

Data Gathering - Gender and Sex

In recent years there has been much greater appreciation of gender identities outside of the male/female binary, as well as transgender identities. In some cases, this has led to confusion over how to gather data relating to sex and/or gender. This guide aims to help you decide what you need to ask and how you need to ask it. This may be as part of academic research, undergraduate projects or internal university projects like staff and student surveys.

Q1 Do you need this data at all?

Ask yourself whether you need data on sex and/or gender at all. UK GDPR states that you should only collect personal data which is necessary for the task. It is easy to think of a question on sex or gender as “standard” when collecting data on people, but you should not do so if it is not relevant or required for your work.

Some personal data relating to sex and/or gender should be treated with greater care, if the context of their use could create significant risks, interfere with individuals’ fundamental rights, or subject someone to discrimination. If you need to use sensitive personal data about sex and/or gender, consider the possible risks associated with its use and put in place additional safeguards. Completing a Data Protection Impact Assessment (DPIA) will assist in identifying potential risks and appropriate solutions to reduce or remove risk.

Q2 Do you need to know about sex, or gender?

Sex and gender are not easily separated. Some conceptions of gender believe it to be based on sex. Some conceptions of sex believe it to be based on gender. Both are socially constructed – shorthand simplifications of complex matters. Below we offer one simplified way to try and distinguish the two.

Sex usually describes someone’s biology – their genes, hormones, internal and external genitalia and secondary sex characteristics, like body hair, breasts etc. Combinations of these five biological components are used to define the social construct of sex. For some people these components all fit either the male or female archetype, for others they do not. Trans people who have undergone medical interventions such as surgery and hormone therapy may have a different combination of these components, as do people with intersex variations. It is also worth reflecting on the fact that most people have not had their genes sequenced, hormones measured, or internal genitalia examined. As a result, their recorded legal sex is generally assumed based on external genitalia and secondary sex characteristics – only two out of the five components. In most countries legal sex can only have one of two values, male or female, despite around 1.7% of the population having intersex variations which means they do not perfectly fit either category.

Gender describes how someone feels (gender identity), how they appear (gender expression) and how they are perceived and treated in the world. This is normally based on those components of sex which are known, and on variable definitions of what is perceived in a given time and place as masculine or feminine. Gender is therefore also a social construct, the bounds of which vary considerably between cultures, location and time.

Ask yourself which of these is relevant to your work, and be clear on why you need this personal data. In most cases gender will be most relevant, in some cases it may be both.

Q3 How should you phrase the question?

In most contexts It is useful to acknowledge that it may not be possible to perfectly capture someone's sex and/or gender. Good examples that allow for this would be:

“Which of the following best describes your sex?”

Or

“Which of the following best describes your current gender identity?”.

This also allows for the fact that gender identity can change over time.

Q4 What categories should you include?

You should think carefully about the work you are carrying out to decide how best to capture this data. Providing an open text field allows the greatest flexibility and inclusivity but may then present challenges around large numbers of respondents and/or the need for you to code these responses into categories. If you are going to prescribe fixed responses, then you should consider more than 2 options whether you are asking about sex or gender. If sex or gender is not essential data it is good practice to additionally provide a “Prefer not to say” option in addition to those discussed below, although in this case you may want to re-assess Q1 – do you need this data at all?

It is worth bearing in mind that the language in this area, around gender identity in particular, is evolving quite fast and so the more you commit to specified options the greater danger that your text becomes out of date.

Sex

Participants from most countries, including the UK, will have a legally recorded sex which can only be male or female. [Some countries](#) allow a third sex marker for people born intersex. You should therefore offer suitable categories depending on where your participants may be.

If it is relevant to capture whether people have an intersex variation, whether or not this is reflected in their legal sex, you should ask a separate question. We suggest:

“Are you intersex and/or have a variation of sex characteristics (VSC)?”

This should have yes and no options and “prefer not to say”. If relevant you may want to include a text box to allow participants to give further details.

Gender

There are many gender labels, and your decision on how to capture them will depend on what you are doing, but as general guidance in order of preference we recommend the following.

1 Free text box for everyone

This is the most inclusive option as it treats everyone the same and allows everyone to self-describe.

2 Three options with text box

You should include male, female (see note below on “man” and “woman”), and a third option which then allows the participant to specify their identity in a text box. You should not label a third field as “Other”, as this can imply a lack of regard or respect. The label will need to correspond to the way you've asked the question.

For example:

Heading “Which of the following best describes your current gender identity”
Label “Another term”.

Heading “How would you describe your gender”
Label “Another way”.

Heading “Gender”
Label “If you describe your gender with another term, please enter it here”.

This will reduce the need for manual coding of responses but is not as inclusive of people who are not male or female.

3 Three fixed options

If it is not possible to provide a free-text option, for example because you expect a very large number of responses, we recommend including male and female (see note below) and labelling a third option descriptively. Again this will depend on how you’ve phrased the question. For example:

Heading “Which of the following best describes your current gender identity”
Label “Another term (for example, but not limited to nonbinary, genderfluid, or agender)”.

Heading “How do you describe your gender?”
Label “I describe my gender with another term (for example, but not limited to nonbinary, genderfluid, or agender)”.

Heading “Gender”
Label “I describe my gender with another term (for example, but not limited to nonbinary, genderfluid, or agender)”.

This further reduces the need for re-coding of data but is the least inclusive option as some participants are likely to find the answers available unsatisfactory.

Q5 What about people who are transgender?

Transgender means a person has a gender identity which differs from the sex they were assigned at birth. “Transgender” is **not** a gender in itself. It should not be listed alongside options such as male and female or man and woman. The same goes for labels like “Trans man” and “Trans woman”.

If knowing whether a person is transgender is relevant to your work you should ask a separate question: “Is your gender different from the sex you were assigned at birth?” with Yes and No options and, in most cases, a “Prefer not to say” option. This category is sometimes described as “Gender Modality”, however this is not a commonly used term so think about your audience if you plan to use it.

Many, but not all, people whose gender identity is outside of the man/woman binary identify as transgender, or trans, as this identity is different from the sex they were assigned at birth. You should not assume people who identify outside of the binary also identify as transgender.

Note on Male/female vs man/woman

In some academic areas the terms “man” and “woman” are preferred over “male” and “female” as gender labels to differentiate from sex labels. If this is the case in your field you may want to use these labels however consider the following points:

Are those the labels you expect your participants would use/associate with gender in everyday language?

Man and Woman are nouns, whereas other identities such as Non-binary and Genderfluid are adjectives and most do not have a noun equivalent. It may feel more consistent therefore to use adjective terms “male” and “female” alongside other adjective labels.

The use of nouns can imply gender is the defining feature of our identity which for many people is not the case.

Further information

Sex, gender and terminology

Advance HE advice for capturing gender as part of equality monitoring:

<https://www.advance-he.ac.uk/guidance/equality-diversity-and-inclusion/using-data-and-evidence/monitoring-questions/gender>

The Office of National Statistics has some excellent further guidance available at:

<https://www.ons.gov.uk/economy/environmentalaccounts/articles/whatisthedifferencebetweensexandgender/2019-02-21#definitions-and-differences>

And you can find more guidance on gender-terminology at

<https://www.stonewall.org.uk/help-advice/faqs-and-glossary/glossary-terms>

<https://genderedintelligence.co.uk/projects/kip/transidentities>

Data protection

Information and guidance on working with sensitive ‘special category data’ is available from the ICO at:

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

Templates and guidance on conducting a Data Protection Impact Assessment (DPIA) can be found on the University website at:

<https://www.lboro.ac.uk/data-privacy/resources/dpia/dpia-process>

Still need help?

Students can seek further advice from their academic supervisor.

Queries related to research can be directed to the ADR or EDI Director for the school.

For work internal to the University (for example surveying staff or students), or if the channels above need to seek further guidance please email the EDI team at EDI@lboro.ac.uk

Appendix: Glossary of gender terms (non-exhaustive)

New gender-related terms are being created all the time but this short glossary should help you understand some of the most common.

Agender: a person who does not identify with any gender. Often feels they are neither masculine nor feminine.

Genderfluid: a gender identity which changes or fluctuates over time

Genderqueer: a person who does not subscribe to conventional gender distinctions but identifies with neither, both, or a combination of the masculine and feminine. Note that “queer” has a history of being used as a slur so while some people identify as queer or genderqueer it is not a label you should apply to people unless you are sure they apply it to themselves.

Non-binary: an identity which falls outside of the male/female gender binary. Non-binary is used both as a specific identity and often as an umbrella-term for other identities which fall outside of the binary. However not everyone who uses one of these other identities would also identify as non-binary.

Two-spirit: an identity which exists in some native North American cultures which combines elements of masculine and feminine. This label should not be used by people outside of those cultures.

A longer list of identities (also non-exhaustive) is at https://nonbinary.miraheze.org/wiki/List_of_nonbinary_identities