

Working group about a person's sex and gender
Data collection and publication – draft guidance
Author: Roger Halliday, Chief Statistician.

This is a draft document that has been put together by the Chief Statistician. It has been shared in draft for the purpose of seeking stakeholder feedback and is not to be used whilst in draft format.

It does not represent the views of Scottish Ministers.

Purpose of paper

To provide a first draft of guidance on collection and publication of data by Scottish public bodies.

Overview

I was asked by the Cabinet Secretary for Social Security and Older People to bring together a working group to look at the way data on sex and gender is collected and published, and put together guidance for public bodies.

This was announced at the same time as the Cabinet Secretary updated Scottish Parliament on plans to reform the Gender Recognition Act 2004 in Scotland. While the two are not directly related, some groups had, in response to the proposals to reform gender recognition, raised concerns about the collection and use of data by sex and gender.

The book “Invisible Women” by Caroline Criado Perez, and work by a number of organisations, has also drawn attention to the frequency with which data is neither collected, aggregated or used in a way that takes account of the differences between men and women, and their impact in areas such as transport, health and access to services.

I have brought together a group of experts in collecting and presenting data from across the UK. I have listened to the varied views and drawn my own conclusions from these. An important part of this work is making sure that this draft guidance is informed by evidence from a wide range of individuals and organisations with views on a lived experience of these issues; to do this, I met a number of external stakeholders to listen to their views, as well as holding public engagement events to give everyone with an interest the opportunity to have their views heard.

I gathered together these insights and considered them when putting together this draft guidance. I am now seeking feedback on this draft.

Given the importance of this topic, it was important for me to carry out this work in a transparent way, so that people can see how it has been put together. To this end, I

have been posting regular blog updates on my Statistics blog, as well as publishing all minutes from the working group meetings on the Scottish Government website.

My role as Chief Statistician brings with it responsibility for the coordination and implementation of professional statistical standards that help maintain trustworthiness in the use of data, its quality, and delivering improved outcomes for people in Scotland. In putting together this draft guidance I have, therefore, rooted the work in a set of widely accepted statistical principles. Whilst the concepts behind definitions are important, so too is having data that is high quality, and can be used to drive changes and improvements that will save time, money and lives – for the benefit of everyone.

DRAFT

Guidance

1. Introduction, purpose and context

The Scottish Government clearly sets out its purpose in Scotland's National Performance Framework: 'to focus on creating a more successful country with opportunities for all of Scotland to flourish through increased wellbeing, and sustainable and inclusive growth', and makes clear the values underpinning this: 'we are a society which treats all our people with kindness, dignity and compassion'. To fully understand the experience of individuals and groups we need data that will allow for comparisons to be made.

The Cabinet Secretary for Social Security and Older People, in her statement to Scottish Parliament on 20 June 2019, highlights the importance of "having disaggregated data...to help show where there is discrimination and indicate where further work needs to be done, in any part of Government". This need also extends to the public sector more widely.

Having high quality data is the backbone to having a public sector that can design services that meet, and are responsive to, the needs of all people in Scotland. This data is an important source of information for those involved in delivering public services, including planning, monitoring and reviewing of decisions in relation to these services.

As such, my expectation of public bodies in Scotland is that they routinely gather and publish information on socio-demographic characteristics of people in Scotland, using this information to design, plan, monitor and evaluate services that are sensitive to the needs of all of Scotland and create the conditions where there are opportunities for all of Scotland to flourish. This includes understanding not just the issues on sex and gender, but on the intersectionality between this and other socio-demographic characteristics.

How data on sex and gender is collected and published is an important part of achieving this wider aim. This guidance sets out things for public bodies to consider when they are collecting data about sex and gender, whether in a survey or an administrative system, so that data collection has a clear purpose and is rooted in the organisations' needs and informs the design, targeting and delivery of public services.

It is equally as important to consider whether the way data is collected as a whole introduces bias against particular parts of society, and building skills across the public sector amongst those collecting data is important to mitigate any risk of bias.

Indeed, the United Nations, in their Gender Statistics Manual¹ and the European Institute of Gender Equality Gender Statistics Database² identify good practice in data and statistics about sex and gender as:

- a) Data are based on concepts and definitions of gender that adequately reflect the diversity of women and men and capture all aspects of their lives; and
- b) Data collection methods take into account stereotypes and social and cultural factors that may induce gender bias in the data.
- c) The presentation of data on sex and gender should help illuminate meaningful differences and similarities between women and men.

In undertaking this work, it is clear that while most public bodies are collecting data about sex or gender, it is fair to say that these terms are sometimes used interchangeably in guidance about data collection, and there is typically limited guidance on exactly what is being asked for. This has the consequence that in reality people answering this question may be interpreting it in different ways.

It was clear that many organisations weren't making conscious decisions about the right question or questions to ask.

And very few organisations were collecting data to help better understand experiences and outcomes of trans people.

I found that even when data on sex or gender is collected, it is not always analysed, published or used in decision making. This clearly impacts on efforts to eliminate discrimination and promote equality.

This matches the UN's Committee on the Elimination of Discrimination against Women who looked at the UK last year and found a "...lack of systematic collection of data, disaggregated by sex, gender, ethnicity, disability and age, in particular with regard to intersecting forms of discrimination, to identify areas in which women lack substantive equality with men, inform policymaking and assess the impact of measures taken."³

What is clear is that there isn't a standard way that data about sex and gender is being collected, either in Scotland or the UK. Internationally this varies with approaches currently developing in a few other countries, and others more established, for example, the Australian Government published guidelines in 2013 on the recognition of sex and gender, which covered data collection⁴.

¹ UN Statistics Division. 2020. "UN Gender Statistics Manual." Accessed February 17. <https://unstats.un.org/unsd/genderstatmanual/Default.aspx>.

² European Institute of Gender Equality. 2020. "Gender Statistics Database." European Institute for Gender Equality. Accessed February 17. <https://eige.europa.eu/gender-statistics/dgs>.

³ UN Committee on the Elimination of Discrimination against Women. 2019. "Concluding Observations on the Eighth Periodic Report of the United Kingdom of Great Britain and Northern Ireland." United Nations. https://tbiinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CEDAW/C/GBR/CO/8&Lang=En.

⁴

<https://www.ag.gov.au/Publications/Documents/AustralianGovernmentGuidelinesontheRecognitionofSexandGender/AustralianGovernmentGuidelinesontheRecognitionofSexandGender.pdf>

2. Underpinning statistical principles

An important starting place for this work are the set of relevant statistical principles that will underpin the work to develop guidance on collection and publication of data on sex and gender identity.

The following are taken from the [official statistics code of practice](#)

- Those producing statistics should be protected from any undue pressure that may influence the production or presentation of the statistics (Practice T1.1)
- All statutory obligations and internationally endorsed guidelines governing the collection of data, confidentiality, data sharing and release should be followed (Practice T2.3)
- The privacy of individuals and business information must be protected in the production and release of statistics and data, ensuring legal obligations are met (Principle T4)
 - The identity of individuals or organisations must be protected at all times. Appropriate disclosure control methods, including the nature of any consent given, should be applied when releasing statistics (Practice T4.2)
- Transparent judgements about statistical definitions and methods, together with judgements about the strengths and limitations, are essential in supporting confidence in the quality of the statistics (Principle Q2)
- Risks to data quality should be minimised – both operationally, such as using insight from audits and inspections of the data – and in the collection and data preparation processes (Practice Q3.1.ii)
- Statistics should be consistent and comparable, while remaining relevant to society (Principle Q4)
- In collecting data, organisations should be clear about the variety of uses and potential uses of statistics (Principle V3)

The following is taken from the [European Statistics code of practice](#)

- The overall methodological framework used for European Statistics follows European and other international standards, guidelines and good practices (Practice 7.1)

Dignity and respect

- Data collection needs to be carried out in a way that treats people with dignity and respect

As with all official statistics, they are collected for the purpose of improving decision making, so another principle would be that

- if statistics produce an unexpected results, they should be investigated further, to see if there is a reason for these results

3. Definitions used

While the terms sex and gender are used by some people interchangeably, a number of organisations have set out a set of definitions⁵. The definitions stated are broadly similar across these organisations.

They define **sex** as generally whether you are male or female, though for some people this does not reflect their innate sense of themselves. There are different aspects to a person's sex:

- Biological aspects of an individual as determined by their anatomy, which is produced by a combination of their chromosomal, hormonal, genital and gonadal characteristics, and their interactions.
- The aspect of **legal sex**, which can change if a full Gender Recognition Certificate (GRC) is obtained. For a trans person with a GRC, their legal sex is their acquired sex. For a trans person without a GRC or other people, their legal sex is their sex registered at birth.⁶

The Equality and Human Rights Commission (EHRC)'s "Statement on Sex and Gender Reassignment: legal protections and language"⁷ explains that in UK law sex is understood as binary and a person's legal sex is determined by what is recorded on a person's birth certificate. A trans person can change their legal sex by obtaining a GRC and a trans person who does not obtain a GRC retains the (legal) sex recorded on their birth certificate for legal purposes.

Organisations such as the World Health Organisation and the Royal Statistical Society⁸ define **gender** as:

- a social construction relating to behaviours and attributes based on labels of masculinity and femininity;

gender identity as:

- a personal, internal perception of oneself and so the gender category someone identifies with may not match their sex at birth;
- what an individual experiences as their innate sense of themselves as a man, a woman, as having no gender, or as having a non-binary gender –

⁵ For example the [World Health Organisation](#), the [Royal Statistical Society](#), and the [USA Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys](#).

⁶ <https://www.gov.scot/binaries/content/documents/govscot/publications/minutes/2019/09/sex-and-gender-in-data-working-group-meeting-september-2019/documents/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/govscot%3Adocument/EHRC%2Bsubmission%2Bon%2Bcollecting%2Band%2Bpresenting%2Bdata%2Bon%2Bsex%2Band%2Bgender..pdf>

⁷ <https://www.equalityhumanrights.com/en/our-work/news/our-statement-sex-and-gender-reassignment-legal-protections-and-language>

⁸ For example the [Office for National Statistics](#), the [World Health Organisation](#), the [Royal Statistical Society](#), and the [USA Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys](#).

where people identify as somewhere on a spectrum between man and woman

And define **transgender** as:

- anyone whose gender identity differs from their sex at birth.

For most people, their sex and gender identity will be the same. For some people, there may be differences. People generally may not necessarily answer a question about their “sex” by thinking about their biology.

The Equality Act 2010 (“the 2010 Act”) legally protects people from discrimination when receiving a service, in the workplace and in other contexts. In relation to the protected characteristic of “sex”, section 11 of the Act provides that:

- (a) a reference to a person who has a particular protected characteristic is a reference to a man or to a woman; and*
- (b) a reference to persons who share a protected characteristic is a reference to persons of the same sex.*

Man and woman are defined for the purposes of the 2010 Act at section 212; ‘man’ means a male of any age and ‘woman’ means a female of any age.

Section 7 of the 2010 Act sets out the protected characteristic of “gender reassignment” and an individual has this protected characteristic if they are ‘proposing to undergo, is undergoing or has undergone a process (or part of a process) for the purposes of reassigning the person’s sex by changing physiological or other attributes of sex.’

Although the 2010 Act doesn’t purport to define sex, section 11, in relation to the protected characteristic of sex, takes a binary approach – someone is either a man or a woman.

The EHRC’s most recent position is that “public bodies are not required to collect information on legal sex and can enable employees and service users to self-identify their sex”⁹. Indeed, Scotland’s Census 2022 will ask a binary sex question with guidance that this can be self-identified, and this is followed by a question on trans status/history. Previous censuses have also asked a self-identified sex question.

On the other hand, a voluntary sex question with guidance that this should be answered according to a person’s sex registered at birth recognises that a trans person can change their legal sex by obtaining a GRC and a trans person who does not obtain a GRC retains the (legal) sex registered on their birth certificate for legal purposes, as per the EHRC’s “Statement on Sex and Gender Reassignment: legal protections and language” that was published on 30 July 2018¹⁰.

⁹ <https://www.gov.scot/binaries/content/documents/govscot/publications/minutes/2019/09/sex-and-gender-in-data-working-group-meeting-september-2019/documents/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/govscot%3Adocument/EHRC%2Bsubmission%2Bon%2Bcollecting%2Band%2Bpresenting%2Bdata%2Bonn%2Bsex%2Band%2Bgender..pdf>

¹⁰ <https://www.equalityhumanrights.com/en/our-work/news/our-statement-sex-and-gender-reassignment-legal-protections-and-language>

The EHRC notes in the statement above that a trans person is protected against discrimination under sections 7 and 11. They are protected under section 11 on the basis of their legal sex. It gives the example of a trans woman who does not hold a GRC being protected against discrimination for being male and a trans woman with a GRC being protected against discrimination for being female.

While the 2010 Act protects against discrimination on the basis of sex, many people take the view that the term 'gender' also raises important issues in current inequalities. For example, Caroline Criado Perez in her book "Invisible Women" writes *"Although I talk about both sex and gender throughout, I use gender data gap as an overarching term because sex is not the reason women are excluded from data. Gender is. In naming the phenomenon that is causing so much damage to so many women's lives, I want to be clear that the root cause and, contrary to many claims you will read in these pages, the female body is not the problem. The problem is the social meaning that we ascribe to that body, and a socially determined failure to account for this"*.

As such, having evidence on sex alone may not give a full picture of inequalities.

Section 13 of the 2010 Act provides that a person discriminates against another if, because of a protected characteristic, they treat that person less favourably than others. This includes less favourable treatment based on a perceived protected characteristic (e.g. sex, gender reassignment or sexual orientation), even where that perception is incorrect. Section 13 can also include less favourable treatment by association. Neither of these points affects the interpretation of the protected characteristic of "sex" under the 2010 Act.

A person may also suffer sex discrimination under the 2010 Act because of social constructions more generally considered to relate to "gender". For example, a woman may suffer discrimination at work because the employer makes unwarranted assumptions about a woman's role in society and what tasks a woman can carry out.

Therefore, a public body may decide to collect data on a person's gender (or their gender identity) as well as on sex, according to their data needs.

4. Considerations when deciding what data to collect

For data collection, the starting point for any organisation is why does it need to collect data. Data will be collected for various specific purposes, and therefore organisations must decide on the most useful definition(s) to use to capture the data that they need.

It is important to think through the way you are collecting data i.e. who you are seeking data from, the arrangements for collecting data (e.g. time of day, or mode of collection) and nature of the questions you use. How you set up the data collection can introduce biases that reduce the value or lead to decisions that unknowingly increase inequalities. So, consider any data collection as a whole and plan to avoid any potential biases.

From the principles above, it is important to only collect a specific item of data from someone where there is a clear need for this, to minimise the burden on an individual and to comply with the legal requirements under the General Data Protection Regulation (GDPR). In addition, data that is used for operational decisions is required to be able to be provided to an individual and rectified. So those collecting data for operational decisions need to take steps to make sure it remains correct.

There may be a number of reasons for collecting data on sex or gender identity. Some important reasons are

- Operational decisions by public sector services about the service people receive e.g. certain pension and benefit calculations may require knowledge of a person's legal sex, or whether to offer a single sex service.
- Designing more inclusive better services. This includes identifying opportunities to improve or transform services, decisions on how best to do that, monitoring key performance measures or societal trends
- An evidence based understanding of service need, to for example understand funding and planning decisions.
- To meet legal obligations to collect certain data.

For the vast majority of people, sex and gender identity questions will provide the same response, namely female and woman/girl, or male and man/boy. This also means that for many uses, whether data is collected using a sex question or a gender identity question, will in reality give you data that will work for uses where you need data on either concept.

However, while there are currently no official figures, estimates suggest that 1 in 250 people in the UK identify as neither male nor female¹¹. The Gender Identity Research and Education Society estimated that between 0.6% and 1.0% of the population were trans¹². As such, for most groups of interest and issues one may

¹¹ Fiona Glen and Karen Hurrell (2012), Technical note: Measuring Gender Identity, Equality and Human Rights Commission. Available at: https://www.equalityhumanrights.com/sites/default/files/technical_note_final.pdf

¹² Gender Identity Research and Education Society (2011) The number of gender variant people in the UK <https://www.gires.org.uk/information-on-prevalence-incidence-and-monitoring/>

want to measure, the inclusion of non-binary or trans people will not skew the statistics disaggregated by sex or gender.

Collecting and presenting aggregate data on any group containing small numbers of people is not straightforward. For population surveys, it is very likely the survey won't go to enough people to get an accurate picture of every person in the population. For administrative data collections, they need to be designed to get very high levels of completeness of this data to give an accurate picture of the people who use, work or are served by a public body. So, you should consider how important using this data will be.

If you are looking for information specifically to plan, design, or fund services, it is important to consider whether it would help you design and run a better service:

- To ask about the concept of gender identity; or
- To ask about the concept of sex; or
- It makes no difference either way.

In a small number of instances, it may be necessary to record a person's legal sex but this would be on an individual basis for a very specific purpose and it would be up to public bodies who need this data to develop the best approach to do this.

Questions about a person's biology should not be asked, except potentially where there is direct relevance to a person's medical treatment. Such a question is likely to breach an individual's human privacy. Whilst medical requirements for the trans population are not simply related to their sex at birth the numbers involved will not impact on resource planning.

This guidance is not about these day to day operational data collection decisions. It will be for public bodies who need to collect this data to establish the best approach to collect this data in their individual institutional settings. Data collection should remain respectful to the dignity and rights of individuals but allow public bodies to operate appropriate safeguarding guidance and monitor outcomes.

The Public Sector Equality Duty [\[PSED\]](#) requires some public bodies in Scotland to gather employment information on the protected characteristics of their employees, including sex and gender reassignment.

The PSED is set out at section 149(1) of the Equality Act:-

(1) A public authority must, in the exercise of its functions, have due regard to the need to—

- (a) eliminate discrimination, harassment, victimisation and any other conduct that is prohibited by or under this Act;
- (b) advance equality of opportunity between persons who share a relevant protected characteristic and persons who do not share it;
- (c) foster good relations between persons who share a relevant protected characteristic and persons who do not share it.

To enable better performance of that duty, the Equality Act 2010 (Specific Duties)(Scotland) Regulations 2012 imposes specific duties on listed authorities. For example, listed authorities in Scotland (including health boards, colleges, and councils) have specific duties under this legislation to gather and use equalities information about their employees, to help them understand the diversity of their workforce, and to reduce any inequalities that exist.

The EHRC, who are responsible for ensuring compliance with the PSED, have stated that the requirement to gather employment information does not mean that employment information on the basis of legal sex (see p. 7 above, on the Equality and Human Rights Commission definition) must be recorded¹³. Outside of the employment context, there may be circumstances when collecting data on sex on a legal sex basis may be required in order for a public body to comply with the PSED. Additionally, they have said that public bodies need to find a balance between gathering appropriate data for a specific purpose and people's right to privacy.

As mentioned above, recording data on gender identity can support the equality monitoring and service development under the 2010 Act. In their evidence to the working group, the Equality and Human Rights commission wrote "*The language used by a body does not dictate whether the PSED requirements have been met. Public bodies are not required to collect information on legal sex and can enable employees and service users to self-identify their sex.*"

Asking individuals to disclose their sex or gender identity raises privacy issues. Public bodies should give consideration to whether an intrusion into someone's private life has a legitimate purpose and is proportionate. This is particularly true when asking trans people to disclose their trans status. As such, making sure questions offer an option of "prefer not to say" is very important. This has been considered when developing the standard questions proposed by this guidance.

For the majority of people their legal sex and gender identity are the same and they do not understand why they are asked the same question twice. Engagement has shown that trans people prefer to answer a self-identified sex or gender identity question followed by a trans status question. Asking this combination can support the principle of asking questions respectfully. This in turn leads to greater levels of response to the questions and better quality data that enables better decisions to be taken.

¹³ <https://www.gov.scot/binaries/content/documents/govscot/publications/minutes/2019/09/sex-and-gender-in-data-working-group-meeting-september-2019/documents/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/ehrc-submission-on-collecting-and-presenting-data-on-sex-and-gender/govscot%3Adocument/EHRC%2Bsubmission%2Bon%2Bcollecting%2Band%2Bpresenting%2Bdata%2Bon%2Bsex%2Band%2Bgender..pdf>

5. Data standards

There are no international standards for how to ask questions around sex and gender, though currently the vast majority of official data is collected on the basis of gender identity and there is very little data collected on the basis of sex registered at birth, or legal sex. With gender identity, it is important that people can express their own deeply held sense of their own gender, hence the suggested questions offering the opportunity to describe this in their own words.

However, as outlined in the statistical principles, using approaches that are consistent across countries is desirable. Data and statistics are developed to meet specific needs. Because of this, different statistics will often measure similar concepts using different definitions and classifications. When considering each set of statistics in isolation, this approach is OK. However, many uses of data require bringing together a number of sources. Having consistent definitions, language and question structure in data collection makes it much easier to interpret these different sources together.

Within the UK, there are some questions about these issues currently asked on face to face (sometimes in a non-private situation), phone and self-completion surveys. These are not currently done in a uniform way. A consideration here is that collecting data can happen through a few different approaches: via a face to face interview, on the telephone, or a paper form or online. The approach needed where someone asks a question (i.e. face to face or telephone interviewing) is subtly different from a situation where someone completes a form/written survey by themselves. Surveys and administrative systems can be completed in different ways. These should be covered in guidance on the different ways data can be collected

A person's legal sex or gender identity should be given by the individual concerned or their advocate (e.g. the parent of a child). It should not be inferred by an organisation based on a person's name or by observation.

This guidance provides standard questions, responses and guidance for gender identity, legal sex and trans status. Separate questions are given for interview led or written self-completion situations. These have been developed from existing questions used somewhere around the world. Data standards such as those shown below usually undergo a process of cognitive testing to make sure the questions work with those who are likely to complete them. While this has happened with some of what is proposed below in a Scottish/UK context, my intention is to make sure this happens for all questions included here.

The following questions that I am proposing have not all been fully tested and I am seeking feedback on them in their current form before undertaking any cognitive testing. I expect that this guidance will evolve and develop over time.

Interviewer led (Face to face/phone) interview – gender identity

This question should be used for describing the gender identity of the person being interviewed and if they are responding on behalf of someone else.

Question wording	How would you describe your gender? Or “How would you describe the gender of (name of respondent)?” when responding for another person
Question responses	<ol style="list-style-type: none"> 1. Man/Boy 2. Woman/Girl 3. In another way (if you would like to, please tell me what other words you use) 4. Prefer not to say <p>To not include in the question, but to code</p> <p>9. Not known</p>
Question guidance	<p>Ideally present a show card with options 1-4 Otherwise read out options 1-4</p> <p>For people who respond “3 - in another way”, their response to the self-declared description of that person’s gender identity should be recorded.</p> <p>The value “Not known” indicates that gender identity has not been recorded. This covers when someone has refused to answer the question or when the question has not been asked.</p>

Written (paper or online) questionnaire – gender identity

Question wording	How would you describe your gender? Or “How would you describe the gender of (name of respondent)?” when responding for another person
	Tick ONE box
Question responses	<ol style="list-style-type: none"> 1. Man / Boy 2. Woman / Girl 3. In another way 4. Prefer not to say <p>To not include in the question, but to code</p> <p>9. Not known</p>
Question guidance	<p>The value “in another way” will open a text field that records the self-declared description of that person’s gender identity</p> <p>The value “Not known” indicates that gender identity has not been recorded. This covers when someone has refused to answer the question or when the question has not been asked.</p>

Interviewer led (face to face/phone) interview – gender identity “in another way” description

For those who respond to the gender identity question “in another way”

Question wording	If you would like to, please describe the words you would use below:
Question responses	This is an open text response 9. <i>Not Applicable</i>
Question guidance	The value “Not applicable” indicates that the non-binary or trans gender identity question was not asked, i.e. the gender identity question has a response of 1, 2, or 4.

Written (paper or online) questionnaire – gender identity “in another way” description

For those who respond to the gender identity question “in another way”

Question wording	If you would like to, please write in the other words you would use below:
Question responses	This is an open text response 9. <i>Not Applicable</i>
Question guidance	The value “Not applicable” indicates that the non-binary or trans gender identity question was not asked, i.e. the gender identity question has a response of 1, 2, or 4.

Interviewer led (Face to face/phone) interview – legal sex

Question wording	What is your sex? Or “what is the sex of (<i>name of respondent</i>)?” when responding for another person
Question responses	<ol style="list-style-type: none"> 1. Female 2. Male 3. Prefer not to say <p>To not include in the question, but to code 9. Not known</p>
Question guidance	<p>Ideally present a show card with options 1-3. Otherwise read out options 1-3. At no time should the interviewer make a judgement of their own.</p> <p>If prompted, guidance should be given “this is sex registered at birth, or acquired sex for those with a Gender Recognition Certificate”</p> <p>The value “Not known” indicates that sex has not been recorded. This covers sex of an unborn child, when someone has refused to answer the question or when the question has not been asked.</p>

Written (paper or online) questionnaire – legal sex

Question wording	What is your sex?
Question responses	<ol style="list-style-type: none"> 1. Female 2. Male 3. Prefer not to say <p>To not include in the question, but to code 9. Not known</p>
Question guidance	<p>Guidance should be given “this is sex registered at birth, or acquired sex for those with a Gender Recognition Certificate”</p> <p>The value “Not known” indicates that sex has not been recorded. This covers sex of an unborn child, when someone has refused to answer the question or when the question has not been asked.</p>

Information on why this proposed sex question differs from the approach taken by the Census 2022:

In proposing a voluntary legal sex question (i.e. one that gives the option of 'prefer not to say') that asks whether someone is male or female, with guidance that this should be according to the sex registered at birth, or for trans people with a Gender Recognition Certificate, their acquired sex, I acknowledge that this is different to the approach taken for the 2022 Census in Scotland, which will ask a compulsory binary sex question, with guidance that people can self-identify their sex. Previous censuses have also asked a self-identified sex question.

As I have said before, I see the Census as a distinct data collection, whereas the guidance that I am putting together intends to be applicable to a wide range of data collections across the Scottish public sector. One of the key differences is that Scotland's Census is underpinned by legislation¹⁴ (which requires that a person must answer a sex question. The detail of the sex question in Scotland's census (a binary sex question asked on a self-identification basis) was recommended by NRS as being the best balance in meeting the diverse range of user needs. Their position is that a binary sex question with self-identification guidance enables census participation for all people and clarifies to data providers and data users the basis of the question. This recommendation was based on extensive testing and feedback from stakeholders.

¹⁴ [Census \(Scotland\) Order 2020](#), [Census \(Scotland\) Regulations 2020](#), and the [Census Act 1920](#)

Interviewer led (Face to face/phone) interview – trans status

Question wording	Do you consider yourself to be trans, or have a trans history?
Question responses	<p>1. No 2. Yes 3. Prefer not to say</p> <p>To not include in the question, but to code</p> <p>9. Not known</p>
Question guidance	<p>Here we use trans is as a term used to describe people whose gender identity is not the same as their sex registered at birth.</p> <p>To only be asked for people aged 16 or over</p> <p>Ideally present a show card with options 1-3. Otherwise read out options 1-3.</p> <p>Responding with the answer 2-Yes opens the question on trans status description</p> <p>The value “Not known” indicates that trans status has not been recorded. This covers when someone has refused to answer the question or when the question has not been asked.</p>

Interviewer led (Face to face/phone) interview – trans status description

For those who respond to the trans status question “yes”

Question wording	If you would like to, please describe your trans status (for example trans man, trans woman):
Question responses	<p>This is an open text response</p> <p>9. <i>Not Applicable</i></p>
Question guidance	The value “Not applicable” indicates that the trans status description question was not asked, i.e. the trans status question has a response of 1, 3 or 9.

Written (paper or online) questionnaire – trans status

Question wording	Do you consider yourself to be trans, or have a trans history? Tick ONE box only
Question responses	1. No 2. Yes 3. Prefer not to say To not include in the question, but to code 9. Not known
Question guidance	Here we use trans as a term to describe people whose gender identity is not the same as their sex registered at birth To only be asked for people aged 16 or over The write-in box for the trans status description should be visible underneath the trans status question The value “Not known” indicates that trans status has not been recorded. This covers when someone has refused to answer the question or when the question has not been asked.

Written (paper or online) questionnaire – trans status description

For those who respond to the trans status question “yes”

Question wording	If you would like to, please describe your trans status (for example trans man, trans woman):
Question responses	This is an open text response 9. <i>Not Applicable</i>
Question guidance	The value “Not applicable” indicates that the trans status description question was not asked, i.e. the trans status question has a response of 1, 3 or 9.

6. Information to accompany data collection

It is important that people understand why they are being asked to disclose certain information, and what a public body will use the collected data for. We recommend that an overview of this background information is provided in a short paragraph before the questions are asked. For example, question testing by the ONS on their interim gender identity standard found that question acceptability increased when people understood that the information was being collected to carry out equalities monitoring.

Sometimes people may be concerned that their answers won't be confidential and other people may not understand why they're being asked for certain information. This can result in a large number of people responding 'prefer not to say' to a question. To encourage high response rates, you should provide more information on why the data is being collected and what it will be used for. Depending on the purpose of data collection, you may want to ask a person's consent to disclose or use the data they have provided for a specific purpose.

For example, a public body could explain:

- Why you are collecting the data
- How the data will be stored (whether anonymously or confidentially), and who will have access to this
- How the data will be used
- Examples of how previous data collections have had a positive impact

For some people, for example trans people with a Gender Recognition Certificate (or trans people without), it will be particularly important that there is a legitimate reason for asking their sex, gender identity (or trans status), and whether this information will be disclosed, and what it will be used for. Maintaining an individual's human rights, legal rights and right to privacy is an important consideration. For example, collecting information anonymously, or having explicit consent from an individual to disclose their information would allow potentially sensitive questions to be asked in a way that maintains the legal rights of a person (including a trans person under the Gender Recognition Act (Scotland) 2004).

7. Analysing, disaggregating and publishing data that has been collected

My expectation is that organisations across Scotland should be publishing disaggregated data that illuminates the situation for men and women, and actively looking to review their data collection to do more to both collect and publish disaggregated data, therefore helping to understand where there are differences and where there are not. This is both just for gender or sex, and its intersectionality with other socio-demographic characteristics whilst preserving privacy.

A public body will need to consider precisely what disaggregation is required. There are a number of considerations that are relevant to reporting this data:

Legal

An organisation should also consider whether it has any specific duties under the public sector equality duty in relation to disaggregating and reporting data.

Public bodies who are subject to the PSED must give due regard to the need to eliminate discrimination, advance equality of opportunity and foster good relations in relation to people who share protected characteristics, when they are providing services. While there is no specific duty to collect information to do this, it would be difficult, if not impossible to meet the duty without adequate data relating to who uses their services.

Public bodies who are listed for the purposes of the Equality Act 2010 (Specific Duties) (Scotland) Regulations 2012 will also require adequate information to prepare Equality Impact Assessments.

Those public bodies must also publish a set of equality outcomes which they consider will enable them to better perform the Public Sector Equality Duty (PSED). A public body must publish a report on the progress made to achieve its equality outcomes every two years and publish a fresh set of equality outcomes within four years of publishing its previous set. Again, information from data must underpin the development and monitoring of these equality outcomes.

Information needs

A key starting point in thinking about how to disaggregate between men and women should be to consider why the data is being collected, and what [equality] objectives the organisation is trying to achieve. The guidance on data collection covers this and gives some details of the reasons why a public body may need to collect data that is disaggregated between men and women.

Statistical

The [Code of Practice for Statistics](#) provides a framework for the production of statistics, to make sure that the things that most need to be measured in society are collected and reported in a way that maintains trust, quality and value of the data. The Code or Practice places “above all, the confidentiality of individuals and of business information”, and ensuring that this is protected. This is very relevant to publishing and reporting statistics on small groups, where there is a risk that an individual’s data may be identifiable in breach of their right to privacy.

The Code is not concerned only with official statistics. It provides a framework that can apply to a much wider range of data that have not traditionally been described as official statistics. Providers of these other types of data, and this includes data held by public bodies, can draw on the Code as they judge appropriate to help support public confidence, and serve the public good. For example, it is accepted good practice that if the analysis of data collected shows results which are unexpected, the reasons for those results should be investigated

It is important that when reporting statistics that the strengths and limitations of the data are communicated clearly, and this includes making clear how the data was collected, whether face to face, whether in writing, over the phone, and the wording of question used (e.g. was it a legal sex question, or a gender identity question with a trans history question). This allows users to take a decision on whether the data is suitable for their needs and uses.

[The Government Statistical Service](#) produces a number of resources related to communicating statistics, and this guidance can support the visual presentation of the statistics and data to conform to ways in which the user can easily digest and understand, so that your statistics and data are accessible.

8. Intersectionality

An intersectional analysis gives an insight into how a combination of socio-demographic characteristics might create discrimination. There is an increasing awareness that taking an intersectional approach to research, policy making and operational decisions is important, as intersectionality can give insight into the experiences of different groups in society, and how particular characteristics can combine to impact on an individual's experiences. However, there is not always disaggregated data available to support such an approach. This is likely to mean that the information on which important decisions are made is not representative of the population it is intending to measure.

For example, data on the protected characteristics may be collected by a public body, but not disaggregated in an intersectional way due to issues around sample size and risk of disclosing an individual's identity. In these cases, organisations should not risk disclosing information about a person's sex or gender identity. However, where sample size and quality allows, data should be disaggregated by socio-demographic factors, including sex, gender identity, gender reassignment, religion, age, disability, ethnicity and sexual orientation, where combinations of these factors can result in discrimination, disadvantage and inequality. Being able to identify cases where combinations of factors are resulting in disadvantage enables policies to be developed and action taken to address these issues.

9. Harmonisation

There is a need to apply a more consistent approach to the collection of data around sex and gender. The Working Group has built on work by Scottish Government Equalities analysts in this area.

Many uses of data require bringing together a number of sources. Having consistent definitions, language and question structure in data collection makes it much easier to interpret these different sources together. The proposed questions in this guidance aim to facilitate better harmonisation across the Scottish public sector in collecting data about sex and gender identity.

Harmonisation is important to allow data collected in different sources, for example on a survey or in an administrative system, to be combined or compared, and still reported in a meaningful way. Harmonisation is an important consideration in terms of publication, for example, it is generally desirable to have harmonised questions and definitions so that data can be aggregated, linked, and disaggregated to smaller levels. This also allows for comparability between groups and other data sources when this data is reported.

Increasing sample size of data about particular groups should in theory increase the usefulness of the data, and allow for an intersectional analysis to be carried out while minimising the risk of disclosure issues. This could be achieved by better linkages between datasets and data sharing between public services, and an important determinant of being able to link data is harmonisation.

The underlying principles that the working group has adopted, based on professional codes of practice and frameworks, mean that a number of presentational issues when publishing this data will ultimately be determined by consideration of professional statistical matters (for example, disclosure control and data reliability, etc.).

10. Disclosure control

Individual data records are important for operational decision making, as well as for research purposes; however, due to the sensitive nature of the information contained in an individual's data, their direct release would violate an individual's right to confidentiality and privacy, and be at odds with data protection legislation.

This is also central to the Code of Practice for Statistics, where the majority of the statistical principles are relevant to disaggregating and presenting data, in particular, *Practice T4.2 under the Trustworthiness pillar, which states that:*

- *The identity of individuals or organisations must be protected at all times. Appropriate disclosure control methods, including the nature of any consent given, should be applied when releasing statistics (Practice T4.2)*

And the [UN Principles of International Statistics](#) also emphasise the importance of maintaining confidentiality:

6. Individual data collected about natural persons and legal entities, or about small aggregates that are subject to national confidentiality rules, are to be kept strictly confidential and are to be used exclusively for statistical purposes or for purposes mandated by legislation

Good practices include:

- *Putting measures in place to prevent the direct or indirect disclosure of data on persons, households, businesses and other individual respondents*
- *Developing a framework describing methods and procedures to provide sets of anonymous micro-data for further analysis by bona fide researchers, maintaining the requirements of confidentiality*

Once data is collected, the need to balance confidentiality and utility of the data, mean that statistical disclosure control methods must be used, in conjunction with sound and appropriate methods to ensure that this data can produce robust results.

Statistical disclosure control aims to protect anonymity of the individual who has provided the data, and maintains confidentiality and privacy by using one of a range of disclosure control methods (e.g. data swapping, recoding, suppression, or data identifiers can be removed or encoded and data fields can be modified by means of statistical disclosure controls), while overall the collective features of the resulting de-identified data are preserved.

Changing the data in this way can have an impact on the usefulness of the data. Therefore it is important to build quality assessment into being a central part of disclosure control practice, and to understand the relationship between disclosure control and data utility.

You should consult [Government Statistical Service \(GSS\) guidance](#) on statistical disclosure control before releasing data that could potentially be identifiable.

11. How to present data on non-binary groups

It is a challenge to collect data that fully reflects the variance in a population, and part of this is driven by the questions that are asked, the definitions ascribed to the terminology used, and the interpretation of these by a respondent. This in turn affects the degree to which findings across surveys and countries can be compared.

In terms of administrative data sources, there are potential changes to registration practices needed so that non-binary could be recorded as an option where statistics are currently disaggregated between male and female, as in this situation it is likely to be difficult to report on the population identifying as non-binary.

It is expected that the Working Group on Non-Binary People will provide more clarity on how non-binary people are recorded in data, and therefore how this data can be disaggregated and reported. This guidance will be reviewed with a view to updating once the non-binary working group has reported its conclusions. It is also expected that this guidance will evolve over time.

12. Data on sex and gender key questions

This provides a list of the important questions those collecting data must be able to answer while collecting and using data about sex, gender or trans status.

Purpose

- **Why are you collecting data on sex, gender or trans status?**
- **Are you clear how it is going to be used to improve wellbeing?**
- **Which question or questions best meets your aims from data collection?**

Bias

- **What are the steps you are taking to avoid bias and maximise response in the data collected?**
- **How would you know that there are no inherent biases in the data?**

Once you have the data

- **Does it tell you anything unexpected that can be used to improve people's wellbeing?**
- **How can you share it to maximise its value?**
- **How have you considered the maintenance of people's privacy in how you manage raw data and publish aggregate information?**