



#BeeWell

Evidence Briefing 1

Inequalities in Young People's Wellbeing

Prepared by the #BeeWell Research Team,
University of Manchester

March, 2022



Contents

Executive Summary	3
Introduction	4
About the #BeeWell Programme	4
Inequalities in young people’s wellbeing	4
What did we do?	5
What did we find?	7
Key findings: gender identity	8
Key findings: sexual orientation	9
Key findings: transgender status	10
Key findings: socio-economic status	11
Key findings: ethnicity	12
Key findings: language	13
Key findings: special educational needs and disabilities	14
Key findings: caregiving responsibilities	15
Key findings: age	16
Conclusion	17
References	18
Appendix	19

The following members of the #BeeWell team contributed directly to this report (in alphabetical order): Louise Black, Megan Cutts, Amy Hibbert, Neil Humphrey, Devi Khanna, Jose Marquez, and Huw Spencer.

Executive Summary

About the #BeeWell programme

#BeeWell is a programme that aims to make the wellbeing of young people everybody's business. A key component of this is an annual survey of the domains and drivers of wellbeing among young people across Greater Manchester. The data generated will feed into the development of resources and tools that will be embedded into the GM system to provide new and improved ways to support young people, creating the best conditions for them to thrive in every community, regardless of their background.

Inequalities in young people's wellbeing

Inequalities are systematic, avoidable and unfair differences in outcomes between different populations or groups (1). This evidence briefing focuses on inequalities in adolescent wellbeing, using data from the first annual wave of surveys in the #BeeWell programme. This focus was determined through consultation with key project stakeholder groups, as was our specific emphasis on life satisfaction, psychological wellbeing, stress, and negative affect.

What did we do?

We surveyed nearly 40,000 young people from more than 160 secondary schools across all 10 LAs in GM, without whose efforts in Autumn 2021 this study would not have been possible. Survey responses were linked to demographic data about these young people provided by LAs and/or schools. The data were analysed using robust statistical methods that enabled us to quantify wellbeing inequalities across nine dimensions: gender identity, sexual orientation, transgender status, ethnicity, language, age, socio-economic status, caregiving responsibilities, and special educational needs and disabilities.

What did we find?

It is possible to summarise our findings as follows:

- For most dimensions, we found minimal or mixed/inconsistent evidence of wellbeing inequalities. In cases where inequalities were identified, they were not so pronounced as to be practically meaningful. This statement can be applied to findings relating to ethnicity, language, age, socio-economic status, caregiving responsibilities, and special educational needs. With one exception (the difference in life satisfaction between young people in Years 8 and 10), all of the group differences for these analyses were less than one-fifth of a standard deviation (which could be considered negligible).
- For some dimensions, we found consistent evidence of pronounced wellbeing inequalities that were routinely large enough to warrant significant concern. This statement can be applied to findings relating to gender identity, sexual orientation, and transgender status. All of the group differences for these analyses were larger than one-fifth of a standard deviation, but were often much larger than this. Perhaps most notable were inequalities relating to negative affect, which were the most pronounced in all analyses relating to these dimensions, and reached nearly nine-tenths of a standard deviation in the comparisons between heterosexual and gay/lesbian or bi/pansexual young people.

What next?

We found consistent evidence of pronounced wellbeing inequalities that were routinely large enough to warrant significant concern in relation to gender identity, sexual orientation, and transgender status. Such characteristics are often not examined in large cohort studies of young people, and so these findings are particularly noteworthy. Our findings indicate an urgent need to better understand how and why such marked wellbeing inequalities exist in relation to gender, sexual orientation, and transgender status, and to enact change across the system that can help to 'narrow the gaps'.

Introduction

About the #BeeWell programme

#BeeWell is a programme that aims to make the wellbeing of young people everybody's business. A central component of this is the #BeeWell survey (available [here](#)), which is being completed annually by pupils in secondary schools across Greater Manchester (GM). The research design includes a longitudinal element (tracking pupils from Year 8 to Year 10) and a cross-sectional element (annual survey of pupils in Year 10).

The survey focuses on domains of wellbeing (e.g. life satisfaction, optimism) and drivers of wellbeing (e.g. physical activity, bullying). We link these data to information on pupil characteristics (e.g. ethnicity) shared by Local Authorities (LAs) and/or schools. In the first annual wave of data collection (autumn 2021), responses to the #BeeWell survey were received from nearly 40,000 young people from more than 160 secondary schools across all 10 LAs in GM.

The results of the survey will help us to better understand the development and drivers of young people's wellbeing. The data generated will feed into the development of resources and tools that will be embedded into the GM system to provide new and improved ways to support young people, creating the best conditions for them to thrive in every community, regardless of their background.

Inequalities in young people's wellbeing

Inequalities are systematic, avoidable and unfair differences in outcomes between different populations or groups (1). This evidence briefing focuses on inequalities in adolescent wellbeing, using data from the first annual wave of surveys in the #BeeWell programme. This focus was determined through consultation with key project stakeholder groups, including our Research Advisory Group, Youth Steering Group, and our #BeeWell Young Researchers¹.

For this briefing the following domains were chosen from the wider #BeeWell survey battery: psychological wellbeing, stress, life satisfaction, and negative affect. For brief definitions of these domains, please refer to Table 2. *Psychological wellbeing* was selected because it is being used as a key indicator in the new GM Strategy². *Stress* was chosen because it was voted as the most important of our various wellbeing domains by our #BeeWell Young Researchers. Finally, *life satisfaction* and *negative affect* were selected by the #BeeWell Research Team. The former was nominated as it is often used in large studies as an overarching wellbeing indicator (2). The latter was chosen so that we could examine inequalities in mental health difficulties (e.g. sadness, anxiety), which are a particular concern for this age group (3,4).

We examine nine different inequalities dimensions in the #BeeWell dataset: gender identity, transgender status, sexual orientation, ethnicity, age, socio-economic status, special educational needs, care-giving responsibilities, and language. These were selected because previous research had indicated the presence of inequalities in some of the wellbeing outcomes noted above, and/or they had been highlighted by one or more of our stakeholder groups. For example, an earlier study found that girls were nearly 3 times more likely to experience significant emotional problems (negative affect) than boys (5). Similarly, our #BeeWell Young Researchers nominated ethnicity as a key characteristic that they wanted our analysis to focus on.

¹ The #BeeWell Research Advisory Group comprises academics/researchers, young people, and other stakeholders who support and provide strategic advice on the development of a well planned and executed research plan for the programme. Our Youth Steering Group are a group of young people who meet regularly to provide advice and support to the project team, with particular reference to ensuring that #BeeWell reflects the views and preferences of young people. Finally, the #BeeWell Young Researchers are a group of Year 10 pupils who are learning about the world of academic research, with a focus on wellbeing.

² The GM Strategy is a plan for Greater Manchester written by all 10 councils, the Mayor, and a range of other stakeholders (e.g. the NHS, transport, the police and the fire service, businesses, voluntary, community and social enterprise organisations, and members of the public). It outlines collective ambitions for the future of the city-region and its residents. It spans a range of aspects of people's lives, including their health and wellbeing.

What did we do?

In autumn 2021 we surveyed young people in secondary schools across GM. As noted above, we linked the survey data to information on pupil characteristics (e.g. ethnicity) shared by LAs or schools. This report uses data on 37,978 young people across 165 schools³.

We were able to compare some characteristics of our sample (specifically, sex, age, ethnicity, special educational needs, and language) to those of the population of young people aged 11-16 in GM, using government statistics. We found that there were very few differences, and where these did exist, they were negligible. This gives us confidence that our sample findings can be applied to the population of young people aged 11-16 in GM.

Information about the inequalities dimensions can be seen in Table 1, below.

Table 1. Inequalities dimensions.

Inequalities dimension	Source	Categories	Notes
Gender identity	Survey	Male (inc trans boy) (41.71%) Female (inc trans girl) (40.03%) Non-binary (2.36%) Describe myself in another way (2.77%) Prefer not to say (5.27%) Missing (7.86%)	
Sexual orientation	Survey	Heterosexual/straight (67.50%) Gay/lesbian (2.66%) Bi/pansexual (7.65%) Describe myself in another way (3.73%) Prefer not to say (9.07%) Missing (9.39%)	
Transgender status	Survey and LA	Cisgender (79.73%) Transgender (7.08%) Missing (13.2%)	Correspondence between sex (LA data) ⁴ and gender identity (survey data). Missingness primarily driven by non-response on gender identity item
Ethnicity	LA	White (64.35%) Black (5.38%) Asian (17.74%) Chinese (0.78%) Any other ethnic group (2.26%) Mixed (5.76%) Unclassified (1.86%) Missing (1.89%)	
Age	LA	Year 8 (53.67%) Year 10 (46.33%)	
Socio-economic status	LA	Eligible (24.94%) for Free School Meals in the last six years Not eligible (73.16%) Missing (1.90%)	
Special educational needs	LA	Identified as having special educational needs and disabilities (SEND; in receipt of SEND Support or Education, Health and Care plan, 13.99%) No SEND (84.45%) Missing (1.56%)	
Caregiving responsibilities	Survey	Caregiving responsibilities (38.96%) No caregiving responsibilities (51.54%) Missing (9.59%)	Young people were asked if they care for a family member who has an illness, disability, mental health condition, or drug/alcohol dependency ⁵
Language	LA	English as an additional language (21.23%) English as first language (77.49%) Missing (1.28%)	

³ This is not the total number of schools and young people who took part in the first annual #BeeWell survey because not everyone completed the standard version of the survey that includes the four wellbeing domains that we focus on in this evidence briefing.

⁴ Data provided by LAs from the school census does now enable young people and/or their parents to request that these data are changed. However, for recording and reporting purposes, only 'male' and 'female' are available.

⁵ It is important to note that there is a known discrepancy between self-report and other ways of reporting in relation to young carers. In other words, a much higher proportion of young people self-identify as young carers than national estimates of the proportion of young carers. For this reason, we refer to them having 'caregiving responsibilities' as opposed to being 'young carers'.

Information about the domains of wellbeing can be seen in Table 2, below.

Table 2. Wellbeing domains.

Wellbeing domain	Measure	N of items	Sample item	Response format	Interpretation	Mean score	Standard deviation
Psychological wellbeing (feeling good and functioning well)	Short Warwick Edinburgh Mental Wellbeing Scale (SWEMWBS)	7	"I've been feeling useful"	Strongly agree, agree, disagree, strongly disagree	Higher score = higher wellbeing	23.08 (possible range 7 to 35)	5.82
Life satisfaction (overall quality of life as a whole)	Office for National Statistics (ONS) Life Satisfaction item	1	Overall, how satisfied are you with your life nowadays?	0-10 scale, with 0 = not at all, 10 = completely	Higher score = higher life satisfaction	6.63 (possible range 0 to 10)	2.50
Stress (feeling overwhelmed by the demands of daily life)	Perceived Stress Scale – 4 item version (PSS-4)	4	In the last month, how often have you felt that difficulties were piling up so high that you could not overcome them?	Never, almost never, sometimes, fairly often, very often	Higher score = higher stress	7.29 (possible range 0 to 16)	3.01
Negative affect (experience of negative emotions such as sadness and worry)	Me and My Feelings (M&MF), emotional problems subscale	10	I worry a lot	Never, sometimes, always	Higher score = higher negative affect	6.71 (possible range 0 to 20)	4.73

We calculated young people's score for each wellbeing domain and then transformed these data into z-scores. Z-scores are standard scores that indicate the distance of a given score from the mean in standard deviation units⁶. Using z-scores means we can compare inequalities across our four wellbeing domains with ease, even though they are measured on different scales (that is, they have different response formats and numbers of items - see Table 2 above).

For each wellbeing domain, we ran a multi-level regression analysis, using the inequalities dimensions as predictor variables. Regression analysis enables us to estimate the relationship between an outcome of interest (e.g. life satisfaction) and a set of explanatory variables (e.g. ethnicity). We used multi-level regression because our data are 'nested' (e.g. young people within schools) and it is important to account for the similarities between young people from the same school in our analysis. Failing to do so can lead to incorrect inferences.

For each inequalities dimension, a reference group (typically the largest group) is nominated, and then the difference in outcomes for each other group is calculated. For example, for gender identity, males are the reference category against which we compare outcomes for those who identify as female, non-binary, and describe themselves in another way.

As with any project of this kind, some of our data were missing. Handling missing data is important because it can lead to biased analyses. Accordingly, we used a technique called multiple imputation in which missing data are replaced by multiple sets of plausible values. This enabled us to make use of the full #BeeWell sample in our analyses.

For further detail regarding our analyses, please refer to Appendix 1.

⁶ The standard deviation is a measure of how spread/dispersed the data are in relation to the mean (average). Larger standard deviations are indicative of greater spread/dispersion.

What did we find?

Our findings can be interpreted in the following ways:

1. In the tables below, the value in a given cell is the standardised difference between two groups for an outcome. So, for example, the difference between females and males for life satisfaction is -0.344, meaning that females score approximately one-third of a standard deviation lower than males for this outcome⁷. The values provided in each table have been calculated to take into account all other inequalities dimensions. For example, the gender identity inequalities reported in Table 3 control for the other inequalities reported in Tables 4 to 11⁸.
2. Accompanying each value in 1 above, we use asterisks to denote whether the 'p value' for the difference in scores is statistically significant. The p value tells us how frequently, if we ran our study again many times, we would get data as extreme (or more extreme) than the data we have, if there is no actual difference in the population. It is expressed as a percentage (e.g. 0.08 = 8%). If the p value is small enough (normally less than 5% from a very large number of hypothetical results – hence ' $p < .05$ '), the finding is considered to be statistically significant. This means that the result is therefore considered unlikely to be the result of random noise. That said, very large samples such as that used here lead to increased test sensitivity. This means that some statistically significant results can emerge where the actual magnitude of difference between groups is not meaningful. As a result, our main emphasis in the sections below is on the size of the inequalities (see 1 above).
3. Accompanying each table, we provide a visualisation called a radar plot. This maps out inequalities for the four wellbeing domains for a given dimension (e.g. sexual orientation). The scale for each wellbeing domain runs from -1, through 0 (zero), to +1. The 0 (zero) line represents the average for the whole sample. If the coloured lines for each group are all at or close to 0 (zero) for a domain, this is indicative of equality, on average, between the groups. A good example of this is the radar plot for language, where there is equality of outcome across the four wellbeing domains for young people who speak English as an additional language and those for whom English is their first language. Where the coloured lines for each group deviate substantially from the 0 (zero) line for a domain, this is indicative of inequality, with larger deviations (e.g. greater distance between the reference group and other groups) demonstrative of greater inequalities. A good example of this can be seen in the sexual orientation radar plot, where young people who are gay/lesbian have notably worse outcomes across all four wellbeing domains than young people who are heterosexual, though the biggest absolute difference is for negative affect.

⁷These standardised differences can be converted back to units of the original measurement scale as required by multiplying them by the corresponding standard deviation reported in Table 2. For example, the male↔female standardised difference of -0.344 in life satisfaction equates to $0.344 \times 2.5 = 0.86$ points on the ONS4 Life Satisfaction item.

⁸The exception to this is that we do not control for transgender status and gender identity at the same time, due to the strong correspondence between these variables. Accordingly, for transgender status analyses, we control for all variables noted in Table 1 except for gender identity. For all other analyses, we control for all variables noted in Table 1 except for transgender status.

Gender identity

Key findings: gender identity

Our analysis reveals multiple inequalities by gender identity. Females score significantly lower than males across the four wellbeing domains. This gap is around one-third of a standard deviation in all the domains except for negative affect, where the gap is well above half of a standard deviation. However, the most noteworthy gender differences are between males and non-binary young people. These gaps are similar (about a half to two-thirds of a standard deviation). Most notably, the difference for negative affect is about two-thirds of a standard deviation (just over 3 points difference on the Me and My Feelings emotional problems subscale). A similar pattern is evident for young people who describe their gender identity in another way ('other'). Their wellbeing is also significantly lower than that of males and in terms of the magnitude of inequalities, sits somewhere in between females and non-binary young people in all domains other than negative affect, where it is similar to that of females.

Figure 1. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by gender identity

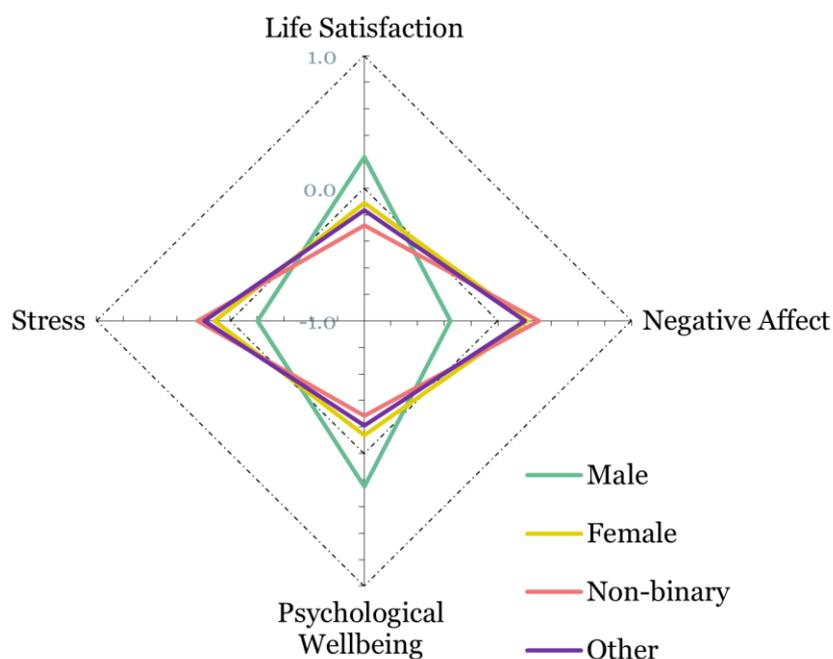


Table 3. Inequalities in wellbeing domains: gender identity comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
Male↔Female	-0.344***	0.576***	-0.384***	0.321***
Male↔Non-binary	-0.513***	0.662***	-0.532***	0.447***
Male↔Other	-0.401***	0.558***	-0.460***	0.400***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Sexual orientation

Key findings: sexual orientation

Our analysis reveals multiple inequalities by sexual orientation. The most noteworthy are the gaps between heterosexual and gay/lesbian young people, and between heterosexual and bi/pansexual young people. Both sets of comparisons are remarkably similar, in that they each reveal statistically significant inequalities of a similar size for a given wellbeing domain. We note in particular that gay/lesbian and bi/pansexual young people score just over eight-tenths of a standard deviation higher than heterosexual young people for negative affect. This is equivalent to around 4 points difference on the Me and My Feelings emotional problems subscale. There are also sizeable inequalities between heterosexual↔gay/lesbian young people and heterosexual↔bi/pansexual young people for life satisfaction, psychological wellbeing, and stress, equivalent to between a half and two-thirds of a standard deviation in each case. A similar pattern is evident for young people who describe themselves in another way, though in each case, the size of the gap is much smaller than the heterosexual↔gay/lesbian and heterosexual↔bi/pansexual comparisons.

Figure 2. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by sexual orientation

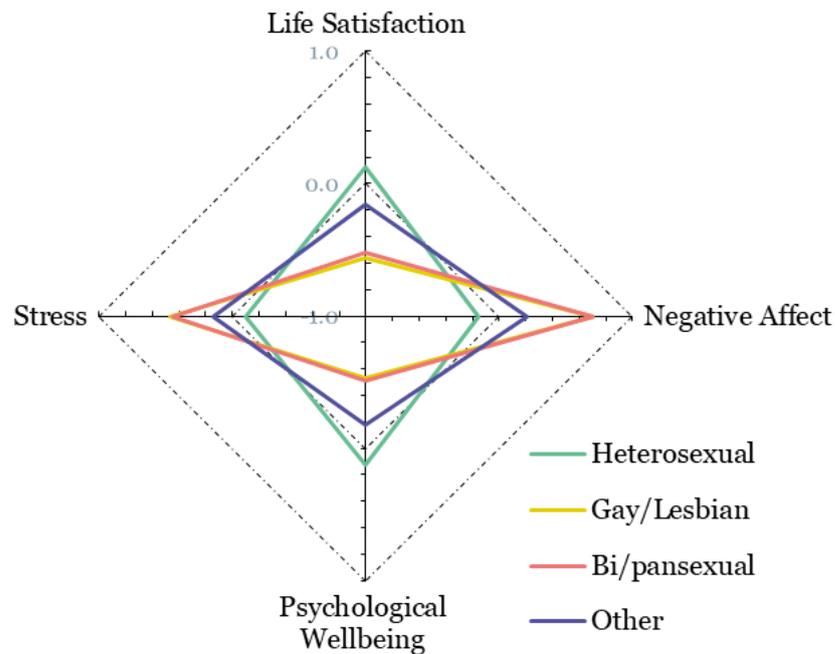


Table 4. Inequalities in wellbeing domains: sexual orientation comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
Hetero↔Gay/lesbian	-0.681***	0.857***	-0.651***	0.570***
Hetero↔Bi/pansexual	-0.649***	0.859***	-0.630***	0.551***
Hetero↔Other	-0.278***	0.362***	-0.307***	0.244***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Transgender status

Key findings: transgender status

Our analysis reveals multiple inequalities by transgender status. Transgender young people report significantly higher levels of stress and negative affect, and lower levels of life satisfaction and psychological wellbeing, than their cisgender counterparts. The size of these differences is between a quarter and a third of a standard deviation in each case. Again, the largest inequality is observed for negative affect, where transgender young people score approximately 1.5 points higher on the Me and My Feelings emotional difficulties subscale.

Figure 3. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by transgender status

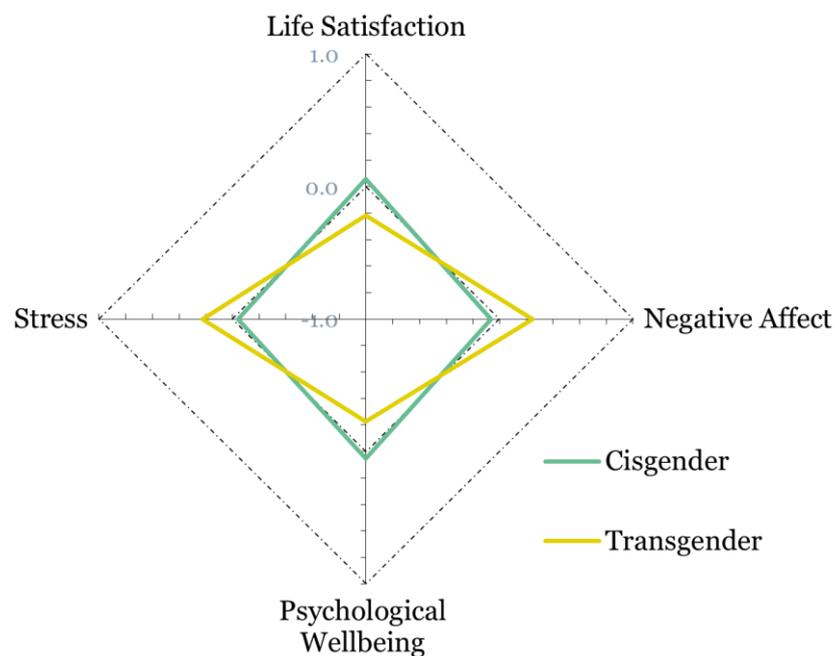


Table 5. Inequalities in wellbeing domains: transgender status comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
Cisgender↔Transgender	-0.274***	0.307***	-0.280***	0.265***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Socio-economic status

Key findings: socio-economic status

Our analysis reveals multiple inequalities by socio-economic status. However, these inequalities are less pronounced than for those already reported in relation to gender identity, sexual orientation and transgender status. Young people eligible for free school meals report significantly lower levels of life satisfaction and psychological wellbeing than their non-eligible counterparts. The size of this difference is, however, rather small, being just over one tenth of a standard deviation in both domains. This equates to approximately a difference of 0.3 points on the ONS4 Life Satisfaction item and 0.6 points on the SWEMWBS scale. Young people eligible for free school meals also report significantly higher levels of negative affect and stress than their non-eligible counterparts. In each case, the size of this difference is less than a tenth of a standard deviation.

Figure 4. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by FSM

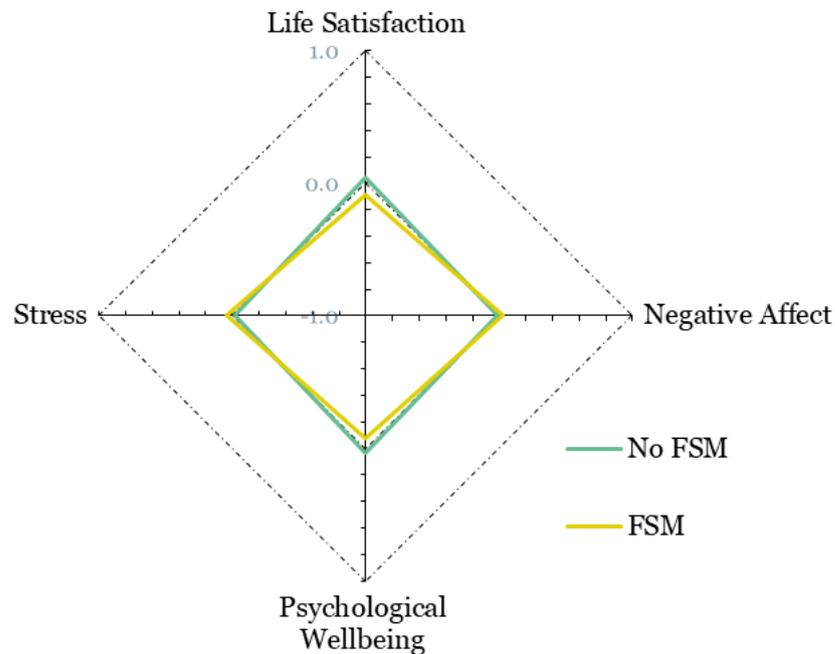


Table 6. Inequalities in wellbeing domains: socio-economic status comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
No FSM ↔ FSM	-0.126***	0.047***	-0.108***	0.063***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Key findings: ethnicity

Our analysis revealed a complex picture in relation to ethnicity that varies by wellbeing domain and ethnic group. First, we see that young people of Chinese and mixed ethnicity report significantly lower levels of life satisfaction than their white peers. Second, young people of black, Asian, mixed and unclassified ethnicity report significantly lower negative affect than their white counterparts. Third, compared to white young people, psychological wellbeing is significantly lower among those of black ethnicity, and significantly higher among those of Asian and any other ethnicity. Finally, Asian young people report significantly less stress than their white peers. As with the socio-economic analysis reported above, ethnic inequalities were much less pronounced than those by gender identity, sexual orientation and transgender status, with all significant differences across ethnic groups and wellbeing domains being less than one fifth of a standard deviation (and, in five of these cases, less than one tenth of a standard deviation).

Figure 5. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by ethnicity

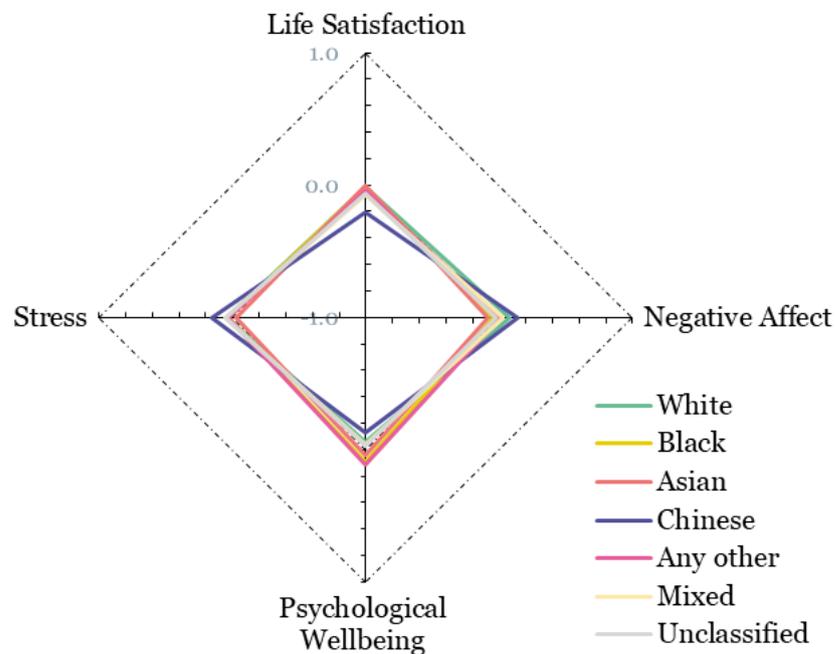


Table 7. Inequalities in wellbeing domains: ethnicity comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
White↔Black	0.004	-0.124***	-0.144***	-0.010
White↔Asian	0.017	-0.157***	0.099***	-0.061**
White↔Chinese	-0.191**	0.065	-0.078	0.118
White↔Any other	-0.029	-0.067	0.163***	-0.004
White↔Mixed	-0.056*	-0.051*	0.014	0.014
White↔Unclassified	-0.049	-0.096*	0.021	0.018

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Key findings: language

Our analysis reveals little evidence on inequalities by language. Indeed, the only significant difference found was in relation to stress, with those speaking English as an additional language reporting significantly lower stress levels than those speaking English as their first language. However, the size of this difference was negligible, being less than one twentieth of a standard deviation, or approximately 0.1 points on the PSS-4 stress measure.

Figure 6. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by EAL

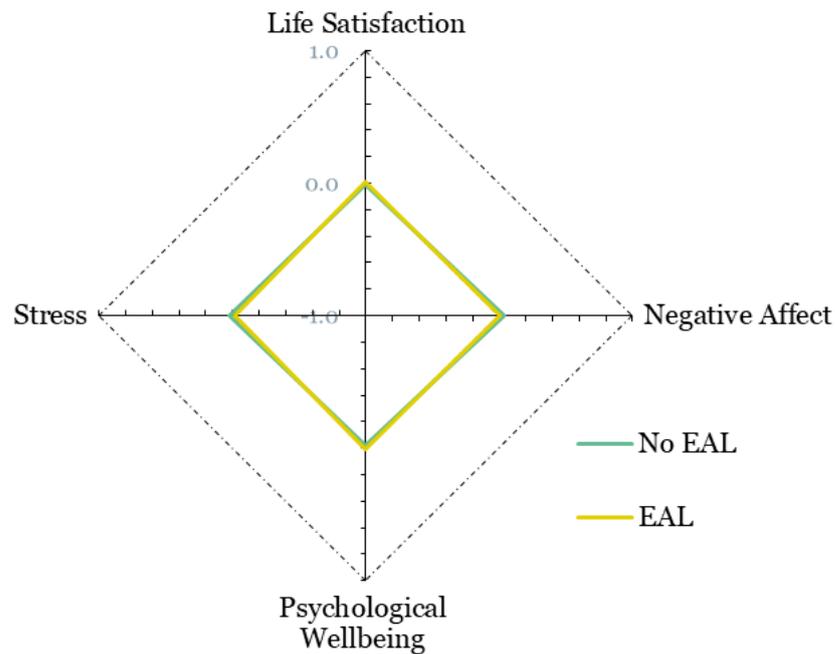


Table 8. Inequalities in wellbeing domains: language comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
No EAL ↔ EAL	0.025	-0.030	0.026	-0.040*

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Special educational needs and disabilities

Key findings: special educational needs and disabilities

Our analysis revealed no inequalities by special educational needs and disabilities for life satisfaction. However, young people with special educational needs (e.g. those in receipt of SEND Support or an Education, Health and Care Plan) reported significantly higher levels of negative affect and stress, and significantly lower levels of psychological wellbeing, than their peers without special educational needs. These inequalities were all relatively small, with each being less than one tenth of a standard deviation in size. For example, the difference in psychological wellbeing equates to just over 0.5 points on the SWEMWBS scale.

Figure 7. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by SEND

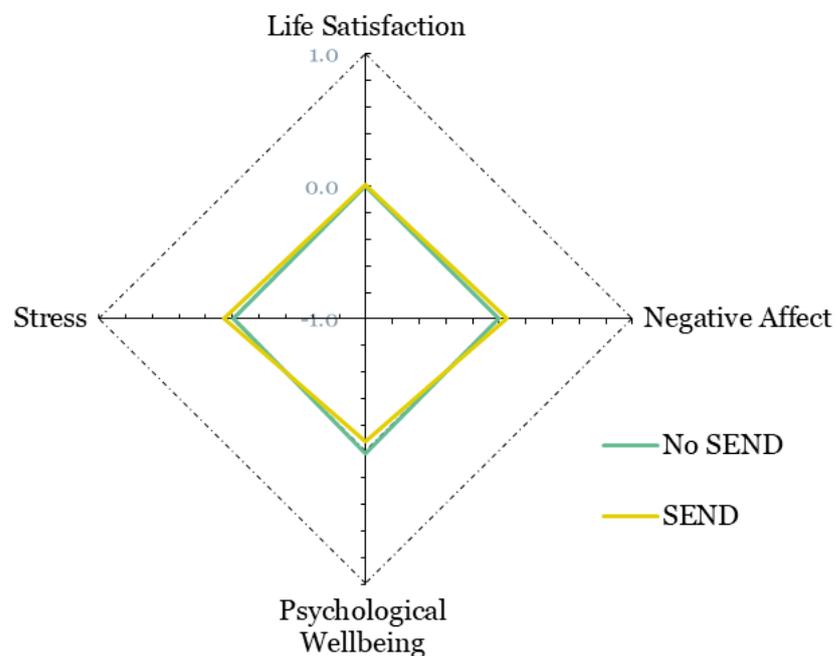


Table 9. Inequalities in wellbeing domains: special educational needs and disabilities comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
No SEND ↔ SEND	0.004	0.057***	-0.094***	0.067***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Caregiving responsibilities

Key findings: caregiving responsibilities

Our analysis reveals multiple inequalities by caregiving responsibilities. Those young people with caregiving responsibilities reported significantly lower levels of life satisfaction and psychological wellbeing, and significantly higher levels of negative affect and stress, than their peers without such responsibilities. However, the magnitude of these inequalities is relatively small, being on average about one tenth of a standard deviation. As an example, on the SWEMWBS psychological wellbeing scale, this equates to a difference of about 0.6 points.

Figure 8. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by caregiving responsibilities

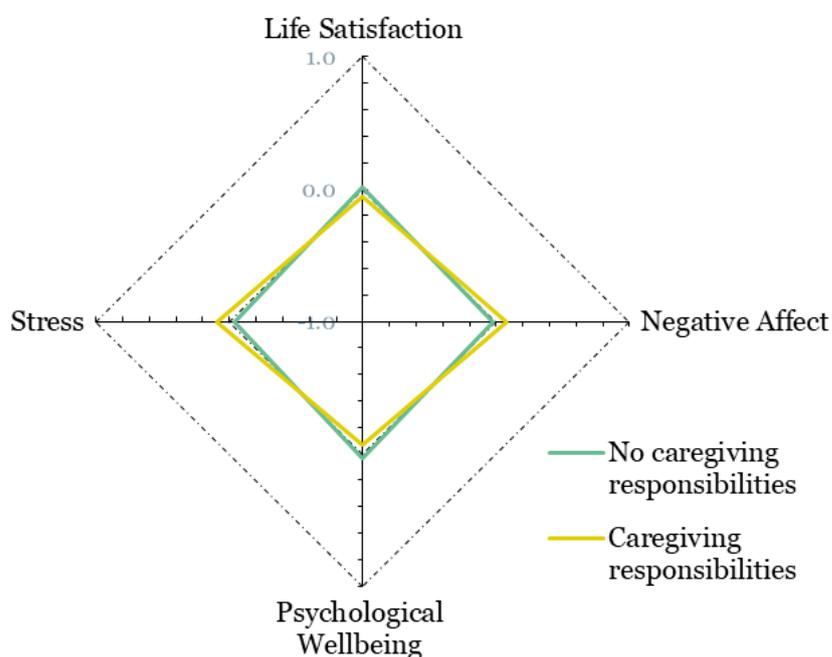


Table 10. Inequalities in wellbeing domains: caregiving responsibilities comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
No caregiving responsibilities ↔ caregiving responsibilities	-0.070***	0.102***	-0.102***	0.136***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Key findings: age

Our analysis reveals multiple inequalities by age. Young people in Year 10 reported significantly lower levels of life satisfaction and psychological wellbeing, and significantly higher levels of negative affect and stress, than their peers in Year 8. The largest gap is observed for life satisfaction – this is just under a quarter of a standard deviation, the equivalent to approximately 0.6 points on the ONS Life Satisfaction scale. The age differences for stress and psychological wellbeing are around one sixth of a standard deviation in both cases. Finally, the magnitude of inequality is much smaller for negative affect, being only around one sixteenth of a standard deviation.

Figure 9. Inequalities in Life Satisfaction, Negative Affect, Psychological Wellbeing and Stress, by year group

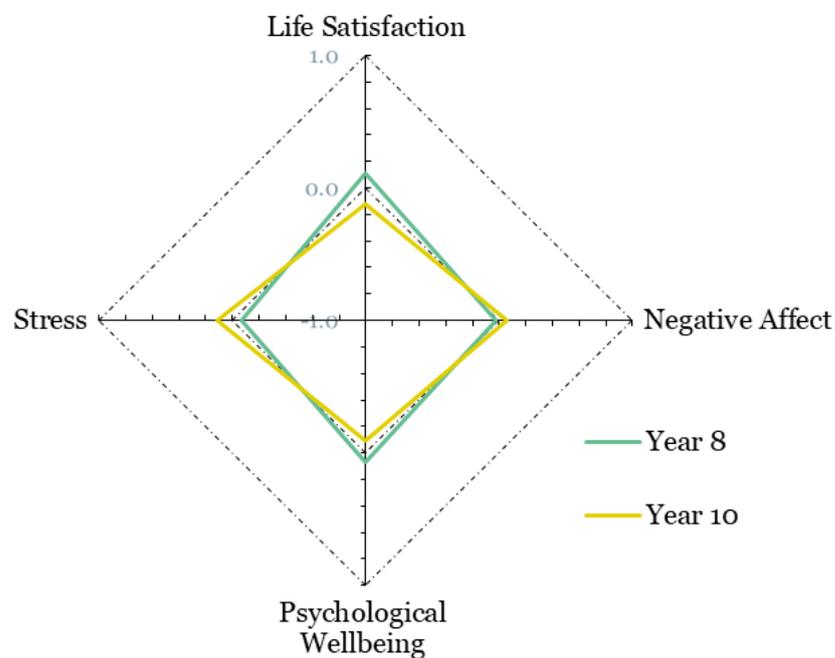


Table 11. Inequalities in wellbeing domains: age comparisons.

Comparison	Life Satisfaction	Negative Affect	Psychological Wellbeing	Stress
Year 8 ↔ Year 10	-0.227***	0.072***	-0.165***	0.178***

Notes: (1) Standardized scores (0 represents GM mean and 1 is the standard deviation); (2) Multi-level regression clustering for school effects and controlling for other inequalities dimensions; (3) *p-value≤0.05, **p-value≤0.01, ***p-value≤0.001.

Conclusion

It is possible to summarise our findings as follows:

- For most dimensions, we found minimal or mixed/inconsistent evidence of wellbeing inequalities. In cases where inequalities were identified, they were not so pronounced as to be practically meaningful. This statement can be applied to findings relating to ethnicity, language, age, socio-economic status, caregiving responsibilities, and special educational needs and disabilities. With one exception (the difference in life satisfaction between young people in Years 8 and 10), all of the group differences for these analyses were less than one-fifth of a standard deviation (which could be considered negligible).
- For some dimensions, we found consistent evidence of pronounced wellbeing inequalities that were routinely large enough to warrant significant concern. This statement can be applied to findings relating to gender identity, sexual orientation, and transgender status. All of the group differences for these analyses were in excess of one-fifth of a standard deviation, but were often much larger than this. Perhaps most notable were inequalities relating to negative affect, which were the most pronounced in all analyses relating to these dimensions, and reached nearly nine-tenths of a standard deviation in the comparisons between heterosexual and gay/lesbian or bi/pansexual young people.

We can have confidence in these findings due to the use of a very large sample whose composition mirrors the population of young people aged 11-16 in GM very well, and the implementation of robust analysis procedures (e.g. use of multi-level regression to account for data nesting, and multiple imputation to take account of missing data).

Perhaps the most significant implication of our findings is the urgent need to better understand how and why such marked wellbeing inequalities exist in relation to gender, sexual orientation, and transgender status, and to enact change across the system that can help to 'narrow the gaps'. In relation to the former, future analyses performed by the #BeeWell research team can be informative (for example, we plan to examine the role of bullying and discrimination as potential underpinning factors), alongside focused consultation with groups of young people who are subject to the inequalities reported here. In relation to the latter, we propose that the wellbeing of said groups should be made a focal priority among those who have committed to acting upon the #BeeWell data (for example the 100+ organisations that have joined the #BeeWell Coalition). There is also an opportunity for schools to use these findings to inform Personal, Social and Health Education and Relationships and Sex Education provision, particularly in relation to discrimination around gender identity and sexual orientation.

References

1. McCartney G, Popham F, McMaster R, Cumbers A. Defining health and health inequalities. *Public Health*. 2019;172:22–30.
2. Children's Society. *The Good Childhood Report 2021*. London: Children's Society; 2021.
3. Rapee RM, Oar EL, Johnco CJ, Forbes MK, Fardouly J, Magson NR, et al. Adolescent development and risk for the onset of social-emotional disorders: A review and conceptual model. *Behav Res Ther*. 2019; 123: 103501.
4. Vizard T, Sadler K, Ford T, Newlove-Delgado T, McManus S, Marcheselli F, et al. *Mental Health of Children and Young People in England, 2020*. London: NHS Digital; 2020.
5. Deighton J, Lereya ST, Casey P, Patalay P, Humphrey N, Wolpert M. Prevalence of mental health problems in schools: Poverty and other risk factors among 28 000 adolescents in England. *Br J Psychiatry*. 2019;215: 565-567.

Appendix 1. Technical details of the analysis

All analyses were conducted using Stata 15.

The estimates reported in the tables are the effects in a series of multilevel model with schools at level 2, students at level 1. Each multilevel model has a specific wellbeing domain as the dependent variable. Intra-cluster correlation coefficients were very small, as follows: 0.015 for life satisfaction; 0.028 for negative affect; 0.0158 for psychological wellbeing; and 0.011 for stress. Despite this, in each case the likelihood ratio test indicated that modelling as multi-level (as opposed to single level) significantly improved model fit. Thus, four multilevel models were estimated, with life satisfaction, negative affect, psychological wellbeing and stress as dependent variables, respectively. All four multilevel models have the same independent variables; that is, all the inequalities variables studied in the analysis (e.g. gender identity, sexual orientation, and so on). The four dependent variables were transformed into z-scores (mean=0, S.D.=1) to facilitate comparisons across different wellbeing domains.

Transgender status was defined as those reporting a gender identity different from the sex data provided by LAs. Given that the overwhelming majority of young people's self-reported gender identity is the same as their sex, there were multicollinearity effects between gender identity and transgender status data (in other words, cisgender was strongly correlated with male/female). Accordingly, these variables were not used together in the multilevel models. Instead, an additional set of four models was estimated using transgender status instead of gender identity in order to obtain the transgender status effects. For all the other inequalities examined, the model controlled for gender identity instead of transgender.

For illustration purposes, to visualize inequalities in the radar plots, the wellbeing standardized score of the reference category (e.g. males) is the constant in a multilevel regression model including this variable (e.g. gender identity) as the only independent variable. Thus, nine additional single-predictor multilevel models, one for each independent variable, were estimated for each of the four wellbeing outcomes. The constant in these models is the equivalent to the mean wellbeing score for this group (e.g. males) after accounting for the hierarchical structure of the data. The score for each of the other categories (e.g. females, non-binary, other) is estimated by summing up the effect estimates for each category (e.g. females, non-binary, other) in the multilevel model that controls for all the other inequalities to the score of the reference category (e.g. males), thus illustrating the distance in standard deviations of each of these categories from the reference category.

In all the models, missing data was handled via multiple imputation. The following variables had missingness at the indicated levels: gender identity (7.86%), transgender (13.20%), sexual orientation (9.39%), caring responsibilities (9.59%), SEN status (1.56%), free school meals eligibility (1.90%), English as a second language (1.28%), life satisfaction (9.64%), negative affect (12.19%), psychological wellbeing (14.07%), and stress (12.47%). To impute these data, we used a multivariate normal regression approach, which employs an iterative Markov chain Monte Carlo method to impute missing values. We performed 20 imputations of the data set. Missing data on ethnicity (1.89%) could not be imputed due to convergence problems with the imputation model. Thus, our final sample went from 37,978 to 37,260.

No weights were applied, since as noted in the main report, our sample was extremely close to the GM population of young people aged 11-16 for available demographic variables. We compared the sample that completed surveys at time one ($N = 37,978$), to published GM-level statistics via the government's education statistics tool (available [here](#)). Table S.1 provides an overview of the similarities between the survey sample and GM rates, as well as to a broader sample for whom demographic data were obtained, but who did not complete surveys ($N = 63272$; this represents the maximum possible sample based on pupils in eligible year groups among schools in GM who initially signed up to participate in #BeeWell). To note is that the data reported below do not take account of missing data; hence there are minor discrepancies with the equivalent data reported in the main document (Table 1), where missing data are reported.

Table S1. Demographic characteristics of the #BeeWell survey sample, broader sample, and GM population of young people aged 11-16.

	Sex (%)	FSM Eligibility (in last 6 years, %)	EAL (%)	SEND (%)	Year Group (%)	Ethnicity (%)
Survey sample	F = 49.17 M = 50.83	N = 74.58 Y = 25.42	N = 78.48 Y = 21.50	N = 85.79 Y = 14.21	8 = 53.67 10 = 46.33	AOEG = 2.31 Asian = 18.08 Black = 5.48 Chinese = 0.80 Mixed = 5.86 Unclassified = 1.89 White = 65.58
Broader sample	F = 49.29 M = 50.70	N = 70.49 Y = 29.51	N = 77.95 Y = 22.04	N = 82.90 Y = 17.10	8 = 50.96 10 = 49.04	AOEG = 2.58 Asian = 17.36 Black = 6.24 Chinese = 0.64 Mixed = 6.15 Unclassified = 2.11 White = 64.92
GM population of young people aged 11-16	F = 48.60 M = 51.40 *	Data not available	N = 78.46 Y = 21.10 **	N = 85.81 Y = 14.19	8 = 51.53 10 = 48.47 ***	AOEG = 2.77 Asian = 17.1 Black = 6.55 Chinese = 0.65 Mixed = 6.32 Unclassified = 1.28 White = 65.4
Max difference	.7	.4	.5	2.9 ****	2.7	1.07

* Note: based on 12/13/15/16 year-olds

** Note: secondary schools, including alternative provision

*** Note: calculated from SEN tables, total N = 67,374

**** Note: bigger discrepancy is between broader sample/GM than survey sample/GM