

Methodology for decision making on the 2021 Census sex question concept and associated guidance

10 February 2021

Helena Rosiecka, Head of 2021 Census Question and Questionnaire Design

Terminology

In this paper, for brevity, the population

- where the gender they identify with is different to their sex registered at birth is referred to as “the trans population”
- who have variation in sex characteristics, sometimes referred to as ‘intersex’ or ‘differences in sex development’, are referred to as the “VSC/intersex population”

ONS is aware that not all members of these populations identify with this terminology.

In addition, in this paper, the concepts of sex evaluated retain the core aspect of being male or female, but take into account aspects of biological, physical, legal, and social identity.

Referencing

A full description of each test undertaken during the development of Census 2021 questions and questionnaires can be found in [summary of testing for Census 2021](#).

References to tests take the form (Year: Test number). “Year” refers to the calendar year the test was undertaken in and the test number is the position of the test within the year considering all testing that took place in that year. For example, the fifth test conducted in 2017 would be referenced as (2017:5).

Purpose

This paper outlines the evaluation of the concept to be collected from the sex question on Census 2021 in England and Wales. It describes what this concept will be.

This work was initiated following the 2019 Rehearsal (2019:15) when stakeholders challenged both the concept of sex ONS intended to collect in the 2021 census and the accompanying guidance.

Executive summary

This paper outlines the evaluation of the concept to be collected from the sex question on Census 2021 in England and Wales. The core purpose of the census is to provide accurate data on the size and characteristics of the population. To achieve this, when considering changes to the questionnaire, demonstrated user need is followed by further work on other considerations, including public acceptability, to ensure as many people as possible across all population groups find it acceptable and easy to answer census questions. ONS has set target response rates of 94% nationally and 80% locally in all local authorities. Meeting these targets are key to the quality objectives of the census.

Section 1: The concepts of sex which could be collected from Census 2021

The Census 2021 question on sex is “What is your sex?” with response options of “Female” and “Male”. Sex can be defined in different ways. This paper considers five concepts of sex that have been identified. All are valid concepts for data to be collected on, this evaluation is in the context of the Census. These are sex as: registered at birth, recorded on birth certificate, recorded on legal/official documents, living / presenting, and self-identified.

Section 2: The Census 2021 data users need

Four key uses of sex data from Census 2021, plus wide-ranging needs for aggregate data tables, were identified during the consultation and engagement. These are for:

1. population estimates and projections and household projections;
2. 2021 Census operational needs, including coverage adjustment process;
3. equality monitoring; and
4. analysis and linkage of ONS Longitudinal Study and microdata sample files.

There are separate needs for data on gender identity, and a separate question “Is the gender you identify with the same as your sex registered at birth?” with response options “Yes” and “No, write in gender identity” is included later in the form. The key Census 2021 sex data user needs have been evaluated in relation to relevance, and comparability and cohesion.

Relevance

Each identified data need has been assessed against current use, and potential use, in relation to existing methodologies, time series and admin data sources. By considering the sources and methods of collecting data, and uses of the data, it is concluded that the most evidenced needs are for sex recorded on legal/official documents and sex as living / perceived.

Comparability and coherence

User consultation responses focussed on a need for continuity over time and across the UK. We would aim for data collected to be comparable with all UK censuses and other social surveys, coherent with the wording of the 2021 census question (which asks ‘what is your sex’) and the other uses of the word “sex” in the census questionnaire.

By considering the approach across UK census offices and on UK surveys, it is concluded that the most evidenced need in relation to this dimension of quality is sex recorded on legal/official documents, with similar, but lower, levels of need for sex as living / perceived.

Section 3: Impact on Census overall

The scale of the impact of collecting any specific concept of sex cannot be determined as multiple factors are at play. Both coverage and response rates are important as low response rates could mean a longer field period or extended estimation processes, and larger coverage adjustment could lead to wider confidence intervals around the estimates. However, 6 per cent of the total

population of England and Wales (3.8 million people) did not respond to the 2011 census and were adjusted for during data processing. In comparison, the trans population is small, estimated to be between 0.3 and 0.75% of the UK population.

The impact on Census 2021 has been assessed in relation to accuracy and reliability. We evaluate different concepts of sex in relation to impact on overall census response, sex question response, and individuals right to privacy.

Accuracy and reliability

This evaluation considers the 2021 Census context, for example that it is a mandatory household survey administered by online or paper self-response by one or more household members. Two strands of additional research were completed using best practice qualitative research methods. These included in-depth interviews with trans people, traditional women's groups and the general population and a review of the evidence base, to help understand what sex concept will:

- maximise Census responses from the trans population / the cisgender population
- maximise sex question responses from the trans population / the cisgender population
- ensure any perceived invasion of privacy is justified by the data need

Sex is a key variable to ensure a genuine census record which, if not completed, could lead to a record being removed during processing.

The highest negative impact is anticipated if sex registered at birth was collected due to reduction in census response. The data need is not considered to be strong enough to justify collecting this.

The least impact is anticipated if sex as living / presenting is collected. However, the adjacent concept of sex recorded on legal/official documents is also anticipated to have low impact.

Section 4: Conclusion

The evaluation of the evidence presented supports the collection of sex as registered on legal/official documents in Census 2021 for England and Wales.

Section 5: Guidance to support collection

Guidance will seek to minimize barriers to completion of the census and the sex question. To do this it needs to balance the following aims:

- direct people in how to answer;
- be considerate of people that have identities other than male or female or have variations of sex characteristics / are intersex;
- recognize that some respondents may have legal/official documents that state male and legal/official documents that state female; and
- use terminology that is acceptable to the whole population.

Contents

Terminology	1
Referencing	1
Purpose	1
Executive summary	2
Contents	4
Introduction	5
Section 1: Sex concept options for Census 2021	6
Section 2: Census 2021 data user needs for sex concept	6
2.1 Background – Consultation and engagement	6
2.2 Relevance	8
2.3 Comparability and coherence	13
Section 3: Impact on Census 2021 overall	16
3.1 Background – Scale of impact	16
3.2 Accuracy and reliability	17
Section 4: Conclusion	20
Section 5: Developing guidance to support collection	21
Section 6: Next steps	21
Section 7: Annex's	22
Annex A: Census 2021 sex question and route to guidance	22
Annex B: Publications and meetings relating to sex and gender identity	23
Annex C: Detailed responses from ONS users of sex data from Census 2021	29

Introduction

Census 2021 impact

The core purpose of the census is to provide accurate data on the size and characteristics of the population. In the 2018 [White Paper](#) we set out our quality objectives, which include:

- response rate targets of 94% nationally and 80% locally in all local authorities
- nationally accurate as measured by a confidence interval of plus or minus 0.2%, with bias less than 0.5% for England and Wales
- high-quality locally with 95% confidence intervals for all local authorities of plus or minus 3%

To achieve this, when considering changes to the questionnaire, demonstrated user need is followed by further work on other considerations, including public acceptability, to ensure as many people as possible across all population groups find it acceptable and easy to answer census questions. This is demonstrated in the [2021 Census topic consultation evaluation criteria](#).

Sex question

“What is your sex?” is a core census question which has been asked since 1801. The questions and response options for Census 2021 have been finalised through the [census secondary legislation](#).

Most people answer ‘What is your sex?’ question on surveys and applications for legal/official documents as a matter of course. As a result, ONS has not previously publicly defined the concept of sex it is aiming to capture. In 2001 and 2011 the population who inquired how trans people should answer were told that they should answer the sex question however they believe is correct. In 2011, this messaging was [included in the online help](#), although this was not highly used. Testing, including the 2019 Rehearsal (2019:15), showed that members of the trans population use a range of answering strategies (e.g. 2017:6, 2017:15, 2018:3, 2018:6, 2019:12, 2020:6).

For Census 2021 the question on sex and response options will remain the same as in 2011, except:

- The two response options are “Female” then “Male”
- There will be a guidance note explaining “A question about gender identity follows if you are aged 16 or over” ([paper form version](#))

The guidance note on the sex question is aimed at making the sex question more acceptable to the non-binary population, rather than the trans population with a binary identity.

The question on gender identity is new for 2021, it reads “Is the gender you identify with the same as your sex registered at birth”. The question is asked of those aged 16 and over and is voluntary.

The research that led to the final Census questions is available in the report “[Sex and gender identity question development for Census 2021](#)”. Census 2021 sex question and route to supplementary guidance is shown in Annex A.

Section 1: Sex concept options for Census 2021

The word “sex” within the question “What is your sex?” can be understood in different ways, as shown in table 1.

Table 1: Concepts of sex

Concepts of “sex”	Potential respondent answers
Registered at birth	Female / Male ¹
Recorded on birth certificate	Female / Male
Recorded on legal/official documents ²	Female / Male
Living / presenting	Female / Male / Other
Self-identified (often referred to as gender / gender identity)	Female / Male / Other

These are listed in order of expected volume of people, from lowest to highest, for whom the answer would be different from sex registered at birth.

For example, the Gender Recognition Act (2004) put in place a process to make it legally possible to change the sex on your birth certificate via a [Gender Recognition Certificate](#). To date, around 6,000 people have done so. Although the gender recognition process seeks to safeguard a person’s privacy, they do not have a right never to disclose the fact that they obtained a GRC.

The last three concepts may not have a single answer, for example because:

- the person has changed some, but not all legal/official documents,
- the person is presenting in different ways in different contexts, or
- the person does not identify as either female or male, so which they most closely identify with may change depending on context

Section 2: Census 2021 data user needs for sex concept

This section considers two dimensions of quality in the [European Statistical System](#).

- **Relevance:** The degree to which statistics meet current and potential user needs in both coverage and content (*here we consider content*).
- **Comparability and coherence:** Comparability is the degree to which the data can be compared over time and domain. Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar.

The background section on consultation and engagement summarises how we identified the key needs for data on sex. We then evaluate the data needs in relation to relevance, including ONS operational needs, and in relation to comparability and coherence.

2.1 Background – Consultation and engagement

Consultation

To ensure we collect the characteristics that are needed by data users, we conducted the [2021 Census topic consultation](#) in 2015. This was a topic level consultation and received 1,095 responses

¹ A very small number of people with VSC/intersex may be classified as “indeterminate” in their first weeks of life. These are reclassified as either male or female unless the infant died before this took place.

² This concept refers to UK issues legal/official documents.

of whom 249 provided a response on their data need for demographic information, which included the sex question.

The consultation responses highlighted the user need for continuity and a high response rate. None of the consultation respondents defined the concept of sex to be collected in their initial response.

The key needs identified were for population estimates and projections and household projections, 2021 Census operational needs, including coverage adjustment process, equality monitoring, and analysis of ONS Longitudinal Study and microdata sample files.

In some cases, they stated that gender rather than sex should be collected.

In most cases, the request was not specific enough to identify a specific concept because cultural as well as biological factors impact the proposed research. Similarly, in the previous [Beyond 2011 consultation \(2014\)](#), and the subsequent [Census 2021 Output Content Design consultation \(2018\)](#) sex is proven as a key variable, but no specific concept has been requested.

Engagement

Since the 2015 consultation, ONS has continued to engage with a range of stakeholders regarding the sex and gender identity questions. Much of this is reported on in the publications listed in Annex B. This annex also provides further information on engagement with women's groups that has been ongoing since September 2017.

In December 2019, 80 social statisticians and users of population level data signed a letter to the directors of UK census authorities with concerns about the sex question guidance. This outlined concerns with the concept of sex to be collected around understanding:

- sex-based discrimination
- different health and socio-economic outcomes between men and women

ONS met with some of the signatories of the letter in January 2020 and subsequently held a roundtable with a wide range of academics who use data on sex in June 2020.

The roundtable showed that there were a range of data needs amongst the expert academic community and no consensus was reached. Therefore, the ONS undertook further research, including follow-up meetings with groups of academics with different views to understand further their data requirements.

2.2 Relevance

Four key uses of sex data from Census 2021, plus wide-ranging needs for aggregate data tables, were identified during the consultation and engagement. These are:

- Population estimates and projections and household projections
- 2021 Census operational needs, including coverage adjustment process
- Equality monitoring
- Analysis and linkage of ONS Longitudinal Study and microdata sample files

Evaluation questions related to relevance

Each identified data need has been assessed against current use, and potential use, in relation to the following questions: What would we aim to collect:

- considering the existing methodologies?
- if we did not already have a time series?
- for comparability with administrative data sources in the future?

Population estimates and projections and household projections

The relevant ONS methodology teams have been consulted for their assessment of the impact of the sex question on the population estimates and projections and housing projections. Their response is provided in full in Annex C. In summary these outputs use a range of information, including:

- (i) Census estimates: This is interpreted as self-identified sex both in 2011 and before, although in 2001 guidance to that effect was only available on request, and prior to that no guidance was provided.
- (ii) Birth registrations: This is the only key source that is explicitly sex registered at birth and is collected directly from the register of births.
- (iii) Death registrations: The sex is provided by the person registering the death, so it may not be sex registered at birth.
- (iv) International Passenger Survey³: The interviewer typically records what sex they believe the respondent to be.⁴
- (v) GP Patient Register: Here applicants are invited to tick whether they are “Male” or “Female”, with no indication of the definition to be used (or even whether it is sex or gender which is being asked for).
- (vi) Annual Population Survey. There is no guidance, but we expect the interviewer typically records what they believe the respondent to be.
- (vii) Ministry of Justice prisoner data. For the purposes of the projections, all people in men’s prisons were counted as male, and all those in women’s prisons as female in the data supplied⁵. [Location within the prison estate is decided](#) by an initial local transgender case board or escalated to other boards if required. All decisions are made on a case by case basis.

³ Data collection for the International Passenger Survey halted in March 2020 due to coronavirus (COVID19) pandemic and restarted in December 2020.

⁴ In future, international migration statistics will be determined by administrative sources. While research on this is ongoing, it is clear that the sex record in these may also not match sex at birth (for example, work visa applications require a passport as evidence but, dependent on where the migrant has come from, we know this may not be their sex at birth).

⁵ Peterborough Prison is the only prison within England and Wales with both male and female residents.

Overall, sex in the source data is largely undefined and can be respondent recorded or recorded by a proxy such as an interviewer. The dominant source is the 2011 Census data and, apart from birth registrations, the other sources may also include people with sex recorded differently from their sex registered at birth.

In addition:

- the projections need the best possible census coverage of the population of England and Wales, with undercount minimised and the aim to meet the national completion target of at least 94%.

ONS plan to develop a revised back series of population estimates for the period 2011 to 2021, which in turn will inform projections. For this, methodological consistency over time, and at either end of the revised series, would be optimal.

Table 2: Assessment of population and household estimates and projections needs

What would we aim to collect:	Concept	Justification
<ul style="list-style-type: none"> • considering the existing methodologies? 	Self-identified OR Living / presenting	The 2011 Census is the main source used and has been interpreted as collecting self-identified sex. However, apart from birth registrations, most other data sources used collect self-reported sex in a context where the individual would not need to disclose their gender history. Where sex is not self-reported, the sex is recorded by proxy, and therefore the sex as living / presenting is most likely to be recorded as the interviewer would not know the respondents' gender history.
<ul style="list-style-type: none"> • if we did not already have a time series? 	Registered at birth	Data is used to project the population size going forward, for this the female fertility rate is calculated. As a trans woman cannot give birth, but a trans man may be able to, using sex registered at birth is optimal. However, due to the range of other social and biological factors impacting population growth any improvement in accuracy is uncertain and likely to be minimal.
<ul style="list-style-type: none"> • for comparability with administrative data sources in the future? 	Recorded on legal/official documents	Binary concept of sex is collected in administrative sources. Whilst not all sources of data collect the same definition of sex, legal/official documents is the nearest equivalent.

2021 Census operational needs, including coverage adjustment process

The relevant ONS methodology teams have been consulted for their assessment of the impact of the sex question on the 2021 Census estimates. Their response is provided in full in Annex C. In summary the need is for:

- **Non-response and multi-ticking:** High levels of either of these would negatively impact the statistical imputation processes as sex is a key predictor variable. There would be a greater risk of bias in imputation of a number of variables. The 'Remove False Persons' process uses sex as a key variable, so high levels of missing information could lead to more records being removed, although the method could be adjusted to use a different variable instead.

- **Collection of the same concept in Census 2021 and Census Coverage Survey:** Matching the Census to the Census Coverage Survey is key for coverage estimation and imputation. The sex variable is used mainly to indicate a non-match when they differ, although matches can be made automatically when sex is different or missing. Our assessment is that there will be no matching bias caused by collecting different sex concepts.

The census coverage survey is interviewer led and the sex question is asked. If queried the interviewer would access the main census question guidance. However, experience from ONS social surveys show that in a face-to-face interview the concept is rarely questioned – respondents answer as they wish. Three instances of interviewers being asked how to answer were raised in the LFS in the last 3 years. In the face-to-face Census coverage survey respondents respond as they wish. It is likely that sex as living / presenting will be collected as no reason to provide sensitive personal data is provided.

Table 3: Assessment of 2021 Census operational needs, including coverage adjustment process

What would we aim to collect:	Concept	Justification
<ul style="list-style-type: none"> • considering the existing methodologies? 	Living / presenting	Binary concept of sex is collected, which is key to existing methodologies, and aligned with answering strategies of non-binary participants in testing. Although some people may have documents with contrasting sexes – this is more aligned with sex as living / presenting, which is how CCS respondents are likely to answer.
<ul style="list-style-type: none"> • if we did not already have a time series? 	Recorded on birth certificate	Existing methodologies assume a binary answer. If redeveloping processes, a third group could be considered, however this group would be very small so the donor pool for imputation process would be small and risk bias.
<ul style="list-style-type: none"> • for comparability with administrative data sources in the future? 	Recorded on legal/official documents	Binary concept of sex is collected in administrative sources. Whilst not all sources of data collect the same definition of sex, legal/official documents is the nearest equivalent.

Equality monitoring

Sex is a protected characteristic, and so data on sex informs Equality Impact Assessments. The [Equality Act \(2010\)](#)⁶ states:

- “In relation to the protected characteristic of sex—
- a reference to a person who has a particular protected characteristic is a reference to a man or to a woman;
 - a reference to persons who share a protected characteristic is a reference to persons of the same sex.”

The status of being a man or a woman is not further defined within the act, however the Equality and Human Rights Commission (EHRC) state, in their statement on [sex and gender reassignment legal protections and language that](#) a person’s sex is what is on a person’s current birth certificate.

The EHRC page on [sex discrimination](#) (February 2020) further explains that:

⁶ Further information on the act is available in the [gov.uk guidance](#)

“The [Equality Act 2010](#) says you must not be discriminated against because:

- you are (or are not) a particular sex
- someone thinks you are the opposite sex (this is known as discrimination by perception)
- you are connected to someone of a particular sex (this is known as discrimination by association)

In the Equality Act, sex can mean either male or female, or a group of people like men or boys, or women or girls.”

Therefore, a member of the trans population should be protected from discrimination due to their:

- sex registered at birth (due to protection against direct discrimination), and
- sex as living / presenting (due to protection against discrimination by perception)

As gender reassignment is a separate protected characteristic, discrimination on those grounds are not considered here⁷.

Table 4: Assessment of equality monitoring needs

What would we aim to collect:	Concept	Justification
<ul style="list-style-type: none"> • considering the existing methodologies? 	Living / presenting. OR Recorded on legal/official documents	Sex discrimination cases are based on what the person committing the act considers the victim to be, this would be as the person is living / presenting. However, this measure may not be binary and sex is defined as binary within the Equality Act, so legal/official documents is the nearest equivalent. In addition, the sources of information Census 2021 would be compared to are likely to be sex as living / presenting.
<ul style="list-style-type: none"> • if we did not already have a time series? 	Recorded on birth certificate AND Living / presenting	Discrimination may be based a person’s sex as registered on their birth certificate or on the sex they are perceived to be by others. Therefore, ideally both concepts would be captured.
<ul style="list-style-type: none"> • for comparability with administrative data sources in the future? 	Recorded on legal/official documents	Binary concept of sex is collected in administrative sources. Whilst not all sources of data collect the same definition of sex legal/official documents is the nearest equivalent.

Analysis and linkage of the ONS Longitudinal Study and census microdata sample files

Detailed analysis of record level census data is conducted by a wide range of data users following linkage processes carried out by ONS. There are two main record level sources, and lists of research projects using each of these are available publicly:

- [ONS Longitudinal Study for England and Wales](#)⁸

⁷ Further guidance is available on gov.uk at [Equality Act 2010: guidance](#)

⁸ There are separate longitudinal studies in Scotland and Northern Ireland.

- [Census microdata files](#)

The ONS Longitudinal Study is a 1% sample of England and Wales and includes linked census and life events data from 1971 to 2017. It was originally set up to enable detailed health research. However, it has since been used for analysis on all Census topics. These include studies looking at differences between the male and female population.

The ONS Longitudinal Study includes a range of data sources, and retains the variables as recorded in each one. This includes:

- Sex on census records
- Sex as registered at birth (for those born in the England and Wales)
- Sex on life events records, for example cancer registrations

The Census microdata files only contain 2011 Census data, which is interpreted as self-identified sex.

Table 5: Assessment of Analysis of ONS Longitudinal Study and microdata sample files needs

What would we aim to collect:	Concept	Justification
<ul style="list-style-type: none"> • considering the existing methodologies? 	Living / presenting	The primary data sources are previous censuses which have been interpreted as self-identified sex. Birth registrations provide sex registered at birth for those born in England and Wales.
<ul style="list-style-type: none"> • if we did not already have a time series? 	Registered at birth AND Living / presenting	Data is used for a variety of purposes, some of which relate to a person's biology, for example work looking at fertility, or cancers of sex characteristics (e.g. breast cancer, prostate cancer, ovarian cancer). Others relate to factors that are culturally influenced, for example women are the predominant givers of unpaid care. Others relate to factors effected by both biology and culture, for example employment trajectories are impacted by cultural views on female roles, plus the impact of periods of time off to have children.
<ul style="list-style-type: none"> • for comparability with administrative data sources in the future? 	Registered at birth AND Recorded on legal/official documents	As above. Binary concept of sex is collected in administrative sources. Whilst not all sources of data collect the same definition of sex legal/official documents is the nearest equivalent.

Many other uses of census data at an aggregate level are quoted, for example in gaining wider understanding of those giving care, education, health, socio-economic status, deprivation, migration, commuting habits, demand for local bus services, domestic abuse patterns, the gender pay gap. In each case insufficient detail was provided to assess the precise need, but inferences can be made for different topics for different concepts. Therefore, table 5 also reflects these needs.

There is no single concept of sex required across the four primary data needs. All concepts of sex are needed to some extent.

The most evidenced needs are for sex recorded on legal/official documents and sex as living / perceived. The evidence for sex recorded on legal/official documents is stronger when looking forward to increased use of administrative data in the future, whereas the evidence for sex as living / perceived is stronger in relation to existing methodologies.

2.3 Comparability and coherence

Evaluation questions related to comparability and coherence

We would aim for data collected to be:

- comparable with other UK Census' in Northern Ireland and Scotland?
- comparable with the 2011 Census in England and Wales?
- comparable to other UK social surveys?
- coherent with the sex question wording agreed for Census 2021?
- internally coherent within the census form with other uses of the word "sex"?

Needs related to comparability and coherence

Consultation responses focussed on a need for continuity over time and across the UK. In reviewing this, assumptions must be made about how people answer in different circumstances, and whether how people answer is aligned with the guidance. This means that in many cases more than one concept has been recommended.

This is an area where internationally the approach being taken is evolving, and examples of all approaches can be found. As such, this evaluation does not consider comparability beyond the UK.

Table 6: Assessment of comparability and coherence needs

What would we aim to collect to be:	Concept	Justification
<ul style="list-style-type: none"> • comparable with other UK Census' in Northern Ireland and Scotland? 	Self-identified OR Recorded on legal/official documents OR Living / presenting	All four UK countries ask the question "What is your sex?" with response options of female and male. In Scotland the concept to be collected is referenced as self-identified sex. Their accompanying guidance for Census 2022 is available in their " sex question recommendation report "
<ul style="list-style-type: none"> • comparable with the 2011 Census in England and Wales? 	Self-identified	2011 census has been interpreted as collecting self-identified sex. The sex question guidance was: "Transgender or transsexual: select the answer which you identify yourself as. You can select either "male" or "female", whichever you believe is correct, irrespective of the details recorded on your birth certificate. You do not need to have a Gender Recognition Certificate."

		This has been interpreted as self-identified sex, however as the guidance was directed specifically at those who were transgender or transsexual it could also be interpreted to be sex as living / presenting depending on the interpretation of the population covered by the terms transgender or transsexual.
<ul style="list-style-type: none"> comparable to other UK social surveys? 	Living / presenting	<p>A review of how sex is collected in UK social surveys was carried out. This review included 20 non-ONS social surveys and 13 ONS social surveys. Amongst the surveys, the most common strategies are:</p> <ul style="list-style-type: none"> Ask or code if the respondent is male or female without referencing sex or gender Ask or code sex, with response options of male and female, without further defining sex <p>Only a minority of surveys identified gathered both sex and gender together. Of those that did, none specified the concept of sex gathered.</p>
<ul style="list-style-type: none"> coherent with the sex question wording agreed for Census 2021? 	<p>Recorded on legal/official documents</p> <p>OR</p> <p>Recorded on birth certificate</p>	<p>The context of the sex question is shown in Annex A. The phrasing and display of the sex question as presented on the EQ and the paper questionnaire imply that sex, not gender should be provided. However, the precise concept of sex to be provided is not stated.</p> <ul style="list-style-type: none"> The term “sex” is used There are only two response options There is reference to a question on gender identity to follow. <p>There is no reference to “registered at birth”, as used in the gender identity question.</p>
<ul style="list-style-type: none"> internally coherent within the census form with other uses of the word “sex”? 	Recorded on legal/official documents	<p>In Census 2021 there are two other questions that reference the word “sex”.</p> <ul style="list-style-type: none"> Marital status: The respondent is asked if their marriage or civil partnership is, or was, to a partner of the same sex, or the opposite sex. When registering a marriage this is recorded by the registrar and checked using a passport, or birth certificate. As the previous stage of the question asks for “legal” marital or registered civil partnership status, the implication is that the question is asking for the information as recorded on the marriage or civil partnership record Gender identity: The respondent is asked if the gender they identify with is the same as the sex registered at birth. The concept here is clearly defined as separate to “sex” as this is required to meet the information need.

Evaluation of relevance showed that the most evidenced needs were for sex recorded on legal/official documents and sex as living / perceived. Consideration of comparability and coherence shows no need for sex as registered at birth.

The most evidenced need in relation to this dimension of quality is sex recorded on legal/official documents, with similar but lower levels of need for sex as living / perceived.

Section 3: Impact on Census 2021 overall

This section considers three dimensions of quality in the [European Statistical System](#).

- **Accuracy and Reliability:** Accuracy is the closeness between an estimated result and the (unknown) true value. Reliability is the closeness of early estimates to subsequent estimated values
- **Relevance:** The degree to which statistics meet current and potential user needs in both coverage and content. (*Here we consider coverage*)
- **Timeliness and Punctuality:** Timeliness is the time gap between the publication and the reference period of the estimate. Punctuality is the gap between planned and actual publication dates

The impact on the latter two dimensions is led by the impacts on accuracy and reliability. This is because, for example, low response could lead to an extended field period, or census estimation processes taking longer. Similarly, a larger coverage adjustment could lead to wider confidence intervals around the estimates. Therefore, all three dimensions are discussed together under the heading of accuracy and reliability.

The next section discusses the expected scale of impact in relation to the trans population against the quality of the 2011 Census data. The potential for impact linked to the wider population cannot be sized. We then evaluate different concepts of sex in relation to impact on overall census response, sex question response, and individual's right to privacy.

3.1 Background – Scale of impact

The scale of the impact of collecting any specific concept of sex cannot be determined as multiple factors are at play. Quantitative testing to work out the size of this was considered and discarded due to:

- The small size of the trans population necessitating a very large sample
- The sensitivity of the topic meaning that the trans population would be unlikely to engage with the survey
- The potential for impact on the general population relates to the Census 2021 context and the anticipated public interest.

However, the impact is expected to be small in comparison to other factors effecting data quality. For example, in 2011, we achieved a response rate of 93.9 per cent, meaning that an estimated 6 per cent of the total population of England and Wales (3.8 million people) did not respond to the census and were subsequently adjusted for during data processing.

In addition, 0.45% (240,000) sex question responses were edited or imputed during [non-response and imputation](#) processes and there was 99.7% agreement between responses in the 2011 census and the [Census Quality Survey](#).⁹

In comparison, the trans population is small, the GEO estimate it to be between 0.3 and 0.75% of the UK population. ONS research during the development of the gender identity question has found between 0.1% and 0.6% of respondents (unweighted) to be trans. Similarly, the [Crime Survey](#) (6 months, unweighted base), also identified 0.5% of respondents as trans. Census 2021 will provide the first robust estimates of the size of this population.

⁹ Further information on the 2011 Census quality and methodology is available in the [2011 Census Quality and Methodology Information report](#).

3.2 Accuracy and reliability

This evaluation takes into consideration the Census 2021 context, for example

- The census is a mandatory household survey administered by online or paper self-response by one or more household members.
- The Census 2021 sex question and response options are final and do not state the exact concept of sex to be collected,
- There is a new voluntary question “Is the gender you identify with the same as your sex registered at birth?” for those aged 16 or over.
- When answering questions, most respondents do not access the guidance, even within the trans community. However, there is potential for members of the public to re-post the guidance on other forums, leading to a wider readership and therefore impact.

To enable this, two stages of work were completed. Additional research, comprising of online in-depth interviews to understand the potential for each concept to impact overall census response (2020:6). Followed by an in-depth review of the evidence base from both prior to, and since the guidance for the 2019 Rehearsal was finalised. This included research and testing and stakeholder engagement undertaken. In addition, we reviewed the legal context of sex, gender recognition certificates, and the census.

Evaluation questions

What sex concept will:

- maximise Census responses from the trans population / the cisgender population?
- maximise sex question responses from the trans population / the cisgender population?
- ensure any perceived invasion of privacy is justified by the data need?

Sex question response can impact overall response as sex is a key variable in the ‘remove false persons’ process. Therefore, if this variable is not completed, a record is more likely to be removed from the dataset during processing.

Each concept will be given a red, amber or green status against each evaluation question. Where a group is not discussed in the evaluation, this means that there was no evidence from testing of the concept under discussion leading to higher non-response for that group.

How to maximise census and sex question response

Testing showed (2020:6) that the potential impacts on overall census response and response to the sex question were largely the same across groups. Therefore, the evaluation considers both together and, in table 7 the term “respond” refers to both responding to the census as a whole and to responding to the sex question.

Outcomes were assigned against the following evaluation criteria:

- **Green:** No evidence of negative impact on overall response or sex question response across all groups either at the individual level or at a societal level
- **Amber:** Evidence of negative impact on overall response or sex question response for subsets of society either at the individual level or at a societal level

- **Red:** Evidence of negative impact on overall response or sex question response across a wide range of groups at the individual level and at a societal level

Table 7: Evaluation of what sex concept will maximise Census responses

Sex concept	RAG	Justification
Registered at birth	Red	<ul style="list-style-type: none"> - This definition raises strong feelings among various groups. - Trans people would be less likely to respond to census, especially those with a GRC. - Testing suggests some individuals would exert pressure on others to disrupt the census by discouraging completion.
Recorded on birth certificate	Amber	<ul style="list-style-type: none"> - Trans people without a GRC would be less likely to respond to census. - Seen as factual, reducing negative reaction across groups, including non-binary people - Some evidence of continuing impact on likelihood of response as obtaining a GRC is a high burden process. - No evidence that this would lead to re-posting of the guidance on other forums
Recorded on legal/official documents	Green	<ul style="list-style-type: none"> - Trans people without a GRC or another official document would be less likely to respond to census. - Seen as factual, reducing negative reaction across groups, including non-binary people, more than birth certificate as some legal/official documents easier to change - No evidence that this would lead to reposting of the guidance on other forums
Living / presenting	Green	<ul style="list-style-type: none"> - Trans people with binary identities more likely to respond if given this guidance - No available option for those not living / presenting as either female or male. However, previous testing (2018:6) showed that this is partly mitigated against by the guidance that a gender identity question will follow. - Potential for pre-existing campaign to escalate may also apply to this concept
Self-identified	Amber	<ul style="list-style-type: none"> - Trans people with binary identities more likely to respond if given this guidance - No available option for those that do not self-identify as either female or male. However, previous testing showed that this is partly mitigated against by the guidance that a gender identity question will follow. - As with 'registered at birth' there is increased potential for individuals to exert pressure on others to discourage completion".

How to minimise perceived invasion of privacy

An individual's right to respect for their private life is set out in Article 8 of the European Convention on Human Rights (1950), which is incorporated into UK law by the Human Rights Act (1998). Article 8 states:

Everyone has the right to respect for his private and family life, his home and his correspondence.

There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests

of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

This does not prevent interference by a public authority with a person’s right to respect for their private life, but it requires certain conditions to be met for that interference to be permitted by the Human Rights Act. Firstly, the interference needs to be in accordance with the law, and secondly it needs to be both a necessary and proportionate means of achieving one of the stated permitted purposes. Further information can be obtained from the [Equality and Human Rights Commission](#).

Outcomes were assigned against the following evaluation criteria:

- **Green:** Data need sufficient to justify collection of data, AND trans respondents would not be forced to reveal their gender history.
- **Amber:** Data need insufficient to justify collection of data, OR trans respondents would be forced to reveal their gender history.
- **Red:** Data need insufficient to justify collection of data, AND trans respondents would be forced to reveal their gender history.

Table 9: Evaluation of what sex concept would ensure any perceived invasion of privacy is justified by the data need

Sex concept	RAG	Justification
Registered at birth	Red	<ul style="list-style-type: none"> - Trans participants with and without a GRC would be asked to reveal their gender history - Risk of harm for those living with people that do not know their gender history. Greater risk for those from low income backgrounds and younger people due to time and cost of obtaining a GRC (2017:6) - Data need evidenced, but insufficient to justify collection of data classified as private by the Gender Recognition Act. - Data use not within defined situations where a person with a GRC would be expected to provide this information
Recorded on birth certificate	Red	<ul style="list-style-type: none"> - Small proportion of trans participants with a GRC would not be asked to reveal their gender history, but those without would. - Risk of harm for those living with people that do not know their gender history. Greater risk for those from low income backgrounds and younger people - Limited data need evidenced for this concept, linked to equalities monitoring. As there is insufficient justification for sex registered at birth then the need for that concept could be the case to this as the nearest alternative.
Recorded on legal/official documents	Amber	<ul style="list-style-type: none"> - Trans participants with at least one changed official document, for example passport, would not be asked to reveal their gender history, but those without would. - Risk of harm for those living with people that do not know their gender history if no legal/official documents aligned with sex as living / presenting. - Evidenced user need for this concept, with associated low risk of negative impact on overall census response.
Living / presenting	Green	<ul style="list-style-type: none"> - Trans respondents would not need to reveal their gender history. - Evidenced user need for this concept, with associated low risk of negative impact on overall census response.

Self-identified	Amber	<ul style="list-style-type: none"> - Trans respondents would not need to reveal their gender history. - No evidenced user needs for this concept, and risk of negative impact on overall census response
-----------------	-------	--

The highest negative impact is anticipated if sex registered at birth was collected due to reduction in census response. The data need is not considered to be strong enough to justify collecting this.

The least impact is anticipated if sex as living / presenting is collected. However, the adjacent concept of sex recorded on legal/official documents is also anticipated to have low impact.

Section 4: Conclusion

Looking at relevance, there is no single concept of sex required across the four primary data needs. All concepts of sex are needed to some extent.

The most evidenced needs are for sex recorded on legal/official documents and sex as living / perceived. The evidence for sex recorded on legal/official documents is stronger when looking forward to increased use of administrative data in the future, whereas the evidence for sex as living / perceived is stronger in relation to existing methodologies.

Consideration of comparability and coherence showed no need for sex as registered at birth. The most evidenced need in relation to this dimension of quality is sex recorded on legal/official documents, with similar, but lower, levels of need for sex as living / perceived.

The highest negative impact on accuracy and reliability is anticipated if sex registered at birth was collected due to the likely negative impact on response rates caused by the perceived invasion of privacy. The data need is not considered to be strong enough to justify collecting this.

The least impact is anticipated if sex as living / presenting is collected. However, the adjacent concept of sex recorded on legal/official documents is also anticipated to have low impact.

This leads to the conclusion that either sex as living / perceived or sex as recorded on legal/official documents should be collected, and other options discounted.

Therefore, as sex on official documents most closely reflects the census 2021 sex question response options, harmonises better with other UK nations, and is aligned with future usage of administrative data, this is the target concept for the sex question within census 2021.

Section 5: Developing guidance to support collection

This section addresses one dimension of quality in the [European Statistical System](#).

- **Accessibility and Clarity:** Accessibility is the ease with which users can access the data. Clarity is the quality and sufficiency of the metadata, illustrations and accompanying advice.

The online route to enable users to reach the guidance is shown in Annex A. If they choose to click on the appropriate links, respondents understand how to navigate within the webpages (2020:5). Ideally, the guidance should help all respondents answer in line with the concept being collected, sex recorded on legal/official documents.

Online help question guidance is aimed at members of population who may need further assistance in answering the question. This is the population who may:

- actively seek out guidance on how we are asking them to answer the question, or
- find providing information on sex a barrier to the completion of the Census.

However, the guidance may also be viewed and further cascaded by those with views or questions on the concepts of sex and gender in the wider population.

It seeks, to minimize barriers to completion of the census and the sex question. To do this it needs to balance the following aims:

- Direct people in how to answer
- Be considerate of people that have identities other than male or female or have VSC/intersex who may not wish to answer as either male or female
- Accept that respondents may have legal/official documents that state male and legal/official documents that state female.
- Use terminology that is acceptable to the whole population

The guidance used in 2019 rehearsal was tested (2019:12), this was further developed as part of this evaluation work and tested (2020:6). This draft guidance will be further refined based on the results of that testing, and the recommendations in this paper.

Section 6: Next steps

We are aware that most people will not look at the guidance and that there are data needs across all the different concepts of sex. Therefore, we will:

- finalise the draft guidance
- consult on the data user needs for derived variables and outputs from the sex and gender identity questions to analyse all concepts of sex. This would include the needs for microdata accessible via the [Secure Research Service](#) and take into account the need to ensure individual anonymity at all times.
- conduct further research into the concept of sex needed across all ONS data collections going forward.
- work with the GSS Harmonisation team, to support appropriate and comparable data collection on sex across government.

The GSS Harmonisation team work across the UK to promote and facilitate harmonisation. Further information on their current work on sex and gender is available on the [GSS website](#).

Section 7: Annex's

Annex A: Census 2021 sex question and route to guidance

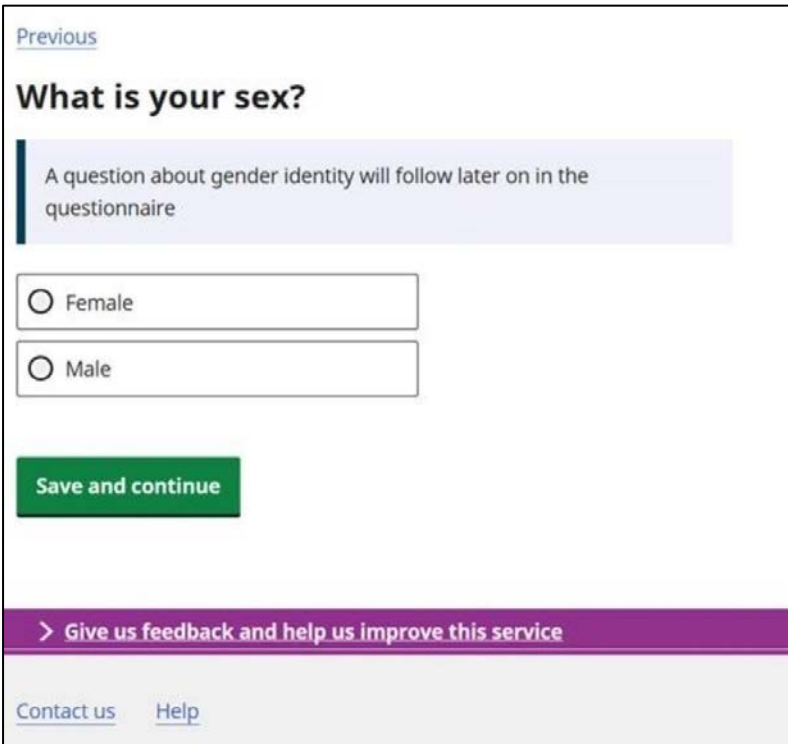
Question context

The phrasing and display of the sex question as presented on the EQ and the paper questionnaire imply that sex, not gender should be provided. However, the precise concept of sex to be provided is not stated.

- The term “sex” is used
- There are only two response options
- There is reference to a question on gender identity to follow.

The question does not have hard validation, as only questions that are required to build the form have this applied to allow respondents to more easily navigate around the form.

Image 1: Census 2021 sex question, online self-completion for those aged 16 and over



The screenshot shows a web form titled "What is your sex?". At the top left is a "Previous" link. Below the title is a light blue box containing the text: "A question about gender identity will follow later on in the questionnaire". There are two radio button options: "Female" and "Male". Below these is a green "Save and continue" button. At the bottom of the form is a purple banner with a right-pointing arrow and the text "Give us feedback and help us improve this service". Below the banner are two links: "Contact us" and "Help".

Route to guidance

On the electronic questionnaire the 'Help' link in the bottom banner brings up the "Online questions help" index page in a new tab.

Paper questionnaire respondents are provided with a link to www.census.gov.uk/help on the front page of the paper questionnaire. From there a respondent must select the link for "Paper questions help" to reach the index page. This is a separate, but very similar page to the online equivalent.

The headings at the top of the page are anchor links that take the user further down the same page to the links for individual guidance. These then take the user to a new page with guidance on the question of interest.

The person must move between tabs, or close the new tab, to return to the EQ. There will be no button on screen allowing the person to quickly return to EQ.

Image 2: Draft Census 2021 online questions help index page

Office for National Statistics Cymraeg

census 2021

Home About the census Keeping your data safe Contact us

Home > [Help with the census](#) > [How to answer questions](#)

Online questions help

If you are completing the paper census questionnaire please use the [paper census guidance](#).

Related content

[Contact us](#)

Contents

- [Your address and people who live here](#)
- [Household and accommodation](#)
- [Name, sex, date of birth and marital status](#)
- [Sexual orientation and gender](#)
- [Country of birth, passports, and not born in the UK](#)
- [National identity and ethnic group](#)
- [Language and religion](#)
- [Health and care](#)
- [Qualifications and full-time education](#)
- [Main job and Armed Forces](#)
- [Visitors](#)

Your address and people who live here

[Who lives here?](#)

[Are there any visitors staying overnight on 21 March 2021 at this address?](#)

[Household relationships](#)

[Do you stay at another address for more than 30 days a year?](#)

[One year ago, what was your usual address?](#)

Household and accommodation

[What type of accommodation is your address?](#)

[Are all the rooms in this accommodation, including the kitchen, bathroom and toilet, behind a door that only this household can use?](#)

[How many bedrooms are available for use only by this household?](#)

[What type of central heating does this accommodation have?](#)

[Does your household own or rent this address?](#)

[Who is your landlord?](#)

[In total, how many cars or vans are owned, or available for use, by members of this household?](#)

Name, sex, date of birth and marital status

[What is your name?](#)

[What is your date of birth?](#)

[What is your sex?](#)

[On 21 March 2021, what is your legal marital or registered civil](#)

Annex B: Publications and meetings relating to sex and gender identity

Research and engagement publications

Information on stakeholder engagement and testing has been provided over time as the project has developed through a series of update reports from the point of the 2021 Census topic consultation

onwards. Census legislation is included in the below listing in italics to show key points of Parliamentary engagement

Sex and gender identity are only the focus of research in the tests listed in the [question development report](#). However, as sex is a key demographic variable, it was included in all tests and feedback incorporated into the sex question decision making as required.

For reference, the paper forms are also available: [Paper questionnaires for England and Wales](#).

These publications evidence our engagement with a range of stakeholders. Further evidence, of more recent engagement, is provided in the next section

1. May 2016, Consultation response [Demography](#) (1.6 MB pdf)
2. August 2016, Gender identity workshop [Gender identity workshop](#)
3. January 2017, we published a [Gender identity update](#)
4. June 2017, [Gender identity update event](#)
5. September 2017, [Qualitative research on gender identity: phase 1 summary report](#),
6. December 2017, [2021 Census topic research report](#)
7. July 2017, 2021 Census Roadshow slides [Gender identity](#)
8. December 2018, 2021 Census topic research update [research and testing of questions and topics for the 2021 Census](#).
9. December 2018, government White Paper [Help Shape Our Future: The 2021 Census of Population and Housing in England and Wales \(PDF 886KB\)](#)
10. February 2019, [how “sex” and “gender” are used in the context of the UN Sustainable Development Goals \(SDGs\)](#)
11. June 2019, The Office for National Statistics (ONS) and Statistics Canada (STC), as part of the UNECE Bureau of the Conference of European Statisticians, collaborated to undertake an [In-depth review of measuring gender identity](#)
12. September 2019, [Measuring sex in the UK censuses: Harmonisation update](#)
13. September 2019, [Guidance for questions on sex, gender identity and sexual orientation for the 2019 Census Rehearsal for the 2021 Census](#)
14. October 2019, [Census \(Return Particulars and Removal of Penalties\) Act 2019](#)
15. March 2020: [Summary of testing for Census 2021](#) (March 2020)
16. May 2020, [Census \(England and Wales\) Order 2020](#)
17. June 2020, [Census \(England\) Regulations 2020, Census \(Wales\) Regulations 2020 and Rheoliadau'r Cyfrifiad \(Cymru\) 2020](#)
18. June 2020, [Exploring existing data on gender identity and sexual orientation](#)
19. June 2020: [Sex and gender identity question development for Census 2021](#)
20. June 2020: [Paper questionnaires for England and Wales](#)
21. 19 October 2020, [Demography question development for Census 2021](#)

The summary of testing, first published in March 2020, has recently been updated with four further tests undertaken in 2020. As such, it contains further details on all the tests referenced in this report. These are summarised in table 10.

Table 10: Summary of testing used in the evaluation of the sex question concept

Reference	Date of testing	Type of testing and sample size
-----------	-----------------	---------------------------------

2017:6	March and April 2017	Qualitative: Four focus groups with 29 cisgender participants; 18 cognitive interviews with transgender participants.
2017:15	August and September 2017	Qualitative: 18 cognitive interviews with cisgender and transgender participants.
2018:3	February 2018	Qualitative: 230 informal interviews at four LGBT+ events.
2018:6	February to May 2018	Qualitative: 40 cognitive interviews with participants who were cisgender, transgender, non-binary and intersex.
2019:12	August 2019	Qualitative: 23 cognitive interviews.
2019:15	September to November 2019	Quantitative: Approximately 300,000 households took part in the 2019 Rehearsal.
2020:5	September and October 2020	Qualitative: 20 cognitive interviews with participants who were aged over 65, had English as an additional language, had dyslexia, or had learning difference/difficulty
2020:6	October and November 2020	Qualitative: 52 cognitive interviews with participants who were cisgender, trans, members of Women's Groups and trans allies

ONS stakeholder engagement on the sex question

Following the consultation in 2015, ONS continued to engage with those stakeholders who had consented to further contact regarding the sex and gender identity questions.

Our September 2017 research report on initial work on the sex and gender identity led to increased levels of engagement, in particular with users of data on the female population. ONS proactively identified women's groups to engage with who had undertaken any research on sex, used any statistics which depended on sex or who appeared to have an interest in the collection of sex data (and possibly gender identity data too).

In April 2018 a stakeholder event 'Understanding Data Collection on Sex and Gender Identity' was held in London. Stakeholders were consulted about the work that ONS were doing on gender identity and why it was being conducted. They were asked for their views on the sex and gender identity question options for the 2021 Census.

Representatives from the following organisations were invited and representatives from the organisations listed in bold attended the event.

- Women's Institute
- **Women's Budget Group**
- Equality Now
- The Equality Trust
- **Women's Aid**
- Gender and Development Network
- **Equality and Diversity Forum**
- **Equality Challenge Unit**
- Demos
- The Women's Resource Centre
- Sex and Gender Ethics Society
- Fawcett Society

ONS confirmed that there will be a mandatory question that will collect male and female information. There was a consensus from most stakeholders that although they understood the user need for gender identity data, they would not want this to impact the data currently collected by ONS on sex as this data is used for various purposes in their organisation.

Some cross-government working / topic groups were regularly attended:

- Harmonisation topic group
- Gender Identity Cross-Government topic group
- Scottish National Statisticians Sex/Gender Working Group

Other organisations were met with individually, these organisations we met with in 2019, when the Census 2021 questions on sex and gender had been agreed, are listed below. Those we met with more than once on this topic are marked with a star (*)

- a:gender
- Centre for Equalities
- Civil Service HR
- Central local information partnership (CLIP) population sub-group
- Care Quality Commission *
- Department for Culture, Media and Sport
- Department for Education *
- Equality and Human Rights Commission *
- Family Resources Survey Conference
- FemQuant
- Government Digital Service
- Government Equalities Office *
- Intersex UK
- Ipsos MORI
- LGBT Foundation Stonewall
- Market Research Society
- Medopad
- Ministry of Justice
- National LGB&T Partnership
- Natsal
- NHS Digital
- NHS England *
- Northern Ireland Statistics Research Agency *
- National Records for Scotland *
- National Statisticians Data Ethics Advisory Committee
- Office for Standards in Education
- ONS social surveys
- Stonewall *
- Universities and Colleges Admissions Service (UCAS)
- Understanding Society
- United Nations Economic Commission for Europe
- Welsh Government

In December 2019, 80 social statisticians and users of population level data sent a letter to the directors of UK census authorities with concerns about the sex question. This outlined concerns with the concept of sex to be collected around understanding:

- sex-based discrimination
- different health and socio-economic outcomes between men and women

These concerns are addressed in the discussion of user needs in section 2. In addition, concerns were raised regarding:

- Size of the population for whom concepts change response (addressed in section 3.1)
- Impact on other data collection exercises (addressed in section 2.3)
- Reference to those with intersex conditions (addressed in section 5)

ONS met with some of the signatories of the letter in January 2020. It was agreed that ONS would then hold a Roundtable with a wide range of academics who use data on sex. Due to COVID, this was delayed until June 2020. The roundtable event included people from the following organisations:

- Fairplayforwomen.com
- LGBT Foundation
- Government Equalities Office
- Equality and Human Rights Commission
- NHS England
- Anglia Ruskin University
- Stonewall
- UCL
- University of Edinburgh
- ESRC
- Imperial College London
- London School of Economics
- King's College London
- University of Kent
- University of Manchester
- University of Essex
- University of Edinburgh

Since the 2015 consultation, ONS has continued to engage with a range of stakeholders regarding the sex and gender identity questions. Much of this is reported on in the publications listed above.

The roundtable showed there were a range of data needs amongst the expert academic community and no consensus was reached. Therefore, the ONS undertook further research, including follow-up meetings with groups of academics with different views to understand further their data requirements:

- July 2020, meeting with data users supporting collecting data on sex registered at birth
- August 2020, meeting with data users supporting continuing to collect data on self-identified sex

ONS did not present any research at any of these meetings, instead allowed the stakeholders to present their views. None of the meetings yielded further evidence of data needs for specific concepts of sex in Census 2021.

Annex C: Detailed responses from ONS users of sex data from Census 2021

Section 2.2 discusses the key user needs for data on sex. The primary users in relation to two of these needs, population estimates and projections and household projections, and 2021 census operational needs, including coverage adjustment process are internal ONS teams.

These teams were asked to provide further detail of their needs. Their responses are summarized in section 2.2 and are reproduced in full below.

Population estimates and projections and household projections
Centre for Ageing and Demography response:

The population projections (and derived household projections) use the previous Census as their base. This is currently the 2011 Census. The 2011 Census sex question guidance indicated that people could self-identify. Exact wording was:

“Select the answer which you identify yourself as. You can select either "male" or "female", whichever you believe is correct, irrespective of the details recorded on your birth certificate. You do not need to have a Gender Recognition Certificate.”

In practice this means that the projections are fundamentally based on respondent-identified sex, noting that:

- *Most respondents will not have looked at the guidance, so (as with earlier Censuses) will have self-identified anyway. The proportion of the population who looked at the sex question guidance was 0.003% (a bit less than 1 in 30 thousand people).*
- *A respondent’s self-identified sex may or may not be their sex as registered at birth. For the cis population, it would presumably be so. However, for the trans population this would be split.*

Projections also take account of population changes over time through births, deaths and migration:

- *Births are projected through a combination of:*

(i) sex ratios based on sex at birth, taken from birth registrations – this is the only key source that is explicitly sex at birth

(ii) female fertility rates. The latter are derived with the number of births as the numerator and the number of females as the denominator. The number of females in that denominator comes from the Census-based population estimates. This means that:

- *It will include some trans people (born male) as well as biological females.*
- *It will also include biological females who are infertile.*
- *It will exclude those who were registered as female at birth, but recorded their sex as male in the Census [This is because they will be counted in the male population].*

- *Deaths are projected through death rates. These are derived from:*

(i) Number of deaths, by sex, taken from death registrations. The sex is provided by the person registering the death, so it may not be sex at birth. For those who lived in a

gender different from their sex at birth, they may well have been recorded in their lived gender (indeed, we cannot assume that the person registering the death will even know the deceased's sex at birth). There is no linkage made with birth records.

- (ii) Population estimates. As noted previously, these are primarily derived from Census and so self-identified.*
- Migration is derived from a range of admin and survey sources. These are also liable to be sex as defined by the person completing the form.*
 - The main international migration source to date has been the International Passenger Survey (IPS), where the interviewer would typically record what they believed the respondent to be. Clearly in most cases this will be the same as sex at birth, but people who have transitioned to living in a new gender may well have been recorded in that gender.*
 - The main internal migration source is the GP Patient Register, where applicants are simply invited to tick whether they are "Male" or "Female", with no indication of the definition to be used (or even whether it is sex or gender which is being asked for). So for many transgender respondents, it may well be lived gender, different from the sex registered at birth.*

*The household projections also use 2001 Census data. This is also self-defined (there was no written guidance for the 2001 Census although we have evidence that ONS advised individual contacts that they could self-identify). They also use the **Annual Population Survey**, which also isn't necessarily sex at birth. In addition (like the population projections) they use **prisoner data** from the Ministry of Justice (MoJ).*

- For the purposes of the projections, all people in men's prisons were counted as male, and all those in female prisons as female – i.e. this is assumed to reflect "lived sex", not necessarily sex at birth.*
- According to published MoJ statistics, in 2019 there were 163 transgender prisoners in England and Wales (out of an overall prison population of over 80,000, so a rate of around 0.2%). Of these, 34 were in women's prisons and 119 in men's prisons.*

In summary, sex in the projections is predominantly respondent-identified. The dominant source is the 2011 Census (self-identified) and, with the exception of birth registrations, the other sources are also likely to include people recorded differently from their sex at birth.

In that context, indicating that people to self-identify in the 2021 Census will be most consistent. Note too that PSD is likely to develop a revised back series of population estimates for the period 2011 to 2021, which in turn will inform projections. On that front too, having self-identification for 2021 will offer most methodological consistency over time, and at either end of the revised series.

In addition, the projections need the best possible coverage of the population of England and Wales, with undercount minimised and aiming to meet the national completion target of at least 94%. Therefore, it could be argued that any guidance which is likely to deter people from answering is disadvantageous.

2021 Census operational needs, including coverage adjustment process:
Census and Population Statistics Hub response:

Edit and Imputation

Missingness

The sex question had very low missingness in 2011 (less than 0.5%). It was a key matching variable used to find nearest neighbours in the donor imputation process because it has been identified as a key predictor variable for other census characteristics.

We are currently investigating scenarios around high item non-response caused by potential controversy around the sex, gender identity and sexual orientation questions. These include higher non-response of the sex question by people or groups unhappy with the published guidance. Potential impacts could be: a smaller donor pool for imputation leading to a greater risk of bias in imputation which could affect a number of other census variables, loss of a key predictor variable for a significant number of records, again leading to a greater risk of bias in imputation for other census variables, and increased time needed to run imputation to resolve missingness leading to more time pressure on the rest of the imputation process and perhaps subsequent processes.

High item non-response from the trans population itself may slightly increase the overall non-response rate for sex. This is unlikely to cause issues with sex distributions in isolation after imputation, but imputation of sex will not take into account the gender identity question, so there is a small risk around the maintenance of the multivariate relationship between sex, gender identity and other variables after imputation for the trans population specifically. No research is planned to investigate this before the census due to the anticipated small magnitude of the impact.

Multi-ticks (where both sexes are selected on the paper questionnaire) could also lead to an inflated missingness rate as sex would be set to missing for these cases, so will also be factored into scenario testing. This could apply to cisgender (where both sexes selected as a protest), trans and intersex populations.

Inconsistencies

Following changes to legislation, there are no hard edit rules around marital/civil partner status and sex so any inconsistencies in the way trans people may answer the question would not have an effect on edit rule failures for this rule.

There are hard edit rules surrounding age of mothers compared with their children – a mother cannot be more than 65 years older than her child. The way transwomen answer the sex question could have a small effect on edit rule failures for this rule but this is likely to affect only a tiny proportion of records so it has been considered unlikely to have any significant measurable effects on imputation.

Census Linkage

In 2011, there was less than 0.5% and 1.8% missingness on the sex variable in the census and census coverage survey respectively. The percentage of matches made where sex was the same, missing or different was 98.2%, 1.4% and 0.4% respectively. In 2021, we again expect sex to be generally well recorded and accurate. Whilst records matching on sex is not a strong indicator of a match, records not matching on sex is a strong indicator of a non-match. Therefore, sex is used as a matching variable in the majority of our deterministic matchkeys and as a probabilistic matching variable. However, it is still possible for matches to be made automatically when sex is different or missing. Cases that are not matched automatically by either deterministic or probabilistic methods, will be

reviewed clerically and could therefore be matched if the records do refer to the same person, despite there being a difference or missingness in the sex variable. We therefore do not anticipate that there will be a bias in matching caused by advice that is, or is not, given regarding the sex question. Lastly, if in 2021 we find that there is a much higher than expected rate of missingness in the sex field then we could react to that by tweaking the matchkeys and/or the parameters for the probabilistic matching.

Census Coverage Estimation

Estimation is based on the collection of data from the Census Coverage Survey. The requirement is that the same concept is captured in both census and CCS. If there happen to be any differences in response between the Census and CCS (this can be due to a number of reasons such as collection, capture or imputation), the census value is used for estimation. Estimation would only be impacted if the guidance affected overall response or matching rates significantly.

Census Coverage adjustment

This imputes records into the census database so that the overall database is consistent with the census estimates from the estimation process. No particular concept is required as long as the same concept is collected in both the census and the census coverage survey, which is identical to the estimation requirement.

Statistical disclosure control

There would be no impact on the SDC methods used. Sex is not a targeted variable in record swapping, and it would have no impact on the rules used as part of the flexible table builder.

Removing false persons

Sex is a key variable in this process, however if levels of missingness were significantly greater than anticipated the methodology could be adjusted to use a different variable with higher item level response.