

Data collection and publication guidance

**Sex
Gender Identity
Trans Status**

September 2021



Scottish Government
Riaghaltas na h-Alba
gov.scot

Overview

I was asked by the then 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 scope of this work is about data collected and used by Scottish public bodies – for operational, statistical and research purposes. It is separate from Scotland’s Census 2022.

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 including biological and physical 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 was making sure that the guidance is informed by evidence from a wide range of individuals and organisations with views based 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. Finally, a public consultation was held on a draft version of this guidance.

I gathered together these insights and considered them when putting together this guidance.

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 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.

Roger Halliday, Chief Statistician

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 needs and experience of individuals and groups we need data that will allow for comparisons to be made.

The then 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 faced by women and men, but on the intersectionality between sex, gender identity and other socio-demographic characteristics (including the protected characteristics in the Equality Act 2010).

How data 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 trans status, 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 European Institute of Gender Equality Gender Statistics Database¹ identify good practice in data and statistics about sex and gender as:

¹ European Institute of Gender Equality. 2020. "Gender Statistics Database." European Institute for Gender Equality. Accessed February 17. <https://eige.europa.eu/gender-statistics/dgs>.

- 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 and analysing data about sex or gender, it is fair to say that these terms are sometimes used interchangeably in the questions used in 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 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 and/or gender identity 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 findings of the UN's Committee on the Elimination of Discrimination against Women who looked at the UK in 2019 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."²

What is clear is that there isn't a standard way that data about sex and gender identity is being collected, either in Scotland or the UK. For example in justice statistics for England there are a number of approaches³. 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.

In addition, it was clear that very little data is collected on people's trans status or history in Scotland, and there are no definitive international standards for how to do this.

² 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://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CEDAW/C/GBR/CO/8&Lang=En .

³ A Technical Guide to Statistics on Women and the CJS (publishing.service.gov.uk) (see pages 9 to 11).

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, gender identity and trans status.

The following are taken from the Code of Practice for Statistics⁵

- The Chief Statistician/Head of Profession for Statistics should have sole authority for deciding on methods, standards and procedures, and on the content and timing of the release of regular and ad hoc official statistics (Practice T2.1)
- Organisations should look after people's information securely and manage data in ways that are consistent with relevant legislation and serve the public good. (Principle T6)
- Statistics should be based on the most appropriate data to meet intended uses. The impact of any data limitations for use should be assessed, minimised and explained. (Principle Q1)
 - Statistics should be based on data sources that are appropriate for the intended uses. The data sources should be based on definitions and concepts that are suitable approximations of what the statistics aim to measure, or that can be processed to become suitable for producing the statistics (Practice Q1.1)
- Methods and processes should be based on national or international good practice, scientific principles, or established professional consensus (Practice Q2.1).
- Statistics, data and metadata should be compiled using recognised standards, classifications and definitions. They should be harmonised to be consistent and coherent with related statistics and data where possible. Users should be provided with reasons for deviations from these standards and explanations of any related implications for use (Practice Q2.2).
- Statistics producers should collaborate with topic and methods experts and producers of related statistics and data wherever possible (Practice Q2.6).
- Users of statistics and data should be at the centre of statistical production; their needs should be understood, their views sought and acted on, and their use of statistics supported. (Principle V1)
 - Statistics producers should engage publicly through a variety of means that are appropriate to the needs of different audiences and proportionate to the potential of the statistics to serve the public good. An open dialogue should be maintained using proactive formal and informal engagement to listen to the views of new and established contacts. Statistics producers should undertake public engagement collaboratively wherever possible, working in partnership with policy makers and other statistics producers to obtain the views of stakeholders. (Practice V1.4).
 - The views received from users, potential users and other stakeholders should be addressed, where practicable (Practice V1.5)..

⁵ <https://www.statisticsauthority.gov.uk/code-of-practice/>

The following is taken from the European Statistics code of practice⁶

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

⁶ <https://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice>

3. Definitions used in this Guidance

While the terms sex and gender are used by some people interchangeably, there is not strong consensus about the meaning of these terms in different contexts. A number of organisations have attempted to set out definitions of the aspects of sex, gender and transgender for the purposes of collecting consistent data⁷. The definitions stated are broadly similar across these organisations.

For the purposes of collecting data, a person's **sex** is generally defined as male or female. There are different aspects to a person's sex:

- **Biological:** as determined by a person's anatomy, which is produced by a combination of their chromosomal, hormonal, genital and gonadal characteristics, and their interactions.
- **Legal:** typically legal sex is their sex registered at birth. However, for a trans person with a full Gender Recognition Certificate, their legal sex is their acquired sex.⁸
- **Self-defined:** a person's innate sense of whether they are female or male

The Equality and Human Rights Commission (EHRC)'s "Statement on Sex and Gender Reassignment: legal protections and language"⁹ 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 Society¹⁰ define **gender** as:

- a social construction relating to a set of norms, roles and relationships that is founded in social mores, laws, processes and policies based on labels of masculinity and femininity. This is time and culture specific;

and **gender identity** as:

- a personal, internal perception of oneself, and so the gender category someone identifies with may not match their sex registered at birth
- what an individual experiences as their innate sense of themselves as a man, a woman, as having no gender identity, or as having a non-binary gender – where people identify as somewhere on a spectrum between man and woman

And define **transgender** as:

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

⁷ For example the [Office for National Statistics](#), [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](#) .

⁸ <https://www.equalityhumanrights.com/en/advice-and-guidance/gender-reassignment-discrimination>

⁹ <https://www.equalityhumanrights.com/en/our-work/news/our-statement-sex-and-gender-reassignment-legal-protections-and-language>

¹⁰ 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](#) .

For the overwhelming majority of people, all the factors involved in determining sex and gender identity are aligned. For some people, there are differences. People will not necessarily answer a question about their sex in exactly the same way: most will need no reflection at all, some will think about their biology, some their legal sex and some their self-defined sex.

4. Considerations when deciding what data to collect

It is a fundamental role for public service in Scotland to eliminate discrimination, foster good relations and advance equality in our society. This helps to meet our purpose in the National Performance Framework of creating “opportunities for all of Scotland to flourish”, as well as meeting legal obligations in the Equality Act 2010 through the Public Sector Equality Duty.

Within this overarching aim, there will be different reasons organisation have for collecting data and the uses should determine the data collected. As such, 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. There is a lot of further guidance on how best to do this online, for example this Government Social Research Quota sampling guidance¹¹.

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 UK General Data Protection Regulation (UKGDPR). In addition, people have the right to access to data about them that is used for operational decisions and that this data is rectified if there are errors. So those collecting data for operational decisions need to take steps to making sure it remains correct.

The Equality and Human Rights Commission have guidance on the public sector equality duty and data protection¹².

The Information Commissioner’s Office (ICO) produce guidance to support organisations to only collect personal data that is needed for a specific purpose, to ensure that sufficient data is collected to fulfil this purpose, and to periodically review the personal data held and processed, checking that it is still relevant and adequate for these purposes, deleting anything no longer needed Principle (c): Data minimisation | ICO¹³.

The ICO also recommend that organisations consider potential risks to the rights of individuals when making decisions about collecting personal data, as well as the risks of not collecting and using information as well as collecting and processing it¹⁴:

¹¹ <https://gss.civilservice.gov.uk/wp-content/uploads/2018/03/Quota-sampling-guidance-4.pdf>

¹² <https://www.equalityhumanrights.com/en/publication-download/public-sector-equality-duty-and-data-protection>

¹³ <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/>

¹⁴ <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/how-do-we-do-a-dpia/#how10>

A Data Protection Impact Assessment aims to consider these points: The ICO provide some guidance on how to put this together¹⁵.

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

- Operational decisions by public sector services about the service people receive
- Designing more inclusive and 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. This includes meeting legal obligations to collect certain data and use it to monitor equalities.
- An evidence based understanding of service need, to, for example, inform funding and planning decisions.
- Helping the public to understand the extent to which progress has been made in the equality and diversity of our nation.

For the vast majority of people, questions on someone's sex and gender identity will provide the same response: 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, a proportion of people will not answer questions on sex by thinking about their biological or legal sex, and instead will respond according to aspects such as their self-defined sex, and their innate feeling of whether they are a woman or a man.

Given that for the vast majority of people sex and gender identity questions will provide the same result, for most issues one may want to measure, whether there is a question about sex or about gender identity, it will not skew the statistics when disaggregated by either concept.

However, the National LGBT survey¹⁶ reported that 13% of respondents were trans, 6.9% were non-binary, 3.5% were trans women, and 2.9% were trans men. This study identified differences according to respondents' age, with younger respondents more likely than older respondents to identify as non-binary, and younger respondents more likely to be trans men, and less likely to be trans women. As these survey results demonstrate, there are likely to be groups in the population where the proportion of trans people is higher, and public bodies should be mindful of possible impacts on the data, for example in terms of quality and comparability, and when considering how to collect and present aggregate data on groups containing small numbers of people in a way that protects privacy and respects disclosure control measures.

¹⁵ "How do we do a DPIA?" <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/how-do-we-do-a-dpia/#how10>

¹⁶ [National LGBT Survey: Summary report - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/612222/national-lgbt-survey-summary-report-2018.pdf)

Where it is not necessary and proportionate, a question requiring the disclosure of a person's biological sex may be an unjustifiable breach of privacy: in some cases this would have the potential to reveal a trans history that otherwise a person may wish to keep private. In a small number of instances, it may be necessary and proportionate to require a person to answer a question on their biological 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. The most likely scenarios where data on biological sex is required would be on a case-by-case basis in a medical context; in a criminal context where a serious sexual offence is being investigated.

This guidance is not prescriptive 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 monitor outcomes and design service improvements.

The Equality and Human Rights Commission (EHRC), who are responsible for ensuring compliance with the Public Sector Equality Duty (PSED), have stated that the requirement to gather employment information does not mean that employment information on the basis of legal sex must be recorded¹⁷. Outside of employee data, there may be a small number of circumstances when collecting data on self-defined sex only could contribute to the failure of a public body to comply with the PSED. Therefore it is important that public bodies find a balance between gathering appropriate data for a specific purpose and people's right to privacy.

Recording data on gender identity and trans status can also support equality monitoring in relation to the 2010 Act protected characteristic of gender reassignment.

Many data collections are used for more than one purpose: for supporting operational decisions, supporting organisational management and improvement, providing accountability, and broader research. As such, many organisations will want to be able to report about sex, gender identity and trans status.

For the majority of people their sex and gender identity are the same and they do not understand why they are asked the same question twice. However, engagement as part of the 2022 Scottish Census has shown that trans people prefer to answer a sex question (on a self-defined sex basis) 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.

¹⁷ <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>

Asking individuals to disclose their sex may raise 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 questions in this guidance.

The Scotland's Census 2022 question on sex and the voluntary question on trans status or history have been thoroughly tested, are respectful and offer a way to collect data on both sex and trans status.

5. Data standards

There are no international standards for how to ask questions around sex and trans status, though currently there is very little data collected explicitly on the basis of sex registered at birth, or biological sex, and it is assumed that the vast majority of official data is collected on the basis of self-defined sex. For example, some questions ask 'what is your sex?', others ask 'are you male or female?' with no guidance on how people should interpret the question.

However, as outlined in the statistical principles, using approaches that are consistent across Scotland and between nations 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 okay. 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.

As such, and recognising that for most organisations and purposes that data on sex, and trans people will be needed, this guidance proposes asking questions on both sex and trans status as standard questions. This includes a question offering people to describe their trans status. This is likely to increase response rate by allowing people to describe things in their own words rather than be prescriptive about categories that people should fit to. Also, terminology is quickly evolving and the possible responses may become out of date. As described in the previous section, there will be cases where collecting both of these data items is not needed.

Within the UK, there are a range of questions currently asked in 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.

A question on a person's sex or trans status should be answered by the individual concerned or their advocate. It should not be inferred by an organisation based on a person's name or by observation. Testing of a gender identity question for Scotland's Census 2022¹⁸ showed that 16% of people found that responding for a person under 16 was not acceptable for them, which is significantly higher than for people responding on behalf of someone 16 and over. As such, the voluntary trans status or history question in Scotland's Census 2022 is only being asked of people aged 16 and over.

This guidance provides standard questions, responses and guidance for 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

¹⁸ <https://www.scotlandscensus.gov.uk/documents/sex-and-gender-identity-topic-report/>

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. This has happened the paper/online questions as part of testing for the 2022 Scottish census. My intention is to make sure this happens for the questions included here in a face to face setting. The outcomes of this testing will be published.

I intend that this guidance will evolve and develop over time alongside developments in the terminology used.

Interviewer led (Face to face/phone) interview – Sex

Question wording	What is your sex? Or “what is the sex of (name of respondent)?” when responding for another person
Question responses	1. Female 2. Male 3. Prefer not to say To not include in the question, but to code 9. Not known
Question guidance	This aligns with the guidance for Scotland’s Census 2022 ¹⁹ with the exception that the respondent may state that they would prefer not to answer the question. The value “Not known” indicates that sex has not been recorded. This covers the sex of an unborn child, when someone has refused to answer the question or when the question has not been asked.

Written (paper or online) questionnaire – Sex²⁰

Question wording	What is your sex?
Question responses	1. Female 2. Male 3. Prefer not to say To not include in the question, but to code 9. Not known
Question guidance	This aligns with the guidance for Scotland’s Census 2022 ¹⁹ with the exception that the respondent may state that they would prefer not to answer the question. The value “Not known” indicates that sex has not been recorded. This covers the sex of an unborn child, when someone has refused to answer the question or when the question has not been asked.

¹⁹ <https://www.scotlandscensus.gov.uk/media/nzljdqvn/sex-question-guidance-30-august.pdf>

²⁰ In proposing a voluntary sex question (i.e. one that gives the option of 'prefer not to say') that asks whether someone is male or female, I acknowledge that this is different to the approach taken for Scotland’s Census 2022, which will ask a mandatory binary sex question underpinned by legislation . Most surveys and data collection are not compulsory, though response for this question is typically very high. This recommendation was based on extensive testing and feedback from stakeholders

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 non-binary, trans man, trans woman):
Question responses	<p>This is an open text response</p> <p>8. Prefer not to say 9. Not Applicable</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 non-binary, trans man, trans woman):
Question responses	This is an open text response 9. Not Applicable
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. I 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 Office of National Statistics (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.

The ICO guidance on transparency gives further details on the type of information that organisations should consider including when collecting personal data, such as a privacy notice that outlines the basis of consent for processing, retention periods, who might have access to the data and those to whom the data might be disclosed: Transparency | ICO²¹

Sometimes people may be concerned that their answers won't be confidential and some 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, how the data will be kept secure and confidential, 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. Again, the ICO have guidance for organisations to consider around consent, and it is recommended that organisations familiarise themselves with this²².

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, it will be particularly important that there is a legitimate reason for asking their sex, and/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 2004).

²¹ <https://ico.org.uk/for-organisations/accountability-framework/transparency/>

²² <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/>

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

My expectation is that organisations across Scotland, as part of their focus on promoting equality, should be publishing disaggregated data that illuminates the situation for people who are female and male, and for trans people, therefore helping to understand where there are differences and where there are not. I'd also expect organisations to actively look to review their data collection to do more to both collect and publish disaggregated data. This is both for sex and trans status, and 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:

Information needs

A key starting point in thinking about how to disaggregate between those who are female or male, and present data about trans people 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 in this way.

Legal

Public bodies who are listed for the purposes of the Equality Act 2010 (Specific Duties) (Scotland) Regulations 2012 will also require adequate information to assess the impact of applying a proposed new or revised policy or practice against the needs of the PSED, for example by completing a Equality Impact Assessment.

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.

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 data²³.

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

²³ <https://www.statisticsauthority.gov.uk/code-of-practice/>

Government Statistical Service has published guidance on how to present data in ways that preserve privacy²⁴

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. See the Ministry of Justice technical guide as an example of doing this²⁵. 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 accessible²⁶.

²⁴ <https://gss.civilservice.gov.uk/policy-store/gssgsr-disclosure-control-guidance-for-tables-produced-from-surveys/>

²⁵ [A Technical Guide to Statistics on Women and the CJS \(publishing.service.gov.uk\)](https://publishing.service.gov.uk/guidance/a-technical-guide-to-statistics-on-women-and-the-cjs)

²⁶ <https://gss.civilservice.gov.uk/guidances/communicating-statistics/>

8. Intersectionality

An intersectional analysis gives an insight into how a combination of socio-demographic characteristics might relate to specific forms of disadvantage. 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 fully representative of the population it is intending to measure.

For example, data on demographic 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 trans status. The Government Statistical Service guidance mentioned in section 7 above can help here.

Where sample size and quality allows, data should be disaggregated, including by sex, gender reassignment, race, religion or belief, age, disability, 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, gender identity, and trans status. The Working Group has built on work by Scottish Government Equalities analysts on 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 trans status.

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 also an important consideration when publishing data. 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).

This is a rapidly evolving area for the collection of data, and for data standards. As such, statisticians in Scottish Government will continue to work with those in the Office for National Statistics and elsewhere to make sure the standards laid out here continue to be relevant.

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 Statistics²⁷ 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

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 integrity and usefulness of the anonymised data is preserved.

You should consult Government Statistical Service (GSS) guidance on statistical disclosure control before releasing data that could potentially be identifiable²⁸.

²⁷ <https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>

²⁸ <https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/policyforsocialsurveymicrodata>

11. How to collect and 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 Scottish Government's Working Group on Non-Binary Equality will consider how better recording, disaggregating and reporting of data could be used to improve equality for non-binary people. This guidance will be reviewed with a view to updating, in light of any relevant recommendations made by the Working Group.

12. Key Questions – Checklist for Data Collectors

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

Purpose

- Why are you collecting data on sex, and/or trans status?
- Are you clear how you are going to actually use the data to improve your services and promote equality?
- Which question or questions best meets your aims and purpose(s) of the data collection?
- Do you need to carry out a Data Protection Impact Assessment (DPIA)²⁹? This is a legal requirement when processing special category data³⁰.

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 that can be used to improve your services and promote equality?
- 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?
- Do any surprising results need to be investigated further?

²⁹ <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/what-is-a-dpia/>

³⁰ <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/#scd1>



Scottish Government
Riaghaltas na h-Alba
gov.scot

© Crown copyright 2021

OGL

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at www.gov.scot

Any enquiries regarding this publication should be sent to us at

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-80201-384-9 (web only)

Published by The Scottish Government, September 2021

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS943446 (09/21)

W W W . g o v . s c o t