



Department for
Science, Innovation
& Technology

Independent report

Review of data, statistics and research on sex and gender: executive summary

Published 19 March 2025

Contents

Background to the review

Executive summary

Recommendations

Full recommendations



© Crown copyright 2025

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.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 <https://www.gov.uk/government/publications/independent-review-of-data-statistics-and-research-on-sex-and-gender/review-of-data-statistics-and-research-on-sex-and-gender-executive-summary>

Background to the review

This [independent review was commissioned \(https://www.gov.uk/government/publications/review-of-data-statistics-and-research-on-sex-and-gender-terms-of-reference\)](https://www.gov.uk/government/publications/review-of-data-statistics-and-research-on-sex-and-gender-terms-of-reference) in February 2024 by the Secretary of State for Science, Innovation and Technology. The aims of the review are as follows:

- Identifying obstacles to accurate data collection and research on sex and on gender identity in public bodies and in the research system
- Setting out good practice guidance for how to collect data on sex and gender identity

All public bodies, as defined by the Cabinet Office, are in scope of the review. The review also considers research institutions and organisations from outside the public sector, where relevant to the aims of the review. The review is UK wide, respecting the devolved nature of areas of responsibility within the research and development landscape and the collection of relevant areas of data and statistics.

This report concerns data and statistics. A further report will examine barriers to research. The review is led by Professor Alice Sullivan, University College London, assisted by policy analysts Murray Blackburn Mackenzie, and Dr Kathryn Webb, University of Oxford.

Approach to the review

- We carried out a review of policies, guidance, datasets and statistics, including administrative data, major flagship surveys, independent academic studies, clinical trials, polling data, and marketing exercises.
- We held over 30 stakeholder interviews with organisations ranging from government departments, regulators and other public sector organisations to fieldwork agencies and those campaigning for the rights of women and LGBT advocacy groups.
- We held an open call for submissions, which ran from 7 May to 1 July 2024. Individuals were invited to submit examples of UK data collection on sex and/or gender identity which they perceived as inadequate or flawed.
- We commissioned a legal opinion to ensure recommendations were compliant with relevant legal frameworks, such as GDPR and ECHR Article 8 rights.

Executive summary

Sex is a key demographic variable and collecting high quality, robust data on sex is critical to effective policymaking across a wide range of fields, from health and justice to education and the economy. It enables policymakers to measure and address disparities between women and men, and girls and boys. The government has a strong interest in promoting high-quality data on sex, both in its role as a funder of research and as a producer and user of statistics.

Accurate record keeping is also vital for operational purposes, for safeguarding and, within the healthcare system, for patient safety and care.

Stakeholders have told us that they recognise the importance of collecting data on sex. Confusion regarding the legal position has posed a barrier to collecting data on sex for some organisations. This report includes specialist legal advice which is presented in appendix 1.

Both people's material circumstances and their identities are important to their lives. We know that sex affects many dimensions of people's lives, and we have much to learn about the ways in which having a transgender identity matters too. Rather than removing data on sex, government and other data owners should collect data on both sex and transgender identities, in order to develop a better understanding of the influence of both factors and the intersection between them.

Collecting data on sex does not reduce people to biological categories, neither does it imply that people should conform with the stereotypes associated with those categories, nor does it deny the existence or experience of people with diverse gender identities. Indeed, people with diverse gender identities are being let down by data collection practices which conflate sex and gender identity, making it impossible to track the outcomes of distinct groups.

Respondents may understandably question why they are being asked for certain information. It is vital that everyone participating in research or providing data for administrative purposes is treated with respect and is informed about the reasons for data collection and reassured about the way their data will be processed and used. Some respondents may be reluctant to provide data on sex while others are reluctant to provide data on gender identity. In both cases, the purposes of the data collection, including potential benefits to the individual and society, should be explained.

Terminology used in the review

For clarity, the term 'sex' when used in this report without any qualifier simply means sex, in other words biological sex, which can also be termed natal sex or sex at birth. These terms have the same ordinary meaning, and this report uses them interchangeably. 'Legal sex' on the other hand is a shorthand for a categorisation which includes holders of a Gender Recognition Certificate (GRC) as the opposite sex. This can also be phrased as 'sex subject to a GRC'.

We have documented arguments against treating sex as a binary variable. These include arguments which instrumentalise the lives of people with Differences of Sex development (DSD) and propagate myths such as the claim that these rare conditions are 'as common as red hair' or that people with DSD do not have a sex. This has resulted in inappropriate and intrusive questions being asked about DSD.

Censuses and surveys internationally are increasingly seeking to capture information about the trans and/ or gender-diverse population. They have approached these questions in different ways and with varying results. It is critical to establish a clear target for any question on gender identity. Stakeholders have expressed a need for clear guidance on data collection in this area. Questions which conflate sex and gender identity have become common, but do not effectively identify the gender diverse population.

In order to test hypotheses regarding the importance of sex and gender identity in any context, distinct variables are required in order to avoid the multicollinearity which arises when a sex/gender hybrid variable is used in the same analysis as a sex variable. Multicollinearity refers to a high degree of correlation between explanatory variables in a statistical analysis. Sex/gender hybrid questions are not useful in disentangling the effects of sex and gender identity because in practice these questions elicit information on sex from most respondents.

In some cases, the conflation of sex and gender identity has been embedded in shared IT systems, posing a barrier to organisations wishing to collect data on sex. For example, a data management system used by rape crisis centres in Scotland uses the following 'gender' fields for both victims and alleged perpetrators: Male/ Female/Intersex/Gender queer/Other.

In the absence of a single source of authoritative guidance on data collection on sex and gender, various organisations have developed their own guidance. We reviewed guidance produced by national governments, non-departmental public bodies including regulators, prominent organisations engaged in market research, and a small number of charities working on issues relating to sex and gender. We found that guidance produced by many of these organisations lacked conceptual clarity, often eliding data on sex with data on self-declared gender identity.

There are well-established principles of question design which should be applied to any topic (we outline these in chapter 2). These principles have sometimes been overlooked in the field of sex and gender due to the politicisation of these questions. It is important to remember that the purpose of survey data collection is to gather data about populations rather than to provide an opportunity for each individual to express the full complexity and richness of their identity.

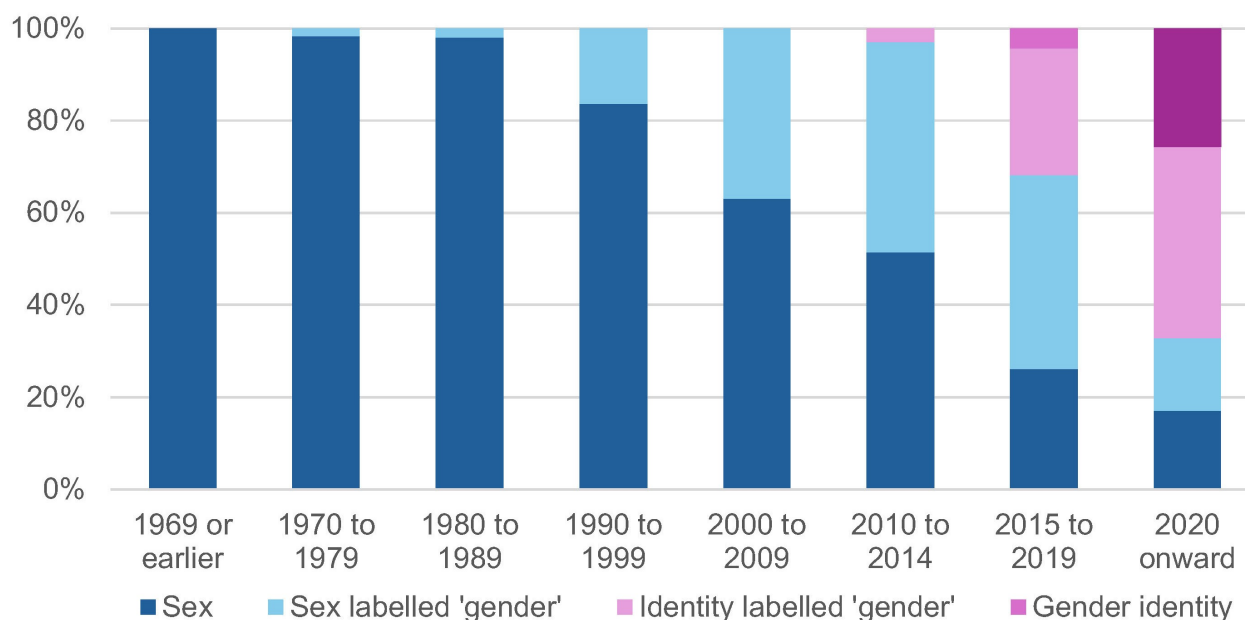
We recognise the high standards which the Office for National Statistics (ONS) typically upholds. However, we have seen evidence of a partisan climate on certain issues, including gender, within the organisation (see chapter 5). Political impartiality (which must be understood broadly, not simply in terms of party politics) is a principle of the Nolan Principles of Public Life. In addition, the UK Statistics Authority Code of Practice states that ‘People in organisations that release statistics should be truthful, impartial and independent, and meet consistent standards of behaviour that reflect the wider public good’.

Review of data collection practices

As part of the review, we examined over 800 administrative and research datasets, surveys, and data collection policies. Our analysis revealed that:

- The meaning of sex is no longer stable in administrative or major survey data. This instability is evident across key policy areas including health and justice. This has led to a widespread loss of data on sex.
- In some cases, the loss of data on sex poses risks to individuals. This is particularly apparent within health and social care. These risks are especially high in the case of minors.
- How sex is defined is rarely made apparent in published outputs, including accredited official statistics. Some publications provide this information in accompanying documentation.
- Some publications present binary data on sex but are underpinned by survey instruments that included additional response options, such as ‘other’.

Figure 1. Sex and gender questions in surveys held by the UK Data Service, 1946 to 2023



View the [data for figure 1 in an accessible table format](#).

- These practices do not meet the standards set out by the Office for Statistics Regulation for collecting and reporting data about sex and gender identity in official statistics. The Standards state that data about sex and gender identity ‘should be explained and defined for the purpose of a particular set of statistics, and terms, including gender, should not be used interchangeably or as a substitute for each other’. [\[footnote 1\]](#)
- The loss of data on sex is a relatively recent phenomenon. Our analysis indicates that in the context of data collection, the term ‘gender’ gained traction as a synonym for sex in the 1990s. More fundamental changes have taken place within the last decade, which has seen both the reframing of ‘gender’ as a synonym for gender identity, and the replacement of sex questions with gender-identity questions. Figure 1 shows these trends, based on a review of over 500 surveys and questionnaires lodged with the UK Data Service. [\[footnote 2\]](#)

Call for evidence

Members of the public submitted a wide range of examples showing a loss of data collection on sex across public and private organisations. These are shown in appendix 3. The submissions further demonstrate the loss of data on sex and its replacement with gender identity data across a diverse range of policy areas, including health and social care; crime and justice; and education, and across different forms of data collection, including patient data, public consultations, equalities monitoring, and visitor surveys.

Taking forward the recommendations of this review

The generation of high-quality data for research and policymaking is a collective cross-government task, covering departments with responsibility for particular areas of government and those in central co-ordinating roles, including Cabinet Office. It also covers specialist statistical bodies, such as the Office for National Statistics and the Office of Statistics Regulation. We hope that this review will provide clarity regarding data on sex and gender identity for those engaged in data collection within government and beyond.

Acknowledgements

The review team are grateful to the stakeholders who agreed to be interviewed and to all those who submitted evidence^{[\[footnote 3\]](#)}. We acknowledge support from the Economic and Social Research Council (ESRC), grant reference ES/X004627/1.

Recommendations

Key recommendations

This section highlights 10 key recommendations before providing the full set of detailed recommendations below.

1. In line with the UK Statistics Authority (UKSA) Inclusive Data Taskforce recommendations, ^{[\[footnote 4\]](#)} data on 'sex, age and ethnic group should be routinely collected and reported in all administrative data and in-service process data, including statistics collected within health and care settings and by police, courts and prisons'.
2. Data on sex should be collected by default in all research and data collection commissioned by government and quasi-governmental organisations. By default, both sexes should be included in all research, including clinical trials, and sex should be considered as a factor in analysis and reporting. As a general rule (with some obvious exceptions), a 50/50 sex ratio is desirable in studies.

3. The default target of any sex question should be sex (in other words, biological sex, natal sex, sex at birth). Questions which combine sex with gender identity, including gender identity as recognised by a Gender Recognition Certificate (GRC) have a mixed target. Sex as a biological category is constant across time and across jurisdictions, whereas the concept of 'legal sex' subject to a GRC may be subject to change in the future and varies across jurisdictions. Using natal sex future-proofs data collection against any such change, ensuring consistency.

4. The form of the question should follow the UK Censuses (England and Wales, Northern Ireland, Scotland) question and response categories.

What is your sex?

Response categories: Female, Male

5. As sex and gender identity are distinct concepts, questions which combine sex and gender identity in one question should not be asked. We have observed a trend for questions which attempt to combine sex and gender diverse identities in one question. Such hybrid questions aim to solicit information on sex from the majority of respondents but on gender identity from some respondents. As such, the target of the question is muddled. Questions that mix sex and gender risk organisations being in breach of the PSED, as they do not identify either the protected characteristic of sex or the protected characteristic of gender reassignment.

6. The word 'gender' should be avoided in question wording, as it has multiple distinct meanings, including:

- a synonym for sex
- social structures and stereotypes associated with sex
- gender identity

If a question targeting gender identity is worded as a question on gender, this is likely to mislead many respondents. Questions on sex have also often been labelled as 'gender'. Change in the use of the term 'gender' means that it is important that questions on sex are labelled explicitly as such.

7. The NHS should cease the practice of issuing new NHS numbers and changed 'gender' markers to individuals, as this means that data on sex is lost, thereby putting individuals at risk regarding clinical care, screening, and safeguarding, as well as making vital research following up individuals who have been through a gender transition across the life course impossible. In the case of children, this practice poses a particularly serious safeguarding risk, and should be suspended as a matter of urgency.

8. Questions on sex and/or gender identity should not contain an additional category for people with DSD conditions, sometimes also known as 'intersex'. People with DSD have a sex, they are not a third sex or sexless category, and to imply that they are is likely to cause offence. DSD is an umbrella term without a single agreed definition, and the question of which conditions are included is contested. Under conventional definitions, people with DSD are estimated to make up 0.018% of births, i.e. fewer than 2 in 10,000. Asking for DSD status is highly intrusive, poses a risk of identifiability, and is unwarranted given the lack of analytical use for data on such a small group. Asking for this information would need to be via a distinct question, not part of a question on sex or gender identity and is likely to be justified only in the context of specialist medical studies.

9. Data providers often default to using ONS Census questions. However, the ONS 2021 Census question 'Is the gender you identity with the same as your sex registered at birth' has been shown to be flawed [\[footnote 5\]](#). The Office for Statistics Regulation (OSR) has stated that the statistics produced by this Census variable do not comply with important quality aspects of the Code of Practice for Statistics and has de-accredited these as official statistics [\[footnote 6\]](#). This question (and variants of it) should not be used.

10. As organisations increasingly seek to collect data on gender identity, the problems identified with the ONS 2021 Census question have left a user need for a simple question which can be used in data collection with the general population. Organisations wishing to collect data on gender identity will need to be clear on the target of their question.

We have identified 3 distinct possible targets for such a question:

- The protected characteristic of gender reassignment
- Trans identification
- Identification as trans and/or gender diverse

Full recommendations

The majority of these recommendations have broad relevance across government. A number of recommendations are directed particularly towards specific bodies. These are:

- recommendation 12 (NHS);
- recommendation 13 (Home Secretary and Police Forces);
- recommendation 23 (Scottish Government and Scotland's Chief Statistician);

- recommendation 24 (EHRC);
- recommendation 51, 53, 54 and 55 (Office for National Statistics);
- recommendation 55, 56 and 59 (UK Statistics Authority).

Sex

1. In line with the UK Statistics Authority (UKSA) Inclusive Data Taskforce recommendations, [\[footnote 7\]](#) data on 'sex, age and ethnic group should be routinely collected and reported in all administrative data and in-service process data, including statistics collected within health and care settings and by police, courts and prisons'.

2. Data on sex should be collected by default in all research and data collection commissioned by government and quasi-governmental organisations. By default, both sexes should be included in all research, including clinical trials, and sex should be considered as a factor in analysis and reporting. As a general rule (with some obvious exceptions), a 50/50 sex ratio is desirable in studies.

3. The default target of any sex question should be sex (in other words, biological sex, natal sex, sex at birth). Questions which combine sex with gender identity, including gender identity as recognised by a Gender Recognition Certificate (GRC) have a mixed target. Sex as a biological category is constant across time and across jurisdictions, whereas the concept of 'legal sex' subject to a GRC may be subject to change in the future and varies across jurisdictions. Using natal sex future-proofs data collection against any such change, ensuring consistency.

4. The form of the question should follow the UK Censuses (England and Wales, Northern Ireland, Scotland) question and response categories.

What is your sex?

Response categories: Female, Male

5. For some purposes, it may be appropriate to make providing information on sex optional for respondents. This can be done by allowing non-response, for example by allowing respondents to move to the next item in an online survey without responding to the sex question, by stating that responding to any question is optional, or more explicitly, via a 'prefer not to say' option. Different approaches may have different implications for non-response in different contexts, and data owners will need to consider these.

6. Omitting or discouraging the option of non- response may be appropriate where failing to collect data on sex would be unsafe, either for the respondent or for others. Examples where a non-response option may not be appropriate would include medical information and data used for safeguarding purposes.

7. Guidance for the sex question should provide clarity on the target of the question as follows.

‘This question is about your sex at birth’.

8. We recommend against using the phrase ‘sex assigned at birth’. This phrasing is inaccurate and misleading, as sex is determined at conception and typically observed in utero or at birth. An individual’s sex is not determined by their birth certificate, it is merely recorded on their birth certificate. In very rare cases an infant’s sex may be inaccurately recorded at birth, but this does not imply that sex is merely an assigned label rather than an inborn characteristic.

9. The concept of ‘legal sex’ is contested and has been subject to change over time and differences between jurisdictions. Therefore, advice on capturing this variable could change. In addition, the concept of ‘legal sex’ has been seen as ambiguous because, in the UK, state- issued documents, such as passports and birth certificates, can record a different sex for the same individual. If data on a person’s sex modified by GRC status, rather than simply their sex, is identified as being required for a specific purpose, we recommend using the England and Wales Census question on sex as above, with the guidance that was used in the Census: ‘If you are considering how to answer, use the sex recorded on your birth certificate or Gender Recognition Certificate.’

10. Whenever sex is recorded, it should be made clear what is intended: whether this refers simply to sex, or to ‘legal sex’ modified for some individuals by a Gender Recognition Certificate (GRC). If a document or record is intended to refer to the latter, this should be subject to change only on provision of a GRC.

11. Organisations that allow sex markers to be changed on official documents should keep records of the number of documents changed annually with basic demographic information attached such as age and sex.

12. The NHS should cease the practice of issuing new NHS numbers and changed ‘gender’ markers to individuals, as this means that data on sex is lost, thereby putting individuals at risk regarding clinical care, screening, and safeguarding, as well as making vital research following up individuals who have been through a gender transition across the life course

impossible. In the case of children, this practice poses a particularly serious safeguarding risk, and should be suspended as a matter of urgency.

13. The Home Secretary should issue a mandatory Annual Data Requirement (ADR) requiring the 43 territorial police forces of England and Wales and the British Transport Police (BTP) to record data on sex in all relevant administrative systems. Relatedly, police forces should cease the practice of allowing changes to be made to individual sex markers on the Police National Computer (PNC).

14. In some cases, changing data on sex held within administrative systems may have been motivated by a desire to ensure that service users are addressed as they wish to be addressed. Service users should of course be treated with respect and addressed by their preferred name and title. It should be possible to store information on forms of address as distinct from information on sex and ensure that relevant people have access to this as required.

15. We are advised that, from a legal perspective, data on sex is close enough to data on sex subject to modification by a GRC to fulfil the public sector equality duty (PSED), even if a GRC is held to affect a person's status under the Equality Act. Given the desirability of a single meaningful and constant target for any question on sex both within and between organisations, a question on sex (i.e. natal, biological sex) is preferable for all purposes. Particularly if data on sex is multi-purpose, for example if it is designed to be used for research and/or operational reasons as well as compliance with the PSED, data on sex, (for the avoidance of doubt, meaning natal/biological sex), only should be collected. Collecting both sex and 'legal sex' would be unduly onerous and would risk identifying individuals with GRCs.

16. As noted by the Office for Statistics Regulation's guidance on 'Collecting and reporting data about sex and gender identity in official statistics', the conflation of terms relating to sex and gender leads to a lack of clarity for both respondents and users of data:

“ Through our work, we have seen instances where there is a lack of consistency and clarity around the term 'gender', both in data collection, and in statistical reporting. In some cases, it is not clear whether producers are using the term gender as a substitution for sex or gender identity. [\[footnote 8\]](#)”

Sex and gender identity are distinct concepts and, in line with the Office for Statistics Regulation guidance, these concepts should not be conflated or combined.

17. As sex and gender identity are distinct concepts, questions which combine sex and gender identity in one question should not be asked. We have observed a trend for questions which attempt to combine sex and

gender diverse identities in one question. Such hybrid questions aim to solicit information on sex from the majority of respondents but on gender identity from some respondents. As such, the target of the question is muddled. Questions that mix sex and gender risk organisations being in breach of the PSED, as they do not identify either the protected characteristic of sex or the protected characteristic of gender reassignment.

18. The word 'gender' should be avoided in question wording, as it has multiple distinct meanings, including:

- a synonym for sex
- social structures and stereotypes associated with sex
- gender identity

If a question targeting gender identity is worded as a question on gender, this is likely to mislead many respondents. Questions on sex have also often been labelled as 'gender'. Change in the use of the term 'gender' means that it is important that questions on sex are labelled explicitly as such.

19. Questions on sex and/or gender identity should not contain an additional category for people with DSD conditions, sometimes also known as 'intersex'. People with DSD have a sex, they are not a third sex or sexless category, and to imply that they are is likely to cause offence. DSD is an umbrella term without a single agreed definition, and the question of which conditions are included is contested. Under conventional definitions, people with DSD are estimated to make up 0.018% of births, i.e. fewer than 2 in 10,000. Asking for DSD status is highly intrusive, poses a risk of identifiability, and is unwarranted given the lack of analytical use for data on such a small group. Asking for this information would need to be via a distinct question, not part of a question on sex or gender identity, and is likely to be justified only in the context of specialist medical studies.

20. In some face-to-face contexts, sex is recorded based on observation rather than by asking a question. Asking for a person's sex in the context of a face-to-face interaction can be perceived as rude. Observed sex is used in operational contexts where asking for an individual's sex may reduce rapport or exacerbate a potentially fractious situation, for example in the context of policing. Similarly, in face-to-face surveys, sex is sometimes recorded based on the interviewer's observation. The potential dissonance and break of rapport generated by asking a person's sex in the context of a face-to-face interaction may be particularly undesirable in surveys which contain sensitive or potentially distressing questions.

21. Data owners should be reassured that it is lawful to collect observational data on sex in both operational and research settings. However, the record must state that this is based on observation only. This is in line with the general principle that the way in which a variable has been captured should

be recorded explicitly in all datasets. Further detail on this point is available in the legal appendix.

22. We have noted some apparent confusion between the concepts of self-reported sex and self-identified gender identity. These are distinct concepts and should not be confused in data collection or guidance. Self-report simply means that the information is reported by the respondent.

23. The Office for Statistics Regulation has written to Scotland's Chief Statistician regarding the Scottish Government's 2021 guidance for public bodies on the data collection and publication of sex, gender identity and trans status, suggesting that this guidance would benefit from clarification taking on board developments since the guidance was published. [\[footnote 9\]](#) Further to this, the Scottish Government guidance should be reviewed to take account of the recommendations of this review, and to consider our legal advice.

24. The Equality and Human Rights Commission (EHRC) should review the material available on its website and either archive or clearly flag documents and guidance that are not consistent with its current view that sex in the Equality Act 2020 refers to 'legal sex' meaning sex subject to modification by a GRC.

Gender identity

25. Data providers often default to using ONS Census questions. However, the ONS 2021 Census question 'Is the gender you identity with the same as your sex registered at birth' has been shown to be flawed. [\[footnote 10\]](#) The Office for Statistics Regulation (OSR) has stated that the statistics produced by this Census variable do not comply with important quality aspects of the Code of Practice for Statistics and has de-accredited these as official statistics. This question (and variants of it) should not be used.

26. Questions on gender identity should recognise that the concept of gender identity as such will be unfamiliar, unclear or irrelevant to some respondents, and that many respondents may not perceive themselves as having a gender identity. Questions should not assume that respondents will agree that they have a gender identity.

27. As organisations increasingly seek to collect data on gender identity, the problems identified with the ONS 2021 Census question have left a user need for a simple question which can be used in data collection with the general population. Organisations wishing to collect data on gender identity will need to be clear on the target of their question. We have identified 3 distinct possible targets for such a question:

- The protected characteristic of gender reassignment
- Trans identification
- Identification as trans and/or gender diverse

28. For organisations wishing to capture the protected characteristic of gender reassignment for the purposes of equalities monitoring, a question on trans status lacks sufficient specificity, and therefore will not assist in compliance with the PSED. To capture the protected characteristic of gender reassignment, we recommend asking a direct question addressed to this target, such as:

Do you have the protected characteristic of gender reassignment?
Response options: Yes/No/ Don't know/Prefer not to say.

29. We acknowledge that gender reassignment will be an unfamiliar concept for many respondents. A guidance note should be included prominently alongside the above question as follows:

'A person has the protected characteristic of gender reassignment if the person is proposing to undergo, is undergoing or has undergone a process (or part of a process) for the purpose of reassigning the person's sex by changing physiological or other attributes of sex'.

30. For organisations wishing to capture trans identification, we recommend asking a question clearly directed towards this target. Questions which meet this specification would follow the following format:

'Are you' [or 'Do you identify as' or 'Do you consider yourself to be'] 'transgender' [or 'trans']?' With response options such as: Yes, Trans woman/Yes, Trans man/Yes, Non-binary/Yes, Other, please specify if you wish [open text]/ No/Don't know/ Prefer not to say.

31. The small differences in the options provided above reflect the fact that different formulations have been used and recommended across different data collection exercises, with apparent success, and we are not in a position to know whether these differences may affect response in any way. In order to settle on a single agreed formulation, it would be desirable to conduct question testing.

32. Further quantitative research should be undertaken to assess what the general public takes as the meaning of key words which may be used in data collection in this area, including 'transgender man', 'trans man', 'transgender', 'trans', 'transsexual', and 'gender reassignment', along the

lines of polling which has already been carried out on the words ‘trans woman’ and ‘transgender woman’.[\[footnote 11\]](#)

33. If guidance is required, we recommend the following: ‘Some people describe themselves as transgender when they do not identify with their sex at birth’.

34. If there is a need and sufficient space for more detailed guidance, we suggest providing the Stonewall definition:

“ Trans people may describe themselves using one or more of a wide variety of terms, including (but not limited to) transgender, transsexual, gender-queer (GQ), gender-fluid, non-binary, gender-variant, crossdresser, genderless, agender, nongender, third gender, bi-gender, trans man, trans woman, trans masculine, trans feminine and neutrois.”

35. It is likely that, for some research purposes, a broader question will be desirable. A question on wider gender-diverse identities may be required, for example to identify respondents expressing a non-binary identity who may or may not identify as trans. If the target of the question is to identify those with gender-diverse identities including, but not limited to, those who identify as trans, we recommend asking a question which makes this clear. For example:

‘Are you [or ‘Do you identify as’ or ‘Do you consider yourself to be’] transgender [or trans], non-binary or gender diverse?’, with response options: Yes, Trans woman/Yes, Trans man/ Yes, Non-binary/Yes, Other, please specify if you wish [open text]/No/Don’t know/Prefer not to say. As this question is novel, it will require full question testing.

36. Organisations considering collecting data on gender reassignment or trans status will need to consider both the fact that this is sensitive personal information and that it identifies a small group. Whether it is appropriate to collect this data will depend on a number of factors, including the:

- size of the dataset
- prevalence of trans identities in the population of interest
- proposed use of the data

Organisations should only collect data which they intend to process. The PSED does not imply a duty to collect data which is unlikely to be useful.

Reporting on data

37. All databases should provide a clear record of how data on sex is defined and collected. All reports using data analysis on sex should provide a clear account of how data on sex is defined and collected, and whether and how this varies between different databases or systems. Any changes to how data on sex is defined over time should be made clear. Data producers should provide a clear audit trail on how data on sex is collected. For example, where appropriate, reporting should include copies of questionnaires and instructions to interviewers. If survey data is collected by a third party using pre-recruited panellists, the source definition should be stated.

38. Analysts must be able to use clear and familiar language in reporting findings on sex. Terms such as women, men, boys and girls are synonymous with (respectively) adult human females and males and children of each sex. Similar considerations apply to terms such as mothers and fathers, sons and daughters. While all language concerning sex and gender has become contested to some degree, those reporting on sex-disaggregated data should not be dissuaded from using familiar sexed terms. Sensitivities which may apply when referring to specific individuals should not apply at the aggregate level. Any guidelines on language use in reporting on data and research should foreground clarity and ease of communication.

Clear language in legislation, guidance and discourse

39. Previously a polite synonym for sex, 'gender' now has multiple distinct meanings. Legislation referring to 'gender' is now open to misinterpretation, even in cases where it may appear clear that, at the time the legislation was enacted, gender meant sex. It is desirable that legislation should refer clearly to sex and/ or to gender reassignment as appropriate rather than using the term 'gender'. This has direct implications where data collection is mandated via legislation. Where organisations feel constrained by the use of 'gender' in relevant governing statutes when it comes to collecting data, the government should consider amending that legislation so that it refers to sex.

40. When reporting on or discussing issues relating to sex, it would be desirable to see a shift to using the term 'sex' instead of gender, given the ambiguity of the term 'gender'. This should be reflected in government language and guidance. For example, guidance for employers on 'gender pay gap' reporting should refer to 'sex pay gap' reporting.

Publishing data on individuals

41. Some organisations may wish to publish data on individual sex. For example, in athletics, race organisers typically publish results according to sex/age categories which are used for competition purposes and to compute sex/ age adjusted performance gradings. In other contexts, the sex of a practitioner may be published so that members of the public can be informed of this information where relevant.

42. Published information on individuals must be accurate and accurately recorded and conveyed. Data on sex should not be reported as being data on gender identity or vice versa, and the 2 concepts must not be combined.

43. Individuals should typically be given the option of not having their sex published.

Opinion research

44. When asking for respondents' opinions, it is vital that respondents can understand what is being asked, so that they can provide accurate information on their views. In cases where sex and gender identity are relevant to the target of the question, clarity on these concepts is important. It should be borne in mind that language which is familiar to some groups of respondents may be unfamiliar or misunderstood by others. For example, it has been shown that some respondents take the term 'trans woman' to mean the opposite of what is actually intended, i.e. some respondents assume that a 'trans woman' is a female who identifies as a man, rather than a male who identifies as a woman. Clear language should be used to identify sex where relevant. For example, 'Should males who identify as women be allowed to compete in female sports categories?' is clearer than 'Should trans women be allowed to compete in women's sports categories?'.

Developing and reviewing questions

45. Some stakeholders perceive a pressure to keep data categories under regular review. We would emphasise that continuity is desirable, and change should only be implemented where there is a good reason. Some categories are subject to real social change, for example, the ethnic

composition of the UK has changed over time. Other categories, such as sex and age, are unaffected by societal change.

46. Consistency throughout the data landscape is desirable, for purposes of comparability and data linkage. Data providers should default to standard ONS categories unless there is a compelling reason to do otherwise. By the same token, the ONS should recognise their role in providing 'gold standard' questions which therefore need to be fit for purpose for the widest possible range of respondents and uses.

47. Where data owners are engaged in question review and development, they should be mindful of the distinction between data users and respondents, and not conflate the needs of these 2 groups, both of which are important.

48. It is often appropriate for data owners to consult with special interest or community groups, but this should be done in a balanced way, mindful of both conflicts between groups and the possibility that groups claiming to represent a particular constituency may not reflect the full range of opinion, or indeed the majority opinion, within that constituency. Community and campaigning groups should not be assumed to have expertise in question design. Data owners should consider carefully which issues these groups are qualified to advise on, and how much weight to give to their input. This also applies to research including focus group work where the participants have been recruited via an organisation with a particular viewpoint.

49. Questions intended for general use should be tested on the wider relevant population, not just on minority groups, including in the case of questions which are primarily intended to identify minority groups.

50. Organisations should consider carefully whether it is appropriate to include internal staff groups in consultations on data collection. This is unlikely to be appropriate when the data collection is external rather than internal and is not without risk in the case of internal data collection. The views on data collection of internal staff networks should not be given undue weight.

51. When considering new questions or substantial changes to questions, the ONS should seek a wide range of views and expertise. This should include engagement with the community of quantitative data analysts from across the relevant academic disciplines.

52. Public bodies should strive for transparency, openness and accountability regarding the development of changes in data collection policy and practice. The leadership and composition of groups working on such questions should be named in the public domain. Organisations should maintain a clear audit trail of who has been consulted on their data collection practices, including organisations and individuals. Anonymity should only be granted to individuals under exceptional circumstances.

Transparency, openness and accountability are particularly important in the case of national statistical bodies which are seen as developing 'gold standard' data.

53. The ONS must publish detailed research reports on all of its question testing and development research. This is publicly funded work and failure to share it can only be detrimental, preventing wider scrutiny of and learning from ONS research. Publishing only basic summaries of research is not sufficient.

54. The ONS should review its file management system and ensure that documentation which has any potential bearing on its data collection practices is kept in good order and not destroyed without appropriate senior level approval.

Organisational culture

55. Government departments and bodies should strive to promote a culture of critical thinking and robust and civil disagreement. Some stakeholders expressed the view that it was difficult to have open and comfortable discussions on data collection on sex and gender. Organisations, including government departments, should strive to tackle any barriers to such discussions taking place, and actively promote a collegial and professional approach. Civil servants have a duty to maintain political impartiality. Objectivity is one of the Nolan Principles of Public Life, 'Holders of public office must act and take decisions impartially, fairly and on merit, using the best evidence and without discrimination or bias'. The UK Statistics Authority Code of Practice principle of Honesty and Integrity states 'People in organisations that release statistics should be truthful, impartial and independent, and meet consistent standards of behaviour that reflect the wider public good'. Government departments and bodies, including the ONS, should review their internal cultures and practices with the goal of upholding these principles.

56. The UK Statistics Authority should consider undertaking a review of activism and impartiality within the civil service, in relation to the production of official statistics.

57. Ministers should consider the vulnerability of government and public bodies to internal activism that seeks to influence outward-facing policy, including through staff networks, and whether stronger safeguards are needed.

Communicating the purposes of research

58. Respondents may understandably question why they are being asked for certain information. It is vital that everyone participating in research or providing data for administrative purposes is treated with respect and is informed about the reasons for data collection and reassured about the way their data will be processed and used. Some respondents may be reluctant to provide data on sex while others are reluctant to provide data on gender identity. In both cases, the purposes of the data collection, including potential benefits to the individual and society, should be explained.

59. We cannot assume that there is universal public understanding of the diverse purposes and benefits of data collection. UKSA's existing aim of advocating for improved statistical literacy should include seeking opportunities to spread knowledge of the wide benefits of accurate data collection, both in the area of sex and gender identity and beyond.

Data for figure 1: Sex and gender questions in surveys held by the UK Data Service, 1946 to 2023

Period	Sex	Sex labelled 'gender'	Identity labelled 'gender'	Gender identity	Total
1969 or earlier	58	0	0	0	58
1970 to 1979	59	1	0	0	60
1980 to 1989	49	1	0	0	50
1990 to 1999	46	9	0	0	55
2000 to 2009	46	27	0	0	73
2010 to 2014	35	31	2	0	68
2015 to 2019	18	29	19	3	69

Period	Sex	Sex labelled 'gender'	Identity labelled 'gender'	Gender identity	Total
2020 onward	12	11	29	18	70

See [Figure 1: Sex and gender questions in surveys held by the UK Data Service, 1946 to 2023](#).

1. Office for Statistics Regulation (2024), [Collecting and reporting data about sex and gender identity in official statistics: A guide for official statistics producers](https://osr.statisticsauthority.gov.uk/publication/collecting-and-reporting-data-about-sex-and-gender-identity-in-official-statistics-a-guide-for-official-statistics-producers/) ([https://osr.statisticsauthori-%20ty.gov.uk/publication/collecting-and-reporting-data-about-sex-and-gender-identity-in-official-statistics-a-guide-for-official-statistics-producers/](https://osr.statisticsauthority.gov.uk/publication/collecting-and-reporting-data-about-sex-and-gender-identity-in-official-statistics-a-guide-for-official-statistics-producers/))
2. The UK Data Service is a digital repository for quantitative and qualitative social science and humanities research data
3. Those organisations whose general impact on data collection practices is considered in the review were also offered the opportunity to comment on the factual accuracy of relevant parts of the report in draft. These were: the Office of National Statistics, the Scottish Government, the Welsh Government, the Home Office, the Equality and Human Rights Commission, NHS England, Arts Council England, the Financial Conduct Authority, Advance HE, Stonewall, IPSOS and the Market Research Society. Responses were received from the majority of those contacted. For the very large further number of organisations for which examples are provided of data collection practices, it was not practical to make direct contact in each case, but references are provided to relevant sources throughout.
4. [UK Statistics Authority \(2022\), Inclusive Data Taskforce recommendations](https://uksa.statisticsauthority.gov.uk/publication/inclusive-data-taskforce-recommendations-report-leaving-%20no-one-behind-how-can-we-be-more-inclusive-in-our-data/) (<https://uksa.statisticsauthority.gov.uk/publication/inclusive-data-taskforce-recommendations-report-leaving-%20no-one-behind-how-can-we-be-more-inclusive-in-our-data/>)
5. Biggs, M., 2024. [Gender identity in the 2021 census of England and Wales: How a flawed question created spurious data](https://doi.org/10.1177/00380385241240441). *Sociology* (<https://doi.org/10.1177/00380385241240441>)
6. [Office for Statistics Regulation \(2024\) Review of statistics on gender identity based on data collected as part of the 2021 England and Wales Census: Final report](https://osr.statisticsauthority.gov.uk/publication/review-of-statistics-on-gender-identity-based-on-data-collected-as-part-of-the-2021-england-and-wales-census-final-report/) (<https://osr.statisticsauthority.gov.uk/publication/review-of-statistics-on-gender-identity-based-on-data-collected-as-part-of-the-2021-england-and-wales-census-final-report/>)
7. UK Statistics Authority (2022), [Inclusive Data Taskforce recommendationS](https://uksa.statisticsauthority.gov.uk/publication/inclusive-recommendationS) ([https://uksa.statisticsauthority.gov.uk/publication/inclusive-](https://uksa.statisticsauthority.gov.uk/publication/inclusive-recommendationS)

[data-taskforce-recommendations-report-leaving-no-one-behind-how-can-we-be-more-inclusive-in-our-data/](#)

8. [Office for Statistics Regulation (2024), [Collecting and reporting data about sex and gender identity in official statistics](#) (<https://osr.statisticsauthority.gov.uk/publication/collecting-and-reporting-data-about-sex-and-gender-identity-in-official-statistics-a-guide-for-official-statistics-producers/pages/2/>).
9. Office for Statistics Regulation (2024), [Ed Humpherson to Alastair McAlpine: Regulatory guidance on collecting and reporting data about sex and gender identity in official statistics](#) (<https://osr.statisticsauthority.gov.uk/correspondence/ed-humpherson-to-alastair-mcalpine-regulatory-guidance-on-collecting-and-reporting-data-about-sex-and-gender-identity-in-official-statistics/>) (29 February 2024).
10. Biggs, M., 2024. [Gender identity in the 2021 census of England and Wales: How a flawed question created spurious data](#). *Sociology* (<https://doi.org/10.1177/00380385241240441>).
11. For similar polling in this area, see: MurrayBlackburnMackenzie (2023), [Clarity matters: how placating lobbyists obscures public understanding of sex and gender](#). 7 August 2023 (<https://%20murrayblackburnmackenzie.org/2023/08/07/clarity-matters-how-placating-lobbyists-obscures-public-understanding-of-sex-and-gender/>).



All content is available under the [Open Government Licence v3.0](#), except where otherwise stated



© Crown copyright