



Final recommended questions for the 2011 Census in England and Wales

Language

October 2009

Table of Contents

1. Summary	4
2. Introduction	7
3. Background	8
3.1 Differing questions and terminology	8
3.2 Welsh language question	8
3.3 Main language and English proficiency questions	8
4. User requirements	9
4.1 User requirement for information on language	9
4.2 User preferences for language concepts	10
5. Methods of question testing	12
5.1 Qualitative testing	12
5.1.1 Cognitive testing	12
5.2 Quantitative testing	12
5.2.1 Lambeth Postal Test	12
5.2.2 2007 Postal Test	13
5.2.3 2007 Census Test	13
5.2.4 2008 Postal Test	13
5.2.5 March 2009 Postal Test	13
5.2.6 July 2009 Postal Test	13
5.2.7 Opinions (Omnibus) survey question testing	13
6. Development of the question	15
6.1 Knowledge questions – which languages can you understand, speak, read or write?	15
6.1.1 Sectional layout	15
6.1.2 Full matrix style	17
6.1.2.1 Full matrix, version 1	17
6.1.2.2 Full matrix, version 2	18
6.1.2.3 Full matrix, version 3	19
6.1.2.4 Full matrix, version 4	20
6.1.2.5 Full matrix, version 5	22
6.1.3 Simplified matrix style	22
6.1.3.1 Simplified matrix, version 1	23
6.1.3.2 Simplified matrix, version 2	24
6.1.3.3 Simplified matrix, version 3	26
6.1.4 General comments on a matrix style language question	27
6.2 New Zealand style conversation questions	28
6.2.1 Conversation question, version 1	28
6.2.2 Conversation question, version 2	29
6.3 US style, two-part question	31
6.3.1 Part 1 primary language	31
6.3.1.1 Primary language, version 1: ‘Is a language other than English used every day in your household?’	32
6.3.1.2 Primary language, version 2: ‘Do you speak a language other than English or Welsh at home?’	33
6.3.1.3 Primary language, version 3: ‘What is your main language?’	34
6.3.2 Part 2 English language proficiency	37
6.3.3 Conclusions for England and Wales language question	40
6.4 Questions for Wales only	42
6.4.1 Knowledge of Welsh– can you understand, speak, read or write Welsh?	42

6.4.1.1	Collecting knowledge of Welsh data in England	44
6.4.2	Frequency of speaking Welsh	44
7.	Recommended location of the question.....	46
8.	Final recommendations.....	47
9.	References.....	49
Annex A:	Summary of suitability of language questions: by type of question	51

1. Summary

Question development for the 2011 Census began in 2005. An iterative and comprehensive process of user consultation, evaluation and prioritisation of user requirements, and qualitative and quantitative question testing has been carried out to inform decisions on the questions to be included in the 2011 Census. This paper presents the final language question recommendations for the 2011 Census in England and Wales.

A question on knowledge of Welsh has been included in every census in Wales since 1891 and it is recommended that this question continues to be asked in Wales in 2011. Questions about languages other than Welsh have never been asked in an England and Wales census.

Consultations with key users of census data in 2005 and 2007 revealed strong demand for the inclusion of questions on language. Overall, 92 per cent of respondents to the 2007 consultation on ethnicity, identity, language and religion expressed a requirement for information on language. Respondents to the consultations recognised a number of ways in which information on language would be used, including:

- to provide services in appropriate languages where English is not understood, including English for speakers of other languages (ESOL) provision
- to monitor the uptake of minority national languages
- to monitor the use of British Sign Language
- to gain a better understanding of the ethno-religious diversity of the UK population
- to understand the impact of English (or Welsh) language ability on employment and other social inclusion indicators
- to identify linguistic resources in the UK for business reasons
- to plan and monitor Welsh language provision in Wales.

A variety of quantitative and qualitative methodologies have been employed to test the questions recommended for inclusion in the 2011 Census. Qualitative question testing on language has been conducted by the data collection methodology (DCM) branch in ONS through a programme of cognitive testing running since February 2005. Quantitative question testing has been conducted at various intervals since June 2006 using a variety of methodologies.

A number of different styles of language question have been considered. Initially, a matrix format was considered, including Welsh, English and other languages in the same question. However, testing suggested the question was difficult for respondents to understand. Separate questions asking about main language and spoken English proficiency were subsequently developed to meet the greatest possible number of user requirements within space constraints. These questions performed well in testing and this paper therefore includes recommendations for the following three questions on language:

- Knowledge of Welsh (asked in Wales only)
- Main language
- Spoken English proficiency

The questions presented below are those for England and Wales that are recommended for the 2011 Census, subject to Parliamentary approval which ONS hopes to attain by the end of 2009.

Figure 1.1: Final recommended language questions for England

17 This question is intentionally left blank ➔ Go to **18**

18 What is your main language?

English ➔ Go to **20**

Other, write in (including British Sign Language)

19 How well can you speak English?

Very well Well Not well Not at all

Figure 1.2: Final recommended English language version of language questions for Wales

17 Can you understand, speak, read or write Welsh?

➔ Tick all that apply

Understand spoken Welsh

Speak Welsh

Read Welsh

Write Welsh

None of the above

18 What is your main language?

English or Welsh ➔ Go to **20**

Other, write in (including British Sign Language)

19 How well can you speak English?

Very well Well Not well Not at all

Figure 1.3: Final recommended Welsh language version of language questions for Wales

17 A allwch ddeall, siarad, darllen neu ysgrifennu Cymraeg?
➔ Ticiwch bob blwch sy'n berthnasol

Deall Cymraeg llafar
 Siarad Cymraeg
 Darllen Cymraeg
 Ysgrifennu Cymraeg
 Dim un o'r uchod

18 Beth yw'ch prif iaith?
 Cymraeg neu Saesneg ➔ Ewch i **20**
 Arall, nodwch (gan gynnwys iaith Arwyddion Prydain)

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

19 Pa mor dda allwch chi siarad Saesneg?
Da iawn Da Ddim yn dda Dim o gwbl

2. Introduction

The next census will take place on 27 March 2011. The proposed topics to be included were announced in the 2011 Census White Paper published in December 2008. To access the White Paper and read further information about how the content of the 2011 Census was determined, please refer to the ONS website at:

<http://www.ons.gov.uk/census/2011-census/2011-census-questionnaire-content/question-and-content-recommendations-for-2011/index.html>

Question development for the 2011 Census began in 2005. An iterative and comprehensive process of user consultation, evaluation and prioritisation of user requirements, and question testing has been carried out to inform the decisions on the questions to be included in the 2011 Census. The questions for England and Wales have been developed with the Welsh Assembly Government and through close collaboration with the statistical offices of Scotland and Northern Ireland. A key aim of this collaboration is to minimise differences between questionnaires across the UK, though it is recognised that differing circumstances will sometimes require different solutions. A paper on the UK differences in census questions will shortly be available on the ONS website.

This paper outlines the development of the language questions. There are equivalent papers which present the recommendations for each of the other questions within the ethnicity, [national] identity, language and religion (EILR) topic area. Questions on the following topics have been recommended for the 2011 Census:

- National identity
- Ethnic group
- Knowledge of Welsh language (Wales only)
- Main language
- Spoken English proficiency
- Religious affiliation

These questions complement the suite of migration questions that have been recommended:

- Country of birth
- Usual address one year ago
- Month and year of arrival to the UK
- Intended length of stay in the UK
- Passports held (as a proxy for citizenship)

The paper will begin by outlining user requirements for information on language in England and Wales, before explaining the testing conducted in developing the recommended questions. The question development process is then explained in detail and the final language question recommendations for the 2011 Census are presented. These recommendations are subject to Parliamentary approval which ONS hopes to obtain by the end of 2009.

3. Background

3.1 Differing questions and terminology

There are many aspects of language use and impact that can be captured in a questionnaire. This paper uses a number of terms to refer to different aspects of language:

- **language ability:** Generic term referring to all aspects of language
- **language knowledge/skill:** Whether a person can understand/speak/read/write a language without any reference to how well they can do it
- **language proficiency:** How well a person understands/speaks/reads/writes a language
- **frequency:** How often a person uses a language

3.2 Welsh language question

A question on knowledge of Welsh has been included in every census in Wales since 1891. The knowledge of Welsh question collects information on the ability of respondents to understand spoken Welsh, speak Welsh, read Welsh and write Welsh.

The *Welsh Language Act 1993* places a duty on public organisations in Wales to treat both the Welsh and English languages equally. The main requirement for collecting data on Welsh knowledge is to inform policy development and for monitoring purposes. Although there is some user requirement to collect information on Welsh speakers in England, this question has not been asked in England as there is also strong user requirement for information on a wide range of languages and singling out Welsh would therefore be unsatisfactory to most users.

3.3 Main language and English proficiency questions

Questions about English and other languages have never been asked in an England and Wales census, though they have been included in other national surveys.

Since 2001 ONS has carried out a programme of research, consultation, testing and analysis to establish the most suitable questions for the 2011 Census in England and Wales. This has included consultation with data users and providers; qualitative and quantitative question testing; carrying out an equality impact assessment; and working with Scotland and Northern Ireland to harmonise the questions where possible.

4. User requirements

The inclusion of questions in the 2011 Census must be supported by a clear user requirement for the information. The user consultation process began in 2005 with a formal three month general topics consultation in which there were around 80 responses received commenting on the subject of language, from a variety of central government, local authority and other data users. A summary of the comments relating to language can be found at

<http://www.ons.gov.uk/about/consultations/closed-consultations/consultation-on-2011-census---responses/index.html>

The Welsh Assembly Government produced a summary of 2005 consultation responses from Welsh organisations or referring to 'Wales specific' topics. The summary is available at <http://wales.gov.uk/topics/statistics/headlines/pop-2007/pop-2006/hdw200603081/?skip=1&lang=en>

An additional detailed consultation exercise was carried out between December 2006 and March 2007 (referred to throughout the paper as the 2007 consultation) to refine the user requirements for information on ethnicity, identity, language and religion from the 2011 Census. The report of the 2007 consultation can be found at

<http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

The summary of Welsh organisations and 'Wales specific' topics from the 2007 consultation can be found at

<http://wales.gov.uk/topics/statistics/headlines/pop-2007/hdw200710312/?lang=en>

4.1 User requirement for information on language

In May 2005 ONS published a consultation document '*The 2011 Census: Initial view on content for England and Wales*'. The consultation document proposed that a question on knowledge of Welsh language should continue to be included in Wales in the 2011 Census, but that there was insufficient evidence of user demand to justify inclusion of English language proficiency. However, the responses to the consultation showed strong user demand for additional questions on language. The 2007 consultation confirmed this demand, with 92 per cent of respondents declaring a requirement for language information. As table 4.1 shows, the greatest requirement for language information came from local and regional government (96 per cent).

Table 4.1 2007 consultation respondents requiring information on language

Organisation type	Yes (%)	No (%)	Total Respondents
Central and devolved government	86	14	21
Experts, community and special interest groups	88	12	121
Local and regional government	96	4	84
Local service providers	93	7	45
All respondents	92	8	271

Respondents to the consultations identified a number of reasons for requiring a question on language, including:

- **to provide services in appropriate languages where English is not understood, including English for speakers of other languages (ESOL) provision:** By far the most common interest in language is to facilitate the provision of public services, both by identifying the need for translation and interpretation services in the short term and for providing English language lessons in the medium term
-

- **to monitor the uptake of minority national languages:** Some respondents required information on the knowledge of UK languages other than English, including Welsh outside of Wales, Cornish, Gaelic and Scots, for which there is an interest in increasing uptake. Information on regional and minority languages is also required to monitor their use for protection and promotion under the European Charter for Regional or Minority Languages
- **to monitor the use of British Sign Language:** Several requests were received for the inclusion of British Sign Language within the language question.
- **to gain a better understanding of the ethno-religious diversity of the UK population:** Some local and regional authority respondents thought that the language question would enhance the data already captured in the ethnic group and religious affiliation questions.
- **to understand the impact of English (or Welsh) language ability on employment and other social inclusion indicators:** Including a question on language proficiency in the census would be useful as it would cover the whole population, including people whose first language is English or Welsh. The information gained could be used to analyse the impact of language ability on employment and other social inclusion indicators
- **to identify linguistic resources in the UK for business reasons:** Some respondents required information about languages spoken in general
- **to plan and monitor Welsh provision in Wales:** The majority (85 per cent) of respondents from Wales needed general language information from the 2011 Census. Although not asked for specific details of their requirements for a Welsh knowledge question, most of these also stated that they required Welsh language data to comply with the *Welsh Language Act 1993*, which places a duty on public organisations in Wales to treat both the Welsh and English languages equally. There was also a strong need from the Welsh Assembly Government and Welsh users for the question to be comparable with the 2001 Welsh language question.

4.2 User preferences for language concepts

There is not enough space to record all aspects of language ability in the census. Thus there was an initial need to be clear about the precise aspect(s) of language ability that were required by census users so that the relevant language information could be collected.

The 2007 consultation document provided a list of 19 different aspects of ability that weren't covered in the 2001 Census. To explore the need for a new question on language, data users were asked to indicate whether they required this information and if so, to what extent (table 4.2).

Table 4.2: 2007 consultation respondents (English and Welsh data users) requiring information on aspects of language ability

Organisation type	Very important (%)	Quite important (%)	Do not need (%)
Ability to speak English	78	15	7
Ability to understand spoken English	77	15	8
Mother tongue or first language	75	20	5
Main language (spoken at home)	75	19	6
Preferred spoken language for communicating with public authorities	74	17	9
Preferred written language for communicating with public authorities	72	18	10
Ability to read English	72	19	9
Which languages are spoken, other than English	70	25	6
Ability to write in English	65	25	10
Which languages are understood, other than English	63	30	8

Which languages are read, other than English	60	31	8
Which languages are written, other than English	50	37	12
Frequency of speaking other languages	45	29	27
Other aspects of language ability	27	6	66
Frequency of speaking Welsh	13	13	74
Ability to speak Welsh (among population in England)	11	14	75
Ability to understand spoken Welsh (among population in England)	10	16	74
Ability to write Welsh (among population in England)	10	13	77
Ability to read Welsh (among population in England)	10	14	76

In addition to this information about importance, ONS also received more detailed responses about user need. The preferences expressed by data users allowed ONS to construct a number of essential and desirable requirements of a new language question:

Essential that the question

- is able to capture spoken languages
- is able to capture sign languages
- allows data users to understand whether respondents had English language difficulties and if so what alternative languages services should be provided in
- consistently excludes people with low English language proficiency when recording English speakers
- is generally clear to respondents and is understood as intended
- does not impact on the knowledge of Welsh language question

Desirable that the question

- consistently excludes people with low proficiency in other languages when capturing languages other than English
 - captures proficiency in English
 - captures proficiency in other languages
 - allows users to distinguish languages used for translation/ learnt at school, etc
 - simultaneously captures multiple languages to gain the best understanding of ethno-linguistic diversity and resource
 - captures where and how languages are used
 - captures written language knowledge
 - captures knowledge of understanding of spoken language
 - capture knowledge of reading languages
 - identifies respondents' main or preferred language
-

5. Methods of question testing

A variety of quantitative and qualitative methodologies have been employed to test the questions recommended for inclusion in the 2011 Census.

5.1 Qualitative testing

Qualitative testing on language has been conducted by the data collection methodology (DCM) branch in ONS through a programme of cognitive testing which began in February 2005.

5.1.1 Cognitive testing

The aim of this testing was to devise questions that collect accurate and meaningful information that meets user requirements, minimises the burden on respondents and are designed to conform as closely as possible to best practice of questionnaire design. The testing was split into four main phases:

- Pre-testing for the 2007 Census Test (February 2005 – April 2006)
- Whole Questionnaire Testing for the 2007 Census Test (June 2006 – August 2006)
- Testing for the 2009 Census Rehearsal
 - **Wave 1** (November 2006 to January 2007)
 - **Wave 2** (April to May 2007)
 - **Wave 3** (July to September 2007)
 - **Wave 4** (October and November 2007)
 - **Wave 5** (January to March 2008)
 - **Wave 6** (April to July 2008)
 - **Welsh language testing, Wave 1** (October to November 2007)
 - **Welsh language testing, Wave 2** (June and July 2008)
- Testing for 2011 Census
 - **Wave 7** (March and April 2009)
 - **Whole questionnaire testing – English language** (June and July 2009)
 - **Whole questionnaire testing – Welsh language** (June and July 2009)

5.2 Quantitative testing

Analysis on the performance of the language question has been conducted for several quantitative tests since June 2006.

5.2.1 Lambeth Postal Test

From June to July 2006, a test of postal enumeration procedures was carried out in the London Borough of Lambeth. Although this was designed to test aspects of the field operation for the 2011 Census, it allowed the opportunity to analyse and evaluate the performance of the questions. A response rate of 25 per cent was obtained and 366 household questionnaires were returned, giving 787 individual respondents. Some of this data was analysed to help inform the development of the language questions prior to the 2007 Census Test.

5.2.2 2007 Postal Test

The ONS questionnaire design and content team ran a postal test of 10,400 households across England in April 2007. Although the main objective of this test was aimed at testing issues related to questionnaire length, it also provided valuable information on the acceptability and understanding of definitions and questions, particularly new questions such as language.

5.2.3 2007 Census Test

A large scale census test was carried out covering 100,000 households in England and Wales on 13 May 2007. The test took place in Bath and North East Somerset, Camden, Carmarthenshire, Liverpool and Stoke-on-Trent. This allowed the opportunity to test new and updated questions.

A detailed evaluation of the 2007 test can be found on the ONS website at:
<http://www.ons.gov.uk/census/2011-census/2011-census-project/2007-test/2007-test-questionnaire-evaluation.pdf>

5.2.4 2008 Postal Test

In July 2008 two postal surveys were carried out, each with a sample of 10,200 households, one across England, and the other in Northampton. The main purpose was to test any impact on response rates of including a question on intended length of stay in the UK.

The test also allowed valuable analysis to be conducted on the performance of the questions developed for the 2009 Census Rehearsal.

5.2.5 March 2009 Postal Test

In March 2009, a postal test was conducted with the primary aim of assisting the development of questions relating to the student population. 20,000 questionnaires were sent to three separate sample areas. 10,000 were sent to a random sample of households in England, of which 5,000 were sent to Durham city and 5,000 were sent to an area of central Norwich. The latter two areas were chosen because they are known to have a high concentration of students, along with a number of other characteristics desirable for testing.

5.2.6 July 2009 Postal Test

In July 2009, a postal test was conducted with the primary aim of testing the questions relating to ethnicity and identity. 27,000 questionnaires were sent to areas of England selected for characteristics that were desirable for the purposes of the test.

5.2.7 Opinions (Omnibus) survey question testing

The Opinions (Omnibus) survey is an ONS run, multi-purpose survey based on interviews with a monthly sample of around 1,200 adults (aged 16 and over) in private households. It currently forms part of the Integrated Household Survey (IHS). One adult is selected from each household to answer the questions. It differs from the census in that interviews are carried out face-to-face by members of the general field force of interviewers in ONS.

As part of the testing process, the questions 'What is your main language?', 'Which of these is true of your main language?' and 'How well do you speak English?' were all included in May and June 2008, as shown in the figures below.

Figure 5.1: Language question on Opinions Survey May-June 2008 (England)

What is your main language? 1. English 2. Other (including sign languages) (please specify)

Figure 5.2: Language question on Opinions Survey May-June 2008 (Wales)

What is your main language? 1. English or Welsh 2. Other (including sign languages) (please specify)
--

Figure 5.3: English proficiency question on Opinions Survey May-June 2008

How well do you speak English? 1. Very well 2. Well 3. Not well 4. Not at all

Figure 5.4: Supplementary language question on Opinions Survey May-June 2008

Which of these is true of your main language? 1. It is the first language I could ever speak 2. I use it most of the time 3. I speak it when I am at home 4. I speak it with my family 5. I speak it with my friends 6. I speak it at school or work 7. I use it when dealing with public authorities
--

6. Development of the question

This section provides a detailed description of how the questions evolved from the beginning of testing through to the finalisation of the questions. It uses evidence gathered from all the strands of research described in the previous sections.

ONS tested the following language questions:

Knowledge questions (Section 6.1)

- What language(s) can you understand, speak, read or write?
- Which of these can you do? (followed by a list of language skills)

Conversation questions (Section 6.2)

- In which language(s) could you have a conversation about a lot of everyday things?
- In which language(s) can you have a conversation about a lot of everyday things?

Two-part questions (Section 6.3)

- Is a language other than English used every day in your household?
- Do you speak a language other than English or Welsh at home?
- What is your main language?
- How well do you speak English?
- How well can you speak English?

Questions for Wales only (Section 6.4)

- Can you understand, speak, read, or write Welsh?
- Do you speak Welsh daily, weekly, less often or never?

6.1 Knowledge questions – which languages can you understand, speak, read or write?

The first series of questions to be tested were based on the knowledge of Welsh questions included in previous censuses and asked ‘Which languages can you understand speak, read or write?’. The response options provided in the question included those for English, Welsh, British Sign Language (BSL) and ‘Other’ languages. All tested versions of the question have carried an instruction to ‘tick all boxes that apply’.

ONS tested a variety of response layouts, which developed iteratively through cognitive testing.

6.1.1 Sectional layout

Initially ONS laid out the response options in a sectional design, arranged in a vertical order on the page. The question consisted of separate parts or sections listed in the order of ‘English’, ‘Welsh’, ‘British Sign Language’ and ‘Other’ language(s). After each section header there were separate tick-boxes for reporting knowledge to understand, speak, read or write that language.

The first question tested allowed just one ‘Other’ language to be reported for each skill (understanding, speaking, reading and writing) (Figure 6.1). ONS then changed the layout to allow an additional two languages, and related abilities in each, to be included (Figure 6.2).

Figure 6.1: Sectional layout language question version 1

4 Which languages can you understand, speak, read or write?

◆ ✓ all boxes that apply.

English

Understand spoken English

Speak English

Read English

Write English

Welsh

Understand spoken Welsh

Speak Welsh

Read Welsh

Write Welsh

British Sign Language (BSL)

Understand spoken BSL

Sign BSL

Other Languages

Understand another language, please write in

Speak another language, please write in

Read another language, please write in

Write another language, please write in

Figure 6.2: Sectional layout language question version 2

Which languages can you understand, speak, read or write?

◆ please ✓ all boxes that apply

English

Understand spoken English

Speak English

Read English

Write English

Welsh

Understand spoken Welsh

Speak Welsh

Read Welsh

Write Welsh

British Sign Language (BSL)

Understand signed BSL

Sign BSL

Other language, please write in

Understand spoken other language

Speak other language

Read other language

Write other language

Other language, please write in

Understand spoken other language

Speak other language

Read other language

Write other language

Other language, please write in

Understand spoken other language

Speak other language

Read other language

Write other language

An advantage of the second version of this question was that it allowed multiple languages (and concepts) to be captured independently, allowing for complex analysis. Version 1 made it easy for people to indicate if they speak and read different languages (for example Punjabi is spoken but written mostly in either Gurmukhī or Shahmukhi). However, in the majority of cases where respondents wish to record an additional language, the format could prove burdensome if the same language has to be written out four times.

In preliminary cognitive testing, respondents generally interpreted the question to be asking about native and learnt languages. When answering the question respondents considered their first language(s), and also languages they had learnt at a later time, perhaps at school.

One disadvantage identified was that respondents found the question quite complex as it asked for a lot of separate information. The question was also far too long for inclusion in the census questionnaire.

6.1.2 Full matrix style

To make the language question fit into the available space ONS redesigned the response options into a matrix style.

6.1.2.1 Full matrix, version 1

The first full matrix version of the question consisted of four parts. It asked respondents about their knowledge of English, Welsh and British Sign Language, and then their knowledge of an ‘Other’ language of their choice. The tick-boxes relating to knowledge ran horizontally from left to right reading ‘Understand’, ‘Speak/Sign’, ‘Read’,¹ ‘Write’,¹ and ‘No Ability’,¹ (Figure 6.3).

Figure 6.3: Full matrix language question version 1

13 What languages can you understand, speak, read or write?
 ♦ Tick all boxes that apply

	Understand	Speak /Sign	Read	Write	No Ability
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Welsh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
British Sign Language	<input type="checkbox"/>	<input type="checkbox"/>			
Other language, please write in	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>				
	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>				
	Understand	Speak /Sign	Read	Write	No Ability
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Some respondents, despite ticking that they could ‘Understand’, ‘Read’ and ‘Write’ English, did not tick that they could ‘Speak/Sign’ it. Generally, if a person can understand, read and write English they will usually be able to speak it as well. Therefore, the number of respondents leaving the ‘Speak/Sign’ box blank was surprising, suggesting that they may have been confused by the tick-box or interpreted ‘speak/sign’ as meaning ‘speak and sign’.

There were five tick-boxes for Welsh, four representing types of language knowledge and one to declare ‘no ability’, so for this part of the question to be completed correctly there should be at least one box ticked. However, many individuals without any knowledge of the Welsh language did not tick the ‘No ability’ box.

¹ Except in the case of British Sign Language

ONS was not able to draw any conclusions about the BSL part of the question because of the low number of respondents who had any knowledge of BSL in the tests of this version.

In relation to the final part of the question, the vast majority of respondents who reported another language both wrote in the language *and* ticked at least one box 'Understand', 'Speak/Sign', 'Read' or 'Write', as intended.

6.1.2.2 Full matrix, version 2

For the next round of testing ONS added an additional tick-box for the reporting of 'No ability' in BSL and the option to report on a second 'Other' language. ONS also removed the 'No ability' in the 'Other' language(s) tick-box as it was not needed (Figure 6.4).

Figure 6.4: Full matrix language question version 2

15 What languages can you understand, speak, read or write?
 ♦ Tick all boxes that apply.

	Understand	Speak / Sign	Read	Write	No Ability
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Welsh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
British Sign Language	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>
Other language, write in	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Other language, write in	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Respondents did not understand this version well. Responses were generally incomplete and inconsistent, and data was unreliable and of poor quality. Some respondents only ticked 'Understand', taking this to cover all language skills (when it was designed to mean 'understand spoken language').

'If you're saying that you understand it, then you can obviously speak and sign it, and read and write it can't you?' (Whole Questionnaire Testing, 2006)

As a result the boxes to the right of the 'Understand' box were often not considered:

'I didn't think they ... Oh, yes. Now it's pointed out, yes I can see. I should have done that shouldn't I?' (Whole Questionnaire Testing, 2006)

Where respondents did not indicate that they had any knowledge of Welsh or BSL, they often did not indicate that they had 'No ability' either.

Respondents generally did not understand what the 'Speak/Sign' option meant, and as a result, many chose to leave it blank:

'I'd leave that blank ... I'd take it to be it's for sign language.' (Whole Questionnaire Testing, 2006)

This under-recording of the ability to speak a language would give the misleading impression that ability to speak languages is lower than the abilities to understand, read and write.

Those who recorded more than one 'Other' language explained that they had entered them in order of fluency. One respondent spoke a language from Nigeria, but did not think to include it until questioned by the interviewer because they were not expecting to be asked about this:

'... you would not expect something like that to come on the census form because you are expecting ... the numbers, to know the numbers of people in this country, I would not expect to go into my roots.' (Whole Questionnaire Testing, 2006)

6.1.2.3 Full matrix, version 3

For the third version, ONS reordered the ability response boxes to remove the emphasis from the 'Understand' tick-box, which had previously come first. From left to right the tick-boxes now read: 'Write', 'Read', 'Speak/Sign', 'Understand' and 'No ability' (Figure 6.5).

Figure 6.5: Full matrix language question version 3

15 What languages can you understand, speak, read or write?
 ♦ Tick all boxes that apply.

	Write	Read	Speak /Sign	Understand	No ability
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Welsh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
British Sign Language			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language, write in	<input type="text"/>				
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Other language, write in	<input type="text"/>				
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Moving the 'Understand' tick-box improved responses to this question, with respondents generally continuing to read across each row after ticking their initial response. However, many respondents continued to answer the question inconsistently and incorrectly and as with earlier testing, often missed the 'No ability' boxes.

Once again, several respondents were confused by the 'Speak/Sign' tick-box. Some thought that it applied only to sign language. Others did not know what it meant:

'Speak and sign? Are you talking about sign language there? I don't know ... I can read and write in English but when I came to that it threw me, that "speak and sign".' (Whole Questionnaire Testing, 2006)

Some respondents felt that to declare an additional language they should have a fairly high standard, possibly even be fluent:

'You would have to be a little bit less than fluency... pretty bilingual.' (Whole Questionnaire Testing, 2006)

However, other respondents felt that a lower standard was acceptable and therefore included an additional language.

As with version 2 some respondents for whom English was not their first language did not record their mother tongue:

'I didn't put my own language on; I didn't see that as an ability, ... because that's just like a natural thing.' (Whole Questionnaire Testing, 2006)

6.1.2.4 Full matrix, version 4

For version 4 (Figure 6.6) ONS moved the 'No ability' box to be the first response option to increase its visibility and encourage respondents to tick it if appropriate. This allowed the position of the 'Understand' category to be changed and the response options to be rearranged so that the order reflected what the question was asking, that is understand, speak, read or write. ONS also renamed the 'Understand' category 'Understand spoken'.

As respondents had difficulty understanding the category label 'Speak/sign', ONS gave BSL its own category labels of 'Understand sign' and 'Sign'. The second 'Other language' option was also replaced with an 'Other sign language' option to accommodate those who use something other than BSL.

This version of the question was included in the 2007 Census Test and in the 2007 consultation.

Figure 6.6: Full matrix language question version 4

Less than one-half (46 per cent) of respondents to the 2007 consultation who required information on language thought that this version would meet their needs. The principal problems identified by data users included:

- the lack of information on a person's main/preferred language of communication
- the lack of space for recording other languages – the question would not provide sufficient detail for European, Asian and African languages, for example

- the need to gather information on different levels of ability
- the need for information on different aspects of language ability, particularly English language proficiency, proficiency in other languages, and levels of illiteracy

The error rates for this question in the 2007 Census Test were quite high. Errors are recorded where respondents fail to answer the question or provide a logically inconsistent answer, such as ticking 'no ability' and 'understand spoken'. The English section of the question had the lowest error rate of the five sections but was still relatively high at 6.7 per cent. The Welsh section of the language question had the highest error rate at 71.2 per cent. The majority of this can be attributed to the high number of non-responses. Many of these non-responses (99.2 per cent) were from the English regions of the census test. While it could be assumed that most people in England don't have any ability in Welsh and therefore left the question blank, the high amount of non-responses would have an adverse effect on the quality of data collected.

For the people who provided a write-in response to the 'Other language' section, 33.5 per cent failed to indicate their ability in this language. The sign language sections also had high error rates; the 'British Sign Language' section had a 61.5 per cent error rate and the 'Other sign language' section had 52.6 per cent. A large proportion of the British Sign Language error is again down to many people failing to provide a response to the question.

The census test evaluation survey found that of the 182 responses that indicated difficulty with the census individual questions, language was the second most problematic question (11.0 per cent of those reporting difficulty). Most of these respondents found it to be badly worded and in other comments, many explained they didn't know the level of expertise required.

This version also performed poorly during cognitive testing, with respondents often failing to tick all relevant boxes. They commonly missed the 'Understand spoken' box. A number of respondents did not notice it, but in some cases they interpreted the four ability boxes as a scale, where 'Understand spoken' was the lowest ability level and 'Write' was the highest:

'It seems to me that it's graded ... speak, read, write means good at it.' (Wave 1)

In these cases the 'Understand spoken' box was often left blank while the rest were ticked. The perception of a graded structure resulted in a proportion of respondents whose responses indicated they were able to speak, read and write, but not understand speech.

The 'Understand spoken' box was also often missed in the 2007 Census Test. Almost one-third of respondents in England test areas indicated that they spoke, read and wrote English but did not tick 'Understand spoken'.

Some respondents with more than one additional language were unsure how to respond using the limited space available. They were inconsistent in how they decided whether to include a language or not. One respondent was fluent in three languages. They wrote in one of the languages, but could offer no reason for selecting the one they had written. Adopting this question would mean that multilingual respondents would be left to enter a language according to their own criteria, leading to inconsistent selection criteria and data that would be difficult to interpret correctly.

The high error rates found throughout testing indicated that many respondents failed to understand this question and that it needed to be simplified.

6.1.2.5 Full matrix, version 5

For the next version (Figure 6.7) ONS redesigned the matrix by removing the 'Welsh language' row (intending to collect this information in a separate question similar to that used in 2001) and removing the 'Other sign language' option.

Figure 6.7: Full matrix language question version 5

15 What languages can you understand, speak, read or write?

◆ Tick all boxes that apply.

	No ability	Understand spoken	Speak	Read	Write
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language, write in	<input type="checkbox"/>				
		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	No ability	Understand sign	Sign		
British Sign Language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		

Cognitive testing found that completion errors remained likely. Many respondents took some time to read and make sense of how to answer the question, suggesting that the question may be too complex. Often respondents did not tick the boxes that were relevant to them, particularly the English 'Understand spoken' box. Many also did not tick any of the boxes in the 'British Sign Language' row.

Respondents with insufficient space to record their other languages tended to declare their languages in order of fluency:

'I could have included French and German but there wasn't enough space ... Russian is the one I'm most fluent in.' (Wave 2, Version B)

Some respondents wrote in their additional languages outside the write-in boxes. This creates problems for processing as such information will not be scanned or will be flagged up to be dealt with by hand, leading to more expensive and time-consuming processing costs.

6.1.3 Simplified matrix style

As a result of the problems identified with the full matrix layout, ONS simplified and shortened the question by:

- changing the question from 'What language(s) can you understand, speak, read or write?' to 'Which of these can you do?' and including a list of language abilities
 - re-designing the matrix so the ability categories were listed vertically, rather than horizontally
 - removing the British Sign Language category in the assumption that the write-in space would be adequate for capturing this information (although ONS appreciated that this would be likely to cause some loss of data on sign languages)
-

- reinstating the 'Welsh' tick-boxes
- focussing the option to report other languages on spoken ability only (as this was a high priority for users and removed the need for a complex tick-box layout)

6.1.3.1 Simplified matrix, version 1

For the first version of the simplified matrix ONS included two rows of write-in boxes where 'Other' languages spoken could be recorded (Figure 6.8).

Figure 6.8: Simplified matrix language question version 1

15 Which of these can you do?
 ◆ Tick all boxes that apply.

	English	Welsh
Speak	<input type="checkbox"/>	<input type="checkbox"/>
Understand	<input type="checkbox"/>	<input type="checkbox"/>
Read	<input type="checkbox"/>	<input type="checkbox"/>
Write	<input type="checkbox"/>	<input type="checkbox"/>

Speak another language. Write in language(s) below

<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Respondents were generally able to understand the question and answer quickly. There were fewer completion errors compared with the full matrix design of the language question.

During cognitive testing the majority of respondents used all four tick-boxes for 'English'.

Like the earlier question, 'What languages can you understand, speak, read or write?' this version of the question did not give respondents any guidance about the proficiency level required for reporting a language. Again, this caused confusion. As one respondent stated:

'If it said "speak another language fluently" I would know what that meant a bit more and I definitely wouldn't even hover over it.' (Wave 2)

'That bit doesn't say to me, can you do a bit of Del-boy French to get you through on holiday, or can you hold a decent conversation or are you fluent?' (Wave 2)

As with earlier testing, responses covered a range of different interpretations and inconsistent decisions about whether or not to include another language. Two respondents who had lived in Spain decided not to include Spanish. One had lived there for 15 years and explained that they decided not to write in the language because:

'I can't speak it like a native.' (Wave 2)

The other felt they did not speak it well enough to consider it to be 'a second language'. Another respondent had a contrasting interpretation, including German because they had a German GCSE and enjoyed visiting Germany (Wave 2).

A further respondent assumed that the questionnaire did not need to know about languages that they felt were not useful in the UK:

‘Ghanaian language, will that be of any help in this country? I don’t think so, that’s why I left it.’ (Wave 2)

6.1.3.2 Simplified matrix, version 2

In the next version of the language question ONS included an additional row of write-in boxes to provide more room for respondents to write in additional languages (Figure 6.9). This version was also tested in Welsh (Figure 6.10).

Figure 6.9: Simplified matrix language question version 2 (English language)

14 Which of these can you do?

◆ Tick all boxes that apply.

	English	Welsh
Speak	<input type="checkbox"/>	<input type="checkbox"/>
Understand	<input type="checkbox"/>	<input type="checkbox"/>
Read	<input type="checkbox"/>	<input type="checkbox"/>
Write	<input type="checkbox"/>	<input type="checkbox"/>

Speak another language. Write in language(s) below

Figure 6.10: Simplified matrix language question version 2 (Welsh language)

14 Pa rai o’r rhain allwch eu gwneud?

◆ Tidiwch bob blwch sy’n berthnasol.

	Cymraeg	Saesneg
Siarad	<input type="checkbox"/>	<input type="checkbox"/>
Deall	<input type="checkbox"/>	<input type="checkbox"/>
Darllen	<input type="checkbox"/>	<input type="checkbox"/>
Ysgrifennu	<input type="checkbox"/>	<input type="checkbox"/>

Os gallwch siarad iaith arall, ysgrifennwch yr iaith/ieithoedd isod

Cognitive testing in English showed that respondents had little trouble in deciding which English language ability tick-boxes to choose. In Welsh language testing all respondents ticked the Welsh language and English language ability boxes for ‘Speak’, ‘Understand’, ‘Read’ and ‘Write’ with most mentioning that their selections were based on being able to confidently ‘speak’, ‘understand’, ‘read’ and ‘write’ in both languages. Only one respondent hesitated before ticking all the boxes:

‘Ro’n i’n meddwl bod “siarad Cymraeg a Saesneg” yn ddigon ... ond wedyn pan welais i’r gweddiill... deall, ac yn y blaen, wnes i sylwi ‘mod i’n gallu gwneud rheina hefyd.’	[I thought that “Speaking Welsh and English” was enough ... but then when I saw the rest ... understand, and so on, I realised that I can do those too.]
---	--

(Wave 1 Welsh language testing)

However, in Wave 1 Welsh language testing, Welsh-speaking respondents had three variations on their interpretation of ‘Deall’ [‘Understand’].

- some interpreted ‘Deall’ [‘Understand’] to refer to understanding spoken Welsh only:

‘O safbwynt bod rhywun yn siarad efo chi ... ac ych chi’n deall beth maen nhw’n ddweud.’	[From the perspective of someone talking to you ... and you can understand what they’re saying.]
--	--

(Wave 1 Welsh language testing)

- others interpreted 'Deall' as referring to being able to understand spoken Welsh and understand written Welsh:

'Hynny yw dilyn beth sy'n digwydd mewn sgwrs neu dilyn yr ystyr wrth ddarllen rhywbeth.'	[That is follow what's happening in a conversation or following the meaning when reading something.]
--	--

(Wave 1 Welsh language testing)

- in some cases respondents interpreted 'Deall' as referring to being able to speak, read and write the language.

'Os ych chi'n gallu siarad e, darllen e, a ysgrifennu e, wedyn ych chi'n gallu deall e – buase dim rhaid i chi gael 'deall' fan hyn.'	[If you can speak it, read it and write it, then you can understand it – so there'll be no need for 'understand' here.]
---	---

(Wave 1 Welsh language testing)

As with earlier findings, some respondents in English language testing seemed to find it difficult to decide whether to write in additional languages, or which Welsh tick-boxes to choose, because there was a lack of proficiency indicators:

'It's not giving you the option to say oh I can speak a little, a lot, or my Welsh is very good or whatever.' (Wave 3)

With reference to their knowledge of the Italian language, one respondent said:

'[I can] speak a little bit, understand a little bit, I can read, like ... but there's not space to write "mother tongue" or levels of proficiency.' (Wave 3)

In Welsh language testing generally, respondents believed that fluency in another language, rather than holding a qualification or the ability to speak 'a little', was necessary to declare a knowledge of that language:

Achos dw i'n ystyried bo fi'n eitha rugl ynddyn nhw. Dw i ddim yn eu defnyddio nhw yn aml...ond fi'n weddol rhugl ynddyn nhw	[Because I consider myself to be quite fluent in them. I don't use them often ... but I'm quite fluent in them]
--	---

(Wave 1 Welsh language testing)

'Gwnes i TGAU Ffrangeg ... ond ddim yn gallu siarad yn rhugl...felly'n teimlo mod i ddim yn ddigon da i nodi hwnna.'	[I did GCSE French ... but I'm not fluent ... so I felt that I wasn't good enough to note that.]
--	--

(Wave 1 Welsh language testing)

The provision of a third row of write-in boxes generally provided adequate space for respondents to write in additional languages although a few respondents wanted to be asked about different levels of language ability:

'Ond dyw e ddim cweit yn rhoi darlun cyflawn ... gall unrhyw un dweud bod nhw'n siarad Cymraeg os ydyn nhw'n gallu cyfri i bump neu ddweud bore da. Mae angen	[But it doesn't quite give a full picture ... anyone can say that they speak Welsh if they can count up to five, or say good morning. It should be included whether you
---	---

<p>cynnwys a ydych chi'n gallu siarad yn achlysurol, yn rhugl ... ac yn y blaen ... gwahanol lefelau. Mae rhai yn siarad neu wedi colli'r arfer. Mae rhai yn gallu siarad Cymraeg ond ddim yn dewis siarad, a does dim opsiwn ar gyfer dysgwyr ...fel fy ngwraig, sy'n gallu siarad tipyn go lew o'r Gymraeg, ond ddim yn ystyried ei bod hi'n rhugl.'</p>	<p>can speak occasionally, fluently, and so on ... the different levels. Some have lost the use of their Welsh. Some can speak Welsh, but choose not to ... and there's no option for learners, like my wife, who can speak quite a bit, but doesn't consider herself to be fluent.]</p>
--	--

(Wave 1 Welsh language testing)

Respondents who had children under three years old were unsure about which boxes they should tick, if any. One respondent said they would tick 'Speak', commenting:

'Even if it's a few words, even if it's key words, then I'd still consider that she [the baby] can speak English'. (Wave 3)

Other respondents with young children acted differently, such as writing in 'BABY SO IS NOT SPEAKING YET', or recording that the child is able to understand the language but is not able to speak.

Clearly, how respondents answer for their babies could have implications on the accuracy of the data collected in terms of incorrectly ticking boxes or not ticking boxes to answer the question on behalf of the baby. However, it is important to note that for 2001, outputs on the knowledge of Welsh question were based on respondents aged 3 or older. Outputs for the 2011 Census have not yet been determined, but the same filter is likely to be applied for young children answering the language questions.

6.1.3.3 Simplified matrix, version 3

In light of this testing, for version 3 of the simplified matrix ONS included a 'Baby too young to talk' tick-box (Figure 6.11). To make space available for this, ONS removed the third row of write-in boxes. ONS preferred to add a tick-box rather than an instruction (for example 'Do not complete for children under 3') as this would reveal whether people who did not indicate any ability did so because they were a child or because they had mistakenly missed the question.

ONS also added a proficiency indicator to guide respondents about whether they should report other languages. This was done by changing the wording for the other language category to 'If you are fluent in any other languages, please write in below'.

Figure 6.11: Simplified matrix language question version 3

16 Which of these can you do?
 ♦ Tick all boxes that apply.

	English	Welsh
Speak	<input type="checkbox"/>	<input type="checkbox"/>
Understand	<input type="checkbox"/>	<input type="checkbox"/>
Read	<input type="checkbox"/>	<input type="checkbox"/>
Write	<input type="checkbox"/>	<input type="checkbox"/>

If you are fluent in any other languages please write in below

or Baby too young to talk

All respondents ticked the English language ability boxes for 'Understand', 'Speak', 'Read' and 'Write'. As with the other simplified matrixes, cognitive testing indicated that respondents find this design easier to understand than the full matrix.

The introduction of the word 'fluent' in relation to other languages helped to clarify for what respondents should include. This led to respondents correctly deciding not to include languages in which they had some ability, but were not fluent in:

'I do have small amount of ability in French and German from school but not enough to actually have a command of the language, therefore I didn't write it.' (Wave 4.1)

Testing of this version of the question with respondents who had babies too young to talk was limited.

Stakeholders from the Census Diversity Advisory Group were concerned about the loss of reference to signing and the Welsh Assembly Government were concerned about ensuring consistency with 2001. The introduction of English tick-boxes in the first column, and the use of the word 'fluent' in the second part of the question may influence respondents' interpretation of the knowledge of Welsh tick-boxes.

6.1.4 General comments on a matrix style language question

Space constraints on the census questionnaire are a significant factor in the development of questions. The matrix format was initially appealing as it offered the most concise way of asking about English, Welsh and other languages.

However, testing suggested that respondents found all versions of the matrix style language question difficult to complete. The complexity of the question resulted in various completion errors which gave inconsistent and incomplete data. The two-language (English and Welsh) format gives the respondent more information to process and increases the burden on respondents.

A significant problem identified in testing was the lack of clarity about the level of proficiency required for respondents to report that they could understand, speak, read or write a language. Those with English as a first language, and with some knowledge of another language, particularly struggled with this question. Respondents often applied variable standards of language proficiency when choosing whether or not to report another language.

It is clear that respondents indicating competency in a language will sometimes have a lower level of skills than others who have chosen not to declare a language. The range of interpretations about what this question was asking meant that answers were inconsistent and therefore not comparable, which would lead to statistics that provide an unreliable picture of language ability.

International research and best practice also advises that a matrix format is not appropriate for a question in a self-complete census questionnaire (for example, Dillman, 2000). Questions using a matrix or tabular format are known to be difficult to understand and interpret for people with poor comprehension skills or English as a second language. They also cause confusion among respondents and add to the challenge of getting respondents to follow a questionnaire in the predefined order.

More fundamentally, after a review of consultation responses ONS was concerned that the matrix format would not be able to meet key user requirements. Many data users need language information to understand how many people do not have sufficient English language ability to access public services. In addition to cognitive testing findings, there is evidence that capturing people who can speak a language without any indication of the level of proficiency includes those

with low ability. For example, a survey of south Asians in England in 1994 found that of those people who said they could speak English the proportion that could speak ‘just a little English’ (as opposed to ‘very well’ or ‘fairly well’) ranged from 12 per cent amongst Indian men to 33 per cent amongst Bangladeshi women (Johnson 2000).

Analysis by Haseldon (2003) of the knowledge of Welsh language question found that ‘definitions of ‘understanding’, ‘speaking’, ‘reading’ and ‘writing’ vary from person to person and depend upon the design of the question.

Based on findings from cognitive testing, quantitative testing, international best practice and external research ONS decided that a matrix style language question should not be adopted in the England and Wales census. Instead, efforts were made to develop a new question on English and other languages that could be asked separately from the question on Knowledge of Welsh language.

6.2 New Zealand style conversation questions

ONS developed and tested a conversation question adapted from the language question used in New Zealand’s census, which asked ‘In which language(s) could you have a conversation about a lot of everyday things?’ (Figure 6.12).

Figure 6.12: New Zealand language question, 2006 Census

13 Mark as many spaces as you need to answer this question. In which language(s) could you have a conversation about a lot of everyday things?

Remember to mark English if you can have a conversation in English.

- English
- Māori
- Samoan
- New Zealand Sign Language
- other language(s), for example GUJARATI, CANTONESE, GREEK.

Print the language(s):

or none (for example too young to talk)

The focus of the question was on collecting data about languages that respondents could speak and understand. There was no option to report written or reading ability as spoken language was the priority for users.

6.2.1 Conversation question, version 1

In early testing ONS had briefly trialled a version of the conversation question addressing English, Welsh, British Sign Language (BSL) and ‘Other’ languages. Cognitive testing was undertaken when the tick-box for ‘Other’ languages included examples of other languages (‘such as Cornish, Urdu, Hindi’) and also when these examples were removed (Figures 6.13 and 6.14).

Figure 6.13: Conversation question version 1a

Figure 6.14: Conversation question version 1b

Respondents interpreted the question as asking about languages they could speak and understand:

‘... you can’t have a conversation unless you can understand [the language].’ (Preliminary testing)

When answering the question respondents took into account their mother tongue and also any additional language(s) they had learnt later on. They understood it to be asking about the language they usually used:

‘... that’s about the languages you most of the time speak, I live here, I most of the time speak English.’ (Preliminary testing)

However, testing of this question at this stage was limited. This question was put to one side while testing of the matrix versions continued.

6.2.2 Conversation question, version 2

ONS reconsidered using a conversation approach when the matrix style question was rejected and the decision was taken to develop a question that could be asked separately from the Welsh language proficiency question. The Welsh tick-box was therefore removed from this question.

To make the question less hypothetical than ‘...could you have a conversation...’ ONS amended it to ‘In which languages can you have a conversation about a lot of everyday things?’

ONS first tested the question with a tick-box for ‘Other’ language with ‘write in main language’ (Figure 6.15). The question was further tested after it was amended to ‘write in main other language’ (Figure 6.16) to avoid confusion if English or British Sign Language (BSL) were the respondent’s main language.

Figure 6.15: Conversation question version 2a – ‘Other, write in main language’)

18 In which languages can you have a conversation about a lot of everyday things?

- ➡ Tick all that apply.
- ➡ Remember to tick English if you can have a conversation in English.
- English
- British Sign Language
- Other, write in main language

--	--	--	--	--	--	--	--	--	--	--	--	--	--

Figure 6.16: Conversation question version 2b - ‘Other, write in main other language’)

18 In which languages can you have a conversation about a lot of everyday things?

- ➡ Tick all that apply.
- ➡ Remember to tick English if you can have a conversation in English.
- English
- British Sign Language
- Other, write in main other language

--	--	--	--	--	--	--	--	--	--	--	--	--	--

Respondents generally considered ‘A lot of everyday things’ to mean talking about domestic and day to day issues, such as housework, shopping, going to the doctor, the weather, car repairs and work, rather than being able to engage in complex discussions or any kind of technical or philosophical debate.

The choice of whether or not to record a language was ultimately influenced by what respondents considered to be the required level of proficiency in order to ‘have a conversation about a lot of everyday things’. Testing identified that there were mixed opinions about the level of language ability required. The majority of respondents suggested that in relation to this question a relatively good to high standard of a language was necessary. For example:

‘Which languages I can have a conversation in for some hours, about almost everything, ... so in which languages I have a very good level.’ (Wave 5)

But other respondents thought that the level of ability required was more basic:

‘... just if you can communicate perhaps say if you went somewhere on holiday and you knew a bit about the language, which perhaps if you could just get by really.’ (Wave 5)

Reasons respondents gave for not reporting a language included that they were not fluent, or that they believed their ability was not of a high enough standard, including that they were not as proficient as they used to be, to justify including it in the answer to this question.

Reintroducing the BSL option was an advantage to the question, as without it, it may not be obvious that BSL should be written as another language. Several respondents had some knowledge of BSL, but like the verbal languages, people often chose not to record them because they did not feel that they had sufficient ability:

‘I know bits and pieces of sign language, but again not enough to feel that I can answer that question.’ (Wave 5)

Overall ‘main’ language was defined as ‘mother tongue’, a ‘first’ language or fluent language. However, for others ‘main language’ referred to the languages they used on a daily basis, which meant that they did not record their ‘mother tongue’ or ‘first’ language if they did not use it in the UK. A person who spoke Swahili at home with their spouse did not report it because when they thought of ‘everyday things’ they thought of English, as they live in Britain and therefore use English to carry out their daily activities. Another respondent who was from Kenya and whose first language was Somali did not include this and solely ticked ‘English’. They explained that this was because:

'I only speak English here. I don't have anybody to speak Somali with, you know.' (Wave 5)

Respondents with babies highlighted that a language question did not really apply as their child was too young to talk.

These findings indicate that asking this style of question would lead to the collection of inconsistent data. Further development would also be required before users considered it suitable for their needs. Furthermore, the Welsh Assembly Government expressed concern that the question may interact with the Welsh language knowledge question if respondents think they should apply the 'a lot of everyday things' proficiency measure to both questions. It was therefore decided that this question would not be suitable for inclusion.

6.3 US style, two-part question

As none of the previously tested style of questions met all the essential elements of a suitable language question, let alone the desirable elements, ONS decided to test an alternative style of question for meeting user requirements within space constraints. This was a two-part question asking about primary language, followed by proficiency in English. The question was based on the language questions asked on the US census long form questionnaire in 2000 (Figure 6.17).

Figure 6.17: United States language question, Census 2000, long form questionnaire

The image shows a section of a questionnaire with a yellow background. It is numbered '11' in a black circle. The text reads: 'a. Does this person speak a language other than English at home?' followed by two radio button options: 'Yes' and 'No → Skip to 12'. Below this is question 'b. What is this language?' with a long horizontal line for an answer and a note in parentheses: '(For example: Korean, Italian, Spanish, Vietnamese)'. Finally, question 'c. How well does this person speak English?' is followed by four radio button options: 'Very well', 'Well', 'Not well', and 'Not at all'.

Three versions of the primary language question (the first part of the question) were developed and tested. One collected household level data:

- 'Is a language other than English used every day in your household?' (In England)

The other two versions captured individual level data:

- 'Do you speak a language other than English or Welsh at home?' (in Wales) and 'Do you speak a language other than English at home?' (in England)
- 'What is your main language?'

They also guided the respondent to include sign languages, clarifying that the question was not only interested in spoken languages.

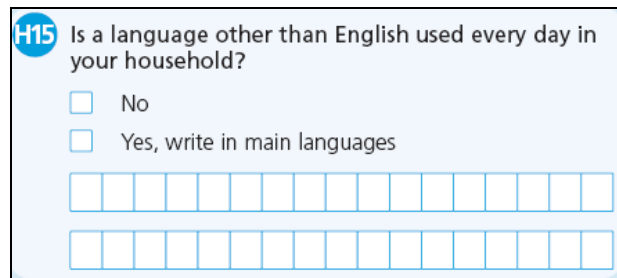
6.3.1 Part 1 primary language

The key purpose of a primary language question is to establish the language that public authorities can communicate with respondents in. The questions were designed to work in conjunction with the second part of the two-part question which asks about English language proficiency.

6.3.1.1 Primary language, version 1: 'Is a language other than English used every day in your household?'

ONS intended to include the first version of the primary language question in the 'household' section of the questionnaire, followed by an individual question on English proficiency in the 'individual' questions. This question asked about languages spoken in the household (Figure 6.18).

Figure 6.18: Primary language question version 1 - Household language (England)



H15 Is a language other than English used every day in your household?

No

Yes, write in main languages

A limitation of this question was that it did not capture other languages spoken well by residents if they are not spoken every day in the household. ONS recognised that 'language at home' will produce an underestimate of linguistic capacity, particularly for respondents who do not speak their 'mother tongue' at home, but may still understand it and have the capacity to revive it, and in some cases may be active users of the language in other contexts (for example at work or at community events and organisations).

Cognitive testing found that compared with the conversation style question (see Section 6.2), this question was easier for respondents to answer and more straightforward, especially as it did not ask about proficiency:

'... it's not concerned with how well you speak it ...' (Wave 5)

However, the inference is that a language would not be spoken in the household every day unless it was of a high standard.

As this question asked about languages spoken in the household respondents sometimes included languages that their family members or housemates who they lived with spoke, even though they did not personally speak them. For example:

'My brother and sister can converse in German so I would write that too.' (Wave 5)

A bilingual respondent felt that the question was context specific, as they spoke different languages at home with different people. However, another respondent who spoke Somali as their first language said they would not include this in response to the question as they did not speak it in the UK as they did not have anybody else to speak Somali with. They would only include it if they had family staying with them in the UK as they would then be using the language.

One respondent made a contrasting decision and included languages spoken by their housemate, even though these languages were not used in the house between those living there. Respondents therefore appeared to be inconsistent in their interpretation of what was meant by language known and languages actually used at home. A key problem with this was that languages would not be attributable to individuals.

Also the question would not provide information for people living in communal establishments as they would not receive the household question. According to the 2001 Census 1 million people (2 per cent of the UK population) were living in a communal establishment. Of those non-staff living in communal establishments there was over-representation of people from 'Irish', 'Other White' 'White

and Black African', 'White and Asian', 'Other Mixed', 'Other Asian', 'Black African', 'Chinese' and 'Any other ethnic group' populations.

ONS therefore ruled out the inclusion of a household question.

6.3.1.2 Primary language, version 2: 'Do you speak a language other than English or Welsh at home?'

ONS then considered replacing the question with 'Do you speak a language other than English at home?' in England and 'Do you speak a language other than English or Welsh at home?' in Wales, on the 'Individual' section of the questionnaire (Figure 6.19). 'BSL' was included rather than 'sign languages' because at the time of designing the question there was not enough space available on the questionnaire to use the longer example ('Including sign languages' would have taken up another line on the page). ONS only tested the Wales version of the question.

Figure 6.19: Primary language question version 2a - language spoken at home (Wales)

17 Do you speak a language other than English or Welsh at home?

No **GO TO** 19

Yes, write in main other language (include BSL)

The Wales version did not make any attempt to distinguish between people whose home language is Welsh from those whose home language is English because user need is mainly to identify those people who cannot function in an official language and there was no space to ask about them separately.

The testing of this version was limited but ONS was concerned that the reference to 'BSL' may confuse respondents.

The question was amended by changing the wording for the 'Yes' tick-box from 'Yes, write in main other language (include BSL)' to 'Yes, write in main other language below (including sign languages)'. This was done for two reasons:

- to clarify what 'BSL' meant, as testing showed that individuals who did not have English as a first language were unsure what 'BSL' was
- the word 'including' replaced 'include' as it was softer, more of a request than a command (Figure 6.20).

Figure 6.20: Primary language question version 2b - including 'sign languages' (English language version, Wales)

18 Do you speak a language other than English or Welsh at home?

No **GO TO** 20

Yes, write in main other language below (including sign languages)

One advantage of this version of the question is that it may be straightforward to complete if the head of household is answering on behalf of other respondents. However, there were a number of problems:

- it was too large for the available space (because the question runs onto two lines). The extra space required would need to be made up, perhaps by removing the instruction in the subsequent religion question
- it was still unable to capture other languages spoken well by residents if they are not spoken every day in the household. One respondent who did not have English as a first language also ticked 'No' as they only spoke English at home. Another respondent who spoke Somali as their first language but did not speak it while living in the UK, chose to tick 'No'
- it may be meaningless for respondents who live alone and do not talk to anyone at home (Siegel *et al* 2001, Extra and Gorter 2008). In 2001, 30 per cent of all households were one-person households (except for 'White Irish' and 'Mixed White and Black Caribbean' ethnic minorities who were under represented in one-person households)
- the concept of languages spoken at home is a complex one and may require further qualifications, such as to whom and when (Extra and Gorter 2008)
- the fact that multiple other languages may be spoken at home appears to overshadow the instruction to write in the *main* other language and since there is limited write-in space it may lead to difficulties processing the data. In cognitive testing not all respondents noticed the reference to include 'main' language. One respondent wrote in their main language of Somali, but then went on to include Arabic and Urdu in their answer.
- analysis of the US version revealed an uncertainty about the level of proficiency required to state another language. (Siegel *et al* 2001)
- 'at home' might be interpreted as a respondent's current home where they live, their childhood home where they grew up, or their home country that they visit occasionally with different answers in each case (Siegel *et al* 2001)

As a result of these considerations ONS decided not to recommend this version for the 2011 Census.

6.3.1.3 Primary language, version 3: 'What is your main language?'

Following concerns that language spoken at home may not be clear to respondents ONS tested an alternative wording, asking 'What is your main language?' in the May 2008 module of the Opinions Survey as well as subsequent postal tests and waves of cognitive testing (figure 6.21).

In the second part of Wave 6 ONS changed 'Other, write in (including sign language)' to the plural 'Other, write in (including sign languages)' to include all possible sign languages and not just one sign language, which may have been confusing (Figure 6.22). Concerns that respondents would write in 'sign languages' rather than a specific sign language led ONS to change the instruction to 'including British Sign Language' for the March 2009 and July 2009 postal tests (Figure 6.24).

For the English version of the question in Wales, ONS tested both the wording 'English or Welsh' (Figure 6.22) and 'Welsh or English' (Figure 6.23) for the first tick-box.

Figure 6.21: Primary language question version 3 – 'main language' (England)

18 What is your main language?

English GO TO 20

Other, write in (including sign language)

Figure 6.22: Primary language question version 3 with wording 'English or Welsh' (Wales, English)

18 What is your main language?

English or Welsh **GO TO 20**

Other, write in (including sign languages)

[] []

Figure 6.23: Primary language question version 3 with wording 'Welsh or English' (Wales, English language)

18 What is your main language?

Welsh or English

Other, write in (including sign languages)

[] []

Figure 6.24: Primary language question with 'including British Sign Language'

18 What is your main language?

English → Go to 20

Other, write in (including British Sign Language)

[] []

One disadvantage of this question is that it does not collect the full range of languages spoken as respondents will only enter a language other than English if it is their main language. While this means that some information on language diversity will be lost, other language questions tested have been found to produce unreliable data on language diversity due to inconsistent choices over whether to include basic abilities, such as a language GCSE.

Main language was considered a useful concept in meeting the essential user need of allowing data users to understand which languages services should be provided in. In order to best meet this need, and to give an indication of language diversity, it is important that respondents report their first or preferred language rather than ticking 'English' because, for example, they speak it more often. It was therefore important to determine whether respondent understanding of the term 'main' language was consistent with the intended meaning of 'first' or 'preferred' language rather than meanings relating to frequency of use or current usage. Table 6.25 shows the results of the June 2008 Opinions Survey for a question that was included asking 'which of these is true of your main language?' The question encouraged multiple responses, so the total percentage of choices is greater than 100.

Table 6.25: Response percentages for multiple response question ‘which of these is true of your main language?’, Opinions Survey June 2008

Response	Percentage of cases
It is the first language I could ever speak	95.9
I use it most of the time	56.9
I speak it when I am at home	58.0
I speak it with my family	58.0
I speak it with my friends	57.7
I speak it at school or work	48.1
I use it when dealing with public authorities	54.1
Total	428.8

Table 6.25 shows that 96 per cent of respondents felt that their main language was the first that they could ever speak. There was significantly less agreement that the other concepts were true of main language. This pattern was more exaggerated amongst those who had previously answered that their main language was not English and suggested that they typically saw main language as referring to the first language which they spoke and the language they used at home and with family.

This appears to be supported by cognitive testing. Many respondents with English as their second language who used English on a daily basis interpreted the question to be asking for ‘mother tongue’ or ‘first’ language and recorded the other language they spoke. Respondents often equated this with the language that they used at home with their family or the one that they felt most comfortable with and were able to express their thoughts in:

‘Your first language, the first language you speak ... the one you were taught as a child ... what your family taught you.’ (Wave 6)

‘The language which I speak at home and every day, and which I think in.’ (Wave 6)

A small number of respondents interpreted ‘main’ language in cognitive testing as the language spoken most often:

‘The one that you use most.’ (Wave 6)

However the majority of respondents interpreting the question in this way had English as a first language so the language they spoke most often was their mother tongue. Their answer was therefore consistent with the answers of those interpreting it as ‘mother’ tongue.

One respondent whose mother tongue was Afrikaans understood the question to be asking about frequency of use rather than mother tongue and answered ‘English’ as their main language. Although such answers would reduce the quality of information on language diversity, the information provided would still meet essential user needs appropriately as anyone who mainly used English as their primary mode of communication would have appropriate language skills to be able to access services adequately in English.

Table 6.26 shows response percentages for the main language question in the 2008 Postal Test and March 2009 Postal Test. Response percentages for the two tests were very similar. The question returned a low non-response rate of around 2 per cent in both tests and only 0.5 per cent of respondents chose to tick more than one box. In both tests around 2.5 per cent of respondents indicated a main language other than English, suggesting that the question does collect information on a significant proportion of people who may require language services.

Table 6.26: Response percentages for the main language question in the 2008 Postal Test and March 2009 Postal Test

Response	2008 Postal Test	March 2009 Postal Test
English	95.1	95.1
Other	2.4	2.6
Ticked both boxes	0.5	0.5
Non response	2.0	1.9
Total	100	100

Recent Welsh language Whole Questionnaire Testing found that although most of the respondents answered the main language question correctly, many of the UK born respondents were confused by the way this question is worded, and expected to have to make a choice between 'Welsh' or 'English' as their main language. One respondent was particularly angered by the question, and crossed out English in the response option. Another respondent refused to answer this question, because he felt strongly that there was no such thing as a 'main' language.

The Welsh language Whole Questionnaire Testing recommendation for this question is that there should be a publicity campaign for Welsh speakers in Wales to raise awareness about the purpose of this question to minimise any confusion.

In Wave 6 of cognitive testing, ONS removed the routing and changed the order of 'Welsh' and 'English' in the first tick-box on the English language version for Wales (figure 6.23). Details of the testing conducted to determine whether to include routing is presented in section 6.3.2 below.

There were no strong opinions expressed by respondents in Wave 6 about which language should be listed first. Quotes in relation to the 'Welsh or English' tick-box include:

'I didn't think a lot about it ... does it matter? Will people complain? It doesn't make any difference to me.' (Wave 6)

'The Welsh will be happy, grand, no problem ... it's normally English or Welsh, so it's normally the other way round'. (Wave 6)

Some Welsh-speaking respondents were confused by the question. A few thought initially that the question was asking 'which of these is your main language, "Welsh" or "English"?' and were confused. Others were glad that they were not asked to make the choice between 'Welsh' or 'English' as they felt that they were bilingual and that both languages were their 'main languages'. To minimise this type of confusion, the language in which the questionnaire is written will appear first. The English language version of the questionnaire in Wales will therefore read 'English or Welsh' and the Welsh language version will read 'Cymraeg neu Saesneg' ('Welsh or English').

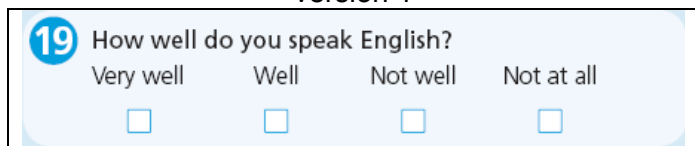
Although the findings suggest there is potential for some confusion in Wales about the 'Welsh or English' tick-box, particularly for Welsh speakers, version 3 does not suffer from some of the problems encountered in previous versions. The question is easily understood by the majority of respondents and understanding of the term 'main language' was found to be comparatively consistent amongst respondents. Unlike other questions considered, testing has suggested that those with only basic ability in a language will not distort the data by writing in the language. The question also fits within the space available on the census questionnaire. For these reasons, version 3 is the recommended primary language question for the 2011 Census.

6.3.2 Part 2 English language proficiency

Initial testing on the proficiency question adopted the wording 'How well do you speak English?' (Figure 6.27). This was based on the US Bureau of Statistics proficiency question 'How well does this person speak English?'. ONS changed the wording to 'How well can you speak English?' (Figure 6.28) to make it clear that the question was asking about ability rather than frequency of

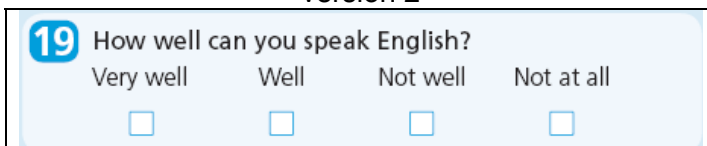
use. Given the wording in version 1, if someone *can* speak English but never *does* they could tick 'Not at all'. It was also changed to be more appropriate for people who may write in a sign language in answer to the main language question, and to reflect the knowledge of Welsh question, which uses 'can' rather than 'do'².

Figure 6.27: Proficiency question version 1



19 How well do you speak English?
Very well Well Not well Not at all

Figure 6.28: Proficiency question version 2



19 How well can you speak English?
Very well Well Not well Not at all

Those who did not speak English as a main language had a clear understanding of this question. When probed about what the question was asking, replies included:

'What is your level ... if you can go to the, I don't know, to the council to another office or something to communicate with the people.' (Wave 6)

ONS noted that respondents were subjective about how to decide which proficiency level to report for their own ability:

'It's only my opinion and it might not be ... it's hard to say.' (Wave 6)

Respondents with young children were asked how they would answer this question on behalf of their children. They replied that they would tick 'Ddim yn dda' [not very well].

Overall, cognitive testing found that respondents interpreted the various proficiency level options in the following ways:

- **Very well**
This category was seen as most relevant to those who had English as a first language. It was thought that if a person did not have English as a first language then to tick this box they should have a high level of English and possibly be fluent.
- **Well**
Respondents felt this level of ability was appropriate for people who were not fluent, but were able to use their language to get by on a day-to-day basis.
- **Not well**
Respondents thought this option referred to having limited English language ability. For example, a person who could follow some of what was being said, but with difficulty, and who may require some words translating.
- **Not at all**
Respondents tended to suggest that this option referred to people who might know the usual greetings or some basic words. They thought that individuals living in the UK would know a couple of words at least, rather than absolutely no English at all. They also thought this box was suitable for babies. Occasionally respondents interpreted this response option as suitable for those who literally spoke no English at all.

Analysis of the US version of this question suggested that a four-part scale is clear to users and allows a two-part distinction in terms of outputs:

² The Welsh question was changed from 'Does the person speak, read or write Welsh?' to 'Can you understand, speak, read, or write Welsh?' for the 2001 Census after testing found that 'Do you speak Welsh' was thus often interpreted as a question of (not necessarily extensive) ability rather than as a question of usage (Jones 2005).

‘The “not at all” and “not well” items come closer to identifying a unique sub-population (one that we might call “in need of English assistance”), than do the “very well” and “well” categories, (which we might say are “English assimilated” to varying extents. These data ... do not indicate “clean” differentiation between each of the four levels of English-speaking ability, but this does not mean that the “how well” item is not a useful discriminator of English ability. The use of four categories instead of two (call them “able” and “unable”) allows individuals (or their proxies) to use a finer discrimination even if this is an arbitrary one to some extent. Erring by one category in responding to “how well” will in many cases not change the final distribution of the two-category able/unable measure.’ (Kominski 1989:5)

As a way of checking the respondents’ self-assessment, ONS asked the Opinions Survey interviewers to give their own assessment of the respondents’ English language ability (Table 6.29). The results suggest a tendency for people to underestimate their English proficiency. However, none of the discrepancies in judgement between respondent and interviewer crossed the boundary between ‘not well’ and ‘well’ which is the most crucial distinction for data users.

Table 6.29: Cross-tabulation of respondents’ self-assessment of their proficiency in English and the interviewers’ assessment of respondents’ proficiency¹
England and Wales

		Percentages			
		How well interviewer thought respondent speaks English			
		Very Well	Well	Not Well	Total
How well respondent thought they speak English	Very well	97	3	-	100
	Well	56	44	-	100
	Not well	-	-	100	100

¹ Based on 326 responses (61 per cent of all respondents answered the question). Data are un-weighted.
Source: Opinions Survey June 2008 module

ONS considered whether to include routing on the main language question to direct people whose main language was English away from the English proficiency question. Being asked a question about English proficiency may seem strange to many respondents in England. However, asking all respondents how well they can speak English would potentially provide useful information on those who have English as a main language but are unable to speak it. The question was tested both with and without the routing.

In Wave 2 of the Welsh language testing the routing was included. Most of the Welsh language respondents followed the routing correctly to question 20 past the English proficiency question onto the religion question. An overseas-born respondent correctly ticked ‘Other’, wrote their main language in the space provided, and correctly answered the proficiency question. A few respondents missed the routing. One respondent saw the routing, but chose to ignore it. This respondent crossed out ‘English’ and wrote a big question mark following the question. They said:

‘Bach wedi cymhlethu efo cwestiwn 18. Ewch i 20, swm i fod i neidio wedyn ... ond be’ di’r pwynt gofyn y cwestiwn ... ond ddim wedi stopio fi rhag ateb y cwestiwn, na bod yn “rebellious”.’	[A bit confused by question 18. Go to 20, I was supposed to jump then ... but what’s the point of asking the question ... but didn’t stop me from answering the question, nor from being “rebellious”.]
--	---

(Wave 2 Welsh language testing)

When asked how they would feel about answering the English proficiency question, most respondents replied that they would not mind answering it. Some were confused as to why they were not asked to answer it. These respondents said that they wanted to answer it, as they thought it was a useful question to ask Welsh speakers.

A couple of Welsh-speaking respondents commented that there might be many Welsh speakers whose English language ability is poor, and they should be able to show this:

'Reit berthnasol i lot o bobl.'	[Very relevant to a lot of people]
---------------------------------	------------------------------------

(Wave 2 Welsh language testing)

'... dw i ddim yn deall pam dw i ddim yn ateb hwnna. Mae 'na Gymry Cymraeg ... hen bob l... yn y wlad heddiw sydd ddim yn dda efo'r Saesneg...maen nhw'n baglu efo'r Saesneg.'	[... I don't understand why I'm not answering that. There are Welsh speaking Welsh people ... old people ... in the country today who aren't very good with English ... they stumble with their English.]
--	---

(Wave 2 Welsh language testing)

ONS removed the routing for the second part of Wave 6 testing. None of the respondents with English as a first language objected when required to complete the proficiency question. However, a few people did laugh or comment when they first saw it. Some were confused about why they were asked how well they spoke English when it was their first language.

For some respondents with English as a first language, this question seemed to reflect a class division between people:

'A posh person would speak it very nicely and well, do you know what I mean? Somebody like me, or even worse [laughs], doesn't pronounce their ts, doesn't pronounce their hs, I mean 'ark at me!' (Wave 6)

ONS also asked the questions without routing in the June 2008 module of the Opinions Survey. This found that 6 per cent (un-weighted) of people with English as their main language identified as 'Well' (Table 6.30).

Table 6.30: Level of proficiency in English¹
England and Wales

		Percentages		
		Main language English	Main language Other	Total
How well respondent thought they speak English	Very well	93	21	90
	Well	6	57	9
	Not well	-	21	1

¹ Based on 532 responses. Data are un-weighted.

Source: Opinions Survey June 2008 module

Because people who speak English as their first language might interpret the proficiency question as a measure of social class, ONS decided to retain the routing on the main language question. This means that those with English as a first language will not be required to answer the proficiency question.

6.3.3 Conclusions for England and Wales language question

The table in Annex A summarises how all the types of language questions tested meet the essential and desirable requirements outlined in Section 2. The two-part main language question was the only one to meet all the essential requirements, and met four of the ten desired elements.

Language question recommendations

It is recommended that the language question is phrased 'What is your main language?' with tick-boxes for 'English' (for England), 'English or Welsh' (for Wales on English language questionnaire) and 'Other, write in (including British Sign Language)' followed by space for a write-in response. Respondents who tick 'English' or 'Welsh or English' will be routed past the second part of the question, which asks 'How well can you speak English?' with tick-boxes for 'Very well/Well/Not well/Not at all'.

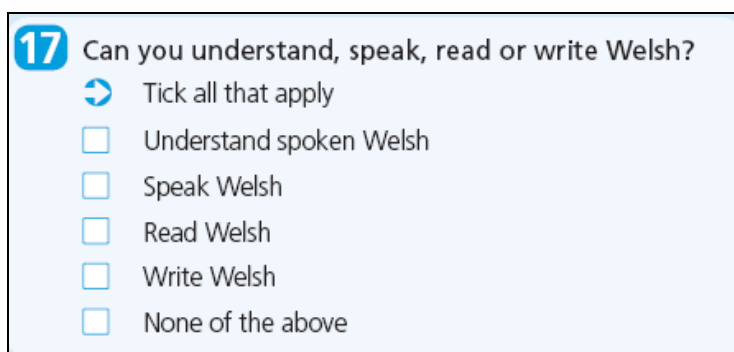
It is recommended that the Welsh language version of the questionnaire uses the same question phrasing as the English language version 'Beth yw'ch prif iaith?' ['What is your main language?'] with tick-boxes for 'Cymraeg neu Saesneg' [Welsh or English] and 'Arall, nodwch (gan gynnwys iaith Arwyddion Prydain)' [Other write in (including British Sign Language)] followed by space for a write-in response. Respondents who tick 'Cymraeg neu Saesneg' will be routed to the next question, skipping the second part of the question, which asks 'Pa mor dda allwch chi siarad Saesneg?' [How well can you speak English?] With tick-boxes for 'Da iawn/Da/Ddim yn dda/Dim o gwbl' [Very well/Well/Not well/Not at all]

6.4 Questions for Wales only

6.4.1 Knowledge of Welsh– can you understand, speak, read or write Welsh?

Overall, the 2001 knowledge of Welsh question worked well in Wales, and the data were suitable for user requirements. ONS cognitively tested the same question in Waves 5 and 6 (Figure 6.31).

Figure 6.31: Knowledge of Welsh question (in English)



17 Can you understand, speak, read or write Welsh?

➔ Tick all that apply

Understand spoken Welsh

Speak Welsh

Read Welsh

Write Welsh

None of the above

The main concerns raised by ONS during the question testing work were that the question did not state what level of proficiency was required. Respondents were unsure how they should answer if they were not fluent but had some ability.

‘Now this is a difficult question to answer for me. I can do a little bit of all of these depending on what it is, but in a census would I answer? It’s almost like I want to be able to qualify this with a “only a bit”. ‘Cos the “None” implies none and I don’t, I don’t know how to answer this question.’ (Wave 6)

Similar difficulties led respondents to provide inconsistent answers, based on whether or not they believed their ability was good enough for them to report it.

Some respondents, for example, mentioned their knowledge of reading basic Welsh, including road signs. One respondent decided to tick the ‘read Welsh’ box after commenting:

‘I can read a bit of Welsh, shall I put that? Only, like road signs [laughs], and “police” and stuff like that.’ (Wave 5)

Contrastingly, another respondent did not appear to consider that this level of proficiency was sufficient to be able to report it and decided to tick ‘None of the above’ after commenting:

‘By reading “Understand, speak, read”, I think I take that as being able to do it at a level where you can pick up a very basic newspaper and understand what’s going on, rather than pick up certain words, ‘cos that’s not really reading’. (Wave 5)

As a result of the inconsistent responses, ONS data collection methodology advised that this question collected inconsistent data.

The Welsh language version of the question was also cognitively tested (Figure 6.32).

Figure 6.32: Knowledge of Welsh question (in Welsh)

17	A allwch ddeall, siarad, darllen neu ysgrifennu Cymraeg?
<input checked="" type="checkbox"/>	Ticiwch bob blwch sy'n berthnasol
<input type="checkbox"/>	Deall Cymraeg llafar
<input type="checkbox"/>	Siarad Cymraeg
<input type="checkbox"/>	Darllen Cymraeg
<input type="checkbox"/>	Ysgrifennu Cymraeg
<input type="checkbox"/>	Dim un o'r uchod

All but one respondent ticked all four answer categories. This one respondent ticked all the options apart from 'writing Welsh', and had not realised that they had left out the last option.

Generally, Welsh speaking respondents felt that most people ticking the boxes would be 'fluent'. Some respondents felt that there should be a tick-box for learners who can do all four answer categories, but who might not tick these boxes as they don't consider themselves to be 'fluent'. Others commented that the question does not ask for level of ability, and that this might result in low statistics for Welsh language use:

<p>'Dyw e ddim yn gofyn pa mor dda ych chi'n gallu gwneud nhw ... neu i ba safon. Jyst beth ych chi'n teimlo sy'n iawn ... wi'n credu bydd pobl sydd ddim yn hyderus ddim yn ticio'r bocsys yna ... a bydd yr ystadegau'n is.'</p>	<p>[It doesn't ask how well you can do them ... or to what standard. Just, what you think is right ... I think people who are lacking in confidence won't tick those boxes ... and the statistics will be lower.]</p>
--	---

(Wave 2 Welsh language testing)

Most of the Welsh-speaking respondents understood the reason for including all four modes of communication – understanding, speaking, reading and writing:

<p>'Ma 'na bobl sy'n gallu siarad Cymraeg sydd ddim yn gallu ysgrifennu Cymraeg.'</p>	<p>[There are people who can speak Welsh who can't write Welsh.]</p>
---	--

(Wave 2 Welsh language testing)

<p>'Dw i'n meddwl fod splitio nhw ...yr opsiynau ... yn beth iawn i wneud achos bod mam fi'n deall Cymraeg, ond ma hi ddim yn siarad Cymraeg.'</p>	<p>[I think that splitting them ... the options ... is the right thing to do, because my mum understands Welsh, but doesn't speak Welsh.]</p>
--	---

(Wave 2 Welsh language testing)

'Deall Cymraeg llafar' [understanding spoken Welsh] was generally understood to mean understanding a 'sgwrs go iawn' [a real conversation], and 'dilyn sgwrs yn iawn' [following a conversation properly], and being able to communicate fluently. One respondent was not familiar with the term 'llafar' [oral/spoken], but guessed the meaning to be 'deall rhywun yn siarad efo ti' [understanding someone talking to]. They said that their friends would probably not understand 'llafar' either.

A couple of respondents with young children (aged three) said that they would tick that their children could speak and understand Welsh, but could not read and write Welsh. In the recent Welsh language Whole Questionnaire Testing, respondents with very young babies were hesitant about which option to tick. A respondent with a three month old baby decided to opt for 'None of

the above', as although the respondent talks to the baby in Welsh, and the baby responds and smiles, the baby isn't able to do much else. Whereas a respondent with a year old baby ticked the option 'Understands spoken Welsh', as this respondent felt that the baby understood everything said in Welsh. In 2001 children under the age of three were coded out of Welsh language data in outputs (Office for National Statistics 2004).

The main priority and user requirement in terms of the knowledge of Welsh question was to ensure consistency with the information collected in 2001, and therefore there was strong support to continue asking exactly the same question in 2011.

It is important to note that the introduction of the national identity question in the 2011 Census, two questions before the knowledge of Welsh language question, may make respondents more likely to say they have Welsh language skills. Estimates from the Welsh Household Panel Survey (WHPS) and the Welsh Local Labour Force Survey (WLLFS) since 1997 and the November 2000 module of the Welsh Omnibus survey were appreciably higher than from the 2001 Census. On the WLLFS and Omnibus Survey, respondents had been given the explicit opportunity to say they were Welsh in the national identity question earlier on in the questionnaire and on the WHPS respondents had just been given the opportunity to state their parents' country of birth. Haseldon (2003) suggests that it is possible that having just said they were Welsh or that their parents were born in Wales, the respondents felt more inclined to say that they had Welsh language skills.

6.4.1.1 Collecting knowledge of Welsh data in England

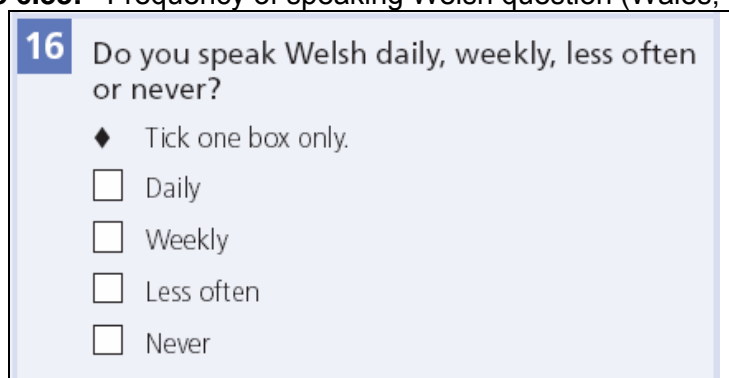
Information on Welsh language knowledge in England has never been collected in previous censuses. Historically, while the Wales questionnaire has included a question about Welsh, the equivalent space on the England questionnaire was left blank.

As part of its 2007 consultation, ONS asked whether users needed information on the Welsh language in England. There was not a strong user need for this data from English users, but several respondents from Wales strongly supported collecting such data as there is an ever-increasing mobile population, mainly in order to assess demand for Welsh language resources. There is similar, and in some cases greater, demand for collecting information on languages other than Welsh. Therefore, information on Welsh language will not be collected specifically on the England questionnaire which will have a blank question in the corresponding space. However, the main language question will collect some information in England on Welsh as a main language.

6.4.2 Frequency of speaking Welsh

ONS tested a question about the frequency of speaking Welsh.

Figure 6.33: Frequency of speaking Welsh question (Wales, England)



16 Do you speak Welsh daily, weekly, less often or never?

◆ Tick one box only.

Daily

Weekly

Less often

Never

Respondents answered the question without any problems. However, many seemed puzzled to be asked this. One respondent, in the Welsh language testing, who was fluent in Welsh, ticked 'Less

often', and commented that if there had been an 'Every fortnight' option, they would have ticked that. All other respondents ticked that they spoke Welsh 'every day'.

ONS introduced the same question to the Labour Force Survey in Wales in April 2007. As a result information on the frequency of use of Welsh will, in due course, become available but only at a local authority level. The higher prioritisation of other questions meant that this question was not recommended for inclusion in the 2011 Census.

Language question recommendations

It is recommended that the question 'Can you understand, speak, read or write Welsh?' is asked in Wales only. It is recommended that a 'Tick all that apply' instruction is included along with tick-boxes for 'Understand spoken Welsh', 'Speak Welsh', 'Read Welsh', 'Write Welsh' and 'None of the above'. The equivalent space on the questionnaire in England will read 'This question is intentionally left blank'.

It is recommended that the Welsh language version of the question asks 'A allwch ddeall, siarad, darllen neu ysgrifennu Cymraeg?' ['Can you understand, speak, read or write Welsh?']. It is recommended that a 'Ticiwch bob blwch sy'n berthnasol' ['Tick all boxes which are relevant'] instruction is included, along with tick boxes for 'Deall Cymraeg llafar' ['Understand spoken Welsh'], 'Siarad Cymraeg' ['Speak Welsh'], 'Darllen Cymraeg' ['Read Welsh'], 'Ysgrifennu Cymraeg' ['Write Welsh'] and 'Dim un o'r uchod' ['None of the above'].

7. Recommended location of the question

The questions on national identity, ethnic group, language and religious affiliation are all linked and therefore should be presented together as a suite of questions. To maintain respondent familiarity with the questionnaire, it is recommended that the question order used for the topics of ethnicity, identity, language and religion in the 2001 Census is retained unless this conflicts with any recommendation:

1. Ethnic group
2. Knowledge of Welsh language (Wales only)
3. Religious affiliation

The main language and English proficiency questions will be placed adjacent to the knowledge of Welsh language question in accordance with the aim of situating similar topics together. ONS have also considered the order of the three language questions. If the new questions were placed first they could potentially influence responses to the knowledge of Welsh language question and thereby impact on the comparability of the latter with 2001 data. For this reason it is recommended that the new questions are asked immediately after the knowledge of Welsh language question. The recommended order for the topics of ethnicity, identity, language and religion in the 2011 Census is:

1. National identity
 2. Ethnic group
 3. Knowledge of Welsh language (Wales only)
 4. Main language
 5. Spoken English proficiency
 6. Religious affiliation
-

8. Final recommendations

The language question development process has produced questions that are recommended for the 2011 Census subject to approval by Parliament through the legislative process. The images below present the recommended questions for England, Wales (in English) and Wales (in Welsh).

Figure 8.1: Final recommended language questions for England

17 This question is intentionally left blank ➔ Go to 18

18 What is your main language?

English ➔ Go to 20

Other, write in (including British Sign Language)

19 How well can you speak English?

Very well Well Not well Not at all

Figure 8.2: Final recommended English language version of language questions for Wales

17 Can you understand, speak, read or write Welsh?

➔ Tick all that apply

Understand spoken Welsh

Speak Welsh

Read Welsh

Write Welsh

None of the above

18 What is your main language?

English or Welsh ➔ Go to 20

Other, write in (including British Sign Language)

19 How well can you speak English?

Very well Well Not well Not at all

Figure 8.3: Final recommended Welsh language version of language questions for Wales

17 A allwch ddeall, siarad, darllen neu ysgrifennu Cymraeg?
➔ Ticiwch bob blwch sy'n berthnasol

Deall Cymraeg llafar

Siarad Cymraeg

Darllen Cymraeg

Ysgrifennu Cymraeg

Dim un o'r uchod

18 Beth yw'ch prif iaith?

Cymraeg neu Saesneg ➔ Ewch i **20**

Arall, nodwch (gan gynnwys iaith Arwyddion Prydain)

19 Pa mor dda allwch chi siarad Saesneg?

Da iawn Da Ddim yn dda Dim o gwbl

9. References

- Bourne S, Davies C A, Fearn V, Maclean E, Nicholls Z and Potaka L (2006) *Whole Questionnaire Testing for the 2007 Census Test*. Office for National Statistics, August 2006
- Daffin J, Dalton S, Davies C A, Draper K, Mathias J, Monkman M, Wallis R and Wardman L (2008) *Wave 6 Testing Report*, Office for National Statistics, July 2008
- Davies C A, Nicholls Z, Monkman M, Potaka L and Tinkler L (2007a) *Wave 1 Testing Report*. Office for National Statistics, February 2007
- Davies C A, Nicholls Z, Tinkler L, Potaka L, Monkman M and Palmer K (2007b) *Wave 2 Testing Report*. Office for National Statistics, June 2007
- Davies C A, Keenan C, Jezard K, Tinkler L, Potaka L, Mathias J, Monkman M and Wallis R (2007c) *Wave 3 Testing Report*. Office for National Statistics, October 2007
- Davies C A, Keenan C, Jezard K, Tinkler L, Mathias J, Monkman M and Wallis R (2007d) *Wave 4 Testing Report*. Office for National Statistics, November 2007
- Davies C A, Jezard K, Tinkler L, Mathias J, Monkman M, Wallis R, Wardman L and Watson J (2008) *Wave 5 Testing Report*. Office for National Statistics, March 2008
- Dillman D A (2000) *Mail and Internet surveys: The tailored Design Method*. New York: John Wiley & Sons
- Extra G and Yağmur (eds) (2004) *Urban Multilingualism in Europe: Immigrant Minority Language at Home and School*
- Extra G and Gorter D (eds) (2008) *Multilingual Europe: Facts and Policies*. Berlin/ New York: Mouton de Gruyter
- Fearn V and Nicholls Z (2006) *Lambeth Small-scale Test: An Analysis of Respondent Behaviour*. Office for National Statistics, September 2006
- Haselden L (2003) *Differences in estimates of Welsh Language Skills*. Office for National Statistics. Available at http://www.statistics.gov.uk/downloads/theme_compendia/fow/WelshLanguage.pdf
- Johnson M (2000) *Black and Minority Ethnic Groups in England: The Second Health and lifestyles Survey*. Health Education Authority
- Jones H (2005) 'Ability to speak Welsh in the censuses of population: a longitudinal analysis' in *Population Trends* 122. Winter 2005. Available at www.statistics.gov.uk/downloads/theme_population/PopTrends122v1.pdf
- Kominski R (1989) *How Good is 'How Well'? An Examination of the Census English-Speaking Ability Question*. US Bureau of the Census. Available at <http://www.census.gov/population/socdemo/language/ASApaper1989.pdf>
- Maclean E (2006) *Summary Testing Report: Ethnicity, Identity, Language and Religion (Preliminary Testing)* Office for National Statistics, April 2006
- Mathias J, Davies C A, Jezard K, Monkman M, Tinkler L, Wallis R, Wardman L (2008a) *Welsh Language Questionnaire Development: Wave 1 Testing Report* Office for National Statistics, January 2008
- Mathias J, Davies C A, Dalton S, Jezard K, Monkman M, Wallis R, Wardman L (2008b) *Welsh Language Questionnaire Development: Wave 2 Provisional Testing Report*. Office for National Statistics, July 2008
- McPake J (2002) *Mapping the languages of Edinburgh*. Scottish National Centre for Languages. Available at <http://www.scilt.stir.ac.uk/publications/>
-

Office for National Statistics (2004) *Census 2001: Report on the Welsh language*. Available at www.statistics.gov.uk/downloads/census2001/Report_on_the_Welsh_language.pdf

Office for National Statistics (2007a) *2011 Census ethnic group, national identity, religion and language consultation: Summary report on responses to the 2011 Census stakeholders consultation 2006/07*. Available at <http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

Office for National Statistics (2007b) *2011 Census ethnic group, national identity, religion and language consultation: Experts, community and special interest group responses to the 2011 Census stakeholders consultation 2006/07*. Available at <http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

Office for National Statistics (2007c) *2011 Census ethnic group, national identity, religion and language consultation: Central and devolved government responses to the 2011 Census stakeholders consultation 2006/07*. Available at <http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

Office for National Statistics (2007d) *2011 Census ethnic group, national identity, religion and language consultation: Local and regional government responses to the 2011 Census stakeholders consultation 2006/07*. Available at <http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

Office for National Statistics (2007e) *2011 Census ethnic group, national identity, religion and language consultation: Local service provider responses to the 2011 Census stakeholders consultation 2006/07*. Available at <http://www.ons.gov.uk/census/2011-census/consultations/eth-group-nat-iden/index.html>

Rudat K (1994) *Black and Minority Ethnic Groups in England*. Health Education Authority

Siegel P, Martin E and Bruno R (2001) *Language Use and Linguistic Isolation: Historical Data and Methodological Issues*. US Census Bureau February 12, 2001. Available at http://www.census.gov/population/www/socdemo/lang_use.html

Welsh Language Board (2008) *The Welsh Language Use Surveys of 2004-06*. Available at <http://www.byig-wlb.org.uk/English/publications/Publications/Welsh%20Language%20Use%20Surveys%202004-06.pdf>

Annex A: Summary of suitability of language questions: by type of question

Requirement		Sectional	Full matrix	Simplified matrix	Conversation	Two-part: home language	Two-part: main language
Essential	Fits available space		✓	✓	✓		✓
	Captures spoken language ¹	✓	✓	✓	✓	✓	✓
	Captures signing ¹	✓	✓	✓	✓	✓	✓
	Allows data users to understand whether translation is needed and if so in which language ¹				✓	✓	✓
	Consistently excludes people with low English language proficiency when capturing English speakers ¹				✓	✓	✓
	Question is generally clear to respondents and is understood as intended			✓			✓
	No impact on comparability of Welsh language question	✓	✓	²		✓	✓
Desirable	Consistently excludes people with low language proficiency when capturing speakers of other languages ¹						✓
	Captures English proficiency ¹					✓	✓
	Captures proficiency in other languages ¹						
	Allows users to distinguish languages used for translation/ learnt at school, etc ¹						✓
	Captures multiple languages known ¹	4 or 6 ³	3, 4 or 5 ³	4 or 5 ³	3 or 7 ³	2 or 3 ⁴	2 ⁵
	Captures written language knowledge ¹	✓	✓	✓			
	Captures knowledge of understanding spoken language ¹	✓	✓	✓			
	Captures knowledge of reading languages ¹	✓	✓	✓			
	Identifies respondents' main or preferred language						✓
	Captures where and how languages are used ¹						

¹ If answered as intended – but may be compromised if question is not clear

² Version 3 thought not to be comparable with 2001 data

³ Including English and British Sign Language, and in most cases Welsh

⁴ Including English

⁵ Second language only recorded if it is main language or language spoken at home