

Background Quality Report

Active Lives Children and Young People Survey – Academic year 2019/20

Relevance

The degree to which the statistical product meets user needs in both coverage and content.

Active Lives was developed in response to the new Government strategy – ‘Sporting Future: A New Strategy for an Active Nation’; and Sport England’s new strategy – ‘Towards an Active Nation’. We have designed Active Lives to give us and anyone working in this field detailed and reliable insight into the physical activity habits of the nation.

Active Lives consists of two similar, yet distinct, research projects: The Active Lives Adult Survey and the Active Lives Children and Young People Survey. Both surveys are administered by leading market research company IPSOS-MORI. This Background Quality Report focuses on the Active Lives Children and Young People Survey, which is an online-administered school-based survey.

The Children and Young People survey provides data for those aged 5-15/16 years across England, about the sport and physical activities they undertake and their attitudes to physical activity. The survey contains the following questionnaires:

- Standalone simple questionnaire for Year 1-2 (age 5-7 years) – collecting some activity data but mainly attitudes and wellbeing.
- Questionnaire for Year 3-11 (age 7-15/16 years) – with variations by year group within it and use of images for Year 3-6.
- Questionnaire for parents of Year 1-2 pupils – to gather detailed information on their child’s activities. This follows the same approach as the Year 3-11 questionnaire and was part of the same questionnaire program, so in practice it’s treated as one questionnaire with the Year 3-11 pupils.
- Teacher questionnaire – one per school to gather key data about the school such as time spent doing PE, facilities, active travel and healthy eating/school food standards to reduce burden on pupils for answering questions about facilities at school.

To develop the survey, an evidence review and focus groups were carried out to inform the design of the questionnaire and survey implementation. Questionnaires were tested using cognitive interviews with children and young people, parents and teachers. Cognitive interviews involve administering survey questions and exploring how respondents understand the questions, recall the information,



make judgements about their responses and then respond to the questions. A pilot test of the survey was also conducted in June to July 2017, and relevant changes were made as a result of this.

Each year, around 100,000 responses to the survey are targeted, with a minimum target sample size of 300 for each English local authority. The large sample size provides reliable and robust data for all questions, enabling subgroup analysis by demographics, by geographic area and by school term (excluding the summer holidays which isn't covered by the survey due to the school-based approach). In academic year 2019-20, the summer term sample was reduced in response to the Covid-19 pandemic, and as such sample targets per local authority were removed.

The Active Lives Children and Young People Survey collects data for key metrics including:

- Participation in sports and physical activity (done in the last week)
- Whether participated at school or outside of school in the last week (or indoors vs. outdoors)*
- Levels of activity (active, fairly active and less active based on average minutes activity per day)
- Swimming capability and confidence
- Sports spectating (watching live sport events at least twice in the last 12 months, for years 5 to 11)
- Volunteering (taking part in sports and physical activity volunteering 2+ times in the last 12 months, for years 5 to 11)
- Outcome and wellbeing measures (including happiness for years 3-11, worthwhileness and life satisfaction (and loneliness from academic year 2019-2020) for years 7-11. There is also a separate happiness measure for year 1-2 pupils.
- Attitudes towards sport and physical activity (enjoyment, confidence, competence, understanding for years 3-11 and knowledge for years 7-11). Year 1-2 pupils have separate enjoyment and competence measures.

* At and outside school activity was adapted to during and outside of school hours for those completing at home / not attending school during the pandemic. For more detail on questionnaire changes please see the Active Lives 2019-20 Technical Note.

The Active Lives Children and Young People survey achieves a high degree of relevance, providing the largest and most detailed physical activity survey for England, and covering whether the Chief Medical Officer's guidelines for physical activity are being met.

Accuracy

Accuracy is the proximity between an estimate and the unknown true value.

The Active Lives Children and Young People survey also achieves a high degree of accuracy. The data from both the Adults and Children and Young People surveys provides official statistics, complying with the Official Statistics Code of Practice. The Active Lives Children and Young People survey is published once per year, and results are released in December. There have been two previous Children and Young People releases: academic years 2017-18 and 2018-19 (published in December 2018 and December 2019 respectively). More information about publications can be accessed via this link: <https://www.sportengland.org/research/active-lives-survey/>

The Active Lives Children and Young People survey uses random sampling, which means that data can be analysed with smaller margins of error; as random sampling is more likely to be reflective of the population allowing accurate analysis. The sampling is conducted in three stages: 1) a sample of schools in England is selected using the Department for Education's School Census, 2) a sample of year groups within each school is randomly assigned using patterns of selection, and 3) Kish grids are used to randomly select classes within the selected year groups. All schools are eligible for the survey, using a probability proportional to size basis (with a few exceptions for: special schools, hospital schools, pupil referral units, very small schools, and schools with only one-year group, which aren't currently included in the survey). The main sample schools are also allocated to a term (autumn, spring or summer) using systematic random sampling. The sample is stratified by local authority, urban/rural indicator and pupil count. Schools can also opt into the survey, but these are not included in the national data for publication or archived data.

Response rates are also monitored to ensure the number of completes and partial completes by Local Authority, individual school, school type, year group, survey type and by demographic groups are in line with targets set for the term and completion rates from previous terms. Break off rates are also monitored. Through achieving 100,000+ completes each year, a high degree of accuracy is achieved, and the sample size enables collection of more accurate mean values (with larger sample sizes the identification of any outliers is easier and their effect diminished). Larger sample sizes also result in smaller margins of error, as it is more likely the sample achieved will be representative of the population at large. Additionally, during data editing, continuous variables within the data (e.g., minutes of activity) are checked for extreme values and outliers, both of which are set to missing to avoid skew within the mean calculations. This is done after grouped breakdowns have been generated (e.g., 30 to 149 minutes grouped, or 150 minutes+) as they do not impact on threshold allocation and we wish to maximise the data used for those headline metrics. As detailed in the 'Relevance' section above, in summer term, sample targets were removed and therefore responses in 2019-20 were fewer due to the coronavirus



pandemic (e.g. 89k responses instead of just over 100k in the main dataset). The response size was still deemed sufficient for national analysis, however the decision was taken not to publish local authority (LA) data due to smaller response numbers by LAs.

The Active Lives Children and Young People survey involves detailed quality assurance processes, and thorough checking is carried out at all stages of the project. Any questions considered for the survey are subject to cognitive testing prior to being added, to ensure the question is interpreted correctly and not misleading or overly taxing for respondents. The questionnaire is also checked by multiple people and across multiple devices. Throughout the fieldwork, a coding scheme is used for back-coding of activities – any which cannot be done automatically using this coding scheme are checked manually by Sport England. At the end of the process, all 'other' answers are assigned a code which indicate which type of activity it is, which can then be used in the derivation of the participation and composite variables. Data checking is also carried out to remove illegitimate responses and ensure patterns of responses make sense against what would be expected. Further information about the data checking, alongside derived variable creation and checks, can be found in the survey's full technical report also provided on UK Data Archive.

Sport England's checking processes include various stages. When the dataset arrives from IPSOS-MORI, there is time set aside for quality assurance. We make sure all variables and survey responses are contained as expected, checking all cases are included within the datasets, and that the file size and structure are similar to previous years. The term variable is checked to ensure an even distribution of survey responses across the year, and the Region / Child Indices of Multiple Deprivation / Local Authority variables are checked to ensure a representative spread of survey responses across England. All variables are reviewed, to make sure they are named and labelled correctly, and missing values are correctly set where appropriate. It is also checked that there aren't too many missing values at questions where there should be data. All headline measures are compared to previous years, and composite / derived variables are checked against the 'building block' variables to ensure data accuracy. Where activity disciplines are grouped, we make sure that all sub-disciplines are correctly feeding into the activity (e.g. Cycling for travel is checked to be included in the overall Cycling category). During academic year 2019-20, due to the Covid-19 pandemic there were reduced response numbers in summer term as a result of reduced sampling. This created an uneven distribution in the spread of responses across the year by both term and school phase – in addition to reduced summer responses, there was also a greater proportion of secondary responses as at home completion was easier for this age group. Therefore, it is advised when analysing academic year 2019-20 data, to analyse within term and/or school phase too. Extra profiling analysis within term was also undertaken to understand the impacts of this change.

IPSOS-MORI provide the weighting plan file, using the latest ONS statistics, to ensure results are nationally representative and therefore as accurate as possible for the population of England, reducing risk of bias in survey estimates and to enable fair comparison of results between survey years. The



weighting is done separately for the three sets of year groups (years 1-2, years 3-6 and years 7-11). The weighted and unweighted demographic profiles are reviewed by Sport England, and compared to previous years and national population statistics to ensure the profile is as expected. The range for weights is reviewed to ensure that any single responses aren't give too much emphasis within the final weighted data. More detail about the weighting procedures can be found within the survey's full technical report. Whilst the results are weighted to the demographic profile of England, as well as by geography of England, it is not also possible to weight each academic year by term and geography. Given the sampling is done by local authority, priority is given to geography within the survey weighting.

Sport England and the University of Edinburgh have worked together to fully check the analysis and published tables. A quality assurance checklist is adhered to. Weighted and unweighted bases and counts, rates, shares, confidence intervals, standard errors, population estimates, and significance tests are checked to be correct against the raw SPSS data outputs and significance test calculations. We make sure any population estimates calculated are using the most up-to-date population figures available. Published tables are checked to ensure rates are shown to 1 decimal place, time periods are shown separately, column headings are correct and showing respondent numbers, etc. We also make sure any new local authorities are included where appropriate.

Additionally, for academic year 2019-20, due to the Covid-19 pandemic and school sites closing to most pupils, the survey was adapted to allow at home completion (rather than at school completion). As a result of this it was important that we enabled the survey to be potentially completed by smartphone (as opposed to just PCs and tablets). Analysis was undertaken to review the impact of completing the survey using a smartphone, to ensure that the data still provided the same level of detail, reliability and validity. This analysis did not present any concerns from the increased usage of smartphones. Analysis was also undertaken to review the impact of the changed situation due to Covid-19 (reduced holidays, increased usage of laptops at home etc.) on our measure of Family Affluence provided by the survey. As the Covid-19 pandemic came about towards the end of the survey year, primarily only affecting the summer term responses, it was deemed that the impact on the family affluence measure was minimal for the academic year 2019-20 (given the questions have a one-year reference period), and therefore no changes were made.

Reliability

Reliability is the closeness of early estimates to subsequent estimated values.

Through the use of a continuous tracking survey, dating back to academic year 2017-2018, the Active Lives Children and Young People survey displays a high degree of reliability. Repeated measures for key metrics have been taken over two time points using independent samples, measuring and assessing



change, at regular yearly intervals. As there have been two previous Children and Young People releases, it is currently too early to talk about change over time, for which we recommend using four years' worth of data; however, between the two Children and Young People releases major shifts in findings and unexplainable results have not been found.

The Active Lives Children and Young People survey also uses detailed statistical tests. With every publication, confidence intervals are included for any percentage rates shown. The Complex Samples module in SPSS used takes account of the survey design effects. In addition to this, other statistics such as shares, standard errors, weighted and unweighted bases and counts are analysed. When reviewing change over time, statistical significance tests are used and only significant differences are reported.

Timeliness and Punctuality

Timeliness refers to the time gap between publication and the reference period. Punctuality refers to the gap between planned and actual publication dates.

The time between the reference period e.g., academic year 2019-20, to publication, provides time for quality assurance processes, back coding, data preparation, data analysis and production of the final report (ahead of publication). Active Lives release dates are outlined in the DCMS Official Statistics Release Calendar. Active Lives releases have always been delivered punctually, meeting the agreed dates outlined in the DCMS Calendar. For 2019-20, the publication date was changed to 14th January 2021 (instead of the 3rd December 2020); this was because of the extra time required to quality assure and analyse the data as a result of the Covid-19 pandemic.

Accessibility and Clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

The Active Lives Children and Young People survey data is made accessible to users. The Sport England website contains detail about the scope and background, data collection methodology, survey development and sampling, survey measurements, supporting information for schools, and data and reporting archive. With each release, documents published on the Sport England website include the full written report, published tables and technical note.

The published tables are made as easy to navigate as possible for users, with an index page taking users to relevant tabs and grouped columns to review relevant survey time points. The total number of respondents unweighted is shown for each breakdown variable, and rates, confidence intervals and population estimates are included. Significant changes are also highlighted. In addition to this, the technical note contains further supporting information about survey design, weighting, confidence intervals, significance testing, population estimates, questionnaire content and definitions from the report.

The Active Lives survey provides open data, and datasets are shared with anyone who requests a copy for analysis and signs the user agreement. When datasets are shared, supporting documents are provided including the Active Lives user guide and data code book. These documents are also uploaded to UK Data Archive, alongside the full Active Lives technical report, where users can download the files themselves if they have created an account.

Comparability and Coherence

Comparability is the degree to which data can be compared over time and domain. Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar.

As the Active Lives Children and Young People survey has been running since academic year 2017-2018, we have now produced three publications. Whilst it is still too early to conduct time series analyses, we are able to compare results between these three time points. Results so far have proven to be highly comparable, with no major or unexplainable shifts in the findings.

Results are compared to the NHS Health Survey for England, which monitors trends in the nation's health and care. It provides information for children aged 0 to 15.

The Health Survey for England (HSE) also provides comparable statistics around the %s of Children and Young People who are active or less active based on their minutes of moderate or vigorous activity each week (although including activities such as gardening which fall outside of Sport England's remit). The activity measure is similar to Active Lives, which also classifies those doing less than 30 minutes of moderate equivalent activity per week as 'less active' (or low activity for HSE).