



# **Growing Up in Scotland: Birth Cohort 2, Sweep 3**

## **User Guide**

**Date:** June 2018

**ScotCen**  
Social Research that works for society



<b>1</b>	<b>Overview of the survey</b>	<b>3</b>
1.1	Study Design	3
1.2	Sample Design	4
1.3	Data collection	5
1.4	Response rates	5
1.5	Objective assessments and measurements	6
1.5.1	Cognitive assessments: British Ability Scales (BAS-3) – Naming Vocabulary and Picture Similarities	6
1.5.1	Height and weight measurements	8
<b>2</b>	<b>Using the data</b>	<b>10</b>
2.1	Documentation	10
2.2	The data file	10
2.3	Variables on the data file	10
2.4	Variable names and labels	11
2.4.1	Variable naming convention	11
2.4.2	Variable labels	11
2.4.3	Derived variables	11
2.5	Survey weights	12
2.5.1	Sweep 3 weights: longitudinal and cross-sectional	12
2.5.2	Longitudinal weights for sweep 3	12
2.5.3	Cross-sectional weights for data from the BC2 Sweep 3	14
2.5.4	Sample efficiency	14
2.6	Comparing data between cohorts	14
2.7	Multicoded questions	15
2.8	Dropped Variables	15
2.9	Missing values conventions	15
<b>3</b>	<b>Data content</b>	<b>16</b>
3.1	Information about the household	16
3.2	Main socio-economic variables	17
3.2.1	National Statistics Socio-economic Classification (NS-SEC)	17
3.2.2	Equivalised household income (annual)	18
3.3	Area-level variables	18
3.3.1	Area-level variables: Scottish Government Urban/Rural Classification	18
3.3.2	Scottish Index of Multiple Deprivation	19
3.4	Topic measurements and instruments	20
3.4.1	Child Development: Strengths and Difficulties Questionnaire	20
3.4.2	Parental Health: Medical Outcomes Study 12-Item Short Form (SF-12)	20
3.4.3	Socioeconomic characteristics: Selected items from Index of Material Deprivation	21
3.4.4	Gender Role in Children: Selected items from the Pre-school Activities Inventory	22

---

3.4.5	Child’s self-control: Child Behaviour Questionnaire .....	22
3.4.6	Parental Styles: The Parenting Practices Questionnaire .....	23
3.5	Cognitive assessments: British Ability Scales (BAS3) – Naming Vocabulary and Picture Similarities.....	24
3.5.1	Score variables in the dataset.....	24
3.5.2	Influences on test scores.....	25
3.6	Height and weight and BMI data.....	26
3.6.1	Overview: BMI.....	26
3.6.2	GUS BC2 sweep 3 BMI measures .....	26
<b>4</b>	<b>More about GUS .....</b>	<b>29</b>
4.1	Contact details .....	29
<b>5</b>	<b>References .....</b>	<b>30</b>
<b>6</b>	<b>Appendix .....</b>	<b>31</b>

---

# 1 Overview of the survey

Growing Up in Scotland (GUS) is a longitudinal research study aimed at tracking the lives of three cohorts of Scottish children from the early years, through childhood and beyond. Funded by the Scottish Government Education Directorate, its principal aim is to provide information to support policy making, but it is also intended to be a broader resource for secondary analysis.

The overarching aim of the study is set out in its purpose, which is:

To generate, through robust methods, specifically Scottish data about outcomes throughout childhood and into adulthood for children growing up in Scotland across a range of key domains:

- Cognitive, social, emotional and behavioural development
- Physical and mental health and wellbeing
- Childcare, education and employment
- Home, family, community and social networks
- Involvement in offending and risky behaviour

Such data will encompass, in particular, topics where Scottish evidence is lacking and policy areas where Scotland differs from the rest of the UK.

## 1.1 Study Design

GUS was designed to provide cross-sectional and longitudinal data at a national level about children who were born in a particular year and who had survived until the first point of data collection.<sup>1</sup>

GUS was initially based on two cohorts of children. The first (known as ‘birth cohort 1’) involved 5217 children at the first sweep of data collection who were aged approximately 10 months at the time of first interview. The second (known as the ‘child cohort’) involved 2858 children at the first sweep of data collection who were aged approximately 34 months at the time of first interview. The second birth cohort (known as ‘birth cohort 2’) was recruited in 2011, and involved 6127 children at the first sweep of data collection who were aged approximately 10 months at the time of first interview.

All cohorts were named samples drawn from Child Benefit records. The configuration of cohorts and sweeps for all sweeps of data collection launched to date is summarised in Table 1.1.

Throughout the text BC1 refers to the older of the two birth cohorts (‘birth cohort one’), and BC2 to the most recent birth cohort (‘birth cohort two’).

---

<sup>1</sup> Because the study does not include children who were stillborn or who died prior to the first contact, and because it includes children who were born outside of Scotland, the sample is not representative of *all births* in Scotland the reference period. Rather it is representative of children of the relevant age living in Scotland at the time of data collection who also lived in Scotland at the age of 10 months.

Child's age at data collection	Cohort/Year of data collection		
	Child cohort	Birth cohort 1	Birth cohort 2
10 months		2005/06	2011
1-2 yrs (22 months)		2006/07	
2-3 yrs (34 months)	2005/06	2007/08	2013
3-4 yrs (46 months)	2006/07	2008/09	2014*
4-5 yrs (58 months)	2007/08	2009/10	2015
5-6 yrs (70 months)	2008/09	2010/11	
7-8 yrs		2012/13	
9 yrs (Primary 5)		2013/14*	
10 yrs (Primary 6)		2014/15	
11 yrs (Primary 7)		2016/17*	
12 yrs (Secondary 1)		2017/18	

\*All data collection conducted face-to-face except in years indicated with a \* when data was collected via web and telephone surveys.

A key aim of using multiple cohorts is to allow the study to provide three types of data:

- Cross-sectional time specific data – e.g. what proportion of 3 year old children lived in single parent families in 2013?
- Cross-sectional time series data – e.g. is there any change in the proportion of 3 year old children living in single parent families between 2007/08 and 2013?
- Longitudinal cohort data – e.g. what proportion of children who were living in single parent households aged 0-1 are living in different family circumstances at age 4-5?

## 1.2 Sample Design

The sample design for BC2 was similar, but not identical, to that used for the two previous cohorts (BC1 and the child cohort). The initial area-level sampling frame of Primary Sampling Units (PSUs) was created by aggregating Data Zones. Data Zones are small geographical output areas created for the Scottish Government. Data Zones are used to release data from the Census 2001 and other neighbourhood statistics. Each PSU was roughly equal in size ('size' = expected number of births, based on the average number of births in each Data Zone for the preceding 3 years). 160 PSUs were selected with equal probability.

Child Benefit records held by HMRC were used to identify ALL children who were within the eligible age range (born between 1<sup>st</sup> March 2010 and 28<sup>th</sup> February 2011) and living in the selected PSUs. In households with twins only one child was selected to reduce burden on the parents.

---

A letter was sent to the Child Benefit recipient in all eligible households informing them that they had been selected for inclusion in the study. Recipients were able to opt-out of the study at this stage. Those cases which did not opt out were deemed eligible for interview.<sup>2</sup>

Further information on the BC2 sample design is provided in the BC2 sweep 1 User Guide.

## 1.3 Data collection

Data collection for BC2 sweep 3 involved a face to face interview with the child's main carer, cognitive assessments conducted with the child and height and weight measurements taken of the child.

Interviews were carried out in participants' homes by trained social survey interviewers using laptop computers (otherwise known as CAPI – Computer Assisted Personal Interviewing). The interview was quantitative and consisted almost entirely of closed questions. Overall, the average (median) interview (including the cognitive assessments and height and weight measurements of the child) lasted 77 minutes. In addition to the interviewer-led sections there was a brief self-complete section in the interview in which the respondent, using the interviewer's laptop, input their responses directly into the questionnaire program.

Further details of the cognitive assessments and height and weight measurements are provided in section 1.4.

At sweep 3, primarily because of the use of data fed forward from previous sweeps, interviewers were instructed as far as possible to undertake the interview with the person who took part in the sweep 2 interview. Where the previous respondent was not available, interviews were undertaken with another resident carer. In the vast majority of cases, interviews were conducted with the child's mother (97%). In most of the remaining cases the respondent was the child's father. A small number of interviews were conducted with the child's grandparent.

Fieldwork was undertaken over a fourteen month period commencing in January 2015. The sample was issued in twelve monthly waves at the beginning of each month and each month's sample was in field for a maximum period of two and a half months. For example, wave 2 was issued at the beginning of February 2015 and remained in field until mid-April 2015.

To ensure that respondents were interviewed when their children were approximately the same age, each case was assigned a 'target interview date'. This was identified as the date on which the child turned 4 years and 10.5 months old. Interviewers were allotted a four-week period based on this date (two weeks either side) in which to secure the interview. In difficult cases, this period was extended up to and including the child's subsequent birthday which allowed a further four weeks.

The majority of interviews took place when the child was aged 59 months (70%). A further 24% took place at 58 months and 3% at 60 months. Just 2% of interviews took place when the child was aged 61 months. The remainder of interviews (around 2%) took place when the child was aged between 62 and 64 months.

## 1.4 Response rates

---

<sup>2</sup> Details of the number of eligible cases identified, the number of opt-outs and so on, is provided in section 1.4

Details of the number of cases issued and achieved and the response rates are presented in Table 1.3.

Table 1.2 Response	
Achieved interviews at sweep 1	6127
Achieved interviews at sweep 2	5020
Cases issued at sweep 3	
All issued to field	5742
Eligible (i.e. child is alive and still living in Scotland)	5678
Cases achieved at sweep 3	4434
Response rate at sweep 3	
% of all eligible cases at sweep 3	78%
% of all sweep 2 cases	88%
% of all sweep 1 cases	72%

## 1.5 Objective assessments and measurements

### 1.5.1 Cognitive assessments: British Ability Scales (BAS-3) – Naming Vocabulary and Picture Similarities

#### Overview

As part of the data collection for sweep 3, the cohort children were asked to complete two cognitive assessments. The assessments – Naming Vocabulary and Picture Similarities – were taken from the Early Years battery of the ‘British Ability Scales Second Edition’ (BAS-3). The British Ability Scales (BAS) is a battery of individually administered tests of cognitive abilities and educational achievements suitable for use with children and adolescents aged from 2 years 6 months to 17 years 11 months. The assessments are normally employed by educational psychologists in a classroom or clinical setting but have been adapted for use in a survey setting, and modified to be administered with the help of a CAPI program pre-determining the complex set of rules for routing children through each assessment. The purpose and method of each assessment is described in Table 1.2. The data is used to estimate an approximate score for each child.

**Table 1.3** Details of cognitive assessments used at sweep 3

Assessment name	Assesses	Method	Max no of items
BAS – Picture Similarities	Non-verbal reasoning	Child is shown a row of 4 pictures and is given a card with a 5 <sup>th</sup> picture. The child places the card under the picture which shares an element or concept with the card.	33
BAS – Naming Vocabulary	A verbal task which concerns knowledge of names	Child is shown a picture and asked to say its name.	36

### Comparison with BC1

BC1 children also carried out Naming Vocabulary and Picture Similarities exercises when they were the same age (58 months). However, different editions of the assessments were used: For BC1, the 2<sup>nd</sup> edition assessment was used (BAS-II), whereas for BC2 the 3<sup>rd</sup> edition was used (BAS-3). Whilst the assessments are almost identical, there are a small number of differences – for example in the individual items, the order of the items and the stopping points – which would introduce caveats when making a straightforward comparison of ability scores.

To allow for comparison between the cohorts, the assessment authors provided the GUS team with a calibration formula to be applied to the BC1 scores. Once applied, the revised scores can be used in comparisons between the cohorts. No adjustment is required to the BC2 scores.

***Note that any comparisons of cognitive ability scores across the cohorts MUST use the adjusted BC1 scores. If you are using an old BC1 dataset this variable may not be available. If you need the revised BC1 score variables and these are not available in your BC1 dataset, please download the latest version from UKDS.***

### BAS Naming Vocabulary

Naming Vocabulary assesses the spoken vocabulary of young children. The test items consist of a booklet of coloured pictures of objects which the child is shown one at a time and asked to name. The scale measures expressive language ability, and successful performance depends on the child's previous development of a vocabulary of nouns. Picture recognition is also crucial; however, the pictures are large and brightly coloured and are unlikely to cause problems except for children with major visual impairments or with no experience of picture books. The items require the child to recall words from long-term memory rather than to recognise or understand the meaning of words or sentences.

Naming Vocabulary score may reflect:

- Expressive language skills
- Vocabulary knowledge of nouns
- Ability to attach verbal labels to pictures

- 
- General knowledge
  - General language development
  - Retrieval of names from long-term memory
  - Level of language stimulation

Low scores may reflect reluctance to speak.

## **BAS Picture Similarities**

Picture Similarities measures the reasoning ability of young children. The test items consist of a booklet with four images on each page and a set of cards each with a single image printed on. The child is shown the row of pictures, given a corresponding card and asked to place the card under the image on the page which shares an element or concept with the image on the card. To undertake the task, the child must identify various, potentially relevant, features of the images and determine which feature the target picture on the card shares with only one of the four possible images on the page. Whilst speech is not required, good verbal-encoding may well help the child solve the problems.

Picture Similarities scores may reflect:

- Non-verbal problem solving (inductive reasoning)
- Visual perception and analysis
- Ability to attach meaning to pictures
- Ability to develop and test hypotheses
- Use of verbal mediation
- General knowledge

Low scores may also reflect impulsiveness (responding without checking the response).

## **Further information**

For more information about the development, administration, scoring and interpretation of the BAS assessments see:

- Elliott, C.D., Smith, P, and McCulloch, K (2011) British Ability Scales Third Edition
- BAS3 British Ability Scales Technical Manual. London: GL Assessment.

## **Obtaining consent for child assessments**

Before undertaking the assessments with the child, the child's parent or carer was required to give informed consent. A bespoke information leaflet detailing the object and content of the assessments was given to parents by the interviewer. After reading the leaflet, parents were then asked to sign a consent form permitting the assessments to go ahead. Levels of consent to undertake the assessments were very high at 97% for Picture Similarities and 98% for Naming Vocabulary.

## **1.5.1 Height and weight measurements**

### **Overview**

---

The relationship between general build and health is of great interest to the Scottish Government, especially in relation to children. This is particularly so, as both the height and the weight of the population appear to have been changing very rapidly over the last two decades. These changes reflect things like changes in children's diet and lifestyle. Although many parents do know their child's height and weight, these measurements are not always up to date or are not known with the precision required for detailed analyses. By including objective measurements of height and weight, GUS provides a reliable and extensive source of data for researchers interested in matters related to children's health.

## Carrying out the measurements

Interviewers were asked to measure the height and weight of all children in the study. However, in some cases it may not have been possible or appropriate to do so, for example if it was clear that the child was unwilling or that the measurement would be far from reliable.

It was recommended that height and weight measurements be taken on a floor which was level and not carpeted. If all rooms in the household were carpeted, a floor with the thinnest and hardest carpet was chosen (usually the kitchen or bathroom).

For the weight measurements, there was an option to weigh the child whilst being held by an adult. In this case, the adult was weighed on his/her own first and then the adult and the child were weighed together. Both weights were entered into the interviewer's laptop which calculated the child's weight.

The interviewer was asked to code whether they experienced problems with the height and/or weight measurements and, if they did, to indicate whether they felt the end result was reliable or unreliable at WeXhei14 and WeXwei19. As a rough guide, if the measurement was likely to be more than 2 cms (3/4 inch) from the true figure for height or 1 kg (2 lbs) from the true figure for weight, it was coded as unreliable.

If the adult respondent was not willing to allow the sample child to have his/her height or weight measured, for example saying that they were too busy or already knew their measurements, a Refusal code was entered for the measurements variables WeXhei01 and WeXwei01, with the reason for refusal at WeXhei02 or WeXwei02.

If the height or weight was refused or not attempted, the respondent was asked to estimate their child's height or weight, in metric or imperial measurements.

Detailed protocols of how to take height and weight measurements are included as appendices to the main interviewer instructions deposited with the dataset and available from the data archive website.

The data has been used to estimate an approximate BMI (Body Mass Index) score for each child. Details on the data and variables associated with the height and weight measurements can be found in section 3.6.

---

## 2 Using the data

### 2.1 Documentation

The documentation is provided as 5 PDF files organised into the following sections:

- A representation of the CAPI questionnaire with variable names added and the show cards used during the interview.
- The Data Documentation comprising
  1. list of variables in the dataset (including derived variables)
  2. list of derived variables with their SPSS syntax
- Project instructions containing interviewer and coding instructions.

### 2.2 The data file

The GUS BC2 sweep 3 data consists of one SPSS file:

GUS_BC2_SW3	4434 cases	Birth Cohort 2
-------------	------------	----------------

### 2.3 Variables on the data file

The data file contains questionnaire variables (excluding variables used for administrative purposes) and derived variables. The variables included in the file are detailed in the “Variable List” document in the data section of the documentation. As far as possible, they are grouped in the order they were asked in the interview. As such, following the questionnaire design, they are grouped according to topic. This document is the best place to look/search for variables when planning your analysis.

Once you have decided which variables to include in your analysis, you should look up details of the question wording using the questionnaire documentation (all variables on the data file are given by name in the copy of the interview schedules provided), or use the data documentation to find the syntax which produced the derived variables. You cannot rely on the individual variable and value labels to always capture the detail of the question asked, or the answer categories used, so reading the interview documentation is essential.

For variables with answers following a scale, such as ‘Strongly agree’ to ‘Strongly disagree’ it must be noted that the order of the answer categories may not follow systematically an ascending or descending scale throughout the list of variables. Also the answers may equally refer to positive or negative statements. The phrasing of the question and the list of answers provided on the show cards - if any - shape the variables. The user must therefore take these variations into account when creating derived variables.

## 2.4 Variable names and labels

### 2.4.1 Variable naming convention

Variables names are normally made up of 8 characters, the first indicates the source of the variable, the second the sweep of data collection and the rest is an indication of the question topic. Therefore where the same question was asked in the different sweeps, the names will usually be the same apart from the second character. If the same question was asked of BC1 respondents at the equivalent sweep<sup>3</sup>, the variable name will be identical. If a variable name has changed substantially between sweeps or cohorts this is marked in the variable list. The naming convention is summarised in Table 2.1

Source of data		Sweep/Sweep		Key theme prefix	Sub-theme stem	Variable number/ abbreviation
Non- sequential Capitals: M, D, AL		Sequential lower case: a, b, c.		Non-sequential Capitals: e.g. C, P, N...	Abbreviated lower case: e.g. hea	e.g. 01 – 99 or lp, nc, cb
Source code	Details	Sweep code	Details			
AL	Area Level variable	a	BC2 Sweep 1 (2011)			
D	Derived variable	c	BC2 Sweep 2 (2013)			
M	Main carer interview	e	BC2 Sweep 3 (2015)			

### 2.4.2 Variable labels

The variable labels have been shortened to be 80 characters or fewer as far as possible; the first two show the source and year of the data (as in the variable name). Although the labels give an indication of the topic of the question it is essential to refer to the questionnaire to see the full text of the question and the routing applied to that variable.

### 2.4.3 Derived variables

Derived variables included in the dataset are listed with the questionnaire variables for the same topic. The SPSS syntax used to create them can be found in the “Derived Variables” section of the documentation.

<sup>3</sup> At sweep 3, the children in BC2 were the same age as the children in BC1 were at sweep 5.

## 2.5 Survey weights

Weighting has been used to correct for different selection probabilities (in households with twins) and for non-response. After applying the weights, results from the GUS sample will be representative of all children who were aged just under 5 years old during the fieldwork period (2015) and who were living in Scotland at age 10 months.

### 2.5.1 Sweep 3 weights: longitudinal and cross-sectional

Of all sweep 3 respondents, 96.1% had taken part in the sweep 2 interview. This means that there is a wealth of information available for both respondents and non-respondents at sweep 2 that can be used to model response behaviour.

The sweep 3 sample of respondents can be split into two components. For the purposes of describing the weighting these have been named Sample A and Sample B and are defined as follows:

- Sample A – mothers who had responded at all three sweeps
- Sample B – mothers who had responded at sweep 1 but had missed interview in sweep 2

The two samples are treated separately during the weighting. This is because the Sample B respondents are likely to have different response behaviour to those in Sample A, as suggested by their much lower response rates (compared in the table below).

Table 2.2 Response rates for different samples of main interview respondents

	Issued	Responding	Response rate
Sample A	4973	4260	85.7%
Sample B	762	174	22.8%
Combined (A+B)	5735	4434	77.3%

Two sets of weights were developed for the responding adults: a cross-sectional weight and a longitudinal weight. Only members of Sample A (who had responded to all three sweeps) have a longitudinal weight.

All BC2 sweep 3 respondents will have a cross-sectional weight (Sample A + B).

### 2.5.2 Longitudinal weights for sweep 3

A model-based weighting technique was used to develop the sweep 3 longitudinal weights, where response behaviour is modelled using data from previous two sweeps. This is the same method used to generate weights for adults who completed the main interview at sweep 2. Ineligible households (deadwood) were not included in the non-response modelling.

The first step was to carry out bivariate analysis. Crosstabs and chi square tests were used to identify variables that were related to response behaviour. All variables were entered into a non-response model.

Non-response behaviour was modelled using logistic regression. This is a method of analysing the relationship between an outcome variable (in this case response to the sweep 3 interview) using a set of predictor variables (characteristics about the respondent and their household taken from the sweep 2 survey data, plus urban/rural and deprivation indicators). The model takes account of the relationship of the predictor variables to the outcome and the relationships of the predictor variables to each other.

Variables found to predict response for the child cohort at sweep 3 are given in the Table below. All of them were entered in the main non-response model and used to calculate non-response weights.

**Table 2.3 Selected household derived variables**

Highest Education level of Respondent
Household own or rent accommodation
Respondent age (banded)
Scottish Index of Multiple Deprivation 15% most deprived datazones
Level of household income
Household employment and family type
Number of children in household
Respondents religion

The model was used to generate a predicted probability for each respondent. This is the probability the respondent would take part in the interview, given the characteristics of the respondent and the household collected at sweep 2. Respondents with characteristics associated with non-response (such as being younger mother) are under-represented in the sample and will receive a low predicted probability. The non-response weights are then generated as the inverse of the predicted probabilities; hence respondents who had a low predicted probability get a larger weight, increasing their representation in the sample. The characteristics related to response behaviour at sweep 3 are given in Appendix A.1.

The final longitudinal sweep 3 weight is the product of the sweep 3 non-response weight and the sweep 2 interview weight. The sweep 2 weights correct for different selection probabilities and non-response to the first interview and non-response to the second interview, hence the final sweep 2 weights correct for non-response at sweep 1 and 2 plus any unequal selection probabilities.

The final weights were scaled to the responding sweep 2 sample size to give a mean weight of one. This makes the weighted sample size match the unweighted sample size. As bias on key variables was similar between trimmed and not-trimmed weights and efficiency of trimmed weights was higher (85.1% vs. 84.7%) 10 cases with highest weights were trimmed. This is expected to be beneficial for weighting of future sweeps.

## 2.5.3 Cross-sectional weights for data from the BC2 Sweep 3

Cross-sectional weights were generated for all respondents at sweep 3 (the combined A and B samples) and should be used for any cross-sectional analysis of BC2 SW3 data.

Calibration weighting methods were used to create the cross-sectional weights. This method takes the pre-calibrated weighted combined sample and adjusts the weights using an iterative procedure. The resulting weighting factors, when applied to the combined data, will make the survey estimates match population estimates for a set of key variables. The population estimates in this instance are survey estimates from Sample A, weighted by the longitudinal weight. Since the longitudinal weight corrects for sampling error and non-response bias at two previous stages of GUS Birth Cohort, the weighted Sample A estimates are the best population estimates available. The choice of the variables to use in the calibration was dictated by the bias remaining in the data after the SW3 longitudinal weights were applied to Sample A and the cross-sectional weight from the first wave (last completed sweep) for Sample B. The key variables used in the weighting were Simd15 (deprivation index), respondent age (DeHGag2) and household tenure.

The calibration corrects for any differences due to differential non-response between Sample A and Sample B.

## 2.5.4 Sample efficiency

Adding weights to a sample can affect the sample efficiency. If the weights are very variable (i.e. they have both very high and very low values) the weighted estimates will have a larger variance. More variance means standard errors are larger and confidence intervals are wider, so there is less certainty over how close the estimates are to the true population value.

The effect of the sample design on the precision of survey estimates is indicated by the effective sample size (neff). The effective sample size measures the size of an (unweighted) simple random sample that would have provided the same precision (standard error) as the design being implemented. If the effective sample size is close to the actual sample size then we have an efficient design with a good level of precision. The lower the effective sample size, the lower the level of precision. The efficiency of a sample is given by the ratio of the effective sample size to the actual sample size. The range of the weights, the effective sample size and sample efficiency for both sets of weights are given in Table 2.4.

Table 2.4	Minimum	Maximum	Mean	N	Neff	Efficiency
<b>Birth Cohort 2</b>						
Longitudinal	0.45	3.31	1	4260	3625	85%
Cross-sectional	0.46	3.26	1	4434	3829	86%

## 2.6 Comparing data between cohorts

Sweep 3 of BC2 is equivalent to sweep 5 of BC1. At both sweeps the cohort children were aged approximately 58 months old. It is therefore possible to compare data between the cohorts where the same questions have been asked.

The variable name does not identify the cohort it was asked of in any way but does identify the sweep/age at which the data was collected (see Section 2.4.1). Therefore where the same

---

question was asked of both cohorts, or the same variable was derived for each, the names will be the same. For example, DeHGnp04 indicates whether the respondent is living with a spouse or partner at the time of the interview. This variable is available both in the BC2 sweep 3 and the BC1 sweep 5 data.

Cohorts can be distinguished by Idnumber: BC1 case Idnumbers begin with a '1'; BC2 Idnumbers begin with a '2'.

## 2.7 Multicoded questions

Some questions in the survey enabled participants to give more than one answer. In the final dataset each of the answer options has been converted into a binary variable with the people who selected that option coded 1 and the rest coded 0.

## 2.8 Dropped Variables

All variables in the questionnaire documentation with '[not in dataset]' next to their name have been deleted from the archived dataset (or have been transformed into derived variables instead).

The following types of variables have been deleted or replaced with a derived variable coded into broader categories in order to reduce the potential to identify individuals:

- Those containing text
- Those which contained a personal identifier (e.g. name/address)
- Those considered to be disclosive, such as:
  - Detailed ethnicity
  - Detailed religion
  - Language spoken at home
  - Full interview date
  - Full date of birth
  - Timing variables

There are no geographical variables in the archived dataset beyond area urban-rural classification, and the Scottish index of multiple deprivation summary variable. Some other geographical variables may be available on request, including via the UKDS Secure Data Service.

## 2.9 Missing values conventions

- 1 Not applicable: Used to signify that a particular variable did not apply to a given respondent, usually because of internal routing.
- 8 Don't know/Can't say.
- 9 No answer/ Refused

These conventions have also been applied to most of the derived variables.

## 3 Data content

### 3.1 Information about the household

In addition to questions asked about the child and parents, the respondent was also asked to provide basic demographic information for each household member including their gender, age and marital status, along with their relationship to each other and to the cohort child (see the questionnaire for details of the relationship variables). Each person was identified by their person number, which they will retain through each sweep of the survey.

A set of derived summary household variables is also included in the data. Amongst other things these detail the number of adults, number of children or number of natural parents in the household. A list of these variables is included in Table 3.1. A set of variables which allow identification of the respondent and their partner (if present) in the household grid are also included. These permit easier analysis of respondent's and partner's age, marital status and relationship to other people in the household. The age variables have been banded for all persons in the household except the study child.

Variable name	Variable label
DeHGrsp01	De - Whether respondent is natural mother
DeHGrsp02	De Whether respondent is natural father
DeHGrsp05	De Resp is childs mother? (incl. adopt./foster/step-mothers)
DeHGrsp06	De Resp is childs father? (incl. adopt./foster/step-fathers)
DeHGrsp07	De Respondent relationship to study child
DeHGrsp08	De Respondent partner relation to the child
DeParTyp	De Indicate whether respondent is child's mother / father / other
DeRspGpa	De - Whether resp is grandparent of child
DeHGnmad	De Number of adults (16 or over) in household
DeHGnmad2	De Nbr of adults other than resp in hhold - banded
DeHGnmkd	De Number of children in household
DeHGnmk2	De Nbr of children in household - Banded
DeHGhsiz	De Household size
DeHGnmsb	De - Number of siblings in household
DeHGnp01	De - Number of natural parents in hhold
DeHGnp02	De - Natural mother in household
DeHGnp03	De - Natural father in household
DeHGrsp04	De - Family Type
DeHGnp04	De - Respondent living with spouse/partner
DeHGprim	De Whether child was mothers first-born

Table 3.1 Selected household derived variables	
DeHGbord	De - Study child's birth order
DeHGmag3	De Child's nat mother's age at interview (banded)
DeHGmag5	De Nat mother age at birth of cohort child (banded)
DeRespAg	De - Respondent age (banded)
DeRPage	De - Respondent partner age (banded)
DeRsex	De - Respondent's sex
DeRPsex	De - Respondent partner's sex
DeMoThID	De - Mother's ID (person number in household grid)
DeFathID	De - Father's ID (person number in household grid)
DeRespID	Me - Respondent ID (person number in household grid)
DePartID	De - Partner ID (person number in household grid)

## 3.2 Main socio-economic variables

### 3.2.1 National Statistics Socio-economic Classification (NS-SEC)

(NS-SEC) is a social classification system that attempts to classify groups on the basis of employment relations, based on characteristics such as career prospects, autonomy, mode of payment and period of notice. There are fourteen operational categories representing different groups of occupations (for example higher and lower managerial, higher and lower professional) and a further three 'residual' categories for full-time students, occupations that cannot be classified due to a lack of information or other reasons. The operational categories may be collapsed to form a nine, eight, five or three category system.

The Growing Up in Scotland dataset includes the five category system, in which respondents and their partner, where applicable, are classified as managerial and professional, intermediate, small employers and own account workers, lower supervisory and technical, or semi-routine and routine occupations. A sixth category 'never worked' is also coded on this variable. The decision on whether or not this category should be included as a separate category, incorporated with category 5 'Semi-routine or routine' or set to 'missing' is dependent on the particular analysis to which it is being applied.

Further information on NS-SEC is available from the National Statistics website at:

<http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html>.

The relevant variables in the GUS BC2 sweep 3 dataset are given in Table 3.2

Table 3.2 NS-SEC variables on the dataset	
Variable name	Description
DeMsec01	De Respondent NSSEC - 6 Category
DeYsec01	De Partner NSSEC - 6 Category
DeMsec10	De Household NSSEC - 6 Category

## 3.2.2 Equivalised household income (annual)

The income that a household needs to attain a given standard of living will depend on its size and composition. For example, a couple with dependent children will need a higher income than a single person with no children to attain the same material living standards.

'Equivalisation' means adjusting a household's income for size and composition so that we can look at the incomes of all households on a comparable basis. Official income statistics use the 'Modified OECD' equivalence scale, in which an adult couple with no dependent children is taken as the benchmark with an equivalence scale of one. The equivalence scales for other types of households can be calculated by adding together the implied contributions of each household member from the table below.

Household member	Equivalence scale
Head	0.67
Subsequent adults	0.33
Each child aged 0-13	0.20
Each child aged 14-18	0.33

For example, a household consisting of a single adult will have an equivalence scale of 0.67 - in other words he or she can typically attain the same standard of living as a childless couple on only 67 percent of its income. In a household consisting of a couple with one child aged three, the head of the household would contribute 0.67, the spouse 0.33, and the child 0.20, giving a total equivalence scale of 1.20. In other words this household would need an income 20 percent higher than a childless couple to attain the same standard of living.

The distribution of income for the population of the United Kingdom as a whole is taken from the most recent available data from the Family Resources Survey. The data and methodology are the same as those used by the Government in its annual Households Below Average Income publication.

GUS collects a banded version of total net household income from all sources in the main CAPI interview. This income data is adjusted, using the above equivalence scale, according to the characteristics of the household, to produce an equivalised annual household income value. Variables with the full equivalised income scale (DeEqvinc) and quintiles of the scale (DeEqv5) are available in the datasets.

## 3.3 Area-level variables

### 3.3.1 Area-level variables: Scottish Government Urban/Rural Classification

The Scottish Government Urban Rural Classification was first released in 2000 and is consistent with the Government's core definition of rurality which defines settlements of 3,000 or less people to be rural. It also classifies areas as remote based on drive times from settlements of 10,000 or more people. The definitions of urban and rural areas underlying the classification are unchanged.

The classification has been designed to be simple and easy to understand and apply. It distinguishes between urban, rural and remote areas within Scotland and includes the

categories shown in Table 3.4. The variables indicating the urban-rural classification of the child's home address is ALaURin2.

Classification	Description
Large Urban Areas	Settlements of over 125,000 people
Other Urban Areas	Settlements of 10,000 to 125,000 people
Accessible Small Towns	Settlements of between 3,000 and 10,000 people and within 30 minutes' drive of a settlement of 10,000 or more
Remote Small Towns	Settlements of between 3,000 and 10,000 people and with a drive time of over 30 minutes to a settlement of 10,000 or more
Accessible Rural	Settlements of less than 3,000 people and within 30 minutes' drive of a settlement of 10,000 or more
Remote Rural	Settlements of less than 3,000 people and with a drive time of over 30 minutes to a settlement of 10,000 or more

For further details on the classification see the Scottish Government website:

<http://www.scotland.gov.uk/Topics/Statistics/About/Methodology/UrbanRuralClassification>

### 3.3.2 Scottish Index of Multiple Deprivation

The Scottish Index of Multiple Deprivation (SIMD) identifies small area concentrations of multiple deprivation across Scotland. It is based on 37 indicators in the seven individual domains of Current Income, Employment, Health, Education Skills and Training, Geographic Access to Services (including public transport travel times for the first time), Housing and a new Crime Domain. SIMD is presented at data zone level, enabling small pockets of deprivation to be identified. The data zones, which have a median population size of 769, are ranked from most deprived (1) to least deprived (6,505) on the overall SIMD and on each of the individual domains. The result is a comprehensive picture of relative area deprivation across Scotland.

SIMD is regularly updated reflecting local changes in the various indicators. The classificatory variables contained in the BC2 sweep 3 dataset is based on the 2012 version of SIMD. It should be noted that prior GUS datasets contain variables which use earlier versions of SIMD.

Two variables are included in the dataset. In the first, the data zones are grouped into quintiles. Quintiles are percentiles which divide a distribution into fifths, i.e., the 20th, 40th, 60th, and 80th percentiles. Those respondents whose postcode falls into the first quintile are said to live in one of the 20% least deprived areas in Scotland. Those whose postcode falls into the fifth quintile are said to live in one of the 20% most deprived areas in Scotland. The variable is ALaSNim2.

Further details on SIMD can be found on the Scottish Government Website:

<http://www.scotland.gov.uk/Topics/Statistics/SIMD/Overview>

## 3.4 Topic measurements and instruments

### 3.4.1 Child Development: Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ)<sup>4</sup> is a brief behavioural screening questionnaire designed for use with 3-16-year-olds (Goodman, 1997). The scale includes 25 questions which are used to measure five aspects of the child's development – emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and pro-social behaviour. A score is calculated for each aspect, as well as an overall 'difficulties' score which is generated by summing the scores from all the scales except pro-social. For all scales, except pro-social where the reverse is true, a higher score indicates greater evidence of difficulties. The dataset includes the constituent items, and the derived variables including the various composite scores and total score. Details of these variables are included in Table 3.7, with syntax illustrated in the derived variables documentation.

Variable name	Description
DeDsdem1	De SDQ: Emotional symptoms score
DeDsdco1	De SDQ: Conduct problems score
DeDsdhy1	De SDQ: Hyper-activity or inattention score
DeDsdpr1	De SDQ: Peer problems score
DeDsdps1	De SDQ: Pro-social score
DeDsdto1	De SDQ: Total difficulties score

### 3.4.2 Parental Health: Medical Outcomes Study 12-Item Short Form (SF-12)

Health-related quality of life was measured by the Medical Outcomes Study 12-Item Short Form (SF-12). This has also been used in the Scottish Health Survey, and has previously been used in population surveys on many occasions (for example, the Health Survey for England and the National Survey of NHS Patients). The SF-12 is a widely used self-reported generic measure of health status, yielding both a physical component (PCS) and a mental health component (MCS) summary scale score. It is tailored for use in large health surveys of general populations. Higher scores on both the physical and mental health component scales are indicative of better health-related quality of life, the indicator is based on informants' self-reports of their own physical and mental functioning and as such are subjective. This may lead to differential reporting between informants with equivalent status.

<sup>4</sup> Further details on the SDQ can be found at: <http://www.sdqinfo.com/>

Table 3.6 Constituent and derived variables associated with SF-12	
Variable name	Description
MeHpgn01	Me - How is resp health in general
MeHlmt01	Me - Resp health limits moderate activities
MeHlmt02	Me - Resp health limits climbing stairs
MeHlmt03	Me - Resp health limited accomplishments past 4 wks
MeHlmt04	Me - Resp health limited reg activities past 4 wks
MeHlmt05	Me - Resp mental health limited accomplishments past 4 wks
MeHlmt06	Me - Resp mental health limited quality of work/activ past 4 wks
MeHlmt07	Me - Resp physical pain limited normal work past 4 wks
MeHpgn02	Me - Time resp felt calm in past 4 wks
MeHpgn03	Me - Time resp felt energetic in past 4 wks
MeHpgn04	Me - Time resp felt down in past 4 wks
MeHpgn05	Me - Time resp health interfered socially in past 4 wks
DeSF12ph	De - Physical PCS - 12 Scale
DeSF12mn	De - Mental MCS - 12 Scale

### 3.4.3 Socioeconomic characteristics: Selected items from Index of Material Deprivation

Selected items from the DWP Index of Material Deprivation, as first used in the 2004/2005 Family Resources Survey (McKay and Collard, 2004), were included in the sweep three questionnaire. Most material deprivation measures generally ask respondents about the ownership of items regarded as ‘necessities’ by a majority of the population. People are then classified as ‘deprived’ if they go without some of these items. Poverty measures based on this type of information are also known as consensual poverty measures. Essentially, the absence of items is taken to reflect deprivation and the greater the number of items absent, the greater the degree of deprivation. In most research nowadays, respondents are asked to clarify if they do not have or consume an item whether this is because: a) they do not ‘need’ it, or b) they ‘cannot afford’ it. It is, therefore, possible to distinguish between ‘unenforced’ and ‘enforced’ hardship.

Table 3.7 Variables associated with the Index of Material deprivation	
Variable name	Description
MeEmd01	Me - Holiday away from home one week+ /yr
MeEmdb01	Me - Celebrations at special occasions
MeEmdb02	Me - Night out once a month
MeEmdb03	Me - Family car or van
MeEmd04	Me - Enough money for house decoration
MeEmd05	Me - Household contents insurance

MeEmd06	Me - Regular savings of £10+ / month
---------	--------------------------------------

### 3.4.4 Gender Role in Children: Selected items from the Pre-school Activities Inventory

Selected items from the Pre-School Activities Inventory (PSAI) were included in the sweep 3 questionnaire. The Pre-School Activities Inventory (PSAI) is a psychometric scale for the assessment of gender role behaviour in young children (Golombok and Rust, 1993). It is used primarily by psychologists and has been normed on a large and representative sample of children between the ages of two-and-a-half years and five years of age. The full scale contains 24 items which concern the child's characteristics and toy and activity preferences, 12 of which were included in the questionnaire. Details of these variables are included in Table 3.8.

Table 3.8 Variables associated with the Pre-school Activities Inventory	
Variable name	Description
MeSPsai1	Me - How often has child played with the following toys during the past month? Guns (or used objects as guns)?
MeSPsai2	Me - How often has child played with the following toys during the past month? Jewellery?
MeSPsai3	Me - How often has child played with the following toys during the past month? Dolls, doll s clothes or a doll s pram?
MeSPsai4	Me - How often has child played with the following toys during the past month? Swords (or used objects as swords)?
MeSPsai5	Me - Child engaged in activities in past month? Pretending to be a female character (e.g.princess)?
MeSPsai6	Me - Child engaged in activities in past month? Playing at having a typically male occupation (e.g. soldier)?
MeSPsai7	Me - Child engaged in activities in past month? Play fighting?
MeSPsai8	Me - Child engaged in activities in past month? Playing at taking care of babies?
MeSPsai9	Me - Child engaged in activities in past month? Dressing up in girlish clothes?
MeSPsai10	Me - Child engaged in activities in past month? Enjoys rough and tumble play?
MeSPsai11	Me - Child engaged in activities in past month? Avoids getting dirty?
MeSPsai12	Me - Child engaged in activities in past month? Likes pretty things?

### 3.4.5 Child's self-control: Child Behaviour Questionnaire

The Child Behaviour Questionnaire (CBQ) (Rothbart, Ahadi and Hershey, 1994) is an assessment of the temperament of children between the ages of 3 and 7 years. It includes sub-scales that have been used to measure self-control in children as young as age 4.

In the sweep 3 questionnaire self-control was assessed using the ‘attention focusing’ and inhibitory control’ subscale items of the CBQ, which each have 6 items. These scales capture the child’s ability to concentrate and his/her capacity to plan and to inhibit inappropriate responses.

Table 3.9 Variables associated with the Child Behaviour Questionnaire	
Variable name	Description
MeSFsc1	Me - Child has a hard time keeping mind on an activity.
MeSFsc2	Me - Child will move from one task to another without completing any of them.
MeSFsc3	Me - Child can wait before entering into new activities if is asked.
MeSFsc4	Me - Child prepares for trips and outings by planning things will need.
MeSFsc5	Me - Child has trouble sitting still when told to (at movies, church etc).
MeSFsc6	Me - Child shows strong concentration drawing or colouring in a book
MeSFsc7	Me - Child is good at following instructions.
MeSFsc8	Me - Child becomes very involved in, and works for long periods, putting something together.
MeSFsc9	Me - Child approaches places they’ve been told are dangerous slowly and cautiously.
MeSFsc10	Me - Child can easily stop an activity when told no .
MeSFsc11	Me - Child is easily distracted when listening to a story.
MeSFsc12	Me - Child sometimes becomes absorbed in a picture book and looks at it for a long time.

### 3.4.6 Parental Styles: The Parenting Practices Questionnaire

The Parenting Practices Questionnaire (PPQ) (Robinson et al., 1995) is an assessment of 4-12 years old children's mother's and father's parenting styles. This measure was developed to identify continuous scales of authoritative, authoritarian, and permissive parenting styles.

Table 3.10 Variables associated with the Parenting Practices Scale	
Variable name	Description
MePsr01	Me - I encourage child to talk about their troubles.
MePsr02	Me - I find it difficult to discipline child.
MePsr03	Me - I give praise when child is good.
MePsr04	Me - I joke and play with child.
MePsr05	Me - I show sympathy when child is hurt or frustrated.
MePsr06	Me - When child misbehaves, I punish them without really saying why.
MePsr09	Me - I yell and shout when child misbehaves.
MePsr11	Me - I allow child to annoy someone else.
MePsr12	Me - I scold and criticise to make child improve.

MePsr1	Me - I show patience with child.
MePsr14	Me - I threaten to punish child and then don t actually do anything.
MePsr15	Me - I allow child to have a say in the family rules.
MePsr16	Me - I argue with child.
MePsr17	Me - I feel confident about my parenting abilities.
MePsr18	Me - I give child reasons why rules should be obeyed.
MePsr20	Me - I take childs desires into account before asking them to do something.
MePsr21	Me - I explode in anger towards child.
MePsr22	Me - I threaten child with punishment more often than actually giving it.
MePsr23	Me - I express affection by hugging, kissing and holding child.
MePsr26	Me - I give in to child when they causes a commotion about something.
MePsr27	Me - I allow child to interrupt others.
MePsr28	Me - I explain to child how I feel about their good and bad behaviour.
MePsr29	Me - I use threats as punishment without really saying why.
MePsr30	Me - I take into account childs preferences in making plans for the family.
MePsr31	Me - When child asks why they have to conform, I say: because I said so, or I am your parent and I want you to .
MePsr32	Me - I am unsure how to solve childs misbehaviour.
MePsr33	Me - I explain the consequences of childs behaviour to them.
MePsr34	Me - I demand that child does things.
MePsr35	Me - I emphasize the reasons for rules.

## 3.5 Cognitive assessments: British Ability Scales (BAS3) – Naming Vocabulary and Picture Similarities

As noted in section 1.4.1, cognitive assessments were carried out with the cohort children. Details about the variables provided in the dataset are given below.

### 3.5.1 Score variables in the dataset

The dataset provides the following scores for each assessment:

- **Ability score** – Derived from a so-called ‘raw’ score (not provided) and the item set administered (using tables provided on the BAS Naming Vocabulary and Picture Similarities Score Sheets) this is an estimate of child’s level on the ability being measured. It reflects the raw score and the difficulty of the items administered. The ability score is not a normative score. The numbers used are arbitrary and simply provide a common scale of performance level, regardless of the items a child was given.
- **Normative scores (t-scores)** – Derived from standard BAS tables and defined with reference to the standardisation sample used in developing the assessment. T-scores (with mean=50 and standard deviation=10) are provided.

**Table 3.11 British Ability Scales – Naming Vocabulary and Picture Similarities**

Variable name	Description
DePicSAS	De Picture Similarities Ability Score
DePicSTS	De Picture Similarities T-Score
DePicSPt	De Picture Similarities Percentile Equivalent
ZDePicSAS	De Picture Similarities Z-Score
AncB	CAe - Consent obtained for Picture Similarities test
DeNamVAS	De Naming Vocabulary Ability Score
DeNamVTS	De Naming Vocabulary T-Score
DeNamVPt	De Naming Vocabulary Percentile Equivalent
ZDeNamVAS	De Naming Vocabulary Z-Score
AncC	CAe - Consent obtained for Naming Vocabulary test
CAePrb01	CAe - No difficulties during assessments
CAePrb02	CAe - Difficulties: Assessment interrupted
CAePrb03	CAe - Difficulties: Child was ill
CAePrb04	CAe - Difficulties: Child was tired
CAePrb05	CAe - Difficulties: Parent interfered
CAePrb06	CAe - Difficulties: Something else

### 3.5.2 Influences on test scores

It is important to note that the child’s performance may have been affected by influences extraneous to those that the assessment is intended to measure. The conditions listed below can lead either to a higher or lower score than would normally be obtained.

- Non-standard administration of the scale
- Non-standard scoring algorithms used ensure standard scoring in all cases
- Administration disrupted by noise or other interruptions
- Difficulty in establishing rapport with the child
- Child has difficulty in concentrating on the tasks or is easily distracted
- Child is excessively anxious to the extent that concentration/flexibility of thought seem impaired
- Child is reluctant to respond and/or refuses to persevere on more difficult items
- Child has permanent/temporary sensory impairment (particularly vision/hearing) or motor impairment
- Child is on medication of a type that could affect performance
- Child is over-tired or ill.

In anticipation of these issues, the specification of CAPI program and the training of interviewers were designed to ensure standard administration of the assessment. The training of interviewers was also designed to ensure that risks were minimised. To allow for the

consideration of such issues when analyzing the data, interviewers were asked to record details of any interruptions, distractions, behaviours or health circumstances in CAPI. Table 3.9 contains details of the relevant variables where this information is recorded.

Variable name	Label
CAePrb01	CAe - No difficulties experienced during assessments
CAePrb02	CAe - Difficulties experienced because: Assessment was interrupted
CAePrb03	CAe - Difficulties experienced because: Child was ill
CAePrb04	CAe - Difficulties experienced because: Child was tired
CAePrb05	CAe - Difficulties experienced because: Parent interfered
CAePrb06	CAe - Difficulties experienced because: Other reason

## 3.6 Height and weight and BMI data

### 3.6.1 Overview: BMI

Body Mass Index (BMI), i.e. weight divided by height squared, is a score that adjusts your weight for your height. Taken as a number in isolation, the BMI it does not actually represent anything medically. It is only meaningful in the context of a distribution of values for a population. Individuals are placed into bands to show where they stand in relation to the rest of the population, in particular whether they have unusually high or low BMI.

In adults BMI stays fairly constant on average as people get older. Therefore BMI categories for adults ignore age and calculate the same BMI for two people with the same weight and height regardless of the differences in their ages. However, among young children in particular, BMI changes quite significantly as the child ages. Since to have a certain BMI at one age may be the norm but be unusually high or low at another age, different centiles are calculated for different ages.

While the BMI measure has come under some scrutiny for not always being accurate, it remains the best non-invasive measure for obesity. Furthermore, a review of the measure by (Reilly et al., 1999) in the British Medical Journal suggests that the BMI is more likely to understate, rather than overstate, the true levels of obesity, as has been discussed by Prentice (Prentice, 1998) and Barlow and Dietz (Barlow & Dietz, 1998).

### 3.6.2 GUS BC2 sweep 3 BMI measures

The main child overweight and obesity variables available for GUS BC2 sweep 3 are outlined in Table 3.13. These variables were produced using the International Obesity Taskforce cut-offs. These cut-offs are based on BMI reference data from six different countries around the world (over 190,000 subjects in total aged 0 to 25 from UK, Brazil, Hong Kong, the Netherlands, Singapore, and the United States). In summary, the BMI percentile curves that pass through the values of 25 and 30 kg/m<sup>2</sup> (standard adult cut-off points for overweight and obesity, respectively) at age 18 were smoothed for each national dataset and then averaged. The averaged curves were then used to provide age and sex-specific BMI cut-off points for children and adolescents aged 2 to 18. By averaging the distribution curves from each

reference country, the international cut-offs for children purport to be representative of the countries but independent of the overweight or obesity level in each country. One of the benefits of using these international standards is the possibility of making international comparisons. However, the international classification is not without problems: international reference data differ from those for the UK population, and this is reflected in the sex-specific overweight and obesity estimates produced by the International classification.

In light of this lack of consensus on its use, variables have also been produced using the 85th (overweight cut-off) / 95th (obesity cut-off) BMI percentiles of the UK reference curves (referred to as the National BMI percentiles classification). The National BMI percentiles classification has been used in the past to describe childhood overweight and obesity prevalence trends in the UK and the 85th / 95th cut-off points are commonly accepted thresholds used to analyse overweight and obesity in children (detail on relevant cut-offs and their descriptions are included below). The National BMI percentiles classification has been shown to be reasonably sensitive (i.e. not classifying obese children as non-obese) and specific (i.e. not classifying non-obese children as obese). A key issue to bear in mind, however, is that the National BMI percentiles classification are based on the arbitrary assumption that the prevalence of overweight and obesity at the point when the reference data was compiled was 15% and 5%, respectively. Furthermore, there seems to be no indication that these cut-off points relate directly or indirectly to any physiological outcomes or health or disease risks. It is worth noting that the UK component of the international classification used the same sample as that used to construct the UK reference BMI data.

In addition to these International and National BMI classifications, the Information Services Division (ISD) at the Scottish Government uses an alternative method to produce BMI centiles, Cole's LMS method, which takes into account the fact that BMI data does not follow a normal distribution. Further information can be found at <http://www.isdscotland.org/isd/3640.html>

Note that only those height and weight measurements considered by the interviewer to be reliable were used to calculate the BMIs.

**Table 3.13 BMI cut-off points**

Percentile cut-off	Description
At or below 5th percentile	Underweight
Above 5th percentile and below 85th percentile	Healthy weight
At or above 85th percentile and below 95th percentile	Overweight
At or above 95th percentile and below 98th percentile	Obese
At or above 98th percentile	Morbidly obese

Table 3.14 Child Derived BMI variables	
Variable name	Description
DeBMI	De BMI (reliable child weight measurements only)
DeBMI_b	De BMI (including unreliable measurements)
DeUKbmi	De UK BMI national classification standards
DeINTbmi	De International BMI cut-offs
DeINTbmi2	De BMI status (ovrwt inc. obese) - international cut-offs
DeINTbmi3	De BMI status (non-obese vs obese) - international cut-offs
DeISDbmc	De ISD BMI centiles classification
DeISDbmi	De ISD BMI 5 group classification
DeISDHWt	Dc Study child weight within/outwith ISD healthy range
DeISDovW	Dc Study child overweight, including obese (ISD)

---

## 4 More about GUS

All publications using GUS data are available from the study website:

[growingupinScotland.org.uk/publications](https://growingupinScotland.org.uk/publications)

Users might also be interested in viewing the rest of the GUS website:

[growingupinScotland.org.uk](https://growingupinScotland.org.uk). This contains a large amount of useful information including the background to the study, and a wide range of publications using existing data.

### 4.1 Contact details

You can email the study on [gus@scotcen.org.uk](mailto:gus@scotcen.org.uk) or contact a member of the ScotCen Research team:

- Line Knudsen (Senior Researcher): [line.knudsen@scotcen.org.uk](mailto:line.knudsen@scotcen.org.uk)
- Jackie Palmer (Data Manager): [jackie.palmer@scotcen.org.uk](mailto:jackie.palmer@scotcen.org.uk)
- Paul Bradshaw (Study Director): [paul.bradshaw@scotcen.org.uk](mailto:paul.bradshaw@scotcen.org.uk)

Alternatively you can call the ScotCen office on 0131 240 0210 and ask for a member of the GUS team.

---

## 5 References

- Barlow, S.E., and Dietz W.H. (1998) Obesity evaluation and treatment: Expert Committee recommendations, The Maternal and Child Health Bureau, Health Resources and Services Administration and the Department of Health and Human Services. *Pediatrics*, 102(3): E29.
- Golombok, S. and Rust, J. (1993) 'The Measurement of Gender Role Behaviour in Pre-school Children: a Research Note', *Journal of Child Psychology and Psychiatry*, 34(5), pp. 805-811.
- Goodman, R. (1997) The Strengths and Difficulties Questionnaire: a research note, *Journal of Child Psychology and Psychiatry*, 38, pp581-586.
- Prentice, A.M. (1998) 'Body mass index standards for children' *British Medical Journal*, 317, pp1401-2.
- Robinson, C. C., Mandleco, B., Olsen, S. F., & Hart, C. H. (1995). Authoritative, authoritarian, and permissive parenting practices. Development of a new measure. *Psychological Reports*, 77, 819–830.
- Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly*, 40, 21-39.
- Reilly, JJ., Dorosty, A.R., and Emmett, P.M. (1999) Prevalence of overweight and obesity in British children: cohort study. *BMJ*, 319, pp1039-1039.

## 6 Appendix

Table A.1 Full model for non-response to sweep 3 interview						
Variables in the equation	B	S.E.	Wald	df	Sig.	Exp(B)
<b>Highest educational level of respondent</b>						
Degree or equivalent			35.962	5	0.000	
Vocational qualification below degree	-0.258	0.127	4.097	1	0.043	0.773
Higher Grade or equivalent	-0.402	0.171	5.493	1	0.019	0.669
Standard Grade or equivalent	-0.591	0.144	16.965	1	0.000	0.554
Other	0.320	0.322	0.985	1	0.321	1.377
No Qualifications	-0.830	0.171	23.639	1	0.000	0.436
<b>Tenure</b>						
Owner occupier			22.584	2	0.000	
Rents HA/council	-0.133	0.121	1.207	1	0.272	0.875
Rents private	-0.521	0.120	18.973	1	0.000	0.594
<b>Mother's age at child's birth</b>						
Under 29			19.697	2	0.000	
30 to 39	0.345	0.096	12.904	1	0.000	1.413
40 or older	0.605	0.161	14.064	1	0.000	1.832
<b>SIMD</b>						
in 15% most deprived	-0.344	0.100	11.836	1	0.001	0.709
<b>Household income</b>						
<15599			12.248	5	0.032	
15600-25999	0.263	0.125	4.418	1	0.036	1.301
26000-36399	0.382	0.155	6.041	1	0.014	1.465
36400-51999	0.381	0.176	4.697	1	0.030	1.464
52000+	0.499	0.179	7.774	1	0.005	1.648
Missing info	0.059	0.133	0.195	1	0.658	1.060
<b>Respondent employment status</b>						
Information not available for at least one parent			13.295	4	0.010	
Lone parent in paid work for >=16 hours	0.267	0.271	0.972	1	0.324	1.306
Information not available for at least one parent	-0.154	0.182	0.716	1	0.398	0.857
Lone parent unemployed or in paid work for <16 hours	-0.118	0.203	0.335	1	0.563	0.889
Couple family either resp and partner working >=16 hours or one of them	0.174	0.183	0.906	1	0.341	1.190
<b>Number of children in the household</b>						
1			5.493	2	0.064	
2	0.110	0.095	1.338	1	0.247	1.117
3+	-0.135	0.112	1.448	1	0.229	0.874

Table A.1 Full model for non-response to sweep 3 interview

<b>Respondent religion</b>						
None			7.360	3	0.061	
Protestant/other Christian	0.206	0.115	3.196	1	0.074	1.229
Catholic	-0.174	0.111	2.442	1	0.118	0.841
Other	-0.067	0.200	0.111	1	0.739	0.936
Constant	1.665	0.247	45.361	1	0.000	5.288

---

ScotCen Social Research  
Scotia Bank House  
6 South Charlotte Street  
Edinburgh  
EH2 4AW  
T 0131 240 0210  
[www.scotcen.org.uk](http://www.scotcen.org.uk)

A company Limited by Guarantee  
Registered in England No.4392418  
A Charity registered in England and Wales (1091768) and Scotland (SC038454)