Title: Mainly Heterosexual, Other, and Straight: the measurement of sexual minority status and its impact on sample size, demographic distribution and health outcomes

Authors: Evangeline Tabor [1], Dylan Kneale[1], Praveetha Patalay[1,2]

Corresponding author: Tabor, Evangeline, evangeline.tabor.18@ucl.ac.uk

[1]Social Research Institute, University College London, London

[2] MRC Unit for Lifelong Health and Ageing, University College London, London

Abstract

Background: Sexual orientation has been measured in a wide variety of ways which reflect both theoretical and practical considerations. However, choice of sexual orientation measure and recoding strategy can impact sample size, as well as demographic and health profiles, in analyses of sexual minority populations. We aimed to examine how choice of sexual orientation dimension and recoding strategy decisions impact estimates in the sexual minority population in two population-based studies in the UK.

Methods: We used data collected at age 17 (2018) in the UK Millennium Cohort Study and at wave six (2012-13) and eight (2017-18) of the English Longitudinal Study of Ageing. Descriptive statistics were used to examine the impact of choice of sexual orientation dimension (i.e. identity, attraction and experience) and data recoding strategy on achieved sample size and composition by selected demographic and health measures within and between datasets.

Results: Dimension choice and recoding decisions resulted in variation in sample size. For example, more respondents reported some same-sex sexual attraction than reported a non-heterosexual identity (MCS: 20.77% vs 8.97%, ELSA: 4.77% vs 1.04%). Demographic distributions varied, but not substantially by dimension choice or recoding strategy. Overall, in both datasets sexual minority respondents were more likely to be White and in the highest quintiles for income and education than heterosexual respondents. Health status did not vary substantially by dimension choice or recoding strategy, however sexual minority respondents reported worse health than their heterosexual peers.

Conclusions: This study explores a range of practical and theoretical considerations when analysing sexual minority respondents using survey data. We highlight the impact recoding

strategy decisions may have on sample size and demographic and health distributions in this understudied population. We also demonstrate the benefits of including multiple dimensions for capturing mechanisms of interest, elucidating ambiguous responses and exploring sexual diversity.

Introduction

Recent research using large observational datasets has shown sexual minority people in the UK have poorer mental and physical health outcomes than heterosexual people (1,2). This research has been facilitated by the growth in inclusion of relevant measures, exemplified by the recent introduction of routine collection of sexual identity by the NHS and the inclusion of a sexual identity question in the 2021 England and Wales Census (3,4).

However, the dimension of sexuality captured by the included measures and therefore used by researchers is varied. For example, studies have used respondent sexual identity, sexual attraction, and sexual behaviour (5), each collected through different questions and allowing for different response options. These different measures represent the result of a range of theoretical and practical considerations, including survey design, queer and sexuality theory, and public policy priorities (see Sell, 2007; Savin-Williams, 2009; Wolff *et al.*, 2017; Matsuno *et al.*, 2020 for more information), but have important consequences for research into inequities by sexual minority status.

Of particular interest here, health outcomes vary by which dimension of sexual orientation is selected for measurement, and focusing on only one dimension may miss important health disparities and limit the effectiveness of policy recommendations (5). Similarly, researchers are encouraged to use the dimension of sexual orientation which is most relevant to their research question, however this is constrained by which dimensions have been captured (9).

In the following sections we will examine four ways sexual orientation has been measured and analysed: sexual attraction, sexual behaviour, sexual identity, and through composite measures.

Attraction and Behaviour

Sexual attraction and behaviour have been used to identify sexual minority individuals who may not identify with the LGBTQ+ community (8).

For some researchers, attraction is the most important dimension of sexual orientation as it can be considered the primary source of other dimensions such as behaviour and identity (6,7). Some have argued it is also the most consistent dimension across time and context, while others have argued it may be stable over many years but fluid across a lifetime (7,10,11).

Likewise, measuring sexual behaviour has been an important tool for monitoring sexual health and risk (12). For example, terms in the public health literature such as ‘men who have sex with men’ (MSM) have been used to identify individuals who may be at risk due to specific behaviours and as a way of avoiding the social and cultural implications of identity labels (13).

However, measuring attraction and behaviour has several limitations. Firstly, while there is considerable overlap between reporting same-sex attraction and behaviour and sexual identity, those who report any same-sex attraction or behaviour is a much larger proportion of the population (12). Many of those who report same-sex attraction or behaviour do not report an LGB+ identity, and reported attraction and behaviour do not overlap neatly (12).

Secondly, the timescale on which the behaviour and attraction is measured has important implications for inferring sexual orientation, and varies widely between measures (5,12). For example, measuring lifetime prevalence of same-sex behaviour may reflect fluidity over time and experimentation during development, while recent behaviour may map more onto current sexual identity (5,12). Additionally, questions on sexual behaviour may not always be appropriate or have a low response rate (12).

Finally, what is meant by sexual attraction and behaviour are often left underdefined with variation in understanding of the term between individuals (5,7,14). Sexual behaviour and attraction measures are often predicated on the existence of two binary genders or sexes which may not match respondents own understanding (5).

Overall, these measures are useful for identifying those who do not align themselves with the LGBT community and therefore may not be reached by existing support systems. However, alone neither can tell researchers much about each other, or the identity and experiences of respondents.

Sexual identity

Sexual identity has been repeatedly recommended as a key measure for population health surveys (5,15,16). Sexual identity is not just a proxy marker for sexual attraction and behaviour, but also a powerful expression of self-identity, community, and political alignment (13,15).

Naming is a way of exercising power and control, and for marginalised groups the right to determine their own name is often hard won (13). Terms such as ‘gay’, ‘lesbian’, and ‘LGBT’ form part of the language of global political movements for rights and freedoms, as well as a global LGBTQ+ community, in ways that public health terms such as ‘men who have sex with men’ do not (13). Sexual identities can invoke a wider conversation about discrimination and disadvantage in a way that measures of attraction and behaviour may obscure (15).

However, as with attraction and behaviour, measuring sexual identity has a number of limitations and considerations.

Firstly, identity categories are historically and culturally specific (13). While terms like gay and bisexual have global salience in rights and equalities landscapes, there are many culturally, temporally, or culturally specific terms which may not map neatly onto these more dominant terms (13). As with any categorisation, important nuance and distinctions may be lost, and those who don’t consider themselves as sitting within such terms may be obscured (13).

Sexual identity intersects with other identities such as ethnicity, age, and gender in ways which both materially impact an individual’s experience of the world and their response to measures of sexual identity (5). For example, studies in the UK and the USA have found the prevalence of

those identifying as lesbian, gay or bisexual varies substantially by age, gender, and generation (12,17,18).

Secondly, as with behaviour and attraction, there has been some question around the stability across time of identity responses. For example, in Diamond's studies of sexual identity, nearly two-thirds of young women in the study changed their identity label at least once in a ten year period (7,19).

Lastly, many studies looking at outcomes in sexual minority populations group lesbian, gay, and bisexual respondents. However, this can obscure meaningful differences between the different identity groups and those who reject labels entirely (5,20). Not only does this practice obscure identities, but also potentially hides real differences in risk profiles and health outcomes (5).

Composite methods

As discussed above, there is often a false assumption of interchangeability or perfect alignment between the different dimensions of sexuality (5,13). For example, a lesbian identified woman is assumed to only experience attraction to and engage in sexual behaviours with women, and consider themselves part of a community of lesbian-identified women (5). However, sexuality can be better understood as a matrix which can be fluid, dynamic and multifaceted (5,13).

While the incorporation of more complex models of sexual orientation are theoretically grounded, translating these models into quantitative measures and empirical analysis is limited (5). There have been many attempts to incorporate multiple dimensions of sexuality or develop complex measures however they have not found widespread use or acceptance, not been properly validated or have been cumbersome to use empirically (5,6).

As a result of these issues, and pressure for single-question measures, researchers are largely restricted to relying on single dimension measures (6). As a result, it is recommended to select

the question which best suits the needs of the research question where additional measures are unavailable (6).

The current analysis

This analysis critically examines the implications of using different measures, and different ways of coding response patterns, using data from two population-based cohort studies from the UK which have captured multiple dimensions of sexuality. We aim to understand the impact of dimension choice and researcher coding decisions on sample size and demographic and health distributions. We offer some recommendations to researchers who either work with these studies or plan to use or collect sexual minority respondent data in the future.

Methods

Study Sample

The study incorporates data from two longitudinal population-based studies: the Millennium Cohort Study (MCS) and the English Longitudinal Study of Ageing (ELSA). These datasets were selected as they measured multiple dimensions of sexuality, collected information on both demographic distributions and health outcomes, and were sampled using probabilistic and representative methods. The two datasets also allow for the comparison of two populations representing distinct age groups and generations.

MCS is a large longitudinal birth cohort study of children born between September 2000 and January 2002 (21). There have been seven data collection sweeps at ages 9 months, 3, 5, 7, 11, 14, and 17 years to date (21). Children living in disadvantaged areas, and children from ethnic minority backgrounds were deliberately oversampled to allow analysis of these populations (21). Further detail is available elsewhere (21).

ELSA is a longitudinal panel study of people aged 50 and over in England (22). The original sample of respondents was selected from households who had previously responded to the

Health Surveys for England, a cross-sectional nationally representative household survey, in 1998, 1999 or 2001, with the sample replenished periodically thereafter (22) (23).

Variables

Sexuality is measured in the datasets in three main ways: sexual identity, same-sex attraction, and same-sex behaviour. Demographic variables examined to describe composition of samples include sex and gender identity, age, socio-economic status, ethnicity, and urbanity. Health outcome variables included are life satisfaction, psychological distress and self-rated general health. Each measure is further described in the supplementary methods section (pS5-6).

Sexual Identity

In ELSA, at wave 8 (2017-18) the following sexual identity question was used: "Which of the following options best describes how you think of yourself?". Options were "Heterosexual or straight", "Gay or Lesbian", "Bisexual", "Other", and "Prefer Not to Say".

In MCS at age 17 the following sexual identity question was used: "Which of the following options best describes how you currently think of yourself?". Options were "Completely heterosexual / straight", "Mainly heterosexual / straight", "Bisexual", "Mainly gay or lesbian", "Completely gay or lesbian", "Other", Don't Know", and "Prefer Not to Say".

Sexual attraction

In MCS at age 17 participants were asked to respond to the statement: "I have felt sexually attracted…". Response options included "Only to <opposite sex>, never to <same sex>", "More often to <opposite sex>, and at least once to <same sex>", "About equally often to <opposite sex> and to <same sex>", "More often to <same sex>, and at least once to a <opposite sex>", "Only ever to <same sex>, never to <opposite sex>", "I have never felt sexually attracted to anyone at all", "Don't Know", and "Prefer Not to Say".

In ELSA at wave 6 (2012-13), respondents were asked “Which statement best describes your sexual desires over your lifetime? Please include being interested in sex, fantasising about sex or wanting to have sex?”. Response options included “Entirely for <opposite sex>”, “Mostly for <opposite sex>, but some desires for <same sex>”, “Equally for <opposite sex> and <same sex>”, “Mostly for <same sex>, but some desires for <opposite sex>”, “Entirely for <same sex>”, and “No sexual desires in lifetime”.

Sexual behaviour

In ELSA at wave 6 (2012-13), respondents were asked “Finally, we would like to know a little about your lifetime sexual experiences and desires. Which statement best describes your sexual experiences over your lifetime? Please include all sexual experiences including sexual intercourse, fondling and petting.”. Response options included “Entirely with <opposite gender>”, “Mostly with <opposite gender>, but some experience with <same gender>”, “Equally with <opposite gender> and <same gender>”, “Mostly with <same gender>, but some experience with <opposite gender>”, “Entirely for <same gender>”, and “No sexual experiences in lifetime”.

No measure of sexual behaviour was collected among MCS respondents.

Analysis

Firstly, we generated descriptive statistics of each sexuality dimension in each dataset individually and their overlap.

Secondly, we generated recoding strategies for each dataset which are described in the supplementary section.

Finally, we generated descriptive statistics of demographic distributions and health outcomes for each sexuality dimension and recoding strategy in each dataset.

Results

This analysis includes cohort members who reported their sexual identity, attraction and/or behaviour in at least one wave. The study sample is 10,103 for MCS, and 7,130 for ELSA.

Descriptives for different dimensions of sexuality

The majority of participants in both adolescent and older adult respondents report their sexual identity as “Completely Heterosexual” or “Heterosexual”, and their sexual attraction and experience as exclusively with the opposite sex. For example, 94.69% of ELSA respondents identify as “Heterosexual”, 92.99% report attraction only to the opposite sex, and 94.76% report sexual experience with opposite sex. While MCS respondents are more likely to report non-Heterosexual identity and same-sex attraction than ELSA participants, it is still the minority of the sample (Table 1).

Dimension Overlap

In ELSA and MCS, respondents who identified as Heterosexual and Gay/Lesbian reported high coincidence between their reported identity and attraction/behaviour. For example, 92.99% of “Completely Heterosexual” respondents in MCS and 94.53% of “Heterosexual” respondents in ELSA reported only opposite-sex attraction (see Figure 1A & 1C). Similarly, in MCS, 75.63% of “Completely Gay or Lesbian” respondents report “Only to same sex” attraction, while in ELSA 66.67% of “Gay or Lesbian” respondents reported only same-sex attraction and 47.06% reported only same-sex experience.

However, we still observe differences between identity and reported attraction/behaviour, particularly among sexual minority respondents (see Figure 1B & 1D). For example, of those who reported any same-sex attraction, 74.04% also identified as “Heterosexual” in ELSA and 17.68% identified as “Completely Heterosexual” in MCS.

The collection of multiple dimensions also provides further context for respondents that might otherwise be dismissed as ambiguous. For example, ELSA respondents who identified as "Other" were more likely to report their sexual attraction (65.22%) and experience (70.83%) as entirely with the opposite sex, whereas in MCS the majority of "Other" respondents were evenly split between "Never sexually attracted" (40.76%) and reporting some same-sex attraction (39.49%).

Recoding strategies

We generated eight recoding strategies for ELSA and ten for MCS. A more detailed description of each recoding strategy can be found in the supplementary section (Tables S17-23)).

Sample size and demographic distribution

Identity

"Gay or Lesbian" and "Completely Gay or Lesbian" respondents in both MCS and ELSA were more likely to identify as male than the sample as a whole. While in MCS there are roughly equal proportions of male and female identified participants in the cohort, in ELSA the sample is 55.34% female (Figure 2).

In MCS, both "Completely Heterosexual" and "Completely Gay or Lesbian" respondents have worse SES status than "Mainly Heterosexual/Gay or Lesbian" and "Bisexual" respondents. However, in ELSA "Gay or Lesbian" respondents have better SES status than either "Heterosexual" or "Bisexual" respondents, with "Bisexual" respondents faring the worst (Figure 2).

Finally, in both ELSA and MCS, lesbian or gay participants were more likely to reside in London or an urban area than their heterosexual peers. However, while bisexual and "Mostly heterosexual/gay or lesbian" MCS respondents were similarly likely to reside in an urban area,

bisexual ELSA respondents were less likely to reside in London than heterosexual respondents (Figure 2).

Attraction

In both MCS and ELSA, those who reported exclusively opposite- or same-sex attraction were more likely to identify as male than those who reported non-exclusive attraction (Figure 3).

In MCS those who reported exclusive attraction had similar SES status, with those who reported non-exclusive attraction reporting higher SES status than the exclusive category respondents. However, in ELSA, those who reported exclusively same-sex attraction reported higher SES status than those who reported exclusively or mostly opposite sex attraction (Figure 3).

Finally, those who reported any same-sex attraction in ELSA were more likely to reside in London than those who reported exclusively opposite-sex attraction. Likewise, in MCS those who reported “More often to same” or “Only to same” attraction were more likely to reside in an urban area than those who reported mostly opposite-sex attraction (Figure 3).

Behaviour

In ELSA, respondents who report same-sex behaviour were more likely to be male, white, younger, slightly higher SES and London resident respondents than those who report opposite-sex behaviour (Figure 3).

“Mostly with opposite”, and “Mostly with same” respondents were younger, slightly higher SES than all other categories. Whereas “Equally” respondents were more likely to be female, older, lower SES and not reside in London than any other category where respondents reported any experience (Figure 3).

Identity vs Attraction vs Behaviour

The proportion of respondents who report any same-sex attraction or experience is larger than the proportion of the sample that report a sexual minority identity. Regardless of dimension

choice or cohort, the sexual minority sample is more likely to have higher SES and identify as White British than the heterosexual sample.

While consistent patterns are observed for sample size, SES and ethnic identity, gender identity and/or sex consistently varied by dimension, cohort and reported identity/attraction/experience. For example, in MCS while over 50% of respondents in other sexual minority identity categories identify as female, only 39.38% of “Completely Gay or Lesbian” participants identify as female.

Combined Recodings

Despite variation in sample size as a result of recoding, little substantial variation in demographic distributions was observed for MCS or ELSA (see supplementary).

Regardless of recoding, sexual minority respondents were in general more likely to be white and higher SES than heterosexual respondents in both ELSA and MCS. However, sexual minority respondents in ELSA were more likely to be male and more likely to live in London than heterosexual respondents, whereas in MCS no difference was observed in urbanity and sexual minority respondents were more likely to identify as female than heterosexual respondents (see supplementary).

Health and wellbeing

Identity

In both MCS and ELSA, respondents who reported a non-heterosexual identity reported worse outcomes in life satisfaction, self-reported health, and psychological distress outcomes (Figure 4). This trend held regardless of recoding, however, was more marked in the MCS sample.

Attraction

In MCS, respondents who reported same-sex attraction reported worse outcomes in life satisfaction, self-reported health, and psychological distress outcomes. In ELSA outcomes by

attraction were more variable, however recoded same-sex attraction categories consistently had slightly worse outcomes than opposite-sex attraction categories (Figure 5).

Behaviour

As with attraction, in ELSA outcomes by attraction were more variable, however recoded same-sex experience categories consistently had slightly worse outcomes than opposite-sex experience categories (Figure 5).

Identity vs Attraction vs Behaviour

Regardless of dimension choice or cohort, the sexual minority sample in general has worse outcomes on life satisfaction, self-rated general health, and psychological distress than the heterosexual sample. However, this trend is much more marked in the MCS sample, with outcomes by identity, attraction, and experience more variable in ELSA.

Finally, there is no substantial variation in outcomes between dimensions.

Combined Recodings

Recoding variables using all available dimensions resulted in variations in sample size, but little substantial variation in health and wellbeing outcomes were observed for MCS or ELSA by recoding strategies (see supplementary).

However, regardless of recoding, sexual minority respondents were in general more likely to have worse life satisfaction and self-rated health than heterosexual respondents in both ELSA and MCS. This pattern was much more pronounced in MCS, with sexual minority respondents in MCS also reporting worse psychological distress outcomes than heterosexual respondents.

Sensitivity analyses

A sensitivity analysis was also performed using sex-assigned-at-birth in MCS but found no difference in outcomes.

Discussion

In this analysis, we provide the example of two longitudinal studies from the UK which have captured multiple dimensions of sexuality. We aimed to demonstrate the impact of dimension choice and researcher coding decisions on sample size, demographic distributions and health outcomes. We offer some recommendations to researchers who work with population based studies or plan to use sexual minority respondent data in the future.

Firstly, we will discuss dimension choice and its impact on sample size, demographic distributions and health outcomes. Secondly, we will discuss patterns of response coincidence between dimensions. Finally, we will discuss the impact of coding decisions on sample size, demographic distributions and health outcomes.

Dimension Choice

Identity

Both ELSA and MCS include a measure of sexual identity, however the chosen measures reflect two different conceptual understandings of the dimension.

The identity measure in ELSA is similar to that present in other UK population surveys and is similar to the measure used in the 2021 Census in England and Wales (15,16). As a result, it is useful for cross-dataset analysis and comparison. The categories included are in common use and their salience to the general population has been verified (17). This measure is constructed of strictly nominal categories with identity labels presented as bounded and independent.

By comparison, the measure in MCS transforms sexual identity into an ordinal categorical variable ranging from “Completely Heterosexual” to “Completely Gay/Lesbian”. This shift, despite the similar categories, represents a fundamental change from considering sexual identity as discrete labels to a continuum or spectrum.

The inclusion of this measure was suggested on the grounds of a more nuanced approach to sexual identity and to allow those who might otherwise select “Other” to expand on their answer (private communication, see Box 3). However, it is unclear what is meant by the qualifiers ‘completely’ or ‘mainly’ in the measure and if understanding is consistent between researcher and respondent, especially as they relate to identities that are employed in real world contexts. In addition, the idea of sexuality as a spectrum with heterosexuality and homosexuality as opposite binary ends has been challenged by some researchers of sexuality (5).

Secondly, the positioning of “Bisexual” as a midpoint between “Completely Heterosexual” and “Completely Gay/Lesbian” potentially reinforces ideas of bisexuality as ‘half-straight/half-gay’, requiring equal attraction to men and women, and/or as a label which inherently supports binary understandings of gender - ideas which have been used by some to erase or discredit bisexuality as an identity (24). Additionally, while there has been increased interest in those that report themselves mostly or mainly heterosexual, it is unclear if equal research interest has been expressed in those identifying as “Mainly Gay or Lesbian” (See Box 2)(25). The measure is difficult to harmonise with other measures of sexual identity making it less useful for cross-dataset analysis and comparison. As demonstrated in this analysis, a binarized sexual minority recoding of the measure does allow for some comparability, however the construction of the binary does require careful consideration of the ordinal item structure.

Finally, the timing of collection of sexual identity in the two studies reflect researcher assumptions about the sexual identity of the participants in each study. For example, that the question was first included at age 17 perhaps reflects concerns around asking about sexual identity at an earlier age, but also the position of MCS as a research tool developed in a UK context which is willing to ask about sexuality and is mindful of an increasingly non-heterosexual identified youth population (26). By contrast, ELSA respondents were only asked about sexual identity in the eighth wave of the study (2017-18), two waves after the first questions on sexual

attraction and experience, with no plans to ask again in the near future (private communication). The timing of the inclusion of the question in ELSA reflects the ways in which sexual identity, in particular non-heterosexual identities, are deprioritised in research of this age group, popular understandings of sexuality as fixed by older age, and perhaps researcher concerns around non-response and question appropriateness for this age group. However, these concerns can result in the erasure and understudy of older sexual minorities and the sexuality of older adults in general.

Overall, MCS and ELSA demonstrate two approaches to the measurement of sexual identity and reflect the importance of careful consideration of how identity is conceptualised and how different populations may receive the measure.

Attraction and behaviour

The conceptualisation and acceptability of a measure is also an important consideration when measuring sexual attraction and behaviour. Unlike sexual identity, sexual attraction and experience measures do not require individuals to have taken on a specific social identity and may have a lower 'barrier to entry' than identity measures. As a result, measuring attraction and behaviour may capture individuals whose sexual identity is developing or flexible, or who may not identify as a member of the LGBTQ+ community.

While ELSA respondents have been asked about their sexuality at multiple waves, a question on the gender of the objects of their sexual attraction and experience was only asked once at wave 6 (2012-13). By contrast, MCS respondents have been asked about their sexual attraction at age 14 and again at age 17. Like sexual identity, the choice to repeat this measure or not reflects assumptions about the target population. For example, that sexuality will be changeable at younger ages but fixed at older, the acceptability of the question to different age groups, and the prioritisation of research of sexuality in older and younger cohorts.

A measure of sexual experience is only included in ELSA. While MCS collects information on romantic partners and questions designed to identify sexual behaviour which carries a risk of pregnancy or STIs, the survey does not include questions on the gender of partners. As well as prioritising a conception of sex centred on risk, the selection of measures limits the usefulness of the dataset for understanding the sexuality and sexual behaviour of the target population.

An important element of measures of sexual attraction and behaviour is the time period participants are asked to consider. Although addressing different ends of the lifecourse, MCS and ELSA both ask about lifetime or 'ever' attraction and experience rather than a specific time frame. While this framing potentially allows for the consideration of the general pattern of attraction and experience in an individual's life, it is subject to issues of inaccurate recollection and a social desirability/ normative bias in reporting for both cohorts. This means the measure may not effectively capture participant's 'normal' or routine attraction/experience patterns, nor attraction/experience patterns at a specific life stage. As a result, it is important to be clear in analysis about what responses to these measures can and cannot tell us.

As well as the time period participants are asked to consider, the presence of examples or explanatory text is an important element in a measure of sexual attraction/behaviour. Sexual attraction and behaviour experience are subjective and can incorporate a variety of meanings and forms (14). Where no further definition or explanation is given, as in MCS, this can open up inconsistency between researcher and participant understandings of what is being measured. In contrast, the measure in ELSA is phrased as follows: "Which statement best describes your sexual desires over your lifetime? Please include being interested in sex, fantasising about sex or wanting to have sex?" which supports a more consistent understanding between researcher and participant.

Likewise, both ELSA and MCS reinforce a gender binary in the items used to report attraction/experience by not providing other gender options for the objects of attraction (e.g.

“only to males, never to females”). ELSA has not yet measured gender identity so uses a self-reported binary sex variable, while MCS respondents at age 17 were able to complete a measure of gender identity immediately before the measure of sexual attraction. However, due to survey routing, a participant who had identified themselves as non-binary but who was assigned female at birth would receive the same question phrasing as someone who reported their gender identity as “Female”. While inconsistent with the MCS survey’s acknowledgement of diverse gender identities, this routing also hampers interpretation in the data release. In the data published by the survey, all attraction responses are recoded into one variable with the terms ‘male’ and ‘female’ replaced with ‘same-sex’ and ‘opposite-sex’. As a result, a non-binary identified participant reporting attraction “only to the opposite-sex” can only be interpreted by referring their sex-assigned-at-birth, which obscures both their gender identity and reinforces a binary they may not identify with.

Overall, MCS and ELSA demonstrate two approaches to the measurement of sexual attraction and behaviour and reflect the importance of careful consideration of how these dimensions are conceptualised and how different populations may receive the measure.

Overlap between dimensions

Any single measure of sexual identity with finite items will miss smaller identity labels, may not accommodate those with more than one label, and overlook the temporal and social context of identification and label use. However, the inclusion of multiple dimensions of sexuality does allow for greater understanding of patterns of identity reporting in the datasets.

Almost all respondents who identified as (Completely) Heterosexual also reported only experiencing attraction to and experience with the opposite sex. Likewise, almost all respondents who reported a non-(Completely) Heterosexual identity also reported experiencing some same-sex attraction and some same-sex experience. However, only 74.04% of ELSA

respondents and 55.05% of MCS respondents who reported any same-sex attraction also identified as Heterosexual or Completely/Mainly Heterosexual demonstrating the non-equivalence of identity, attraction and experience as dimensions of sexuality. Reporting plurisexual attraction/experience does not preclude identification with monosexual identities nor did exclusively monosexual attraction/experience preclude identification as bisexual, demonstrating the non-equivalence of identity, attraction and experience as dimensions of sexuality (see Figure 1A-C).

Impact of coding decisions

As of 2019, the Office for National Statistics has estimated that 2.7% of the UK population over 16 identifies as gay, lesbian, or bisexual(27). As a result, even in representative population surveys, sample size remains a persistent issue in analyses of sexual minorities. One common solution is the grouping together of different identities and responses into umbrella categories. While recoding is often a necessary practical step, it can obscure important differences within the sexual minority population and the inclusion or exclusion of certain responses is often left unjustified by researchers. In addition, the asymmetric nature of heterosexual and sexual minority population sizes means that choices about who is or isn't included become particularly meaningful.

We return to Savin-Williams' question 'what is enough?' when we consider sexuality (see Box 2) (7). While relevant to discussions around sexual identities, we first focus on sexual attraction and experience dimensions and measures. Neither sexual attraction or behaviour measure require association with a social identity nor repeated experience, and as a result have relatively low 'barrier to entry' to reporting. As demonstrated in the analysis above, this results in the proportion and number of individuals reporting any same-sex attraction or experience being greater than those reporting lesbian, gay, or bisexual identities. Likewise, reporting same-sex attraction or experience does not always correspond with reporting a sexual minority identity. As

a result, who should be included in a ‘sexual minority’ category using attraction and experience data is not immediately obvious.

One approach is to consider any respondent who has reported same-sex attraction or experience as sexual minority. While useful for understanding the prevalence of attitudes and behaviours in the general population and even sexual health needs, they are perhaps less useful for identifying those for whom same-sex attraction or experience is the norm rather than the exception. However, generating a cut-off for ‘significant’ same-sex attraction or experience also creates problems. While those who report a more frequent pattern of same-sex attraction and behaviour represent a smaller proportion of the general population and map well onto those who report LGB identity, this decision also creates an arbitrary category of respondents who are considered to have ‘enough’ same-sex attraction or experience to be considered sexual minority. This is problematic as the question of ‘enough’ has been used to police queer individuals’ participation in the LGBTQ+ community, erase or invalidate identities such as bisexuality, and been used to deny refugee status to LGBTQ+ asylum seekers (28).

Where more than one dimension is included it is possible to combine outcomes to generate new categories. In the analysis above, categories have been recoded in three main ways; to clarify ambiguous categories, to enforce coincidence between dimensions, and to boost the sexual minority sample size. Clarification of ambiguous identity categories (e.g. “Other”) using attraction and experience measures allows for the justified inclusion of respondents who might otherwise have been classified as missing in sexual minority categories (See Box 3). By contrast, recoding based on coincidence between identity, attraction and/or experience does ensure a sexual minority sample which meets most definitions of the target population but enforces a rigid and exclusionary criteria on a community characterised by diversity and inclusivity. Nonetheless, more inclusive recoding approaches are open to criticisms of over-inclusiveness and concerns over their applicability to real-world communities.

In the following sections, I will discuss the impact recoding decisions have on sexual minority sample size and demographic distributions.

Impact of recoding on sample size

Choice of dimension can have important impacts on the size of the sexual minority population under consideration and observed outcomes. As in the literature, more participants in MCS and ELSA reported experiencing same-sex attraction than report lesbian, gay, bisexual, etc identities (12). For example, 906 participants in MCS and 74 participants in ELSA used lesbian, gay, or bisexual to describe their sexuality, whereas 2,098 participants in MCS and 340 participants in ELSA reported some same-sex attraction. In ELSA, while fewer than those reporting same-sex attraction, more participants reported same-sex experience than LGB identity.

Using multiple dimensions to recode the sample results in high variation of sample size, however two main trends emerge. Firstly, restricting the heterosexual and sexual minority samples to only those with coincident identity, attraction and/or experience results in the smallest sample sizes. Secondly, where categorisation is based on a respondent reporting an LGB identity and/or any same-sex attraction/experience the size of the sexual minority sample is the largest.

Overall, dimension choice and recoding can have significant impact on the sample size, particularly of sexual minority categories. As a result, careful consideration by researchers is essential.

Impact of recoding on demographic distributions

Little substantial variation in demographic distribution by dimension or recoding was observed in either study. However, some differences in gender, age, ethnicity, socioeconomic outcomes, and urbanity were observed between and within cohorts.

In the younger study, sexual minority respondents were more likely to identify as female than their heterosexual peers, whereas sexual minority respondents in the older study were more likely to identify as male. This pattern holds regardless of dimension choice or recoding strategy in both datasets. One potential cause of this difference between the cohorts is the difference in the proportion of sexual minority respondents identifying as gay or lesbian or reporting exclusively same-sex attraction/experience. In both cohorts, monosexual sexual minorities were more likely to identify as male than bisexual respondents or those who report both same- and opposite-sex attraction/experience. As the sample of sexual minority respondents in ELSA is both numerically smaller and has a higher proportion of gay or lesbian respondents than that MCS, this may explain why the sexual minority sample in ELSA is more likely to be male, despite the higher proportion of female respondents in ELSA.

MCS is a much younger cohort than ELSA and a much higher proportion of participants identify with a non-heterosexual identity (8.97% in MCS vs 1.04% in ELSA) and report some same-sex attraction (20.77% in MCS vs 4.77% in ELSA).

In addition, respondents who reported LGB identity or same-sex attraction/experience in ELSA were younger than those who reported exclusively opposite-sex attraction/experience and a heterosexual identity. These patterns reflect an observed trend of younger adults being more likely or more willing to report sexual minority identity, attraction, and experience (27). Additionally, we may be observing a survivor bias in reporting due to historic and contemporary discrimination and inequality, and the consequences of the HIV/AIDs epidemic.

In both cohorts, sexual minority participants reported higher socioeconomic status than non-sexual minority participants regardless of dimension or recoding. However, in both datasets socioeconomic status varied between monosexual and plurisexual respondents. While in the younger study, monosexual sexual minority participants reported similar characteristics to their non-sexual minority peers, and worse socioeconomic status than plurisexual sexual minority participants, in the older study monosexual sexual minority participants had better

socioeconomic status than both plurisexual sexual minority participants and non-sexual minority participants. In the adolescents, this difference may be attributable to a higher likelihood and/or willingness to report a plurisexual identity or attraction by participants with better socioeconomic status. On the other hand, monosexual sexual minority older adults were more likely to identify as male, which due to the age profile of the cohort, results in participants with higher income/wealth and educational qualification characteristics.

As discussed, while there is little substantial variation by dimension and recoding strategy within datasets, we do observe demographic differences between datasets and between sexual minority participants and their heterosexual peers.

Impact of recoding on health and wellbeing

In general, non-sexual minority participants reported better outcomes in life satisfaction, self-rated health, and psychological distress than their sexual minority peers in both studies. This is consistent with the literature on health in sexual minority populations. While these outcomes did not substantially vary by dimension of sexuality employed or by recoding, there were differences within and between the two datasets.

Across the three outcomes, the disparity between sexual minority and non-sexual minority participants was more acute in the adolescent study than in the older adult study. It is unclear if this is a product of differences between the datasets in terms of sample sizes, measures and sampling framework, or reflects a narrowing of inequality over the lifecourse. One crucial factor may be the relative youth of sexual minority respondents in the older adult study compared to non-sexual minority respondents and suggests adjusting for age may be important in future studies.

Nonetheless, regardless of dataset, dimension, or recoding strategy, sexual minority participants in general report worse outcomes than their non-sexual minority peers.

Limitations

The findings may only be applicable to the populations under analysis, however, the inclusion of two population based studies covering different age groups we hope to illustrate the potential breath of considerations. Additionally, the sample sizes involved throughout this analysis are small, particularly for older respondents, which may have obscured important differences between the groups. This analysis also cannot discuss age or period effects due to question only being asked once in both cohorts; future examinations in other datasets or different timepoints of the same datasets would be valuable.

Conclusion

This study explored a range of practical and theoretical considerations when analysing sexual minority respondents using survey data. We highlight the impact recoding strategy decisions may have on sample size and demographic and health distributions. We also draw attention to the value of including multiple dimensions in population surveys, avoiding overzealous grouping or dropping of respondents, and the limits of coincidence between dimensions. However, we have also demonstrated how practical recoding and measurement decisions can be made thoughtfully and carefully (see Box 1). We hope this analysis will empower and inform researchers interested in working on sexual minority health and social inequalities, as well as encouraging all researchers to critically examine how they construct and imagine their population of study.

Box 1: Key Takeaways

- Sexual identity, behaviour and attraction measures are not interchangeable, and care should be taken when selecting a measure for inclusion in a survey or analysis.

- Dimension choice and data management of sexuality variables have important consequences for sample size and make-up, proposed mechanism of action, and health outcomes.
- Asking sexuality questions more than once can help elucidate age and period effects, allow for up-to-date responses, and improve understanding of sexuality development and change over time.
- Including more than one dimension allows for the exploration of sexual diversity and help clarify otherwise ambiguous responses.

Box 2: “Mainly Heterosexual” and “Mostly to Opposite”

The MCS measure of sexual identity introduces new categories which may not correspond with common sexual identities including “Mainly Heterosexual” and “Mainly Gay/Lesbian”. In sample the “Mainly Heterosexual” identity category is the second largest after the “Completely Heterosexual” category. Evidence from the literature suggests this group is demographically distinct from both heterosexual and bisexual groups, and relatively stable across the lifecourse (25). This is supported by MCS where the “Mainly Heterosexual” group is distinct from the other identity categories, and 71.21% of the category also reported some same-sex attraction. Nonetheless, given their identification with the heterosexual label, they could also be reasonably included under the “Heterosexual” umbrella.

Where items include qualifiers such as ‘mostly’, researchers may then recategorize into fewer groups of ‘true’ gay, lesbians and bisexuals (7). Savin-Williams (2009) argues this raises ‘how much’ of a given dimension is needed for inclusion as non-heterosexual, especially as it interacts with methodological concerns around sample size and representativeness (7). For example, in MCS merging this category doubles the number of sexual minority participants from 906 to 2007 and as a result justification for inclusion should be carefully considered.

If combined with other sexual minority groups into a single category, the “Mostly Heterosexual”/ “Mostly to opposite” results in a slightly more male, higher socioeconomic status, and less urban sample. This group also reports better life satisfaction, self-rated health, and psychological distress outcomes than other sexual minority groups, so when combined, outcomes are improved against categories where they are excluded.

While it is highly appealing to include these groups as sexual minorities as they share many qualities in common, it is important researchers are aware of the potential skew that they introduce.

Box 3: “Other”

Where a sexual identity measure includes the options ‘Other’, “Don’t Know”, and ‘Prefer Not to Say’ those responses are often excluded from analyses or marked as ‘missing’ (e.g. Geary *et al.*, 2018). Their exclusion is justified on two main grounds: small sample sizes, and uncertainty around what respondents are expressing when selecting these options. However, Booker et al. (2017) have argued strongly for their inclusion and consideration as groups worthy of further examination (29). These options may cover those who do not understand the question, who may not feel they fit within the available options, for example people who identify as asexual or queer, and those who reject labels altogether. (5,15).

Response patterns to these items can be illuminated by attraction and experience measures. For example, in the MCS study almost 40% of “Other” respondents reported some same-sex attraction, with a further 40% reporting “Never sexually attracted”. This distribution indicates the “Other” category may be selected by sexual minority participants who use sexual identity labels outside of gay/lesbian/bisexual, such as pansexual, queer, or asexual, and reflects the developing sexuality of the cohort. By contrast, the majority of “Other” older adults reported only opposite-sex attraction and experience. Given “Other” respondents in ELSA were more likely to

be over 70, non-white, and lower SES than the sample as a whole, reporting of “Other” identity may reflect poor measure comprehension by respondents who might otherwise identify as heterosexual or straight. This difference observed between MCS and ELSA “Other” respondents points to the value of multiple dimension measurement, and consideration of target population when choosing a measure and data cleaning.

Additional Information

Funding: ET is funded by the ESRC-BBSRC Soc-B Centre for Doctoral Training (ES/P000347/1).

Acknowledgements:

Author Contribution: ET conceptualized the study, developed the methodology, conducted the analyses, prepared the data visualization, and prepared the initial manuscript draft and subsequent revisions. DK and PP supervised the study conceptualization, the development of the methodology and the analysis, and reviewed and commented on the manuscript. All authors approve the final version and consent to its publication.

Data Availability: MCS and ELSA data are available to all researchers, free of cost from the UK Data Service (https://www.ukdataservice.ac.uk).

Ethics Statement:

ELSA data collection was approved by the London Multicentre Research Ethics Committee (MREC/01/2/91), and informed consent was obtained from all participants.

Ethics approval for the MCS study was obtained from the National Research Ethics Service Committee London— Central (reference 13/LO/1786). All parents gave consent for their children to participate, and young people also provided verbal consent.

Ethics approval for this analysis was received from the Institute of Education, UCL.

Bibliography

1. Semlyen J, King M, Varney J, Hagger-Johnson G. Sexual orientation and symptoms of common mental disorder or low wellbeing: combined meta-analysis of 12 UK population health surveys. BMC Psychiatry , 16 , Article 67 (2016). 2016;

2. Kneale D, Thomas J, French R. Inequalities in health and care among Lesbian, Gay and Bisexual people aged 50 and over in the United Kingdom: A systematic review and meta-analysis of sources of individual participant data. The Journals of Gerontology: Series B [Internet]. 2020 [cited 2020 Apr 6]; Available from: https://doi.org/10.1093/geronb/gbaa071

3. ONS. Sexual identity research and testing plan, The 2021 Census [Internet]. Office for National Statistics; 2016. Available from: https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/sexualorientation

4. NHS. LGBT health [Internet]. NHS; 2020. Available from: https://www.england.nhs.uk/about/equality/lgbt-health/

5. Wolff M, Wells B, Ventura-DiPersia C, Renson A, Grov C. Measuring Sexual Orientation: A Review and Critique of U.S. Data Collection Efforts and Implications for Health Policy. The Journal of Sex Research. 2017 Jun 13;54(4–5):507–31.

6. Sell RL. Defining and measuring sexual orientation for research. In: Meyer IH, Northridge ME, editors. The Health of Sexual Minorities: Public Health Perspectives on Lesbian, Gay, Bisexual and Transgender Populations. Springer; 2007. p. 355–74.

7. Savin-Williams RC. How Many Gays Are There? It Depends. In: Hope DA, editor. Contemporary Perspectives on Lesbian, Gay, and Bisexual Identities [Internet]. New York, NY: Springer New York; 2009. p. 5–41. Available from: https://doi.org/10.1007/978-0-387-09556-1_2

8. Matsuno E, Goodman JA, Israel T, Choi AY, Lin YJ, Kary KG. L or G or B or T: Matching Sexual and Gender Minorities with Subpopulation-Specific Interventions. Journal of Homosexuality. 2020;1–23.

9. Baker KE, Harris AC. Terminology Should Accurately Reflect Complexities of Sexual Orientation and Identity. American Journal of Public Health. 2020;110(11):1668–9.

10. Chung YB, Katayama M. Assessment of Sexual Orientation in Lesbian/Gay/Bisexual Studies. Journal of Homosexuality. 1996 Apr 30;30(4):49–62.

11. Saewyc EM, Bauer GR, Skay CL, Bearinger LH, Resnick MD, Reis E, et al. Measuring sexual orientation in adolescent health surveys: Evaluation of eight school-based surveys. Journal of Adolescent Health. 2004;35(4):345.e1-345.e15.

12. Geary RS, Tanton C, Erens B, Clifton S, Prah P, Wellings K, et al. Sexual identity, attraction and behaviour in Britain: The implications of using different dimensions of sexual orientation to estimate the size of sexual minority populations and inform public health interventions. PLoS ONE. 2018;13(1).

13. Young R, Meyer I. The Trouble With ‘MSM’ and ‘WSW’: Erasure of the Sexual-Minority Person in Public Health Discourse. American Journal of Public Health. 2005;95(7):1144–9.

14. Fahs B, McClelland SI. When Sex and Power Collide: An Argument for Critical Sexuality Studies. The Journal of sex research. 2016;53(4–5):392–416.

15. ONS. Measuring Sexual Orientation: A guide for researchers. Newport: Office for National Statistics; 2009.

16. ONS. Guidance for questions on sex, gender identity and sexual orientation for the 2019 Census Rehearsal for the 2021 Census [Internet]. Office for National Statistics; 2019 [cited 2020 Apr 9]. Available from: https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/genderidentity/guidanceforquestionsonsexgenderidentityandsexualorientationforthe2019censusrehearsalforthe2021census

17. ONS. Measuring Sexual Identity: An Evaluation Report [Internet]. Office for National Statistics; 2010. Available from: https://data.gov.uk/dataset/e6033627-3106-47a9-a3fb-4bcaa34c1bcb/measuring-sexual-identity-evaluation-report

18. Fredriksen-Goldsen KI, Kim HJ. Count Me In: Response to Sexual Orientation Measures Among Older Adults. Research on Aging. 2015 Jul 1;37(5):464–80.

19. Diamond LM. Female Bisexuality From Adolescence to Adulthood: Results From a 10-Year Longitudinal Study. Developmental Psychology. 2008;44(1):5–14.

20. Semlyen J. Recording Sexual Orientation in the UK: Pooling Data for Statistical Power. American journal of public health. 2017;107(8):1215.

21. Connelly R, Platt L. Cohort Profile: UK Millennium Cohort Study (MCS). International Journal of Epidemiology. 2014;43(6):1719–25.

22. NatCen Social Research. User Guide to the Main Interview Datasets: Waves 1 to 8 [Internet]. London: ELSA / NatCen Social Research; 2018. Available from: https://www.ucl.ac.uk/drupal/site_iehc/sites/iehc/files/5050_waves_1-8_interviewer_data_user_guide_v01.pdf

23. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort Profile: The English Longitudinal Study of Ageing. International Journal of Epidemiology. 2012;42(6):1640–8.

24. Hayfield N. Bisexual and pansexual identities: exploring and challenging invisibility and invalidation [Internet]. 2021 [cited 2021 Jan 31]. Available from: https://www.taylorfrancis.com/books/e/9780429464362

25. Savin-Williams RC, Vrangalova Z. Mostly heterosexual as a distinct sexual orientation group: A systematic review of the empirical evidence. Developmental Review. 2013 Mar 1;33(1):58–88.

26. Guyan K. Constructing a queer population? Asking about sexual orientation in Scotland’s 2022 census. Journal of Gender Studies. 2021;1–11.

27. ONS. Sexual orientation, UK - Office for National Statistics [Internet]. 2021 [cited 2021 Jun 21]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/sexuality/bulletins/sexualidentityuk/2019#sexual-orientation-in-the-uk

28. Henley J. LGBT asylum seekers' claims routinely rejected in Europe and UK. The Guardian [Internet]. 2020 Jul 9 [cited 2022 Feb 23]; Available from: https://www.theguardian.com/uk-news/2020/jul/09/lgbt-asylum-seekers-routinely-see-claims-rejected-in-europe-and-uk

29. Booker CL, Rieger G, Unger JB. Sexual orientation health inequality: Evidence from Understanding Society, the UK Longitudinal Household Study. Preventive Medicine. 2017;101:126–32.

Table 1: *Sexuality variables in ELSA and MCS (% excluding missing)*

	MCS	ELSA
	N= 10,103	**N= 7,130**
Sexual Identity N (%)		
Completely Heterosexual	7888 (76.25)	4,513 (94.69)
Mainly Heterosexual	1101 (10.64)	
Bisexual	656 (6.34)	31 (0.65)
Mainly Gay or Lesbian	90 (0.87)	43 (0.90)
Completely Gay or Lesbian	160 (1.55)	
Other	157 (1.52)	27 (0.57)
Don't Know	16 (0.15)	-
PNS	35 (0.34)	152 (3.19)
Missing	-	2,364
Sexual Attraction N (%)		
Only to opposite	7656 (75.78)	5,666 (92.99)
Mostly to opposite	1382 (13.68)	218 (3.58)
Equally	385 (3.81)	57 (0.94)
Mostly to same	194 (1.92)	19 (0.31)
Only to same	137 (1.36)	46(0.75)
No attraction/desires	291 (2.88)	87 (1.43)
Don't Know	11 (0.11)	-
PNS	46 (0.46)	-
Missing	1	1,037
Sexual Experience N (%)		
Only w. opposite	-	5,790 (94.76)
Mostly w. opposite	-	158 (2.59)
Equally	-	23 (0.38)
Mostly w. same	-	27(0.44)
Only w. same	-	33 (0.54)
No experience	-	79 (1.29)
Don't Know	-	-
PNS	-	-
Missing	-	1,020

Table 2: *Sexuality variables in ELSA and MCS*

	MCS Age 17					ELSA Age 50+						
	Compl. Hetero. Identity	**Only opposite-sex attraction**	**LGB Identity**	**Any same-sex attraction**	**Total**	**Hetero. Identity**	**Only opposite-sex attraction**	**Only opposite-sex behaviour**	**LGB Identity**	**Any same-sex attraction**	**Any same-sex behaviour**	**Total**
N	7,888	7,656	906	2,098	10,103	4,513	5,666	5,790	74	340	241	7,130
N (%) Women	3719 (47.15)	3476 (45.40)	582 (64.24)	1408 (67.11)	5091 (50.39)	2504 (55.48)	3080 (54.36)	3216 (55.54)	28 (37.84)	219 (64.41)	117 (48.55)	3946 (55.34)
N (%) White	6233 (79.02)	6072 (79.31)	814 (89.85)	1841 (87.75)	8,140 (80.59)	4410 (97.72)	5536 (97.71)	5658 (97.72)	74 (100.00)	335 (98.53)	234 (97.10)	6937 (97.29)
N (%) Highest Qual Level	1403 (17.79)	1394 (18.21)	192 (21.19)	480 (22.88)	1,913 (18.93)	860 (19.06)	998 (17.61)	1001 (17.29)	17 (22.97)	73 (21.47)	70 (29.05)	1230 (17.25)
N (%) Highest Income Quintile	1750 (22.19)	1732 (22.62)	249 (27.48)	588 (28.03)	2,387 (23.63)	975 (21.60)	1058 (18.67)	1080 (18.65)	19 (25.68)	87 (25.59)	67 (27.80)	1355 (19.00)
N (%) Urban/ London	5841 (74.05)	5653 (73.84)	691 (76.27)	1525 (72.69)	7,459 (73.83)	341 (7.56)	440 (7.77)	449 (7.75)	14 (18.92)	43 (12.65)	37 (15.35)	572 (8.02)
N (%) Aged 70+	N/A	N/A	N/A	N/A	N/A	1,918 (42.50)	2,779 (49.05)	2,855 (49.31)	29 (39.19)	138 (40.59)	85 (35.27)	3390 (47.55)
N (%) Life Satisfaction (Agree)	6,293 (79.77)	6,146 (80.28)	490 (54.08)	1,204 (57.39)	7,598 (75.20)	3,624 (80.30)	4,420 (78.00)	4,516 (78.00)	56 (75.68)	255 (75.00)	185 (76.76)	5,560 (77.98)
N (%) General Health (Excellent or V. Good)	5,519 (69.96)	5,387 (70.49)	495 (54.63)	1,211 (57.72)	6,803 (67.34)	1,928 (42.72)	2,208 (38.97)	2,253 (38.92)	28 (37.83)	137 (40.30)	92 (38.17)	2,722 (38.88)
Mean Psych. Distress (95%% CI)	6.50 (6.40, 6.60)	6.43 (6.33, 6.53)	10.95 (10.61, 11.29)	10.39 (10.17, 10.60)	7.28 (7.18, 7.37)	2.91 (2.87, 2.95)	2.98 (2.94, 3.01)	2.99 (2.96, 3.03)	2.99 (2.68, 3.29)	3.00 (2.86, 3.15)	2.89 (2.72, 3.05)	2.99 (2.96, 3.03)

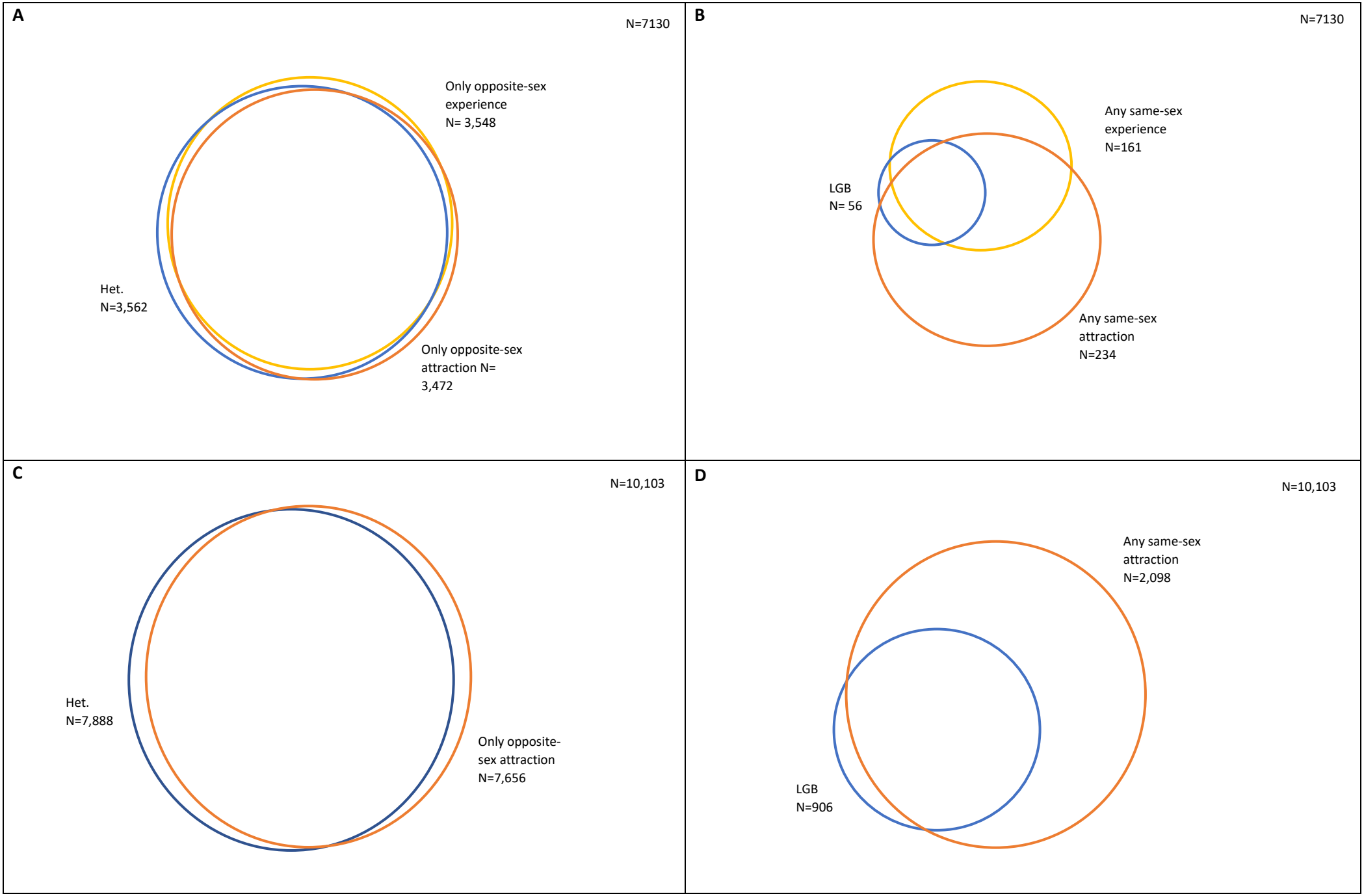
Figure 1: *Overlap between* ***A)*** *Heterosexual identified respondents and those who report only opposite-sex attraction and experience in ELSA,* ***B)*** *LGB identified respondents and those who report any opposite-sex attraction and experience in ELSA,* ***C)*** *Heterosexual identified respondents and those who report only opposite-sex attraction in MCS, and* ***D)*** *LGB identified respondents and those who report any same-sex attraction in MCS. Blue corresponds to sexual identity measures, orange to sexual attraction measures, and yellow to sexual experience measures. All complete case.*

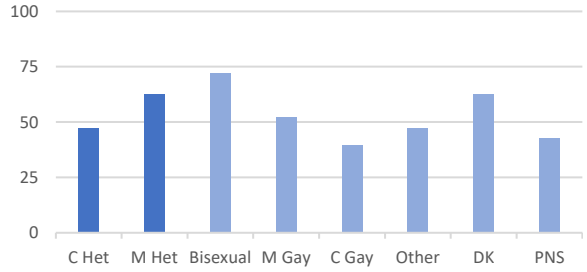
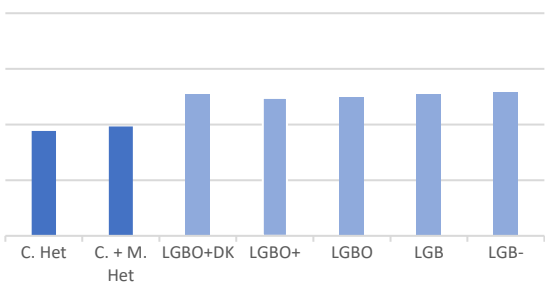
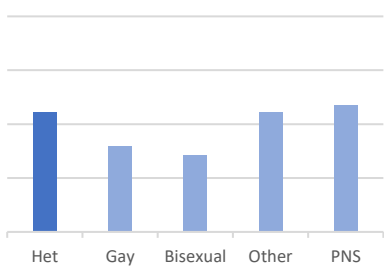
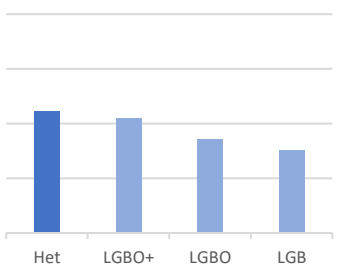
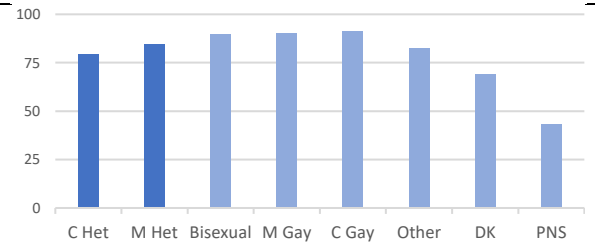
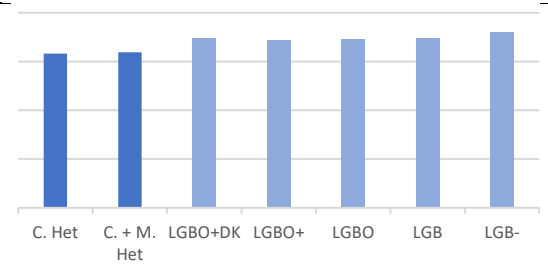
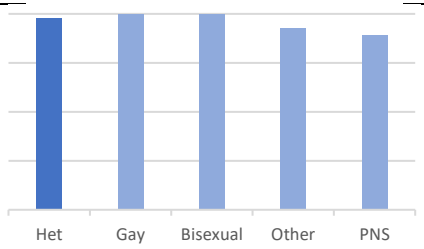
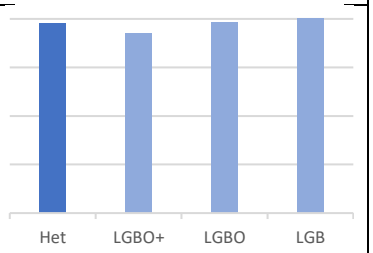
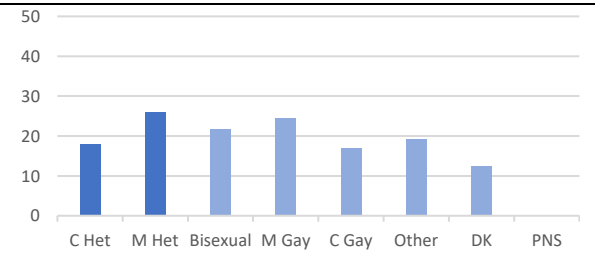
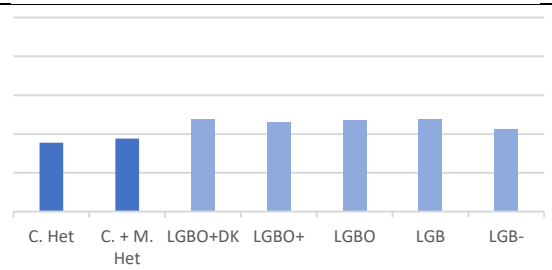
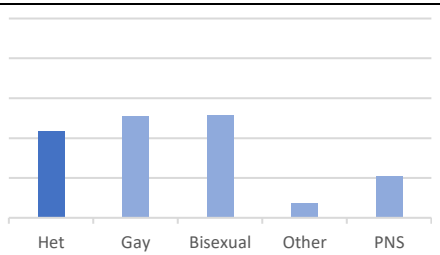
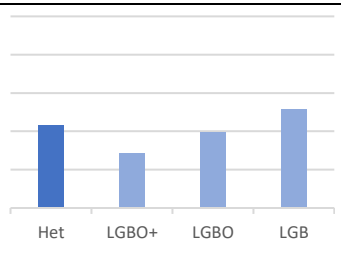
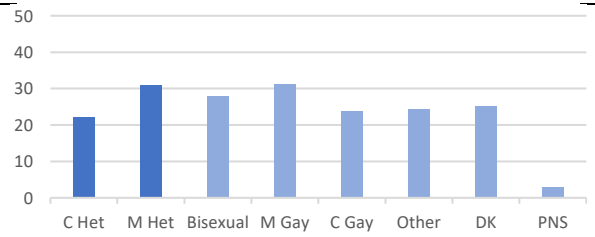
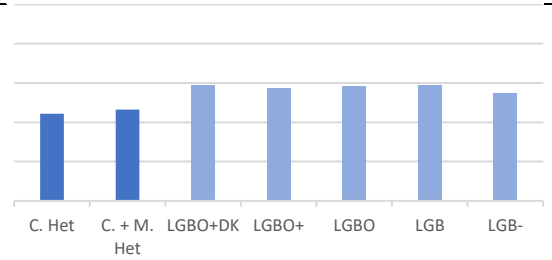
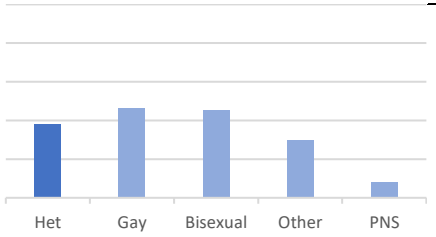
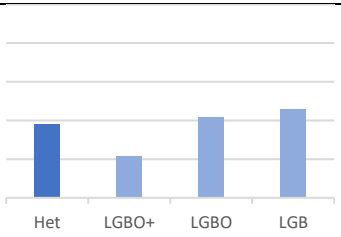
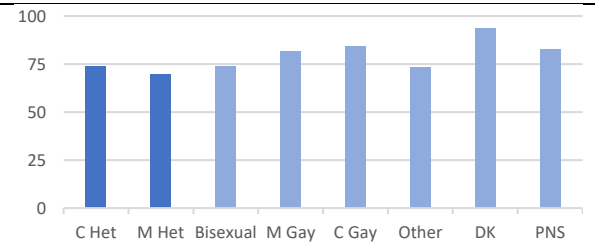
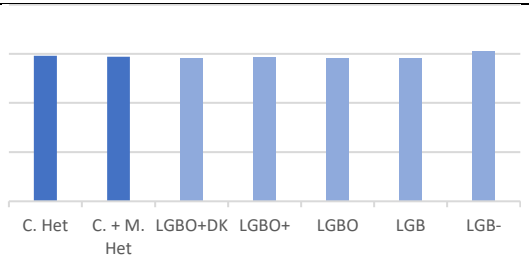
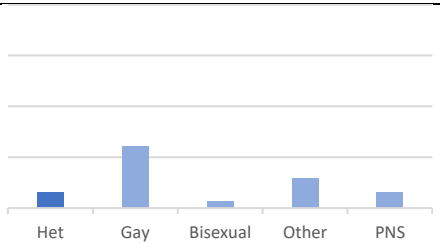
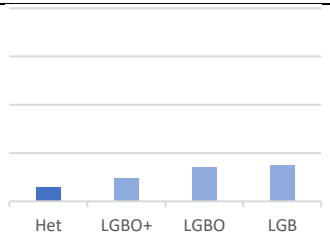
Figure 2: *Demographic distributions by sexual identity in MCS and ELSA*

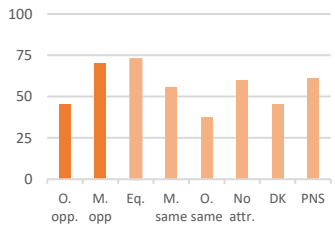
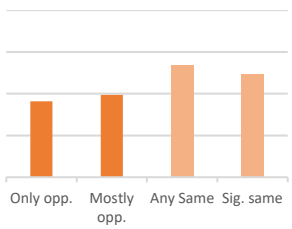
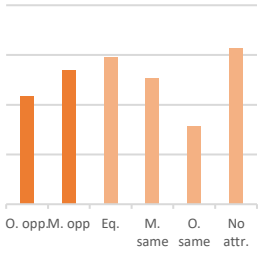
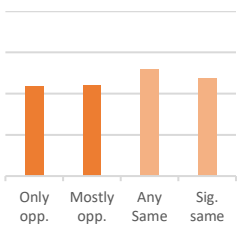
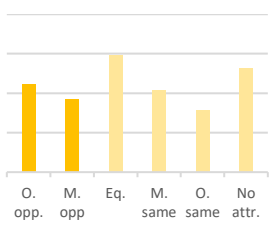
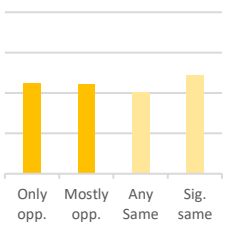
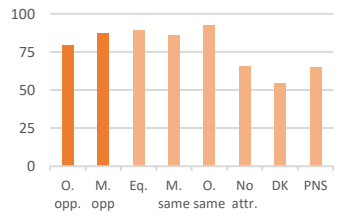
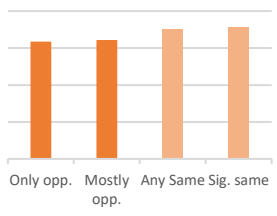
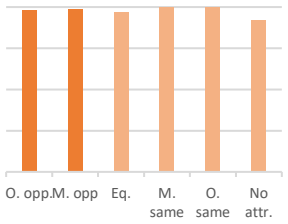
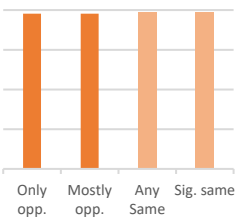
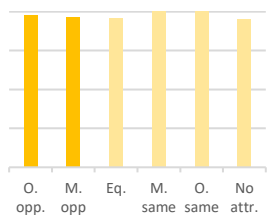
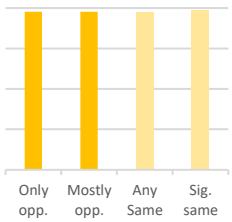
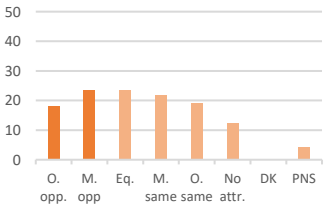
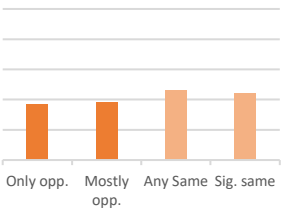
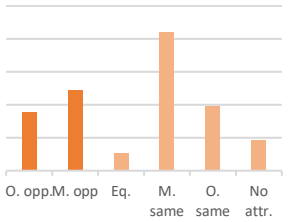
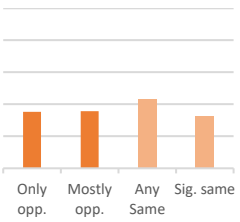
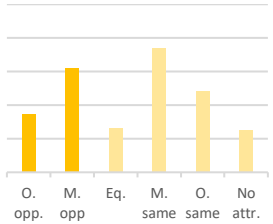
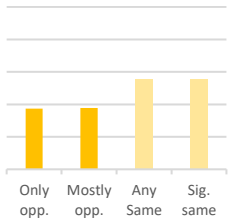
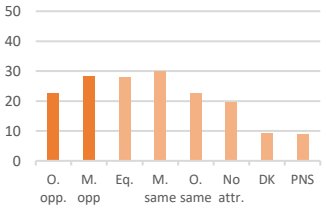
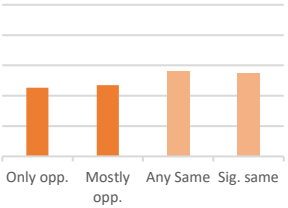
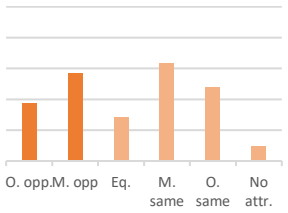
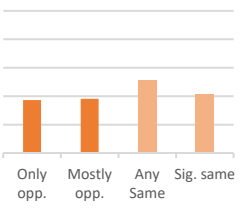
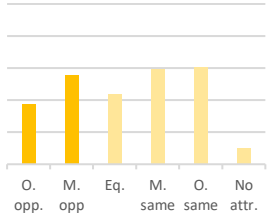
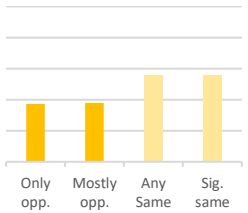
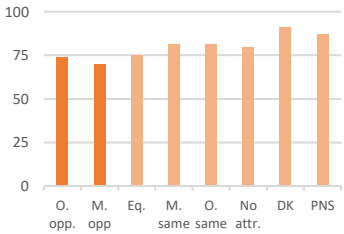
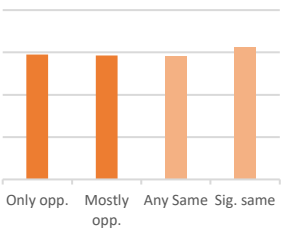
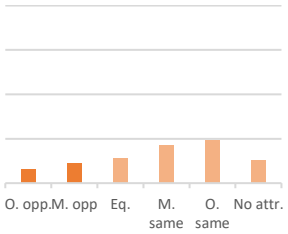
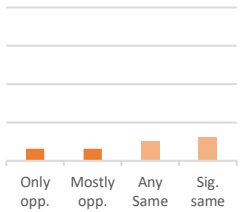
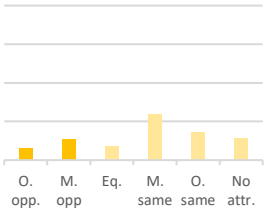
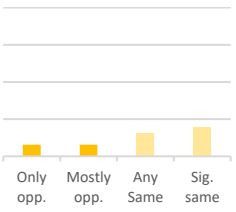
Figure 3: *Demographic distributions by sexual attraction and behaviour in MCS and ELSA*

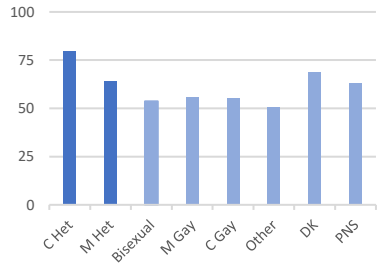
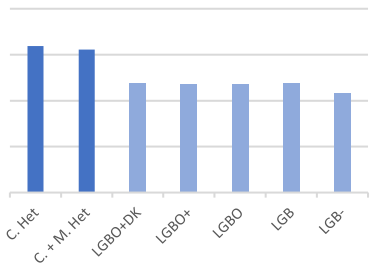
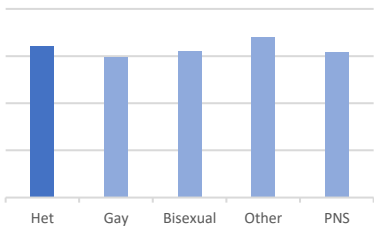
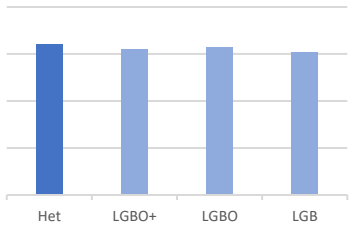
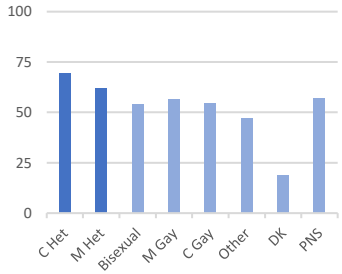
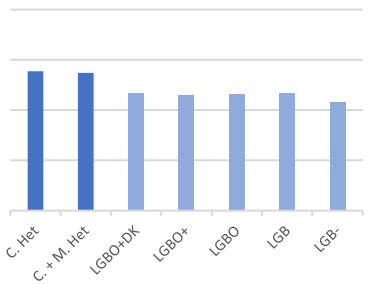
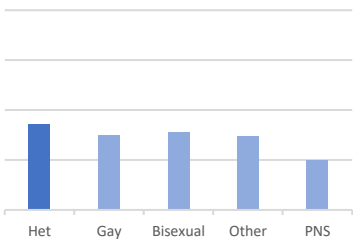
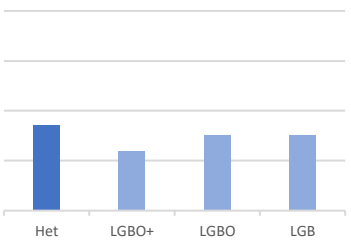
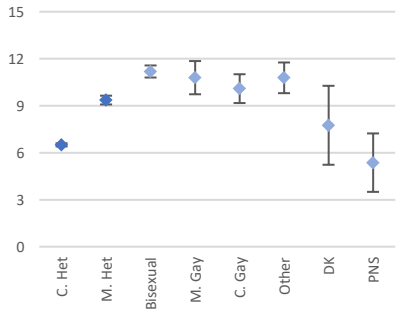
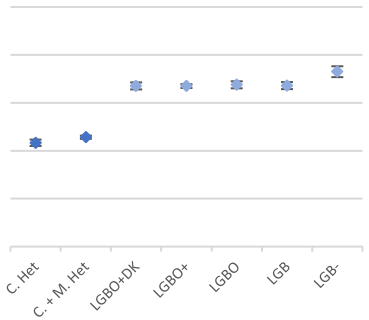
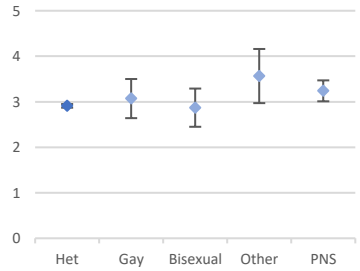
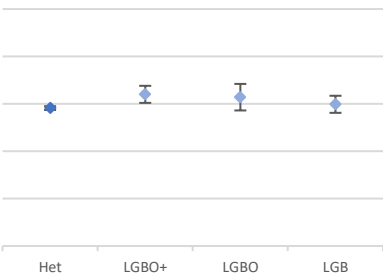
Figure 4*: Health outcomes by sexual identity in MCS and ELSA*

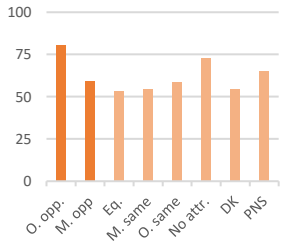
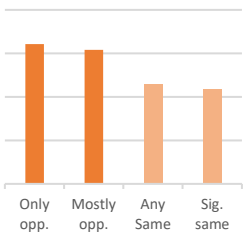
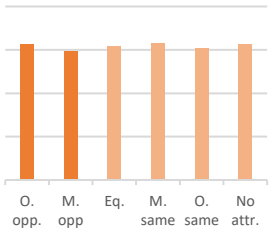
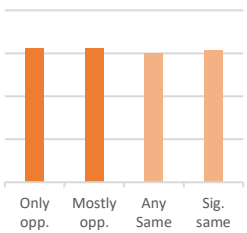
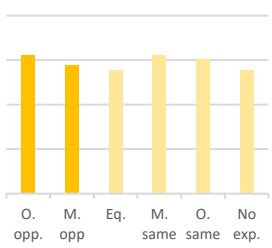
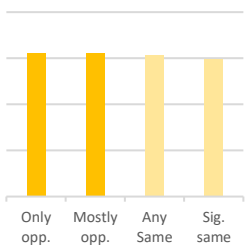
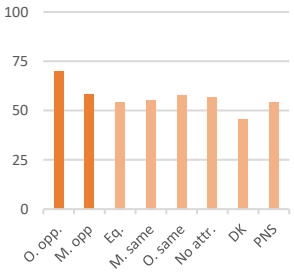
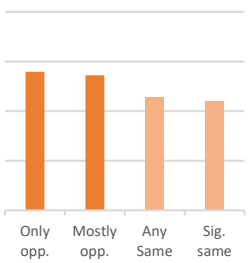
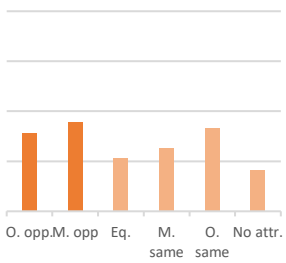
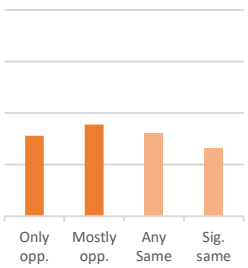
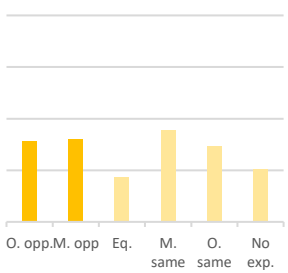
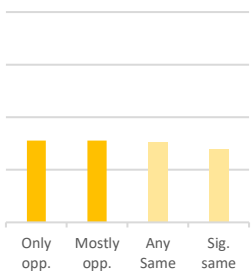
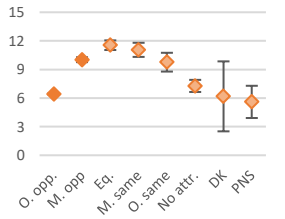
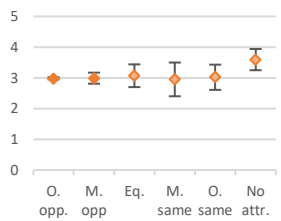
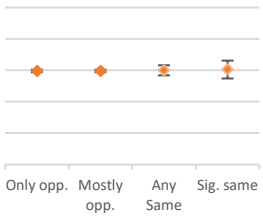
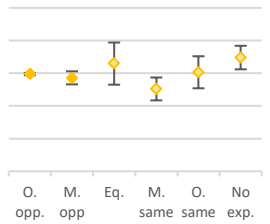
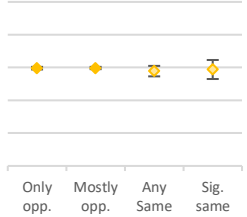
Figure 5*: Health outcomes by sexual attraction and behaviour in MCS and ELSA*


A
N=7130
Only opposite-sex experience N= 3,548
Het. N=3,562
Only opposite-sex attraction N= 3,472
B
N=7130
Any same-sex experience N=161
LGB N= 56
Any same-sex attraction N=234
C
N=10,103
Het. N=7,888
Only opposite-sex attraction N=7,656
D
N=10,103
Any same-sex attraction N=2,098
LGB N=906

Identity	MCS	Identity	ELSA	Identity
% Women				
% White				
% Highest Qual Level				
% Highest Income Quint.				
% Urban/ London				

	MCS	Attraction	ELSA	Attraction	ELSA	Experience
% Women						
% White						
% Highest Qual Level						
% Highest Income Quintile						
% Urban/ London						

	MCS	Identity	ELSA	Identity
% Life Satisfaction (Agree)				
% Self-Rated Gen. Health (Excellent or V. Good)				
Mean Psychological Distress (95% CI)				

	MCS	Attraction	ELSA	Attraction	ELSA	Experience
% Life Satisfaction (Agree)						
% Self-Rated Gen. Health (Excellent or V. Good)						
Mean Psychological Distress (95% CI)