



Sexual Orientation and the Census

January 2024

Tamara Burrows

Sexual Orientation and the Census

An analysis of the problems with the sexual orientation question in the England and Wales 2021 Census

We believe that the Office for National Statistics (ONS) made mistakes in the way it asked the sexual orientation question, in how it analysed the data and in how it presented its findings, especially about lesbians and gay men. In this paper, we examine the ways in which the census does not provide a reliable or accurate measure of sexual orientation and suggest that the ONS failed to follow clear Equality Act 2010 definitions of sexual orientation, allowing blurred boundaries between gender identity, sex, and sexual orientation. We focus on lesbians in particular, although most findings also hold for gay men.

The census sexual orientation question

In 2021, for the first time, the England and Wales census included a question on sexual orientation. The question was optional ('This question is voluntary') and was only for those aged 16 or over.

It asked:¹

Which of the following best describes your sexual orientation?

- *Straight/ heterosexual*
- *Gay or lesbian*
- *Bisexual*

- *Other sexual orientation, write in (with space provided for an answer).*

In response to this question, around three quarters of a million people (748,000 or 1.5% of the total) described themselves as 'gay or lesbian,' 624,000 (1.3%) as 'bisexual', and 165,000 (0.3%) ticked 'other sexual orientation.' A further 3.6 million people (7.5%)—did not answer the question.²

Why do we need data on lesbians, gay men, and bisexuals?

Government bodies collect data about sexual orientation because an understanding of the differences between same-sex and opposite-sex attracted people is needed to develop targeted policies, legislation and public services. In the UK, same-sex marriage is now legal, and public attitudes are improving—but stigma, discrimination and adverse physical and mental health outcomes remain for lesbian, gay, and bisexual (LGB) people. Because these are not evenly distributed, we also need to understand differences *among* LGB people, who are not a homogenous group.

For example, while gay male sex is outlawed in a larger number of countries than is sex between women, lesbians in particular³ experience corrective rape and other forms of sexual violence because of their

¹Duke-Williams, O. *Census Explainer: Sexual orientation and gender identity questions*. <https://ukdataservice.ac.uk/learning-hub/census/census-explainers/census-explainer-sexual-orientation-and-gender-identity-questions/>

² Office for National Statistics (2023). *Sexual orientation, England and Wales: Census 2021*. Statistical Bulletin (retrieved Nov. 1st, 2023)

³ To the extent bisexual women engage in same-sex behaviour this is also true for them

same-sex behaviour.⁴ Gay men are at risk for particular kinds of sexual health problems, while lesbians have some of the lowest rates of sexually transmitted diseases.⁵ When it comes to pay disparities, lesbians (at the individual level) on average earn less than gay men but more than heterosexual women⁶, whereas lesbian *couples* earn less than both gay male and opposite-sex couples.⁷ Gay and lesbian couples who wish to become parents experience different challenges. For lesbians, one partner can carry a baby, but they must use donor sperm. Male same-sex partners need both an egg donor, and a gestational surrogate to carry the pregnancy.

In research, when gay men and lesbians are grouped together (as they often are), or when lesbians are grouped together with bisexual women or transwomen, these complexities are obscured. In fact, perhaps the only truly significant characteristic shared by lesbians, gay men and bisexuals is their minority sexual orientation status. Likewise, bisexual women and lesbians may share few characteristics other than their status as women.

The census

As the only survey that gives a full picture of the population, the census plays a vital role in our understanding of different groups in society, especially those smaller groups for whom it is difficult to obtain representative samples. The 2021 England and Wales census was an opportunity to gather a fuller picture of lesbians, gay men, and bisexual people, and the differences among them.

⁴ Human Dignity Trust (May, 2016). *Breaking the silence: Criminalisation of lesbians and bisexual women and its impacts*.

⁵ Tao, G. (2008). *Sexual orientation and related viral sexually transmitted disease rates among US women aged 15–44 years*. American Journal of Public Health, 98 (6). 1007–1009.

Without basic and accurate descriptive statistics about the differences between these groups, there is no baseline against which government policies, legislation and public services can be targeted, assessed or evaluated. A basic understanding of the differences between the housing arrangements of lesbians, bisexual men and women, and gay men could result in policies and services at the local level that are sensitive to the need of the different constituencies. These policies and services (and take-up of them) could then be measured against information provided within the census.

This paper describes three main methodological areas where we believe the ONS made mistakes in its handling of sexual orientation in the census. First, we discuss the reasons why lesbians have never been treated as a separate category. Second, we describe problems stemming from ONS use of an ‘other’ category. Third, we explain why an overarching sexual minorities (LGB+) category is not an accurate way to measure the LGB category.

1. Lesbians are not treated as a separate category

Given the importance of understanding the differences between LGB categories, it is surprising that during the England and Wales 2021 census exercise, neither lesbians nor gay men were at any stage treated as a discrete group—either in the design of the questionnaire, the sorting

⁶ Drydakis, N. (June 2021). *Sexual orientation and earnings: A meta-analysis 2012–2020*, Discussion paper series/ IZA Institute of Labor Economics.

⁷ Lee Badgett, M.V. & Schreebaum, A. *The impact of wage equality on sexual orientation poverty gaps* (June 2015). The Williams Institute.

of data into categories for analysis, the data analysis, or the reporting of findings.

In the glossary accompanying their reports,⁸ the ONS did not define 'lesbian' (or, for that matter, any sexual orientation category):

We have not provided glossary entries for individual sexual orientation categories. This is because individual respondents may have differing perspectives on the exact meaning.

The closest the census data comes to estimating the number of lesbians in England and Wales is the "percentage of females who identify as gay or lesbian". Buried in an Excel data table, we can see that 288,265 females (1.15%) over the age of 16 in England and Wales "identify as gay or lesbian".⁹ However, because of measurement errors made by the ONS, this figure does not actually represent the number of lesbians in England and Wales.

How the ONS measured the number of lesbians

One reason for not reporting separate findings for lesbians is that lesbians and gay men were included on a single response line in the questionnaire ("Which of the following best describes your sexual orientation?" 'Straight/heterosexual' / 'Gay or lesbian' / 'Bisexual' / 'Other sexual orientation, write in') but were never subsequently separated.

The background to this is that when the Equality and Human Rights Commission (EHRC) carried out early work to develop a standardised approach to measuring the LGB population of Great Britain, they drew on

the ONS Sexual Identity Project focus group finding that some lesbians prefer to describe themselves as gay:

The term 'gay or lesbian' is satisfactory. As all surveys tend to collect information on gender, this wording [collapsing the categories gay and lesbian] is sensitive to those women who prefer the term 'gay' as a self-descriptor to 'lesbian'¹⁰

In other words, it was decided that because 'gender' was always asked about in addition to sexual orientation, lesbians and gay men could be included on a single response line (presumably for reasons of efficiency), with an assumption that they would be disaggregated at the first stage of the analysis. This form of the question became the national harmonised standard with the requirement that all social surveys ask the sexual orientation question in this way.¹¹

As a first step in data analysis, we believe the ONS should have arrived at a 'lesbian' variable (i.e., category for data analysis) by cross-referencing those who responded 'female' to the sex question and 'gay or lesbian' to the sexual orientation question. Separate statistics for gay men and lesbians should then have been reported, including overall figures and socio-demographic cross breaks (as it did for other groups, including those who identified as heterosexual, bisexual, pansexual, asexual and queer). This did not happen, and neither did the ONS cross reference those who responded 'gay or lesbian' with responses to the *gender identity* question. In the end, the ONS did not create or report on a lesbian (or gay male) category at all.

⁸Office for National Statistics (January 2023). *Sexual orientation, England and Wales: Census 2021*. Statistical Bulletin (with Nov. 1st correction), page 6.

⁹ Office for National Statistics (ONS, November 2023) *Sexual orientation (9 categories) by sex, England and Wales: Census 2021 (Dataset)*.

¹⁰ Aspinall, P.J. (2009). *Estimating the size and composition of the lesbian, gay, and bisexual population in Britain*. Equality and Human Rights Commission (EHRC) Research Report 37. Page 44.

¹¹ GSS Harmonisation Team (April, 2019). *Sexual orientation harmonised standard*. <https://analysisfunction.civilservice.gov.uk/policy-store/sexual-orientation/>

If we further investigate the way the ONS constructed its variables, we can begin to see why this might be the case. Looking more closely at who would have been included in this variable representing lesbians, had it been constructed (i.e., females who identify as 'gay or lesbian'), we see that there are a number of errors of classification. In short, the lesbian category would have included both male and/or opposite-sex-attracted people, while excluding some same-sex-attracted females. This perhaps gives the first clue as to why the ONS did not provide separate statistics for lesbians and gay men. The next section describes these problems in more detail.

Problems with how the ONS measured the numbers of lesbians

When they describe "females who identify as gay or lesbian", who exactly is the ONS talking about? This is a question that must be asked, since in addition to those same-sex-attracted females accurately reporting their sex as female, it seems that some male people also fell into this category.

Disaggregation of ONS figures for the gender identity question in the census reveals that of those classified as transwomen (i.e., born male and identifying as women), only 34% had entered 'male' as their response to the sex question.¹² Notwithstanding the possibility that some of these respondents were individuals who misunderstood the gender identity question altogether,¹³ those transwomen who responded to the sex question as 'female' (i.e., according to their gender identity rather than their sex) and who identify as lesbians would have met the criteria for

inclusion, erroneously, in the lesbian category. It is likely that a large percentage of transwomen do indeed identify as lesbian—the 2017 LGBT survey found that of the approximately 3,000 transwomen who took part, only 16% identified as heterosexual.¹⁴ The reliability of the ONS figures is thus compromised, if the definition of lesbian is females who identify as gay or lesbian.

Arguably, where sexual orientation is defined as same-sex-attracted, gender identity should not be used to measure sexual orientation. Defining sexual orientation using gender identity makes it impossible to differentiate same-sex-attracted females from the broader category of gender-diverse male people who identify as women and as lesbians, at both at questionnaire stage and the analysis stage. Regardless of how they identify, it should be possible to differentiate between same-sex-attracted females and transwomen who are attracted to women. These two groups have distinct socio-demographic characteristics, policy and service needs.

Some exclusively same-sex-attracted females would also have been left out of this category; those transmen (females identifying as male) who entered 'male' in response to the sex question and are exclusively same-sex-attracted, but identify as heterosexual, would not have been included.

Looking carefully at those who entered 'female' in response to the sex question and 'gay or lesbian' in response to the sexual orientation question, the potential for further misclassifications becomes evident. For those trans individuals who followed ONS guidance by providing accurate sex information, it would appear that some females identifying

¹² Biggs, M. (2023, V 4.1, 18 April). *Gender Identity in the 2021 Census of England and Wales: What Went Wrong?* <https://osf.io/preprints/socarxiv/yw45p>

¹³ Ibid.

¹⁴ Government Equalities Office (GEO, July 2018). *National LGBT Survey research report, p.15 and Annex 3, Q7-9, Characteristics*. This survey was for 'people who self-defined as lesbian, gay, bisexual or as having another minority sexual orientation or gender identity, or as intersex,' p. 5.

as gay men would have been counted as *lesbians*¹⁵— that is, those transmen who accurately report their sex as ‘female’ but are attracted to men and call themselves gay. Likewise, for those who gave their sex accurately, some males who identify as lesbian would be included in numbers of *gay men*. Regardless of how ‘sexual orientation’ and ‘sex’ are defined, these represent clear errors.

While the number of misclassifications may be small (given the small size of the trans population) the effects on research findings will still be significant, because the lesbian population is also small.

Why did the ONS not measure or report findings separately for lesbians?

It seems possible that part of the reason for the ONS neglecting to categorise lesbians and gay men separately for analysis purposes was to avoid having to address these sources of error. We assume the deeper reason the ONS was at pains to avoid labelling ‘females who identify as gay or lesbian’ as ‘lesbians’ has to do with its unwillingness to offend those who define sexual orientation according to gender identity, rather than sex.

When the sexual orientation question was first developed by the ONS in the 2000s, it was not contentious to define sexual orientation as being same-sex-attracted (rather than same *gender identity*-attracted). But as definitions of sex became muddled with those of gender identity, this became an issue the ONS should have addressed. Despite all the (publicly-funded) testing and quality assurance exercises the ONS conducted in preparation for the census, its formulations seem only to

have further clouded the conceptual foundation of the sexual orientation question.

2. Including an ‘other, please write in’ category

A second problem with the sexual orientation question has to do with the inclusion of an ‘other, please write in’ category. The concepts measured in surveys, as well as the survey questions themselves, are supposed to be clear and precise. How well a questionnaire measures the concept it was designed to evaluate is known as construct validity—a fundamental principle of good questionnaire design.

The ONS census glossary¹⁶ defined sexual orientation as follows:

Sexual orientation is an umbrella term covering sexual identity, attraction, and behaviour. For an individual respondent, these may not be the same. For example, someone in an opposite-sex relationship may also experience same-sex attraction, and vice versa. This means the statistics should be interpreted purely as showing how people responded to the question, rather than being about whom they are attracted to or their actual relationships.

The problem with using identity as a measure of sexual orientation

As we have seen, the ONS treated sexual orientation as an identity category. This is not unexpected because social surveys—which have limited space and so cannot ask multiple questions to arrive at a composite measure—tend to ask about identity, rather than other aspects of sexual orientation. Social scientists agree that there are three

¹⁵ Where ‘females who identified as gay or lesbian’ is the measure used for lesbian.

¹⁶ Office for National Statistics (ONS, November 2023) *Sexual orientation (9 categories) by sex, England and Wales: Census 2021 (Dataset)*. Page 3.

dimensions to sexual orientation: identity (how one identifies and labels their sexual orientation), behaviour (who one has sex with), and attraction (who one is sexually attracted to), and that these three dimensions do not necessarily align within an individual.

Asking about identity has been considered the least intrusive way of measuring sexual orientation, and the social aspect of identity has long been viewed as the best correlate with experiences of discrimination. Until recently, this method has likely yielded responses that correlate reasonably well with other dimensions of actual sexual orientation (i.e., who one has sex with, and who one is attracted to). This is why the measurement of identity is currently considered best practice.

Of course, if identity is the best practice measure for sexual orientation, an 'other, please write in' category is a sensible inclusion in a sexual orientation question, since identities are subjective. The ONS tested the question with focus groups in 2016, and added the 'please write in' component following feedback from those present. Allowing respondents "to report their sexual orientation rather than being classified as 'other' was well-received by participants", and this change was "generally supported by stakeholders at the Sexual Identity Workshop."¹⁷

While using identity as a measure for sexual orientation and including an 'other, please write in' category may, in the past, have been an acceptably approximate measure of sexual orientation, in recent years—as identity categories have become more fluid and as heterosexual people (especially young people) increasingly identify into LGB categories (perceiving this to be a matter of individual identity rather than a

correlate of behaviour and attraction), this measure is increasingly unreliable and lacks the ability to distinguish between people who are, and are not, LGB (i.e., same-sex-attracted). Census designers should have predicted and addressed this issue. Instead, they further confused the categories.

Pansexual, asexual, and queer identities

When census analysts coded the 'other, please write in' responses, they found the three most common responses were 'pansexual', 'asexual', and 'queer'. Figure 1 shows that around 11% (of all who identified as non-heterosexual) identified as one of these groups: pansexual (3%), asexual (2%), and queer (1%). A further 5% ticked the 'other' box and wrote in either some other identity, or nothing at all.

The act of including an 'other' category presented certain personally-favoured identity categories (e.g. 'asexual', 'pansexual', 'queer') as valid alternatives to the three main orientation categories ('straight/heterosexual', 'gay or lesbian', 'bisexual'). As we have seen, sexual orientation categories concern existing patterns of sexual attraction and behaviour, and are not (just) about favoured identity.

The ONS has responsibility to ensure there is statistical information to support the Equality Act 2010¹⁸ which defines sexual orientation as: '*a person's sexual orientation towards persons of the same sex (that is, the person is a gay man or a lesbian), persons of the opposite sex (that is, the*

¹⁷ Office for National Statistics. *Sexual orientation question development for Census 2021: How we researched, developed and tested the Census 2021 question on sexual orientation*.

¹⁸ Hasledon, L. & Joloza, T. (April, 2009). *Measuring sexual identity: A guide for researchers*. Household, Labour Market and Social Wellbeing Division, ONS. Page 3.

person is heterosexual), or persons of either sex (that is, the person is bisexual)¹⁹.

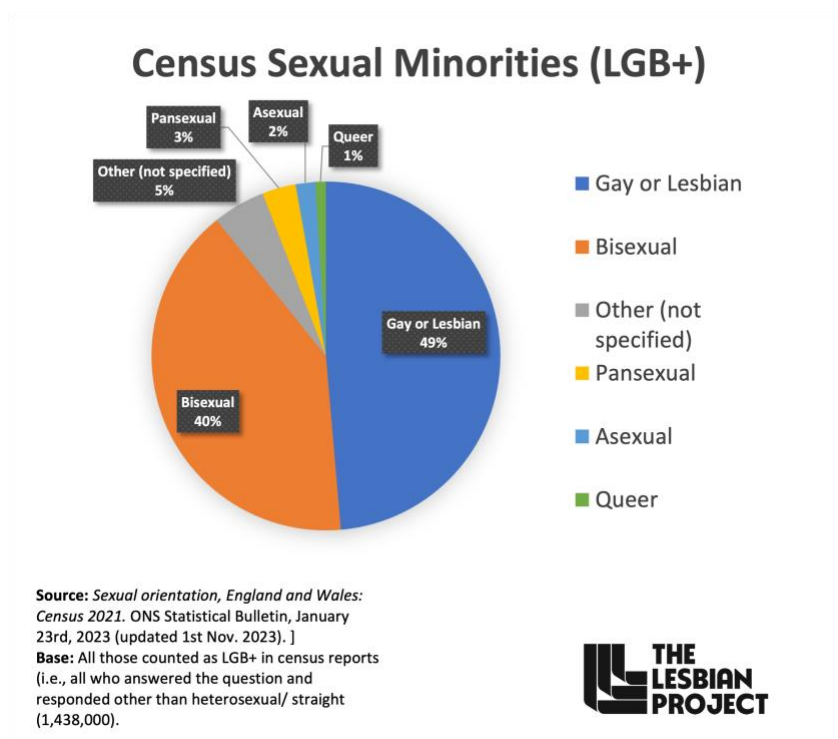


Figure 1. 'Sexual minorities' category in the England and Wales 2021 Census.

Arguably, everyone fits into one of these three categories (gay or lesbian, bisexual, heterosexual), irrespective of their individual favoured identity, since everyone's main pattern of sexual behaviour and attraction will fit into one of the three.

In fact, the ONS need not have included the 'other' category at all, since the phrasing: 'which of these *best describes* your sexual orientation' (our italics) is sensitive to the fact that one option may not perfectly fit the sexual orientation of a person across all dimensions. This both acknowledges that there will always be some subjectivity involved in reporting one's sexual orientation and allows the respondent some agency and choice. Yet, departing from the Equality Act 2010 definitions and including an 'other' category effectively served to encourage some respondents to exclude themselves from the main sexual orientation categories—resulting in data loss.

For instance, 'asexual' tends to be loosely defined by LGBT campaigning groups, and does not mean the total absence of sexual interest. Anyone who (genuinely) falls into one of the first three main orientation categories might also count themselves as asexual.²⁰ Asexuality is not counted as a sexual orientation according to the Equality Act 2010. 'Pansexual' (usually defined as either 'attracted to all genders' or 'attracted to people regardless of their gender') is also not included in the Equality Act 2010. It overlaps with the bisexual category; some bisexuals will also count themselves as 'pansexual'. Meanwhile, 'queer' is extremely loosely defined and has no standard definition. Many gays, lesbians, bisexuals, and heterosexual people call themselves 'queer.'

The design of the census to include an 'other' category (and then to count all those who mark it as members of a single group that rivals the three main sexual orientations) reflects a lack of clarity in the foundational

¹⁹ UK Equality Act 2010, Section 12.

²⁰ Benoit, Y. (2023). *Ace in the UK*, Stonewall.

concept and definition of sexual orientation. It is difficult to avoid concluding that in designing this part of the census, the ONS did not have a good sense of what it was they were measuring.

In its reports, the ONS has compounded the original error, going on to include 'pansexual', 'asexual', and 'queer' as three primary data categories for analysis, alongside 'gay or lesbian', 'heterosexual', and 'bisexual.' While the ONS provided basic demographic breakdowns for the 'other' category, it has failed to do so for the category of 'gay or lesbian'. This makes the statistical findings about lesbians and gay men difficult to understand.

As an example, take the following section from an ONS statistical bulletin.²¹ We can straightforwardly understand the age patterns of pansexual, asexual, and queer categories, as seen below:

Those aged 16 to 24 years were also the most likely [compared to other age groups] to have identified as:

- *pansexual (0.49% or 31,000)*
- *asexual (0.18% or 11,000)*
- *queer (0.10% or 6,000)*

But when it comes to information about lesbians, there is added confusion, because they are always grouped together with (at least) gay men:

People aged 16 to 24 years were twice as likely [compared to other age groups] to have identified as bisexual than as gay or lesbian (4.00% compared with 2.11%).

3. Sexual Minorities (LGB+) category

Census questionnaire designers would have understood that the inclusion of an 'other' category both increases non-response and inflates the amount of unusable data. In developing the census question, ONS research found that those who ticked the 'other' box included "those who were heterosexual/ straight but did not understand the question", people who consider themselves 'curious,' people who object to the concept of labelling in principle, others who wrote in their marital status, and various gender identities.²²

Figure 1 shows that a large proportion (of non-heterosexuals) ticked the 'other' box—around 11%, around half of whom wrote in 'pansexual', 'asexual', or 'queer'. Of the other half of those who ticked the 'other' box, some left the space blank, while others responded saying they had 'some other identity'. It is at least theoretically possible that same-sex-attracted females who identify themselves using the word 'homosexual,' for instance, would simply have noted 'some other identity'. It is unclear how many of these 'some other' identities would map onto the existing lesbian, gay and bisexual options, as well as how many include heterosexual sexual orientations—or indeed other kinds of responses that have nothing to do with sexual orientation.

In their analysis, rather than re-coding 'other' responses back into the primary question categories when appropriate, or removing data where the response did not map onto any primary category, the ONS simply added this data to its count of LGB people. Instead of relying on the relevant sexual orientation categories set out in the 2010 Equality Act

²¹ Office for National Statistics (ONS), (25th January 2023) *Sexual orientation: age and sex, England and Wales: Census 2021*.

²² Office for National Statistics (ONS). *Sexual orientation question development for Census 2021: How we researched, developed and tested the Census 2021 question on*

sexual orientation.

<https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/sexualorientationquestiondevelopmentforcensus2021>

(lesbian, gay, or bisexual), the ONS simply grouped together all of the following people: those who ticked 'gay or lesbian,' or 'bisexual'; those who wrote in 'pansexual', 'asexual', or 'queer'; and all those who ticked the 'other' box and wrote in some other identity (or indeed wrote in no identity at all) into a single overarching category called 'sexual minorities' ('LGB+' for short). The ONS then used this new, much wider and all-inclusive 'LGB+' group as its primary aggregated reporting variable²³ when describing overall sexual orientation findings.

Why did the ONS choose to use LGB+ rather than LGB as its main sexual orientation variable? It appears to have recognised that same-sex-attracted people were identifying out of traditional measurement categories, and over-aggregating the categories was their way of rectifying the situation.

Conclusion

The ONS made a number of errors in its measurement of sexual orientation in the 2021 England and Wales census. Its conceptual basis for the lesbian category was improperly formulated, and its findings were inaccurate and confusing. There were problems in the design of the question itself, in the way the variables for analysis were constructed, in the way the data was analysed, and in the way the statistics were reported.

At no stage in the process were lesbians (or gay men) treated as a separate category. Moreover, some heterosexuals were included in LGB categories—and some same-sex-attracted people were recorded as heterosexual. The primary aggregated sexual orientation variable used

by the ONS was 'sexual minorities' or LGB+ (which included heterosexuals), rather than LGB.

The ONS conception of sexual orientation was at best unclear, and the terms 'lesbian', 'gay', and 'bisexual' were left undefined. The need for absolute clarity over category definitions is fundamental to large scale survey methodology, yet as we have shown the ONS was constructed using extremely muddy definitions of sexual orientation. Lesbians were included only in over-aggregated findings in reports; statistics concerning lesbians were either merged with those concerning gay men, or subsumed into an overarching 'sexual minorities' group ('LGB+'). The ONS did, however, analyse and report findings for those identifying as heterosexual, bisexual, pansexual, asexual, and queer. Socio-demographic descriptions were available for other groups—but not for lesbians or gay men.

For lesbians, who have historically experienced inequality and discrimination based on their sexual orientation, their sex, and the interaction between the two, this erasure is not a novel phenomenon. We know from previous research that lesbians are routinely left out of funding provision, from data collection, and from service provision, usually as a result of being lumped together with other groups.²⁴

The government has long acknowledged its need to better disaggregate public data, and similar concerns have been raised about the disaggregation of ethnicity data. This paper shows that, for lesbians, it is more than a mere disaggregation problem. Lesbians (and gay men) were overlooked in the 2021 England and Wales census at every stage.

²³ Office for National Statistics (ONS, September 2023). *Sexual orientation classification. Census 2021: Use of these groups of categories to research and analyse census 2021 sexual orientation data.*

²⁴ The Lesbian Project (February, 2023). [Lesbian Erasure in the UK](#).