Office for Statistics Regulation

Final report

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

(produced by the Office for National Statistics)

Assessment Report 391

12 September 2024

Executive Summary 3
Required future work by ONS and the GSS 5
Introduction 7
Background to the review 7
What has happened since the interim report was published? 9
Our approach 10
Analysis of concerns about the statistics 12
English language and “none specified” gender responses 13
National vs local level estimates 16
Communicating uncertainty in the data 17
ONS’s position 19
OSR’s regulatory decision 21
Required future work by ONS and the GSS 25
Guidance for use of the census gender identity statistics 25
Gender identity question development 25
Annex A – Census publication timeline 27

Edits to this report

Readers should note that we have made the following edits to one paragraph of this report to improve the clarity of the paragraph. The edits are in the section ‘English language and “none specified” gender responses’ (page14).

Original text - 12 September 2024

The NRS Scotland Census data show that of the 19,987 people who answered “Yes” to whether they consider themselves to be trans or have a trans history, 0.07% did not provide a specific gender identity or trans status (Table 2).

Edited for clarity - 24 September 2024

The NRS Scotland Census data show that of the 19,987 people who answered “Yes” to whether they consider themselves to be trans or have a trans history, 0.07% of census respondents aged 16 and over did not provide a specific gender identity or trans status (Table 2).

Further edit for clarity – 11 June 2025

The NRS Scotland Census data show that 19,987 people answered “Yes” to whether they consider themselves to be trans or have a trans history. 0.07% of census respondents aged 16 and over did not provide a specific gender identity or trans status (Table 2).

Executive Summary

This is the final report on our review of statistics on gender identity based on data collected as part of the 2021 England and Wales Census.

Following the first release of census statistics on gender identity in England and Wales in January 2023, concerns were raised publicly with ONS and with us about the published estimates of the trans population. As additional census data were published, these concerns extended to the relationship between gender identity and proficiency in English. We also received concerns about the level of methodological information published.

Learning from new evidence in Scotland's Census, the Office for National Statistics (ONS) wrote to us on 5 September 2024 to request that the gender identity estimates from Census 2021 should no longer be accredited official statistics and should instead be classified as official statistics in development.

There is great value in ONS having undertaken this work on developing these statistics. We recognise the challenges in collecting data at scale on a sensitive topic for the first time. While there are lessons to be learned, ONS has worked with good intentions aiming to address an important data gap on gender identity.

This commitment is demonstrated by ONS recognising that it can do more to support users to interpret the census gender identity data and to understand any limitations, and has led to its request to cancel the accreditation of the gender identity estimates from Census 2021. ONS's decision to label these data as statistics in development supports its desire to publicly reflect the evolving understanding of gender identity and to undertake further research. We welcome statistics producers approaching us when they have any concerns about the statistics they produce.

We have written to ONS to support its proposal to publish the Census 2021 gender identity statistics in England and Wales as official statistics in development rather than accredited official statistics. We found that the issues highlighted in this report are unique to the statistics on gender identity and therefore all other outputs from the Census 2021 in England and Wales are unaffected by this decision and remain designated as accredited official statistics.

This report sets out the details of our regulatory decision and shares our recommendations to ONS on the steps it must take to help users of the census gender identity statistics understand their strengths and limitations. In addition, we have also set out the development work we consider is required on the Government Statistical Service (GSS) gender identity harmonised standard. This includes a need to publish additional advice for statistics producers who are currently using the standard and for those who are thinking about adopting it. The GSS Harmonisation Team must consider these recommendations as it further develops its workplan. This report has been reviewed and approved by the Authority's Regulation Committee, which oversees the work of OSR.

While we recognise ONS's commitment throughout the census operations to obtaining sound evidence, we have found that that the Census 2021 gender identity statistics published by ONS do not fully comply with the following areas of the Code of Practice for Statistics:

Quality – **Suitable data sources.** As ONS has now recognised, the question developed to determine the size of the trans population of England and Wales did not work as intended. The evidence indicates that people may have found the question confusing and therefore gave a response that did not reflect their gender identity. This appears to be more likely for people who do not speak English as their first language. It is important to note that neither ONS, nor we at OSR, identified in the development or delivery process of the census that there was a significant risk that certain groups of people were more likely than others to misunderstand the question.

Assured Quality. There are a number of areas where ONS could have better communicated with users of the census gender identity statistics. Our interim report found that ONS's communications did not sufficiently manage users' expectations on what areas ONS was exploring and what the research might show. Additionally, until recently, ONS's approach has been characterised by a focus on defending the estimates of the size of the trans population. ONS has not shown sufficient willingness to engage constructively with the emergent evidence. The publication of the Scotland Census data on trans and trans history in June 2024 has been instrumental in a shift in ONS's approach.

We have confidence that those working on the data and research have been working in line with the Code, but consider that stronger leadership from ONS, particularly on clearer public messaging when concerns were raised and on setting expected outcomes for ONS's subsequent research work, was needed.

Communicating uncertainty. We have found that ONS has not done enough to communicate uncertainty and that users who use the census outputs at a national level would struggle to understand the impact of ONS's published research on the use of the gender identity data. This problem is replicated at the local level, where the research indicates that the variability of the estimate may be greater.

ONS has conducted a range of research activities to explore the quality of the estimates, but these efforts have been hampered by a reluctance to communicate the limitations of the gender identity statistics. ONS's headline message in the research report around the validity of the census data, and how users should interpret and use the data at national and smaller area levels, is unclear.

We have also considered three further factors relating to the status of these estimates:

ONS's approach to criticism of the gender identity data. We noted in our interim report that the content and tone of the published communication from ONS has been somewhat closed and at times defensive. Some users have raised concerns that ONS has been captured by interest groups, leading to a lack of objectivity. We found no evidence of this form of bias through any of our work. We consider it regrettable that ONS's defensiveness has created an impression of bias to some external observers.

Whether the statistics are materially misleading. We also considered whether the statistics are materially misleading. The purpose of the Code's Quality pillar is to ensure that statistics are a fair representation of what they seek to measure and are not materially misleading.

At the national level, triangulation with other sources, including data from the Scottish Census, suggests that an estimate of around "1 in 200" triangulates with other sources and is not likely to be materially misleading. While the information on the size and

nature of all the potential biases is incomplete, it is hard to draw the same conclusion for some more-detailed breakdowns, including for local areas where the data indicate a higher concentration of people misunderstanding the question.

Our regulatory approach. When developing new statistics or introducing a new data source or methodology, initial publications are often labelled ‘official statistics in development’ by producers. We would usually assess the official statistics against the Code of Practice once the development is complete.

However, the census is a unique case. Our regulatory accreditation process was undertaken in tandem with the development and first publication of the statistics in the respective countries of the UK. While this process allows us to provide users of these statistics with the reassurance of quality that comes with accreditation at first publication, it means that if we have any subsequent concerns about any aspect of the Code of Practice for Statistics, we may need to reconsider our accreditation decision in light of new evidence.

In this case, new evidence has emerged that has demonstrated that the statistics do not comply with important quality aspects of the Code of Practice for Statistics. With the benefit of hindsight and this new evidence, our conclusion is that the census gender identity statistics, which were collected from a novel question in an area where measurement practices are still emerging, would be better described as official statistics in development, and should not be labelled as accredited official statistics.

We acknowledge that this is a limitation with our approach, and as a result of this review, we are reviewing what changes might be needed to our assessment approach in circumstances where accreditation at first publication is in line with serving the public good.

Required future work by ONS and the GSS

There is great value in ONS having undertaken this work on developing how people are able to describe society. We recognise the challenges in collecting data at scale on a sensitive topic for the first time.

Our recommendations focus on the steps ONS must now take to help users of the census gender identity statistics understand their strengths and limitations and the development work we consider is required on the GSS gender identity harmonised standard.

Guidance for use of the census gender identity statistics

- It is important that ONS communicate clearly across all its census outputs what official statistics in development means in practice for the census gender identity statistics.
- In doing this, the varying levels of information across the gender identity publications and data tables must be addressed by ONS across its suite of publications and data on gender identity.
- ONS should provide more practical information, according to the needs of stakeholders, to help users understand the uncertainty inherent in the data and how this affects their use and interpretation. This information could include sensitivity analysis, and example use cases published at different levels of geography and population highlighting any likely anomalies or implausibilities.

- ONS should consult with us as it develops these communications, which should be treated as a priority workstream.

Gender identity question development

These data were identified as being important for society, and so it is important that the original user need for robust data on gender identity be addressed for future statistics. ONS should consider how these data needs are best addressed in future as part of its further testing of the question.

We consider that there is sufficient correlational evidence that the gender identity question in the census has underperformed, and some users, producers and stakeholders have lost confidence in the suitability of the gender identity question.

This issue is important as the census gender identity question is also the question used in the GSS gender identity data harmonised standard. As we set out in a letter to the Deputy National Statistician with oversight of the Government Statistics Service (GSS) in February 2024, we are seeing in our regulatory work the implications of the lack of confidence in the harmonised question by producers, who are unsure what gender identity question they should use and feel unsupported by the Government Statistical Service. In particular:

- We can confirm that our interim recommendation remains: namely that, as part of the ongoing development of the GSS harmonised standard for gender identity, ONS should carry out further testing of the question. ONS must give consideration to developing and testing a harmonised question that commands the confidence of producers and users. The merits of the Scottish Census question should also be considered in this process.
- This work on the question should be considered as part of the outcome for the Future of Population and Migration Statistics consultation, in determining the future production of these statistics in an administrative data-based framework.
- Given that the census gender identity question is also the GSS harmonised standard, ONS should prioritise this work, ensuring that adequate resources are made available.
- We consider that there is an immediate need for the GSS to publish advice for statistics producers who are currently using the gender identity question and for those who are thinking about adopting the standard. This guidance should set out the limitations of the question, reference the Scottish Census approach and caution against knee-jerk changes to established collections until ONS has undertaken further research.
- We recommend that stakeholders be fully involved and engaged in future work, to help build and restore confidence.

The GSS should publish its initial plans for developing the gender identity harmonised standard within three months of our report's publication.

Introduction

This is the final report on our review of statistics on gender identity based on data collected as part of the 2021 England and Wales Census.

Learning from new evidence in Scotland's Census, the Office for National Statistics (ONS) wrote to us on 5 September 2024 to request that the gender identity estimates from Census 2021 should be no longer accredited official statistics and should instead be classified as official statistics in development.

Additionally, ONS has informed us that it will do more to support users in interpreting the census gender identity data and will work with others across the Government Statistical Service (GSS) on developing the gender identity harmonised standard and on sharing guidance for existing survey questions.

We have written to ONS to support its proposal to publish the Census 2021 gender identity statistics in England and Wales as official statistics in development rather than accredited official statistics.

This report sets out the details of this regulatory decision. It also shares our recommendations to ONS on the steps it must take to help users of the gender identity statistics understand their strengths and limitations, and on the development work we consider is required on the GSS gender identity harmonised standard.

The report is structured as follows:

- This section sets out the background to our review, including setting out the timeline of ONS and OSR outputs on this topic
- The second section sets out an analysis of the concerns about the statistics
- The third section sets out our regulatory judgement and decision
- The fourth section sets out future work required by ONS and the GSS

Note that the term trans is used in this report to describe anyone who stated in the census that their gender identity was different to their sex registered at birth. This includes people who identify as a trans man, trans woman or non-binary, and anyone with another minority gender identity. This terminology aligns with the approach adopted by ONS and the National Records of Scotland (NRS).

Background to the review

A voluntary question on gender identity was asked for the first time in the 2021 England and Wales Census conducted by ONS. This question is also the current GSS harmonised standard (currently in development) for collecting data on gender identity.

The question that people were asked was "Is the gender you identify with the same as your sex registered at birth?". They had the option of selecting either "Yes" or selecting "No" with the option to write in their gender identity.

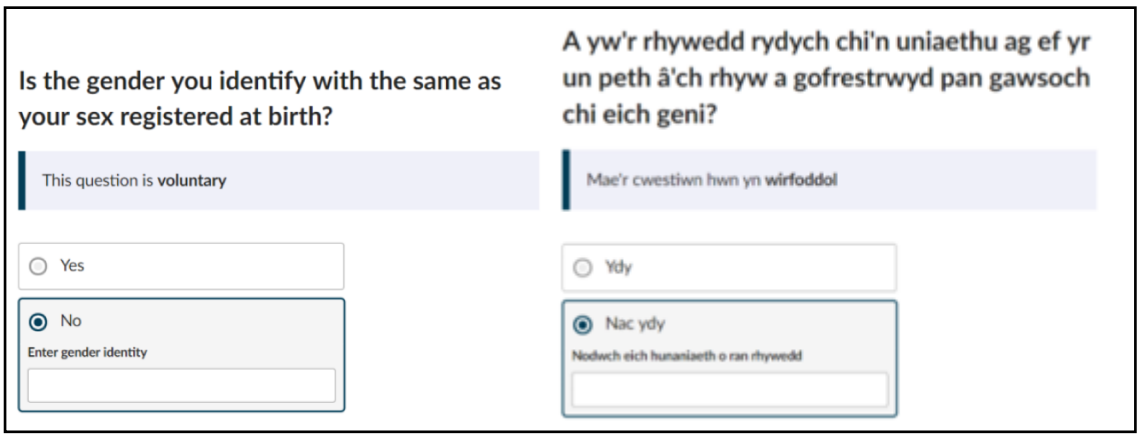

Figure 1. Source: Office for National Statistics – Sex and gender identity question development for Census 2021

The census statistics on gender identity were first released in January 2023 in the statistical bulletin Gender identity, England and Wales: Census 2021.

Following this publication, concerns were raised publicly with ONS and with us about the published estimates of the trans population. As additional data were published, these concerns extended to the relationship between gender identity and proficiency in English. We also received concerns about the level of methodological information published. In response, ONS initiated research work on the topic in April 2023.

Also in April 2023, we commenced a review, based on the principles of the Code of Practice for Statistics, of the census gender identity statistics and ONS's research findings. The review has considered ONS's approach to responding to questions raised about the results of the gender identity data. It has not considered gender identity as a concept or the decision to collect data on gender identity in the census.

On 19 June 2023, ONS published an article which provided additional information on the methodology for collecting and processing data on gender identity in Census 2021. We wrote to ONS to welcome the additional information and to share our first emerging findings from our review to help inform its research.

The publication of ONS's research was postponed twice, having been preannounced for June and then August 2023. ONS published an update on the progress of its research in August 2023.

We published an interim report in October 2023. This set out our expectations for ONS in line with the Code of Practice for Statistics on what we considered was needed to provide assurance on the gender identity census data and the harmonised question. We based the recommendations in the report on evidence that we had gathered on how ONS tested, quality-assured, communicated and engaged with users, both in advance of publishing the results and from the research to date.

In our interim report, we stated that we were 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, which is the level of compliance required to achieve accredited official statistics status. But we stated that we would consider the accreditation of this new subset of the Census dataset once ONS had published its

research findings. Further, we said that we would publish a final report with our findings after we had reviewed, and engaged with users on, the full set of research.

What has happened since the interim report was published?

At the time of writing our interim report, ONS was focusing on developing a comprehensive package of research on gender identity to explain to users what can and cannot be done with the gender identity census data. It told us that it had chosen to delay publishing its findings until it could be more confident in the conclusions it was drawing. We recommended that ONS set out a timeline for when it would complete its research; in response, ONS decided to instead publish the work it had already done.

Given the amount of time that passed since ONS initiated the research in April 2023, users questioned the justification of the publication delays given that the research amounts to six pages of content, much of which repeats text or data from OSR's review or signposts to previous material. The report also does not reflect the four outputs that had been preannounced by ONS, which is what users were expecting (Annex A, timeline). ONS did not acknowledge the change in research output. This approach had the potential to damage trust and confidence in the research.

We have confidence that those working on the data and research have done so diligently and in line with the Code. However, we feel this work could have been led with more openness, transparency and direction.

ONS published its research report, Quality of Census 2021 gender identity data, in November 2023. The report covers the background to the question, respondent error, relationships with other characteristics, comparability with other sources and engagement with users of the data. The publication of this research report demonstrates that ONS has sought to take action to respond to user demand for more information to support use of the data. This effort has included releasing additional data tables and engaging with over 20 organisations, including government departments, local authorities, academia and the non-profit and commercial sectors, to understand their needs for the data and supporting information.

The additional data tables published alongside the Quality of Census 2021 gender identity data report are useful for understanding the nature of responses and the intersectionality with other characteristics. These include data on:

- the proportions of the estimated trans population within each local authority who provided a write-in response (GI01: Census 2021 gender identity, proxy indicators of uncertainty for local authorities)
- distribution across categories for the estimated trans population and additionally the distribution based only on those who provided a write-in response to the question (GI02: Census 2021 gender identity, individual characteristics by gender identity)
- distributions for household characteristics for the estimated trans population in households and additionally the distribution based only on those who provided a write-in response to the question (GI03: Census 2021 gender identity, household characteristics by gender identity)
- the coding index for the gender identity question (GI04: Census 2021 gender identity write-ins)
- counts and percentages for sex by proficiency in English by gender identity (GI05: CT21_0186 - Sex by proficiency in English by gender identity)

- proportions of the estimated trans population by age and country of birth – UK or non-UK (GI06: Census 2021 gender identity by country of birth)

Additionally, in November 2023, ONS updated its methodology document, which presents known quality information affecting sexual orientation and gender identity data: sexual orientation and gender identity quality information for Census 2021. This document includes a link to the quality report noted above.

On 27 June 2024, National Records of Scotland (NRS) published the Scotland's Census 2022 data for sexual orientation and trans status or history. Alongside this statistical publication, NRS published a separate quality assurance document for the variables included in the statistical release.

The voluntary question asked in Scotland to those 16 years or older was "Do you consider yourself to be trans, or have a trans history?". People were asked to tick "No" or "Yes". People who ticked "Yes" were asked to describe their trans status (for example, non-binary, trans man, trans woman).

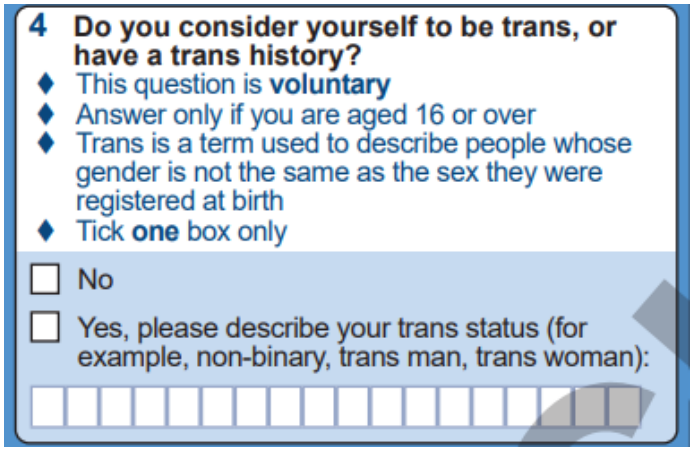

Figure 2. Source: Scotland's Census 2022

The voluntary question asked in the Scotland Census was different to the voluntary question asked in the England and Wales Census, which was "Is the gender you identify with the same as your sex registered at birth?". Other differences between the censuses included the following:

- The Scotland question included a prompt for people to describe their trans status or history and gave examples of non-binary, trans man and trans woman, in that order. The England and Wales question did not include any examples of gender identities for the prompt.
- The Scotland Census included a brief definition of trans on the questionnaire. The England and Wales Census did not include a definition.
- The question in Scotland was asked directly after a question on sex on the census form in Scotland, while it was asked later in the census form in England and Wales.

Our approach

As well as reviewing ONS's research, we carried out stakeholder engagement. The views of those who use, or would like to use, the gender identity statistics helped to direct our research. We engaged with a range of stakeholders, users and potential

users of the data. Our engagement included speaking to some of the stakeholders who had raised concerns with us and ONS during the research for our interim report, as well as several organisations which had not previously contributed to the review.

Our engagement highlighted a wide variety of views and strength of opinion. These stakeholder views helped us to gain a better understanding of the extent to which ONS's post-publication research has responded to user concerns and helped users to better understand and use the data. It was also helpful to hear stakeholders' views on the gender identity question and the potential for its further development as the harmonised standard.

This report has been reviewed and approved by the Authority's Regulation Committee, which oversees the work of OSR.

Analysis of concerns about the statistics

This section sets out the analytical concerns about the statistics.

As we set out in our interim report, measuring gender identity is undoubtedly challenging. Gender identity 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. The concept of gender identity may be unfamiliar to some, others may not feel they have a gender identity, and there are few robust alternative sources of data to cross-check against. These factors create challenges to developing new statistics.

Following the first release of the census statistics on gender identity in England and Wales in January 2023, concerns were raised publicly with ONS and with us about the published estimates of the trans population. As additional census data were published, these concerns extended to the relationship between gender identity and proficiency in English. We also received concerns about the level of methodological information published.

We noted in our interim report that those aged 16 and over whose main language is not English (or Welsh if using the Welsh questionnaire) made up 10 per cent of the overall population in the ONS census data, but 29 per cent of those whose gender identity differed from their sex registered at birth. Additionally, the geographical areas with the highest proportion of people reporting themselves as trans in the census were areas that also had a high proportion of people whose first language is not English (or Welsh). Further, we shared that these data led some users of the census gender identity statistics to conclude that the question on gender identity may not have been understood by some, particularly 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 and an April 2023 research paper, Gender Identity in the 2021 Census of England and Wales: What Went Wrong?. These publications were cited by several stakeholders we spoke to as important pieces of work in informing their views on the quality of the data.

Some stakeholders we spoke to, who also contributed to our interim report, told us that their initial concerns that the question may have been more likely to be misunderstood by those with lower English language proficiency had been justified by the findings set out in the ONS research report, Quality of Census 2021 gender identity data, published in November 2023 and by the information in the additional data tables.

The publication on 27 June 2023 of Scotland's Census 2022 statistics on sexual orientation and trans status or history has provided additional data. The publication of these statistics further supports the view that the question in the England and Wales Census was misunderstood by some respondents, and that this had an impact on the quality of the gender identity data collected in the England and Wales Census.

This section of the report looks at to what extent the publication of the ONS research report in November 2023 addressed these concerns. Additionally, we consider whether its publication has mitigated any remaining concerns, such as by providing increased quality information on uncertainty or additional guidance on using the statistics.

English language and “none specified” gender responses

Census respondents who recorded “No” to the question “Is the gender you identify with the same as your sex registered at birth?” in the ONS England and Wales Census were prompted to write in their gender identity. This write-in box could be left blank (Figure 3).

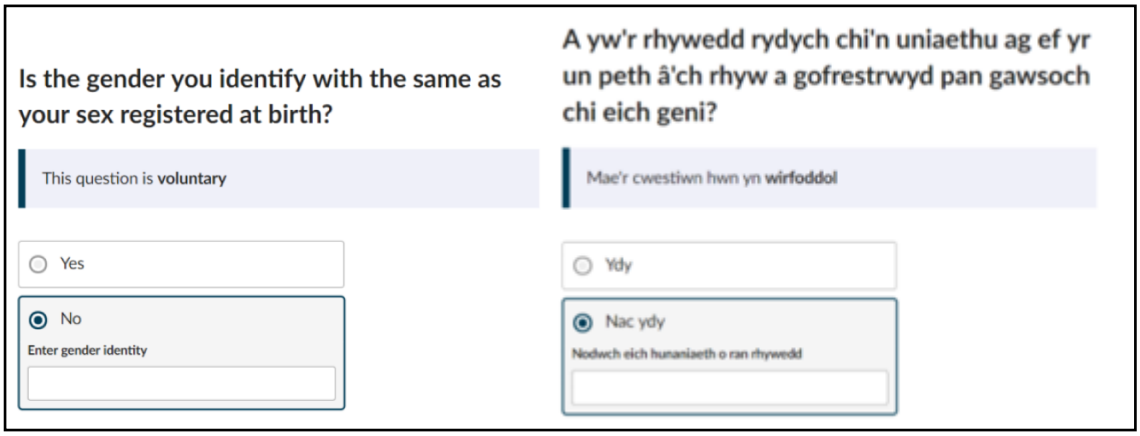

Figure 3. Source: Office for National Statistics – Sex and gender identity question development for Census 2021

Of the 262,112 adults recorded as having a gender identity different from their sex registered at birth, 117,775 people were classified in the census category ‘Gender identity: different from sex registered at birth, but no specific identity given’; they left the box blank; or, if they did write in a response, it was not coded as valid. This figure equates to 0.24% of respondents (Table 1).

	Gender identity the **same as sex at birth**		Gender identity different from sex registered at birth, **but no specific identity given**		Gender identity different from sex registered at birth **and specific identity given**		Not answered	
All usual residents aged 16 or over	45,389,635	93.46%	117,775	0.24%	144,335	0.3%	2,914,625	6.00%

Table 1. Source: ONS census data

In Scotland’s Census, respondents who ticked “Yes” to the question “Do you consider yourself to be trans, or have a trans history?” were asked to describe their trans status. Examples of non-binary, trans man and trans woman were also given. This write-in box could be left blank (Figure 4).

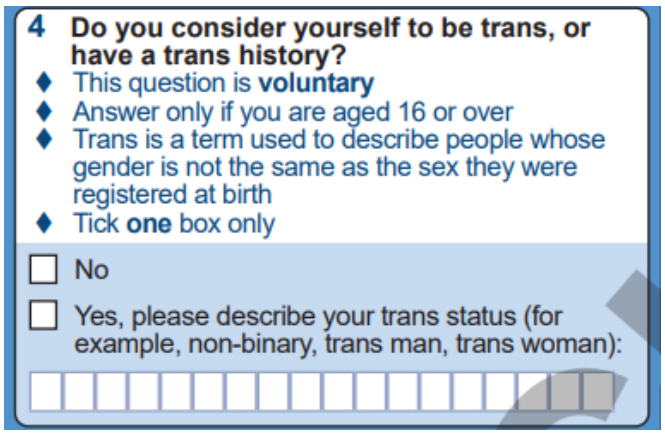

Figure 4. Source: Scotland's Census 2022

The NRS Scotland Census data show that 19,987 people answered "Yes" to whether they consider themselves to be trans or have a trans history. 0.07% of census respondents aged 16 and over did not provide a specific gender identity or trans status (Table 2). They left the box blank, or, if they did write in a response, it was not coded as valid. These responses are recorded in the census category "Yes: Trans status or history not otherwise specified".

	No: Not trans and does not have a trans history		Yes: Trans status or history **not otherwise specified**		Yes: Trans or has a trans history (specified)		Not answered	
All people aged 16 and over	4,259,372	93.64%	3111	**0.07%**	16,876	0.37%	269,230	5.92%

Table 2. Source: NRS census data

The ONS and NRS data are set out comparatively in the Scotland quality assurance document (Figure 5).

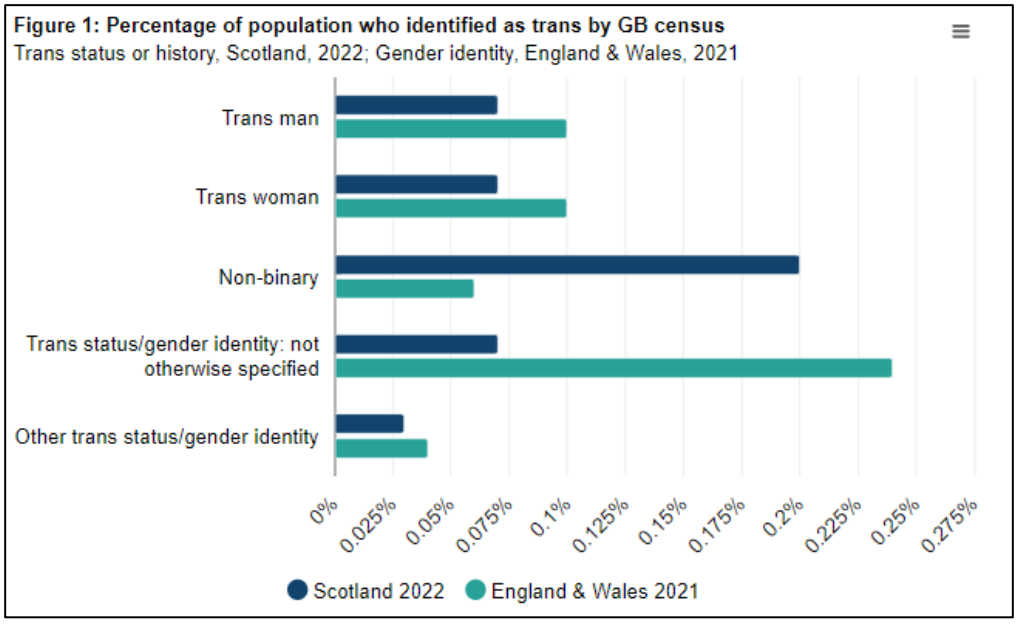

Figure 5. Percentage of population who identify as trans by GB census. Source: NRS quality assurance report

While the quality assurance report for the Scotland Census expresses that caution should be used when interpreting the difference between Scotland Census trans data and the England and Wales Census gender identity question due to the different questions being asked, the Scotland data show that a greater percentage of census respondents 'wrote in a response' in Scotland than in England and Wales. The different focus and wording of the question used in Scotland may have been better understood than the focus and wording in England.

The ONS research report, which was published prior to the publication of the Scotland Census data, investigated patterns of trans identification by proficiency in English. The report noted that *"of those who indicated their gender identity was the same as their sex registered at birth, around 2% did not speak English well. This proportion was around 13% for those who indicated their gender identity was different to their sex registered at birth and did not write-in a specific gender identity. These patterns were still evident after adjusting for the different age profiles of the groups involve."*

Further, ONS noted that *"respondents recorded as trans who were born outside the UK or had a lower proficiency in English were around twice as likely to tick the box and not write in a response compared with other trans respondents*" (Table 3) and that *"This could be attributed to either the tick-box being answered inaccurately, or those with lower proficiency in English not being sure how to accurately describe their identity in English, or simply wanting to minimise the amount they needed to write in their response to this voluntary question."*

Population group	Gender identity the **same as sex at birth**	% of popn group	Gender identity different from sex registered at birth, **but no specific identity given**	% of popn group	Gender identity different from sex registered at birth **and specific identity given**	% of popn group	Not answered	% of popn group
All usual residents aged 16 or over	**45,389,635**	**93.46%**	**117,775**	**0.24%**	**144,335**	**0.30%**	**2,914,625**	**6.00%**
Main language is English (English or Welsh in Wales)	41,296,005	94.05%	69,165	0.16%	116,625	0.27%	2,428,405	5.53%
Main language is not English (English or Welsh in Wales)	4,093,630	87.92%	48,610	1.04%	27,720	0.60%	486,220	10.44%
Of which, can speak English								
Very well	1,786,555	90.65%	9,930	0.50%	9,665	0.49%	164,670	8.36%
Well	1,480,955	86.20%	23,295	1.36%	11,720	0.68%	202,060	11.76%
Not well	703,400	85.44%	13,135	1.60%	5,515	0.67%	101,230	12.30%
Not at all	122,725	85.20%	2,245	1.56%	805	0.56%	18,255	12.67%

Table 3. Source: ONS CT21_0008 – Proficiency in English by gender identity

A breakdown of census data by English language proficiency is presented in the quality assurance report for the Scotland Census. The pattern of data observed the ONS census data – that respondents with a lower proficiency in English were more likely to state their gender identity was different to their sex and birth but not to write in a gender identity – was not apparent in the trans and trans history data from the Scotland Census (Table 4).

	No: Not trans and does not have a trans history	%	Yes: Trans or has a trans history – all	%	Yes: Trans status or history (not otherwise specified)	%	Yes: Trans or has a trans history (specified)	%
Main language is English	4,016,453	93.65%	18,396	0.43%	2858	0.07%	15538	0.36%
Main language is not English – speaks English very well/well	197,700	93.68%	1,365	0.65%	177	0.08%	1188	0.56%
Main language is not English – speaks English not well/ not at all	45,219	92.46%	232	0.47%	82	0.17%	150	0.31%

Table 4. Number and percentage of people by English language proficiency and response to the trans status or history question. Produced by OSR; source data from NRS quality assurance report

National vs local level estimates

ONS has found in its research that it cannot say with certainty whether the census estimate that 0.54% of the population who identify as a gender that is different to their sex registered at birth is more likely to be an overestimate or an underestimate. The report also states that, from ONS's research comparing the census estimates with other available sources, it found no reason to conclude that the census statistics on gender identity were implausible.

The research report recognises that while the GP Patient Survey (GPPS) is a good comparator for the census data, in that it has very good coverage of the population and collected data for a similar point in time, it has a drawback as a comparator in that the wording of the question is very similar to that on the census. The report acknowledges there could be a risk that any respondent errors are correlated across the two sources.

The research report compared the census estimates of the trans population with the 2021 Canadian Census and surveys in the United States. The Canadian 2021 Census results show that 0.33% of Canada's population provided an answer to the gender question that was different to their answer to the sex question. The report shares that studies in the United States suggest that between 0.3% and 0.6% of the country's population identify as a gender that is different to their sex registered at birth.

The publication of Scotland's Census data on trans status or history, which followed the ONS research report, has contributed to a better understanding of the trans population. Scotland's Census 2022 found that 19,990 people were trans or had a trans history. This figure equates to 0.44% of people aged 16 and over (Table 5).

Category	Scotland 2022	England and Wales 2021
Trans people (defined in the Scotland quality assurance report as people with a trans status or history in Scotland and people with a gender identity different from their sex registered at birth in England and Wales)	0.44%	0.54%
Not stated	5.92%	6.00%

Table 5. Percentage of trans people, Scotland 2022 and England and Wales 2021. Source: NRS quality assurance report

While the proportion of trans people identified in England and Wales is 0.1 percentage points or around 23% higher than in Scotland, data from both 'England and Wales' and 'Scotland' are broadly comparable, indicating a trans population of around "1 in 200".

Communicating uncertainty in the data

Clearly communicating any uncertainty around estimates is an important part of ensuring the appropriate use and interpretation of statistics, as it helps users of the data understand the data's strengths and limitations.

Our interim report found that ONS should have communicated the inherent uncertainties in this new data collection when it published the data to aid use, stating that "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."

We commented that had ONS done so, it would have aided users' understanding of the appropriate use of the data and provided relevant context when issues were identified by users. Further, we advised ONS to communicate in its future research how the gender identity data can and cannot be used.

We have looked again, using the principles set out in our guidance on communicating uncertainty, at how ONS has communicated uncertainty. We reviewed ONS's research report, Quality of Census 2021 gender identity data, and the presentation of uncertainty in the gender identity data across a number of census outputs.

ONS's research report provides valuable information for users on the potential sources of uncertainty in the gender identity estimates. To date, the report is the most comprehensive overview of uncertainty ONS has provided for these data. It explores a wide range of possible factors, many of which had not been previously presented in the statistical bulletin or the quality information. The research report also provides ONS's conclusions on whether each factor has likely impacted the quality of the data. The dataset included in the research report provides data tables to allow users to explore the variation in distribution across local authorities and individual and household characteristics.

The report states that there are good reasons to expect higher levels of uncertainty with the estimates relating to the gender identity question than for other census topics. It also shares that users of the data told ONS that they expect higher uncertainty around census estimates for this topic, particularly as it is a voluntary question, and that they would welcome more information to support their use of the data. To meet this need, ONS released additional tables alongside the report and an accompanying

explanation of how these data should be used. The explanation includes the following information and advice:

Proportion of write-in responses for local authorities. A data table that gives proportions of the estimated trans population within each local authority who provided a write-in response (census respondents who recorded "No" to the question "Is the gender you identify with the same as your sex registered at birth?" and also wrote in a gender identity).

With regards to use of these data, ONS advised that, "*If gender identity estimates for a particular local authority are based on a low proportion of write-in responses (therefore a higher proportion of tick-box only responses), then users may wish to exercise greater caution in using these statistics and consider them alongside other data sources and local knowledge for practical purposes."*

Topic distributions by gender identity. These tables provide alternative presentations of previously published data on individual characteristics by gender identity, household characteristics by gender identity, sex by proficiency in English by gender identity and gender identity by country of birth.

ONS advises that *"For each topic, they show the distribution across categories for the estimated trans population and additionally the distribution based only on those who provided a write-in response to the question. Topics where both distributions are similar are likely to be relatively unaffected by any respondent error resulting from simply ticking the wrong box. A greater degree of care would be advisable where the distributions are markedly different."*

Where uncertainties are highlighted, it is not fully explained how they affect the use of the data. We are concerned as to how easy it would be for a member of the public or a casual user to easily understand any limitations with the gender identity data. While ONS states in its research report that the users it spoke to "*understand the relatively high levels of uncertainty in the estimates for this topic*", we consider they are likely to be expert users, and that ONS should do more to support a wider range of users of the data.

Understanding uncertainty is also complicated by the number of separate ONS publications on this topic and the different levels of detail provided in each. For example, the quality report Sexual orientation and gender identity quality information for Census 2021 does not convey the same level of information as the research report.

Many users will likely have complete confidence in the data solely by virtue of them being published by ONS and may not seek to conduct their own detailed analysis of their strengths and limitations or seek out this information in the first place.

We consider the following points:

- 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 by ONS. 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, as most questions in the census have been asked many times before.

- While the ONS research report provides additional information and explores potential sources and impacts of misestimation, ONS's headline message around the validity of the census data, and how users should interpret and use the data at national and smaller area levels, is unclear.
- There is insufficient evidence for any implication that these uncertainties support the central estimate as the best estimate.
- Most users of the data would reasonably expect some estimate of underestimation of the trans population from non-response from data on a sensitive topic – recognising the voluntary nature of the question – but would not reasonably expect the correlations that imply sub-groups misunderstood the question.
- Statisticians are used to dealing with the usual uncertainties that come with sensitive topic survey data, and many regular and expert users are familiar with these as well. However, in this case the potential misunderstanding of the question adds an additional uncertainty that users would not reasonably expect. This scenario is distinctive and uncommon, and the uncertainty information should indicate this.
- Trans identity represents a comparatively small proportion of the overall population. This means the accuracy of estimates is particularly susceptible to over-estimation if the question incorrectly identifies people as trans, as appears to be the case here. This problem of "false positives" should always be considered when collecting data on small populations.
- ONS's attempts to communicate uncertainty, while welcome, focus on relatively high-level messaging that provides only limited help to users, as it does not fully guide use of the data, particularly at smaller area levels.
- The research report notes that the statistics based on write-in responses may be more reliable than those not based on write-in responses, but there is more advice and guidance that could be given to different types of users to highlight this insight.
- The varying levels of information across the bulletin, quality information and research report have resulted in inconsistent messaging for users regarding the uncertainty of the estimates.

ONS's position

ONS's research report Quality of Census 2021 gender identity data published in November 2023 found patterns in the data that are consistent with some respondents not interpreting the question as intended.

In the research report, ONS concludes: *"To summarise our investigation into potential respondent error: there are patterns in the data that are consistent with some respondents not interpreting the question as we had intended. These patterns reflect those respondents with lower proficiency in English and characteristics likely associated with English language proficiency such as ethnicity and country of birth.*

However, for the reasons outlined in this report, we cannot say whether the census estimates are likely to be an overestimate or an underestimate of the true value, given other sources of uncertainty, not least the potential impact of question non-response. Therefore, the overall impact on the data of any misinterpretation of the question

cannot be determined. To assist users in interpreting the detailed breakdowns of the data, we have produced additional supporting tables."

Learning from new evidence in Scotland's Census, the Office for National Statistics (ONS) wrote to us on 5 September 2024 to request that the gender identity estimates from Census 2021 should no longer be accredited official statistics and should instead be classified as official statistics in development.

In its letter, ONS stated that it "*cannot quantify all the uncertainty surrounding responses to the Census 2021 question on gender identity, however the available evidence does show there is potential for bias in how the question was answered by those who responded that they did not speak English well. Changing the designation of the statistics helps to signal the limitations arising from this issue, while also better signalling the fact that our knowledge and understanding of the best approach to collecting data on gender identity continues to evolve."*

We welcome statistics producers approaching us when they have any concerns about the statistics they produce.

OSR's regulatory decision

This section sets out our regulatory decision, based on the timelines set out in the first section of the report, and the analytical concerns in the second section.

As noted in our interim report, a user need for data on gender identity was identified as part of ONS's public consultation on proposed topics for the 2021 Census questionnaire in England and Wales. In particular, the ability to capture data on the trans population was seen as a data gap for service provision that was important to address.

To help establish the extent to which the census gender identity data are being used, we asked all the stakeholders we talked to if they were aware of any use cases of the data. Some stakeholders stated that they believed the data were being used to support decision making, but in these cases, they were likely not the only source of evidence being drawn upon.

Several local authorities, such as Shropshire Council and West Yorkshire Combined Authority, have published dashboards, data tables or reports that summarise or map the trans population within their area. Similar local analysis has been undertaken by the House of Commons Library and some third-sector organisations, such as Trust for London, which produced an interactive map exploring the link between gender identity and deprivation. In the course of this review, we were not able to find examples of these census data being directly cited in a policy decision or action in support of resource allocation.

We acknowledge that ONS's request to label the census gender identity data as statistics in development supports its desire to publicly reflect the evolving understanding of gender identity. It also shows that ONS recognises it can do more to support users in interpreting the census gender identity data and understanding any limitations.

To determine our regulatory decision, we have considered whether the census gender identity data meet the standards set out in the Code of Practice for Statistics. In particular, we have looked at how well the ONS research report published in November 2023 addressed concerns about potential misunderstandings of the gender identity question and how such issues might have affected the quality of the gender identity data in the census. We have also formed a view on whether the publication of the research report has mitigated any remaining concerns by providing increased quality information on uncertainty or guidance on using the statistics. Through our review, we have considered three key areas of the Code:

Quality – **Suitable data sources.** As ONS has now recognised, the question developed to determine the size of the trans population of England and Wales did not work as intended. The evidence indicates that people may have found the question confusing and therefore gave a response that did not reflect their gender identity. This appears to be more likely for people who do not speak English as their first language. The different question adopted for the Scottish Census does not seem to have encountered the same difficulties.

It is important to note that neither ONS, nor we at OSR, identified in the development or delivery process of the census that there was a significant risk that certain groups of people were more likely than others to misunderstand the question.

Assured Quality. A core requirement of the Code of Practice is that statistics producers conduct suitable proportional quality assurance, and they are transparent about the quality assurance approach taken throughout the preparation of the statistics.

There are a number of areas where ONS could have better communicated with users of the census gender identity statistics. Our interim report found that ONS's communications were not sufficiently managing users' expectations on what areas ONS's research was exploring and what the research might show. Upon publishing the research, which was preannounced, delayed twice and cut from four outputs to one, ONS made no acknowledgment of any changes to outputs. This approach had the potential to damage trust and confidence in the research.

Additionally, until recently, ONS's approach has been characterised by a focus on defending the estimates of the size of the trans population. ONS has not shown sufficient willingness to engage constructively with the emergent evidence. The publication of the Scotland Census data on trans and trans history in June 2024 has been instrumental in a shift in ONS's approach.

We have confidence that those working on the data and research have been working in line with the Code, but consider that stronger leadership from ONS, particularly on clearer public messaging when concerns were raised and on setting expected outcomes for the work, was needed.

Communicating uncertainty. The Code requires that the extent and nature of any uncertainty in the estimates be clearly explained. ONS has conducted a range of research activities to explore the quality of the estimates, but these efforts have been hampered by a reluctance to communicate the limitations of the gender identity statistics. ONS's headline message in the research report around the validity of the census data, and how users should interpret and use the data at national and smaller area levels, is unclear.

Our interim report found that ONS should have communicated the inherent uncertainties in this new data collection to aid use when it published the data. We noted that had ONS done so, it would have aided users' understanding of the appropriate use of the data and provided relevant context when issues were identified by users. In addition, we stated that ONS could have been more open and transparent about these novel data and the work that has followed and that ONS's planned research must communicate how the data can and cannot be used. We have found that the research report does not fully meet this need and ONS has not done enough to communicate uncertainty in line with the expectation.

Users who use the ONS census outputs at a national level would struggle to locate any of this research to understand the impact of the research on the use of the gender identity data. This problem is replicated at the local level, where the research indicates that the variability of the estimate may be greater.

We have considered three further factors relating to the status of these estimates:

ONS's approach to criticism of the gender identity data. We noted in our interim report that the content and tone of the published communication from ONS has been somewhat closed and at times defensive. Some users have raised concerns that ONS has been captured by interest groups, leading to a lack of objectivity. We found no

evidence of this form of bias through any of our work. We consider it regrettable that ONS's defensiveness has created an impression of bias to some external observers.

Whether the statistics are materially misleading. The purpose of the Code's Quality pillar is to ensure that statistics are a fair representation of what they seek to measure and are not materially misleading.

ONS confirmed in its letter to us that it is unable to quantify all the uncertainty surrounding responses to the Census 2021 question on gender identity, and that the available evidence does show there is potential for bias in how the question was answered by those who responded that they did not speak English well. The evidence we have considered in this review leads us to the conclusion that considering the understanding of the question alone, this issue is highly likely to have led to an overestimate.

However, other sources of potential bias (for example non-response) in what was a new data collection are not fully known. The ONS research report is correct to say that it "*cannot say with certainty whether the census estimates are more likely to be an overestimate or an underestimate of the total number of trans people aged over 16 years in England and Wales."*

At the national level, triangulation with other sources, including data from the Scottish Census, suggest that an estimate of around "1 in 200" triangulates with other sources and is not likely to be materially misleading. While the information on the size and nature of all the potential biases is incomplete, it is hard to draw the same conclusion for some more-detailed breakdowns, including for local areas where the data indicate a higher concentration of people misunderstanding the question.

Our regulatory approach

It would be reasonable for users to expect that a new question with few data available for triangulation would have extra information provided to support appropriate interpretation and use.

The Code of Practice places a strong emphasises on innovation and improvement. It recognises that the world does not remain static, and that change is an inevitable part of life. We encourage statistics producers to identify opportunities to use new sources and methods and to recognise any fresh needs for information among users of statistics.

When developing new statistics or introducing a new data source or methodology, initial publications are often labelled 'official statistics in development' by producers. We would usually assess the official statistics against the Code of Practice once the development is complete.

However, the census is a unique case. Our regulatory accreditation process was undertaken in tandem with the development and first publication of the statistics in the respective countries of the UK. While this process allows us to provide users of these statistics with the reassurance of quality that comes with accreditation at first publication, it means that if we have any subsequent concerns about any aspect of the Code of Practice for Statistics, we may need to reconsider our accreditation decision in light of new evidence.

In this case, new evidence has emerged that has demonstrated that the statistics do not comply with important quality aspects of the Code of Practice for Statistics. With the benefit of hindsight and this new evidence, our conclusion is that the census gender identity statistics, which were collected from a novel question in an area where measurement practices are still emerging, would be better described as official statistics in development, and should not be labelled as accredited official statistics.

We acknowledge that this is a limitation with our approach, and as a result of this review, we are reviewing what changes might be needed to our assessment approach in circumstances where accreditation at first publication is in line with serving the public good.

Our decision

The correlational evidence set out in this report supports the view that some groups of people were more likely than others to misunderstand the question. This view is also supported by ONS's finding that there are patterns in the data that are consistent with some respondents not interpreting the question as had been intended.

Though the additional data tables published alongside the Quality of Census 2021 gender identity data report are useful for understanding the nature of responses and the their relationships with other characteristics, we consider that the research report and data tables have not fully addressed the concerns raised by users.

Given our findings, we have written to ONS to support its proposal to publish the Census 2021 gender identity statistics in England and Wales as official statistics in development rather than accredited official statistics. We found that the issues highlighted in this report are unique to the statistics on gender identity and therefore all other outputs from the Census 2021 in England and Wales are unaffected by this decision and remain designated as accredited official statistics (called National Statistics in the Statistics and Registration Service Act 2007).

Required future work by ONS and the GSS

There is great value in ONS having undertaken this work on developing how people are able to describe society. We recognise the challenges in collecting data at scale on a sensitive topic for the first time.

Our recommendations focus on the steps ONS must now take to help users of the census gender identity statistics understand their strengths and limitations and the development work we consider is required on the GSS gender identity harmonised standard.

Guidance for use of the census gender identity statistics

- It is important that ONS communicate clearly across all its census outputs what official statistics in development means in practice for the census gender identity statistics.
- In doing this, the varying levels of information across the gender identity publications and data tables must be addressed by ONS across its suite of publications and data on gender identity.
- ONS should provide more practical information, according to the needs of stakeholders, to help users understand the uncertainty inherent in the data and how this affects their use and interpretation. This information could include sensitivity analysis, and example use cases published at different levels of geography and population highlighting any likely anomalies or implausibilities.
- ONS should consult with us as it develops these communications, which should be treated as a priority workstream.

Gender identity question development

These data were identified as being important for society, and it is important that the original user need for robust data on gender identity be addressed for future statistics. ONS should consider how these data needs are best addressed in future as part of its further testing of the question.

We consider that there is sufficient correlational evidence that the gender identity question in the census has underperformed, and some users, producers and stakeholders have lost confidence in the suitability of the gender identity question.

This issue is important as the census gender identity question is also the question used in the GSS gender identity data harmonised standard. As we set out in a letter to the Deputy National Statistician with oversight of the Government Statistics Service (GSS) in February 2024, we are seeing in our regulatory work the implications of the lack of confidence in the harmonised question by producers, who are unsure what gender identity question they should use, and who feel unsupported by the Government Statistical Service. The GSS Harmonisation Team must consider these recommendations as it further develops its workplan.

- We can confirm that our interim recommendation remains: namely that, as part of the ongoing development of the GSS harmonised standard for gender identity, ONS should carry out further testing of the question. ONS must give consideration to developing and testing a harmonised question that commands

the confidence of producers and users. The merits of the Scottish Census question should also be considered in this process.

- This work on the question should be considered as part of the outcome for the Future of Population and Migration Statistics consultation, in determining the future production of these statistics in an administrative data-based framework.
- Given that the census gender identity question is also the GSS harmonised standard, ONS should prioritise this work, ensuring that adequate resources are made available.
- We consider that there is an immediate need for the GSS to publish advice for statistics producers who are currently using the gender identity question and for those who are thinking about adopting the standard. This guidance should set out the limitations of the question, reference the Scottish Census approach and caution against knee-jerk changes to established collections until ONS has undertaken further research.
- We recommend that stakeholders be fully involved and engaged in future work, to help build and restore confidence.

The GSS should publish its initial plans for developing the gender identity harmonised standard within three months of our report's publication.

Annex A – Census publication timeline

6 January 2023:
ONS published the first outputs: Gender identity, England and Wales: Census 2021.

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.

14 April 2023:
ONS released a statement setting out its intention to publish further research on the topic within two months.

8 June 2023:
ONS preannounced four articles due to be published on 19 June 2023. These articles would cover question testing, collection and processing, microanalysis and macroanalysis.

19 June 2023:
ONS published an article, 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.

30 August 2023:
ONS cancelled the release of the remaining research articles and published a statement saying it was conducting ongoing research and undertaking wide user engagement, and would publish further updates in due course.

9 October 2023:
OSR published the interim review, which sets out OSR's expectations for ONS in line with the Code of Practice for Statistics on what it considered was needed to provide assurance on the gender identity census data and the gender identity harmonised question.

1 November 2023:
The last revision of Census 2021 Sexual orientation and gender identity quality information for Census 2021, which sets out the known quality information affecting sexual orientation and gender identity data from Census 2021 in England and Wales to help users correctly interpret the statistics.

8 November 2023:
ONS's research report, Quality of Census 2021 gender identity data, was published. Alongside the report, ONS released an additional dataset.

27 June 2024:
The Scotland Census 2022 data for sexual orientation and trans status or history were published by National Records of Scotland (NRS). As well as the statistical publication, NRS published a separate quality assurance document for the variables included in the statistical release.