

Review of statistics on gender identity based on data collected as part of the 2021 England and Wales Census – interim report

(produced by the Office for National Statistics)

October 2023

Office for Statistics Regulation

We provide independent regulation of all official statistics produced in the UK. Statistics are an essential public asset. We aim to enhance public confidence in the trustworthiness, quality and value of statistics produced by government.

We do this by setting the standards they must meet in the [Code of Practice for Statistics](#). We ensure that producers of government statistics uphold these standards by conducting assessments against the Code. Those which meet the standards are given National Statistics status, indicating that they meet the highest standards of trustworthiness, quality and value. We also report publicly on system-wide issues and on the way statistics are being used, celebrating when the standards are upheld and challenging publicly when they are not.

Contents

Executive Summary	4
Introduction.....	8
Background to the statistics.....	8
Our review	8
What we found so far	10
The question requires more tailored testing.....	10
The question design involved UK wide collaboration	11
The quality assurance was proportionate based on what was known.....	13
ONS's research must communicate how the data can and can't be used	14
Expectations for further work	17
Conclusion.....	19

Executive Summary

Introduction

- ES.1 The gender identity question was asked, on a voluntary basis, for the first time in the 2021 England and Wales Census, conducted by the Office for National Statistics (ONS). The question represents the current [Government Statistical Service harmonised standard](#) in development for collecting data on gender identity.
- ES.2 The [first release](#) of statistics on gender identity based on Census data was in January 2023. When publishing the results ONS said: “Gender identity refers to a person’s sense of their own gender, whether male, female or another category such as non-binary. This may or may not be the same as their sex registered at birth.”
- ES.3 Measuring gender identity is undoubtedly challenging. It is inherently personal and therefore can be a sensitive topic. Those whose gender identity differs from their sex represent a small proportion of the population which creates challenges for data collection. The concept may be unfamiliar to many people and there are few robust alternative sources of data to cross-check against.
- ES.4 Following the first release, concerns were raised with ONS and the Office for Statistics Regulation (OSR) about the published estimates of the transgender population. As additional breakdowns were published, these concerns extended to the relationship between gender identity and proficiency in English (or Welsh). Users also raised concerns about the level of methodological information published.
- ES.5 ONS initiated research in response to these concerns in April 2023. We then announced our plans to carry out a short review of these statistics and the findings of ONS’s research, formed around the principles within the [Code of Practice for Statistics](#). The review considers ONS’s approach to responding to questions raised by users about the results of the gender identity data. It does not consider gender identity as a concept or the decision to collect data on gender identity.
- ES.6 On 30 August ONS cancelled the publications and released a [statement](#) advising that the research is ongoing and that it will publish further updates in due course. This delay is likely to have disappointed and impacted some users.
- ES.7 In light of this delay, we have produced an interim report. This interim report sets out our expectations of what is needed from ONS to provide assurance on the gender identity data and harmonised question in line with the Code. We have based these on what we know so far about how ONS tested, quality assured, communicated and engaged with users, both in advance of publishing the results and where appropriate from the research to date.

What we found so far

- ES.8 The review considers ONS’s approach to responding to questions raised by users about the results of the gender identity data. This interim report looks at four aspects: how ONS tested the question; how this compares with the approach taken in the Scottish Census; the approach to quality assurance; and how ONS responded once some users raised concerns about it, including addressing uncertainty.

- ES.9 ONS worked with good intentions when seeking to address an important data gap on gender identity. These intentions guided its approach to the question design and testing process, and to the processing and dissemination of the resulting data.
- ES.10 ONS considered language proficiency as part of its wider approach to testing the questionnaire and its own research pointed to the need for more targeted language testing on gender identity. ONS should explore ways to improve participation of those with lower English language proficiency as part of any further testing of the harmonised standard.
- ES.11 There were strong working relationships between ONS and National Records of Scotland (NRS) in sharing updates and learning as they went through testing, processing and quality assuring the gender identity data. We recognise that different contexts and user need shaped the choice of questions used by ONS and NRS. However, on some aspects of the questions, ONS and NRS received conflicting feedback from similar user groups. To inform future development and use of the gender identity question, we consider that further investigation is required to understand these different user perspectives and needs in more detail.
- ES.12 The general approach taken by ONS to quality assure the data, when considering it alongside the other variables, was proportionate particularly given the lack of data on this topic area. The approach involved independent scrutiny and expertise when choosing what areas to probe and analyse. In hindsight, given it was a new question, ONS could have prioritised analysis of the agreement rates with the Census Quality Survey as part of its quality assurance process, before releasing the data. Further clarity is also needed to help users understand how the sex and gender identity questions were asked and relate to each other, given the correlation has been perceived by some as an indication that the question has been misunderstood.
- ES.13 Given that this is a contested area, with little data available for triangulation, it would be reasonable for users to have expected clearer messaging on the impacts on quality and uncertainty. This gained added importance when users raised challenges and offered alternative hypotheses to explain the data. In this context ONS now needs to undertake further analysis and be clear on what can and cannot be concluded from the data.
- ES.14 ONS has been keen to deliver public value by releasing these new data on gender identity in a timely manner. ONS has taken steps to publish updates on its research around these data and we have seen positive engagement on these issues in our interactions with ONS. We recognise the complexity and sensitivity of the research ONS is conducting and agree that this should not be rushed. However, the published communication to date does not sufficiently manage users' expectations of what areas ONS's research into these data is exploring and what it then might show. A more open approach would manage users' expectations around the use of the data and lay the ground for any changes in advice on the use of this data resulting from ONS research.
- ES.15 The message that ONS provides upfront in a box on the gender identity data harmonised standard page is a good example of a clear and honest acknowledgment of uncertainty without the need to quantify it. This approach should be adopted more broadly for ONS's outputs and quality information surrounding the gender identity data to ensure the data are not open to misuse and misinterpretation.

Expectations for Further Work

ES.16 Throughout our review, it has been clear that the ONS team leading the quality assurance and additional exploration of the data is committed to understanding the quality of the data and how this can best be communicated to users.

ES.17 ONS's research on gender identity should seek to understand to what extent, if any, has a misunderstanding of the gender identity question in the 2021 England and Wales Census led to a significant misestimation of the transgender population. When the target group is small, the effect of incorrect responses on the data will be amplified.

ES.18 ONS should consider the testing and operationalisation of the question as well as the production and analysis of the results. If there is evidence that there has been a significant misestimation of the transgender population, to a degree that might affect usage, the research should set out both the implications for the use of the results from the Census and the use of the GSS harmonised question in future data collection. We expect ONS to take on board the following recommendations as it works on developing its research.

ES.19 Recommendations to support further use of the data:

- ONS should publish a statement within four weeks of this interim report that sets out a timeline for its ongoing research, covering what it is analysing and what it plans to publish by when. This statement must be open about the potential quality issues it is investigating, even if it cannot quantify the scale of them. This is important both for the use of published estimates, but also the use and further development of the harmonised question.
- ONS should make data available to understand subgroups of the reported transgender population. For example, it should support analysis of differences in responses for those who provided an explicit and unambiguous write-in response indicating that they were transgender, compared with those who provided a possibly ambiguous or tick-box only response.
- ONS should consider any potential misunderstanding of the question or mode effects on the data as part of its analysis into the agreement rates for the gender identity question in the Census and Census Quality Survey. In particular, it should publish analysis of group differences to inform future use of the data and question.
- As part of the [Future of Population and Migration Statistics consultation](#) and programme, the UK harmonisation group should set out its early thinking on gender identity estimates for Great Britain and publish its development plan for these statistics.

ES.20 Recommendations to support further use of the question:

- As part of the ongoing development of the harmonised standard for gender identity, should the question remain the same, ONS should carry out further testing of the question. This should be considered as part of the outcome for the [Future of Population and Migration Statistics consultation](#), in determining future production of these statistics in an administrative data based framework.

- ONS should publish more details on the feedback it received in the Census rehearsal, including any insight about those with lower English language proficiency.
- Once NRS has published its data on gender identity from the 2022 Scotland Census, planned for 2024, ONS and NRS should carry out a lessons learned exercise comparing the performance of the two questions and use this to inform future developments and use of the questions.

ES.21 As this interim report makes clear, there was always good reason to expect a significant degree of uncertainty around estimates of the transgender population derived from the gender identity question in the 2021 Census. A willingness on the part of ONS to provide further context and clarity on the initial estimates would be evidence of the strength of the Census process and not a weakness. ONS providing updates to the interpretation of these estimates should be regarded as a normal part of ongoing statistical production for a new area of data collection. The same would be true of any changes to the initial estimates, as long as they are fully and transparently explained. Such updates should not undermine user confidence in the robustness of the Census results as a whole, in which most questions have been asked many times before.

ES.22 We are satisfied that the overall approach to the 2021 Census complies with the standards of trustworthiness, quality and value in the Code of Practice for Statistics in order to achieve 'accredited official statistics' status. We will consider the accreditation of this subset of the Census dataset when ONS publishes its research findings. We will publish a follow up report with our findings after we have reviewed, and engaged with users on, the full set of research.

Introduction

Background to the statistics

- 1.1 The gender identity question was asked for the first time in the 2021 England and Wales Census, conducted by the Office for National Statistics (ONS). The question represents the current [Government Statistical Service harmonised standard](#) in development for collecting data on gender identity. The question asked, “Is the gender you identify with the same as your sex registered at birth?”, with those who answered ‘no’ being asked to write in their gender identity. The question was voluntary and asked only of respondents aged 16 and over.
- 1.2 Measuring gender identity is undoubtedly challenging. It is inherently personal and therefore can be a sensitive topic. Those whose gender identity differs from their sex represent a small proportion of the population which creates challenges for data collection. The concept may be unfamiliar to many people and there are few robust alternative sources of data to cross-check against.
- 1.3 The [first release](#) of statistics on gender identity based on Census data was in January 2023. This led to some concerns being raised with ONS and the Office for Statistics Regulation (OSR) regarding the correlation of gender identity with other characteristics (in particular English or Welsh language proficiency), as well as the level of methodological information published on how the gender identity question has been processed.
- 1.4 The subsequent datasets showed that those aged 16 and over whose main language is not English (or Welsh if using Welsh questionnaire), made up 10 per cent of the overall population, but 29 per cent of those whose gender identity differed from their sex registered at birth. The data showed that areas with the highest proportion of people reporting themselves as transgender were areas that also had a high proportion of people whose first language is not English (or Welsh). This led some users to conclude that the question on gender identity may have been misunderstood by those with lower English language proficiency. Michael Biggs, Associate Professor in Sociology at the University of Oxford, explored some of these concerns in a [Spectator article](#).
- 1.5 ONS initiated research in response to these concerns in April 2023. We then [announced our plans](#) to carry out a short review of these statistics and the findings of ONS’s research, formed around the principles within the [Code of Practice for Statistics](#). The review considers ONS’s approach to responding to questions raised by users about the results of the gender identity data. It does not consider gender identity as a concept or the decision to collect data on gender identity. We published our emerging findings as part of a letter to ONS on 19 June, this interim report builds on those findings and should be considered our current view.

Our review

- 1.6 Our review was based on desk research of information in the public domain and stakeholder engagement in the form of interviews. These interviews were carried out with teams in ONS involved in the question design, testing and analysis of the gender identity data collected in the 2021 Census. We also spoke to several users of these statistics. Where possible, ONS has also shared updates on its ongoing research to inform our review. ONS’s openness with us has been very useful in

informing our own research, and we are very grateful for the approach it has taken in sharing these updates.

- 1.7 This interim report sets out our expectations of what is needed from ONS to provide assurance on the gender identity data and harmonised question in line with the Code. We have based these on what we know so far on how ONS tested, quality assured, communicated and engaged with users, both in advance of publishing the results and where appropriate from the research to date.
- 1.8 We will publish a final report following publication by ONS of its research reports. These outputs have been postponed twice, having been preannounced for June and then August. We expect ONS to take on board the findings and recommendations set out in this interim report as it works on developing the research reports for publication.

What we found so far

- 2.1 User need for data on gender identity was identified as part of the Office for National Statistics' (ONS) [public consultation on proposed topics for the 2021 Census questionnaire](#) in England and Wales, carried out in 2015. In particular, the ability to capture data on the transgender population was seen as a data gap for service provision that was important to address.
- 2.2 The [2018 White Paper Help Shape our Future: The 2021 Census of Population and Housing in England and Wales](#) highlights the complex issues to consider in designing a question on gender identity, including the need to clarify how the concepts around sex, gender and gender identity are expressed and can be measured.
- 2.3 In developing the gender identity question for the 2021 Census, ONS looked at international best practice and worked closely with other National Statistics Institutes (NSI) to understand what questions were already being used and how they performed. These NSIs faced similar challenges to ONS, and this continues today, there now being multiple approaches used to collect data on gender identity in censuses across the world.

The question requires more tailored testing

- 2.4 Prior to the first publication in January 2023, ONS published a comprehensive article on [Sex and gender identity question development for Census 2021](#) which sets out how it researched, developed and tested the gender identity question. The article provides relevant context on the research that led to the [2018 White Paper recommendation](#) to include a question on gender identity, as well as the research underpinning the question design.
- 2.5 Some users of these statistics told us that the engagement of experts in the question design and testing was not as wide as it could have been. The ONS team involved in the question development told us that, while they endeavoured to balance competing user needs, it was extremely challenging to identify a question that would be acceptable for use and also addressed the findings of the quantitative testing. ONS recognised from the outset that the chosen approach may need to be updated in time, as societal understanding of gender identity evolves and once there is sufficient evidence to evaluate the performance of the question. This rationale aligns to ONS's experience of developing questions to collect data on ethnicity, which are now standardised and used widely.
- 2.6 The transgender and non-binary community represent a relatively small proportion of the population, and therefore the likelihood of this community being picked up in a random sample is low. To address this, ONS carried out some targeted testing of the gender identity question within this community. This included cognitive interviews with transgender and non-binary participants, as well as informal interviews at LGBT+ events. There were no transgender participants in the Welsh language focus groups and cognitive testing. However, ONS noted it received feedback from stakeholders working in Wales who supported the translation based on their experiences.
- 2.7 In our engagement for this review, we heard from a user who felt that those who are transgender and non-binary have been responding to the Census and other surveys historically in the way that best reflects their identity. They also felt that the risk of

miscomprehension is more likely to occur with those not in the trans and non-binary population, potentially including other minority groups. While ONS paid particular attention to capturing those who are transgender correctly, there is a greater risk with false positives or negatives when the target group is small. This is because the effect of incorrect responses on the data will be amplified.

- 2.8 The gender identity question was also tested with respondents whose gender identity corresponds with the sex registered for them at birth through user experience (UX) and panel-based quantitative testing. The 2021 Census questionnaire as a whole was tested through a Census rehearsal in 2019 involving approximately 300,000 households across England and Wales. This included households in areas that have a high proportion of people whose main language is not English. Feedback from the rehearsal showed that some respondents found the gender identity question difficult to answer, but the non-response rate for this question was lower than ONS expected.
- 2.9 As part of an [in-depth review of measuring gender identity](#) carried out by ONS and Statistics Canada in 2019, ONS shared its findings that, in languages where the concept of gender is not well established, it is unlikely that one standard question could be directly translated and advised that question testing should ensure respondents understand what is being asked. The concept of gender is often conflated with the concept of sex and the term 'gender' can be used by some interchangeably with the term 'sex'. The way these terms are used in law also differs across countries. ONS told us that it faced difficulties finding people with lower English language proficiency who were willing to participate in question testing.
- 2.10 Translated copies of the Census questionnaire were [available in nearly 50 languages](#) and a public support team was put in place to support user queries. ONS also carried out extensive community engagement across England and Wales to promote the Census and encourage different communities to complete it. It told us that the number of downloads and queries received were minimal. It is difficult to know the true scale of their use, as individuals may have downloaded a translated copy on behalf of an organisation or group. Moreover, it is possible that some individuals chose to use online translation tools to translate the English or Welsh questionnaires as they went through completing it. ONS told us that it initially had a small project looking at how all questions appeared through Google Translate. However, once it was decided that ONS would not promote the use of Google Translate, this project was stopped.
- 2.11 **Our judgement:** We consider that ONS worked with good intentions when seeking to address an important data gap on gender identity. These intentions guided its approach to the question design and testing process, and to the processing and dissemination of the resulting data.
- 2.12 ONS considered language proficiency as part of its wider approach to testing the questionnaire and its own research pointed to the need for more targeted language testing on gender identity. ONS should explore ways to improve participation of those with lower English language proficiency as part of further testing of the harmonised standard.

The question design involved UK wide collaboration

- 2.13 A UK harmonisation group was set up to ensure developments and learning were shared across the relevant census statistics producers. This was in line with a

[published statement](#) that was updated throughout the course of the censuses. Limited user need was identified for a question on gender identity in the [2015/16 Northern Ireland Statistics and Research Agency \(NISRA\) consultation for the 2021 Census](#). As a result, no gender identity question was included in the 2021 Northern Ireland Census.

- 2.14 ONS and National Records of Scotland (NRS) worked closely together throughout the development and testing of the gender identity question, and some of the testing was jointly funded. NRS told us that the engagement with ONS has been excellent throughout the development, testing and now quality assurance of the gender identity data. With a view to creating coherent statistics, the harmonisation group agreed that it would look to ask the same question (the question ONS ended up with) across the censuses if this was appropriate for user need in each nation.
- 2.15 During the testing, NRS identified specific user need for an explicit transgender question, and one that captured a trans history. This user need was not mirrored to the same extent in [England and Wales](#). Additionally, ONS's testing received some negative feedback from individuals whose gender identity corresponds with the sex registered for them at birth and non-binary respondents on the trans history question.
- 2.16 NRS's testing identified a need for the sex and gender identity questions to be positioned together to provide better understanding by respondents. It also found that a non-binary sex question produced less item non-response than a question set comprised of a binary sex question followed by a gender identity question. By contrast, ONS was guided by the practical consideration of where the question would best fit on paper. It decided that the gender identity and sexual orientation questions should be placed with the questions that only apply to people aged 16 years or over, at the end of the sociocultural questions, rather than immediately after the sex question.
- 2.17 The Equality, Race and Disability Evidence Units in the Welsh Government published their [strategy](#) in 2022. One of the priorities identified was the need to review the extent to which the gender identity question used in the 2021 Census (alongside other options including the NRS approach) meets the needs of Welsh users. The Welsh Government is in the early stages of scoping this work and will be considering how to maintain some level of coherence across the UK. A 2021 report outlining [recommendations for furthering LGBTQ+ equality in Wales](#) recommended that the Welsh Government should take practical steps to provide recognition of non-binary identities where legal gender is not required, including reviewing data management to ensure that open-ended nonbinary options are included. The harmonised standard has therefore not been considered acceptable by some LGBTQ+ stakeholder groups in Wales.
- 2.18 The UK harmonisation group, which includes the Northern Ireland Statistics and Research Agency (NISRA), is working closely to determine how it will collectively produce estimates for Great Britain given the divergence in approach. This work will need to consider the coherence of the data in outputs within nations as well as across Great Britain.
- 2.19 **Our judgement:** There were strong working relationships between ONS and NRS in sharing updates and learning as they went through testing, processing and quality assuring the gender identity data. We recognise that different contexts and user needs shaped the choice of questions used by ONS and NRS. However, on some aspects of the questions, ONS and NRS received conflicting feedback from

similar user groups. To inform future development and use of the gender identity question, we consider that further investigation is required to understand these different user needs in more detail.

The quality assurance was proportionate based on what was known

- 2.20 More topic-based quality assurance was carried out for the 2021 Census than in previous years. Gender identity and sexual orientation (combined) formed one of 13 topics assigned for bespoke quality assurance reports. The team carrying out the quality assurance worked closely with the relevant topic leads in ONS to apply expert insight to the data to test thinking. These quality assurance reports were reviewed by an independent quality assurance panel comprising of experienced researchers to apply additional scrutiny to the data.
- 2.21 The Code of Practice for Statistics states that quality assurance should be proportionate to the data. ONS told us it analysed relationships between the gender identity data and other variables based on areas that were perceived to be of public interest at the time.
- 2.22 The Census Quality Survey (CQS) is a telephone survey carried out following the Census that asks a sample of people the same questions as were asked on the Census. The CQS in 2021 received usable responses from 8,598 households. Agreement rates are calculated by comparing the answers to this survey with those given by the same respondents in the Census. CQS agreement rates can be a helpful tool for understanding the performance of new questions, particularly if they are voluntary. However, the change in survey mode should be considered where it may have an effect on more sensitive questions. ONS only had access to the weighted CQS agreement rates after the first release of gender identity statistics. In hindsight, it could have prioritised analysing the CQS agreement rates before releasing the data given it was a new question. Moreover, a write-in response was not required for the gender identity question in the CQS which could impact the way the gender identity response was then processed and coded.
- 2.23 In June 2023, ONS published information on the [methodology for collecting and processing data on gender identity in Census 2021](#). A write-in response for gender identity could only be provided where individuals recorded their gender identity as being different to their sex registered at birth (i.e. answering the gender identity question with 'no'). Where 'man' or 'woman' were given as the write-in response for those who answered the gender identity question as 'no', these responses were coded to 'trans man' or 'trans woman'. On the online questionnaire, if someone selected 'yes' to the gender identity question, any write-in response would be deleted.
- 2.24 In the article, ONS explains that the coding of write-in responses did not make use of information from the sex question. This was because while the gender identity question refers to sex registered at birth, the sex question referred only to 'what is your sex?' As some transgender people may view their sex as aligned to their gender identity rather than sex registered at birth, they may have chosen to respond accordingly irrespective of the question guidance. This is particularly likely to be the case for those who have legally changed their sex and hold a Gender Recognition Certificate. The 2021 Census [dataset for gender identity by sex registered at birth](#) shows that the majority of people coded as transgender recorded their sex registered as birth the same as their gender identity. For example, 68 per cent of those recorded as trans men recorded their sex registered at birth as male. This

correlation featured in the concerns raised with us and has been perceived by some as suggesting wide misunderstanding of the gender identity question, as opposed to an intentional decision by respondents. These concerns suggest that some users may assume the gender identity question does relate to the sex question.

2.25 **Our judgement:** We consider the general approach taken by ONS to quality assure the data, when considering it alongside the other variables, was proportionate particularly given the lack of data on this topic area. The approach involved independent scrutiny and expertise when choosing what areas to probe and analyse. In hindsight, given it was a new question, ONS could have prioritised analysis of the agreement rates with the Census Quality Survey as part of its quality assurance process, before releasing the data. Further clarity is also needed to help users understand how the sex and gender identity questions were asked and relate, given the correlation has been perceived by some as an indication that the question has been misunderstood.

ONS's research must communicate how the data can and cannot be used

- 2.26 ONS has published a [range of datasets](#) on gender identity data combined with other variables. These datasets represent a mix of tables pre-determined by user need from the Census topic consultation and bespoke tables produced in response to user need. The datasets range from gender identity by other personal characteristics to qualifications and family composition.
- 2.27 Alongside the release of these datasets, ONS published a [blog](#) in April 2023 which advised users to look out for future articles exploring the data, which would help users interpret the data and understand interactions between variables. While this gives some indication that further work was required to understand the data, it does not adequately highlight the inherent uncertainty in data on a new and hard to measure Census concept, in an area where very limited comparable data was available for quality assurance.
- 2.28 ONS's approach to releasing data on gender identity was made in the spirit of transparency and it was keen to deliver value to users by providing new data on gender identity. We recognise that ONS considers it has been transparent about the potential quality issues through the range of updates it has published, set out in the timeline in table 1. The quality information published alongside the first gender identity outputs included an acknowledgement that responses may be influenced by different interpretations of the question. ONS has also introduced more caveats alongside the data following our [letter on 19 June](#), where we recommended that the communication of uncertainty should be strengthened.
- 2.29 When ONS began receiving queries following the release of first outputs in January, it set up triage meetings that allowed the relevant teams across the organisation to join up and share their perspective on the queries, to provide coordinated and coherent responses to users. This also enabled ONS to respond to media enquiries more easily when these concerns began to feed through into the media coverage of the data. In our interactions with ONS as part of this review, we have seen an openness and recognition of these concerns from the teams in ONS.
- 2.30 In a context where there are a range of concerns around the data, it is important that users are kept updated on research plans and findings as they emerge to support their use of the data. We recognise the positive intention and ambition that underpinned ONS's approach to releasing data, but we consider that ONS should

have developed its quality information more fully, including clearly and prominently highlighting areas of uncertainty in this new collection, before releasing the data to support appropriate use. When releasing big data sets for the first time, it is inevitable that users will want to interrogate and analyse the data for their own purposes. ONS could have done more to see this as an opportunity to draw on the analysis of external users and be clearer that it was open to feedback as part of releasing the data.

- 2.31 The content and tone of the published communication from ONS have been somewhat closed and at times defensive. For example, the default position appears to be that the question is the best that was tested, which can still be true while acknowledging the question may have faced some issues when used in practice. It would be reasonable for users to expect a new question with little data available for triangulation to be considered and communicated as experimental and under development. ONS has assured itself that the data appear to triangulate with other sources at the national level, which will be sufficient for some users who only require high level data, but it should consider how it supports the onward use of the granular breakdowns. There remains an absence of an overarching message that explains the nature of this uncertainty and that clearly sets out that ONS is still trying to understand the data through further analysis. There appears to be some reluctance from ONS to communicate uncertainty unless it can be quantified.

Table 1: Timeline of ONS updates on gender identity.

<p>6 January 2023: ONS published the first outputs on Gender identity, England and Wales: Census 2021</p> <p>4 April 2023: ONS published a package of datasets on gender identity data combined with other variables. It also published a blog alongside the release, which said that later in 2023, it would be publishing analytical articles further exploring the data to help users interpret the data and understand interactions between variables.</p> <p>14 April 2023: ONS released a statement setting out its intention to publish further research on the topic within two months.</p> <p>8 June 2023: ONS preannounced four articles due to be published on 19 June. These articles would cover question testing, collection and processing, microanalysis and macroanalysis.</p> <p>19 June 2023: ONS published the article on Collecting and processing Census 2021 data on gender identity, as well as an update on the research into gender identity data. In this update, ONS said it planned to publish the remaining articles later in Summer 2023.</p> <p>30 August 2023: ONS cancelled the release of the remaining articles and published a statement saying it was conducting ongoing research, undertaking wide user engagement, and would publish further updates in due course.</p>

- 2.32 ONS has said it is focusing on developing a comprehensive package of research on gender identity that explains to users what can and cannot be done with the data. This will be informed by conversations it is carrying out with key user groups. It told us that it has chosen to delay publishing its findings until it can be more confident in the conclusions it is drawing. We recognise the complexity and sensitivity of the research ONS is conducting and agree that this should not be rushed. However, given ONS has missed the dates for committed publications, it is reasonable for users to expect a realistic roadmap for the research that maintains the momentum of its original ambition for the research.

- 2.33 Given the volume of work ONS has carried out to investigate these data further, ONS must bring users along with it as its understanding of the data improves. The external communication around the work published so far and the delays to date could give the impression to users that ONS is not fully explaining the potential issues with the data on gender identity. Media coverage suggests these statements may have reinforced the perception of quality of the published estimates. However, in the absence of ONS's full research outputs, they present only a partial picture. This has the potential to undermine confidence in the good analysis that has been carried out.
- 2.34 On the current [gender identity data harmonised standard page](#), ONS has added a box which states "The Census 2021 gender identity question is the question that tested best with respondents prior to the census being conducted in March 2021 and remains the current recommended standard. Given that this is the first time that these data have been collected, the standard's current status is "under development". The banner also advises those considering using the question to contact ONS for support while it carries out further research into the quality of gender identity data. We welcome the clarity of this message and the support that is offered to those considering using the question. It would be beneficial to users for ONS to include this wording or something similar within the published quality information that is signposted from the gender identity datasets for transparency.
- 2.35 **Our judgement:** ONS has been keen to deliver public value by releasing these new data on gender identity in a timely manner. ONS has taken steps to publish updates on its research around these data and we have seen positive engagement on these issues in our interactions with ONS. We recognise the complexity and sensitivity of the research ONS is conducting and agree that this should not be rushed. However, the published communication to date does not sufficiently manage users' expectations of what areas ONS's research into these data is exploring and what it then might show. A more open approach would manage users' expectations around the use of the data and lay the ground for any changes in advice on the use of these data resulting from ONS research.
- 2.36 Given that this is a contested area, with little data available for triangulation it would be reasonable for users to have expected clearer messaging on the impacts on quality and uncertainty. This gained added importance when users raised challenges and offered alternative hypotheses to explain the data. In this context ONS now needs to undertake further analysis and be clear on what can and cannot be concluded from the data.
- 2.37 The message that ONS provides upfront in a box on the gender identity data harmonised standard page is a good example of a clear and honest acknowledgment of uncertainty without the need to quantify it. This approach should be adopted more broadly for ONS's outputs and quality information surrounding the gender identity data to ensure the data are not open to misuse and misinterpretation.

Expectations for further work

- 3.1 Throughout our review, it has been clear that the ONS team leading the quality assurance and additional exploration of the data is committed to understanding the quality of the data and how this can best be communicated to users. The team has sought innovative ways to interrogate the data, including subgroup analysis and data linkage. It has also ensured to keep NRS updated on its findings as it goes along and NRS told us that ONS has been incredibly helpful.
- 3.2 We expect ONS's research on gender identity to seek to understand to what extent, if any, has a misunderstanding of the gender identity question in the 2021 England and Wales Census led to a significant misestimation of the transgender population. When the target group is small, the effect of false positives and negatives on the data will be amplified.
- 3.3 ONS should consider the testing and operationalisation of the question as well as the production and analysis of the results. If there is evidence that there has been a significant misestimation of the transgender population to a degree that might affect usage, the research should set out both the implications for the use of the results from the Census and the use of the GSS harmonised question in future data collection.
- 3.4 Beyond the use of the 2021 results, the [Future of Population and Migration Statistics consultation document](#) highlights issues with lack of coverage of administrative data on gender identity. It is therefore crucial that ONS understands the performance of this question through the additional research and determines how it will look to collect these data in the future. This should be considered alongside its approach for the development of UK wide estimates on gender identity.
- 3.5 We expect ONS to take on board the following recommendations as it works on developing its research.

3.6 Recommendations to support further use of the data:

- ONS should publish a statement within four weeks of this interim report that sets out a timeline for its ongoing research, covering what it is analysing and what it plans to publish by when. This statement must be open about the potential quality issues it is investigating, even if it cannot quantify the scale of them. This is important both for the use of published estimates, but also the use and further development of the harmonised question.
- ONS should make data available to understand subgroups of the reported transgender population. For example, it should support analysis of differences in responses for those who provided an explicit and unambiguous write-in response indicating that they were transgender compared with those who provided a possibly ambiguous or tick-box only response.
- ONS should consider any potential misunderstanding of the question or mode effects on the data as part of its analysis into the agreement rates for the gender identity question in the Census and Census Quality Survey. In particular, it should publish analysis of group differences to inform future use of the data and question.

- As part of the [Future of Population and Migration Statistics consultation](#) and programme, the UK harmonisation group should set out its early thinking on gender identity estimates for Great Britain and publish its development plan for these statistics.

3.7 Recommendations to support further use of the question:

- As part of the ongoing development of the harmonised standard for gender identity, should the question remain the same, ONS should carry out further testing of the question. This should be considered as part of the outcome for the [Future of Population and Migration Statistics consultation](#), in determining future production of these statistics in an administrative data based framework.
- ONS should publish more details on the feedback it received in the Census rehearsal, including any insight about those with lower English language proficiency.
- Once NRS has published its data on gender identity from the 2022 Scotland Census, planned for 2024, ONS and NRS should carry out a lessons learned exercise comparing the performance of the two questions and use this to inform future developments and use of the questions.

Conclusion

- 4.1 As this interim report makes clear, there was always good reason to expect a significant degree of uncertainty around estimates of the transgender population derived from the gender identity question in the 2021 Census, given that this is the first time that the question has been asked, that the concept of gender identity may be unfamiliar to many people, that the proportion of the population who are trans is relatively small and that there are few robust alternative sources of data to cross-check against.
- 4.2 ONS has engaged to develop, test and quality assure a question and data on gender identity to meet user need. While there are lessons to learn, ONS made decisions with the right motivations and in line with the Code of Practice for Statistics.
- 4.3 However, ONS should have communicated the inherent uncertainties in this new data collection when it published the data to aid use. Had ONS done so, it would have aided users understanding of appropriate use of the data and provided relevant context when issues were identified by users. ONS could have been more open and transparent about these novel data and the work that has followed.
- 4.4 A willingness on the part of ONS to provide further context and clarity on the initial estimates would be evidence of the strength of the Census process and not a weakness. ONS providing updates to the interpretation of these estimates should be regarded as a normal part of ongoing statistical production for a new area of data collection. The same would be true of any changes to the initial estimates, as long as they are fully and transparently explained. Such updates should not undermine user confidence in the robustness of the Census results as a whole, in which most questions have been asked many times before.
- 4.5 We are satisfied that the overall approach to the 2021 Census complies with the standards of trustworthiness, quality and value in the Code of Practice for Statistics in order to achieve 'accredited official statistics' status, but we will consider the accreditation of this subset of the Census dataset when ONS publishes its research findings. We will publish a follow up report with our findings after we have reviewed, and engaged with users on, the full set of research.

