. There would, however, be an increased burden on the data cleaning operation.

ANNEX A: THE TOPIC GUIDE

Cognitive question testing the language and health questions

Topic Guide – FINAL VERSION

Introduction

Introduce self, Ipsos MORI

If you have a colleague with you, explain that they are here to observe you.

Research commissioned by the General Register Office for Scotland – the organisation that runs the Census – which involves talking to members of the public to understand how they would answer revised possible questions for Scotland's 2011 Census and other Scottish official statistics.

Information about importance of Census (e.g. it is used by government, health authorities and many other organisations to allocate resources, tackle discrimination and plan services for everyone.)

Thank participants for agreeing to be interviewed; mention should take around 45 minutes to an hour.

Anonymity of respondents and MRS (Market Research Society) code of conduct

Permission to record, explain how it will be used.

Completing the Census questions

We would like you to complete this form. I'd like you to work through it as if this is the day of the Census and you have just received this form through your door. Please complete the form as you would if I was not here. Just work at your own pace. This is not a test.

Respondent completes form for themselves up to and including Q21. If recruited as parent/carer of someone with a long term health condition, they complete form for themselves (up to Q21) and for the household member with the condition (up to Q21).

In general throughout the questionnaire, note any comments made expressions or body language.

Q16 Language matrix

How did you find this question?

How easy or difficult is it to answer?

How did you decide which boxes to tick?

What languages do you think should be included in this question?

If no Scots options ticked:

Did you consider ticking any of the Scots boxes?

Why did you not tick any?

What did you think was meant by the term 'Scots'?

Who do you consider to speak Scots? [*Probe: Anyone they know? People from a certain area?*]

If ticked any Scots options:

Why did you select this/these option(s)?

How sure were you about ticking the Scots box(es)?

What did you think was meant by the term 'Scots'?

Who do you consider to speak Scots? [*Probe: Anyone they know? People from a certain area?*]

If ticked some but not all Scots options:

Why did you not tick [speak/read/write]?

Q17 How well can you speak English?

How did you find this question?

How easy or difficult was it to answer?

How did you decide which box to tick?

What did you think of the response options?

What level of English do you think someone would have who ticked “not well”?

Q18 Do you use a language other than English at home?

How did you find this question?

How easy or difficult was it to answer?

What came into your mind when you were asked this question?

What did you think was meant by ‘at home’?

Did you notice you could tick more than one box?

What did you think of the choice of response options?

Do you think any other languages should be included in the options?

If ticked English Only

Did you consider writing any other language in “other”?

Do you speak any other languages at all? [*Probe: Why did you choose not to include this? To what level do you speak this language?*]

If ticked other

When do you speak [language(s)]?

Did you consider including more than one language in the write-in box?

Do you speak any other languages at all? [*Probe: Why did you choose not to include this? To what level do you speak this language?*]

If ticked Scottish Gaelic at Q16 but did not write in at Q18

Did you think about writing in Gaelic at “other”? [*Probe: Why didn’t you?*]

If ticked Scots at Q16 but did not write in at Q18

Did you think about writing in Scots at “other”? [*Probe: Why didn’t you?*]

THANK RESPONDENT

Is there anything else that you would like to say, that hasn’t already been mentioned?

We would like to thank you for taking part and remind you that the findings will be used by the Scottish Government to improve wording of the next Census.

