Scottish Centre *for* Social Research

Growing Up In Scotland Sweep 3: 2007-2008
User Guide

Authors: Joan Corbett, Louise Marryat, Paul Bradshaw, Mireille Ferrandon

1 Overview of the survey

The data files contain data from Growing Up in Scotland (GUS) Sweep 3, the third year of a longitudinal research study aimed at tracking the lives of two 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 aims of the study are:
- To provide reliable cross-sectional data on the characteristics, circumstances and experiences of children in Scotland aged between 0 and 5.
- To document differences in the current characteristics, circumstances and experiences of children from different backgrounds
- To generate information about longer-term outcomes across a range of key domains and to document differences in those outcomes for children of different backgrounds.
- To identify key predictors of adverse longer-term outcomes with particular reference to the role of early years service provision
- To measure levels of awareness and use of key services
- To examine the nature and extent of informal sources of help, advice and support for parents
- To generate parental assessments of the services accessed and used; and to improve understandings of choice and constraint in service use.

At sweep 3, data collection for the study included two main elements:
1. A face-to-face CAPI interview with the cohort child's main carer
2. Two cognitive assessments undertaken with children in the birth cohort (further details are included in sections 3 and 7)

1.1 Study Design

The survey is based on two cohorts of children: the first aged approximately 10 months at the time of first interview and the second aged approximately 34 months. A named sample of approximately 10,700 children was selected from the Child Benefit records to give an achieved sample of 8,000 overall.

The configuration of cohorts and sweeps for the first four sweeps of data collection is summarised below. BC1 refers to the younger of the two cohorts ('birth cohort') and CC1 to the slightly older cohort ('child cohort').

Table 1.1   Sample design: sweeps 1 to 4

Sweep / Launch year	Age at interview					
	0-1	1-2	2-3	3-4	4-5	5-6
1 / 2005	BC1		CC1			
2 / 2006		BC1		CC1		
3 / 2007			BC1		CC1	
4 / 2008				BC1		CC1

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

1. *Cross-sectional time specific data* – e.g. what proportion of 2-3 year-olds are living in single parent families in 2007?
2. *Cross-sectional time series data* – e.g. is there any change in the proportion of 2-3 year-olds living in single parent families between 2005 and 2007?
3. *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 2-3?

1.2   Sample Design

The area-level sampling frame was created by aggregating Data Zones. Data Zones are small geographical output areas created for the Scottish Government. Data Zones are used by Scottish Neighbourhood Statistics to release small area statistics. The Data Zone geography covers the whole of Scotland. The geography is hierarchical, with Data Zones nested within Local Authority boundaries. Each data zone contains between 500 and 1,000 household residents. More information can be found on the Scottish Neighbourhood Statistics website: http://www.sns.gov.uk.

The Data Zones were aggregated to give an average of 57 births per area per year (based on the average number of births in each Data Zone for the preceding 3 years). It was estimated that this number per area would provide us with the required sample size. Once the merging task was complete, the list of aggregated areas was sorted by Local Authority[1] and then by the Scottish Index of Multiple Deprivation Score. 130 areas were then selected at random. The Department of Work and Pensions then sampled children from these 130 sample points.

Within each sample point, the Child Benefit records were used to identify all babies and three-fifths of toddlers who met the date of birth criteria (see Table 1.2). The sampling of children was carried out on a month-by-month basis in order to ensure that the sample was as complete and accurate as possible at time of interview.

In cases where there was more than one eligible child in the selected household, one child was selected at random. If the children were twins they had an equal chance of being selected. If the eligible children were in different age cohorts the younger child had a higher chance of being selected given that those children had a higher chance of being included in the sample overall.

After selecting the eligible children, the DWP made a number of exclusions before transferring the sample details. These exclusions included cases they considered 'sensitive' and children that had been sampled for research by the DWP in the last 3 years.

---

[1] Local Authority has been used as a stratification variable during sampling, this means the distribution of the GUS sample by Local Authority will be representative of the distribution of Local Authorities in Scotland. However, the sample sizes are such that we would not recommend analysis by Local Authority. The small sample sizes would give misleading results.

Table 1.2   Eligible child dates of birth for inclusion in the Growing Up in Scotland study by cohort

Sample Number	Dates of Birth required	
	Birth Cohort	**Child Cohort**
1	01-June-2004 - 30-Jun-2004	01-June-2002 - 30-Jun-2002
2	01-Jul-2004 - 31-Jul-2004	01-Jul-2002 - 31-Jul-2002
3	01-Aug-2004 - 31-Aug-2004	01-Aug-2002 - 31-Aug-2002
4	01-Sep-2004 - 30-Sep-2004	01-Sep-2002 - 30-Sep-2002
5	01-Oct-2004 - 31-Oct-2004	01-Oct-2002 - 31-Oct-2002
6	01-Nov-2004 - 30-Nov-2004	01-Nov-2002 - 30-Nov-2002
7	01-Dec-2004 - 31-Dec-2004	01-Dec-2002 - 31-Dec-2002
8	01-Jan-2005 - 31-Jan-2005	01-Jan-2003 - 31-Jan-2003
9	01-Feb-2005 - 28-Feb-2005	01-Feb-2003 - 28-Feb-2003
10	01-Mar-2005 - 31 Mar-2005	01-Mar-2003 - 31 Mar-2003
11	01-Apr-2005 - 30-Apr-2005	01-Apr-2003 - 30-Apr-2003
12	01-May-2005 - 31-May-2005	01-May-2003 - 31-May-2003

1.3   Development and Piloting

Policy priorities and key topics of interest for the sweep 3 questionnaire were initially discussed and agreed by the study's Scottish Government Project Managers and Policy Advisory Group. The questionnaire was then developed by the GUS team at ScotCen with input from colleagues at the Centre for Research on Families and Relationships in reference to these priorities and topics. A subset of new questions was included in a small cognitive pilot in September 2006, with a full instrument initially piloted in paper form in November 2006. This instrument was revised and converted into CAPI for the second Dress Rehearsal Pilot in January 2007.

2 Data collection methods

2.1   Mode of data collection

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. There was a brief, self-complete section in the interview in which the respondent, using the laptop, input their responses directly into the questionnaire programme.

At sweep 1, primarily because of the inclusion of questions on the mother's pregnancy and birth of the sample child, interviewers were instructed as far as possible to undertake the interview with the child's mother. Where the child's mother was not available, interviews were undertaken with the child's main carer.

At sweep 2, interviewers were instructed to undertake the interview with the sweep 1 respondent. Where this was not possible or appropriate, interviews were conducted with the child's main carer. In practice, most interviews were undertaken with the sweep 1 respondent and this was usually the child's mother.

At sweep 3, interviewers were instructed to undertake the interview with the same respondent as in the previous sweep, that is Sweep 2 or Sweep 1 if the household was missed at Sweep 2. Where this was not possible or appropriate, interviews were conducted with the child's main carer. In practice, most interviews were undertaken with the previous sweep respondent (98.6% of interviews were with the previous respondent) and this was usually the child's mother (98.2% of interviews were with the child's mother).

2.2 Length of Interview

Overall, the average interview lasted around 67 minutes. The birth cohort interview had a slightly longer average length at 69 minutes, than the child cohort interview at 63 minutes. The median interview length was 64 minutes for the birth cohort and 58 minutes for the child cohort.

2.3 Timing of fieldwork

Fieldwork was undertaken over a **fourteen** month period commencing in April 2007. 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, sample 2 was issued at the beginning of May 2007 and remained in field until mid-July 2007.

To ensure that respondents in both samples were interviewed when their children were approximately the same age, each case was assigned a 'target interview date'. For the birth cohort this was identified as the date on which the child turned 34.5 months old, and for the child cohort the date the child turned 58.5 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 vast majority of interviews were achieved within the four-week target period.

3 Child assessments

As part of the data collection for sweep 3, cohort children in the birth cohort 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 II). 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 programme pre-determining the complex set of rules for routing children through each assessment. The purpose and method of each assessment is described in table 3.1. The data is used to estimate an approximate score for each child. Further details on the data and variables associated with the cognitive assessments can be found in section 7.7.5.

3.1 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.

3.2 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 the child's:
- 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).

3.3 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 (1996) British Ability Scales Second Edition
(BAS II): Administration and Scoring Manual. London: NFER-Nelson.
Elliott, C.D., Smith, P, and McCulloch, K (1997) British Ability Scales Second Edition
(BAS II): Technical Manual. London: NFER-Nelson.

3.4 Obtaining consent for child assessments

Before undertaking the assessments with the child, parents were 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 was very high at 94% for each assessment.

Table 3.1 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 5th 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

4 Response rates

Details of the number of cases issued and achieved and the response rates are detailed in Table 4.1.

Table 4.1 Number of issued and achieved cases and response rates

	Birth Cohort	Child Cohort	All Sample
Achieved interviews at sweep 1	5217	2858	8075
Achieved interviews at sweep 2	4512	2500	7012
Cases to field at sweep 3:			
All issued to field*	4665	2599	7264
Eligible i.e. achievable or 'in-scope'**	4630	2582	7212
Cases achieved at sweep 3	4193	2332	6525
Response rate			
As % of all eligible cases at sweep 3	91%	90%	91%
As % of all sweep 1 cases	80%	82%	81%

* The number of cases issued to the field at sweep 3 is higher than the number of Interviews achieved at sweep 2 because some of the sweep 1 respondents missed at sweep 2 came back at sweep 3.

** Cases which were considered out-of-scope or unachievable were mostly ineligible addresses – usually due to the family having moved away from Scotland.

5 Coding and editing

Additional coding and editing tasks were performed after the interviews were conducted. The GUS Sweep 3 Coding Instructions provide details of the tasks that were conducted.

6 Weighting the data

6.1 Overview

Two sets of weights have been developed for each cohort:

1. A cross-sectional weight that should be used for any cross-sectional analysis of Wave 3 data only. All sample members that responded at W3 have a cross-sectional weight.

2. A longitudinal weight for any analyses that includes more than one wave of data. Sample members that have responded at every wave of GUS have a longitudinal weight.

6.2 Background

- The sampling frame was the child-level Child Benefit records held by the Inland Revenue. Children were selected from 130 sample points in Scotland. The sample points consist of aggregations of Data Zones. [2]

- There are two cohorts of children: the birth cohort and child cohort. Children in the birth cohort were aged approximately 10 months at the time of first interview whereas children in the child cohort were aged around 34 months. Weights for the birth and child cohorts have been generated separately, since these two groups should always be analysed separately.

- The Sweep 3 interview follows up all main carers who responded at the previous interview and gave ScotCen permission to be re-contacted. Some of the respondents of Sweep 1 who had asked not to take part for a year but were willing to be contacted the following year were also included at Sweep 3. There was no sub-sampling. Response rates were good at 91% for the birth cohort and 90% for the child cohort.

- At Sweep 3 we used proxy interviews to gather information on the main respondent's resident partner.

6.3 The sweep 3 sample

The sweep 3 sample can be split into two components. For the purposes of describing the weighting, these two components have been named Sample A and Sample B and are defined as follows:
- Sample A – Wave 3 respondents who responded at all waves
- Sample B – Wave 3 respondents who responded at Wave 1 but not Wave 2.

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

---

[2] Further information on the sample design and the weighting process at sweeps 1 and 2 can be found in the User Guides for those sweeps which are available from the Data Archive or the 'using GUS data' section of

The birth cohort contained 198 Sample B members, 73 (37%) of which gave an interview at sweep 3. Similarly, the child cohort contained 118 Sample B members, of which 52 (44%) responded. The response rates for Sample B were much lower than the response rates for Sample A, which was 92% for both cohorts. The issued and responding sample sizes are given in Table 6.1.

Table 6.1 Response rates for different samples

	Issued	Responding	Response rate
Birth cohort			
Sample A	4467	4120	92%
Sample B	198	73	37%
Combined (A+B)	4665	4193	90%
Child cohort			
Sample A	2481	2280	92%
Sample B	118	52	44%
Combined (A+B)	2599	2332	90%

The longitudinal weight will be used for any analyses that include more than one wave of data. Only members of Sample A (who have responded at every sweep of GUS) will have a longitudinal weight. This weight is described in more detail in Section 6.4.

There will be two sets of weights developed for each cohort; a cross-sectional weight and a longitudinal weight. The cross-sectional weight will be used for any cross-sectional analysis of sweep 3 data. All sweep 3 respondents will have a cross-sectional weight (Sample A + B). These are described in more detail in section 6.5.

6.4   Longitudinal weights

Longitudinal weights were only generated for respondents in Sample A. A model-based weighting technique was used to develop the sweep 3 longitudinal weights. This is the same method that was used to generate weights at sweep 2. All sweep 3 respondents in Sample A had taken part in the previous two interviews; this information could be used to model response behaviour at sweep 3. Ineligible households (deadwood) were not included in the non-response modelling[3].

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. The model takes account of the relationship of the predictor variables to the outcome and the relationships of the predictor variables to each other.

The model generated 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 the previous waves. Respondents with characteristics associated with non-response (such as being a private tenant) 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

---

the Growing Up in Scotland website www.growingupinscotland.org.uk
[3] There were 11 individuals with ineligible outcome codes; these individuals were dropped from the analysis. Ineligible outcome codes include households that were vacant, demolished or derelict and non-residential

predicted probabilities. Hence respondents who had a low predicted probability get a larger weight, increasing their representation in the sample.

The birth and child cohorts were modelled separately, although there were similarities between the two models. The characteristics related to response behaviour for Sample A at sweep 3 are given in Table 6.2. The full models are detailed in tables A1 and A2 in the Appendix.

Table 6.2 Characteristics associated with response behaviour in Sample A

Characteristics associated with response	Characteristics associated with non-response
Birth cohort	
Owner occupiers	Rent from a private landlord
From a white ethnic background	From any other ethnic background
Breastfed the baby	Did not breast feed
At least one parent/carer in full-time employment	No parent/carer working
Mother aged 25 or over	Younger mother aged under 25
Child cohort	
Owner occupiers	Rent from a private landlord
Older mother aged 30 or over	Younger mother aged under 20
Live as a couple	Lone parent family
One child in the household	More than one child in the household
Respondent in managerial and professional occupations	Respondent in lower supervisory and technical occupations

6.4.1   Final sweep 3 longitudinal weights

The final weight is the product of the sweep 3 non-response weight and the sweep 2 interview weight. For each cohort the final weights were scaled to the responding sweep 3 sample size, this makes the weighted sample size match the unweighted sample size. Table A3 in the Appendix shows the distribution of the sample weighted by the sweep 3 and sweep 2 weights illustrating the reduction in bias caused by the sweep 3 weights.

The Sample B respondents had participated in GUS at sweep 1 but refused at sweep 2, they also had lower response rates at sweep 3 than respondents in Sample A, as shown in Table 6.1 above.

6.5   Cross-sectional weights

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 sweep 3 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 a set of 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 each stage of GUS, the

weighted Sample A estimates are the best population estimates available. The key variables used in the weighting were: area level deprivation indicator (measured using the Scottish Index of Multiple Deprivation), respondent employment status, respondent age at interview and whether the respondent was a lone parent.

The pre-calibration weights were the sweep 3 longitudinal weight for Sample A and the sweep 1 weight for Sample B (i.e. the weight from the last completed wave). Prior to calibration these weights were scaled to the achieved sample size, this was done separately for each sample.

The calibration corrects for any differences due to differential non-response between Sample A and Sample B. The weighted distribution of Sample A and the weighted distribution of the combined sample, pre and post-calibration, are given in Table A3 for the birth and child cohorts.

6.6   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 6.3.

Table 6.3 Range of weights and sample efficiency

	Weight values			Unweighted achieved sample size	Neff	Sample efficiency
	Minimum	Maximum	Mean			
Birth cohort						
Longitudinal weight	0.68	2.31	1.08	4120	3829	92.9%
Cross-sectional weight	0.69	2.31	1.07	4193	3909	93.2%
Child cohort						
Longitudinal weight	0.65	2.40	1.06	2280	2146	94.1%
Cross-sectional weight	0.65	2.39	1.06	2332	2200	94.4%

6.7   Applying the weights

The cross-sectional weights should be used for any cross-sectional analysis, i.e. any analysis of sweep 3 data only. All sample members that responded at sweep 3 have a cross-sectional weight.

The longitudinal weight should be used for any analyses of more than one wave of data. Sample members that have responded at every wave of GUS have a longitudinal weight.

6.8 Weighting variables

The final interview sweep 3 weights are described in Table 6.4.

Table 6.4 Description of weight variables in the data file

Variable name	Label
DcWTbrth	Dc Birth cohort Sw3 weight (cross sectional sample)
DcWTchld	Dc Child cohort Sw3 weight (cross sectional sample)
DcWTbth2	Dc Birth cohort Sw3 weight - longitudinal
DcWTchd2	Dc Child cohort Sw3 weight - longitudinal

Separate weights are provided for each cohort because analysis should always treat each cohort as a distinct population. However, key analysis using this data may involve comparison between the cohorts. It is usually more convenient to undertake this analysis by combining the two cohort datasets into a single dataset and then ensuring that subsequent analysis is either filtered to select a single cohort, or that output is nested by cohort type (variable name = 'SampType'). On merging the datasets it is possible to create a combined weight variable in order that nested analysis uses just a single weight variable. The value of the combined weight is equal to the value of the corresponding cohort weight variable for that child. Syntax to create the combined main interview weight is included below:

Compute DcWTbrch = DcWTbrth.
If (SampType = 2) DcWTbrch = DcWTchld.

7 Using the data

The GUS Sweep 3 data consists of two SPSS files

GUS_SW3_B.sav	4193 cases	Birth cohort

GUS_SW3_C.sav	2332 cases	Child cohort

7.1 Variables on the files

Each of the data files contain 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.

7.2 Variable naming convention

Variables names are made up of 8 characters, the first indicates the source of the variable, the second the year of 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 a variable name has changed substantially between sweeps this is marked in the variable list. The naming convention is summarised in Table 7.1

Table 7.1 GUS variable naming conventions

Character No:					
1	**2**		**3**	**4, 5 & 6**	**7&8**
Source of data	Sweep/Wave		Key theme prefix	Sub theme stem	Question/Variable number
Non- sequential Capitals: D,M, P, S	Sequential lower case: a, b, c..		Non-sequential Capitals: C, P, N…	Abbreviated lower case: e.g. hea,	01 - 99
Source code	**Details**	**Sweep code**	**Details**		
AL	Area Level variable				
D	Derived variable	a	Sweep 1 (2005/06)		
DP	Derived variable from partner int	b	Sweep 2 (2006/07)		
DWP	DWP variable	c	Sweep 3 (2007/08)		
M	Main carer/adult interview				
P	Partner's interview				
W	Weights and Heights				

7.3 Variable labels

In the Sweep 3 dataset the variable labels are restricted to 40 characters as far as possible; the first 2 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. The variable list shows the page numbers of the relevant questionnaire section.

7.4 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.

7.5 Household data

In addition to the questions asked about the child and parents, the respondent was also asked about each household member. The gender, age and marital status of each household member was collected along with their relationship to each other and the cohort child. Each person was identified by their person number, which they will retain through each sweep of the survey. The variable McHGSl(n) can be used to see whether a person who was in the household at sweep 1 or 2 is still in the household at sweep 3.

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 7.2. 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 and partner age, marital status and relationship to other people in the household.

Table 7.2 Key household derived variables

DcHGnmad	Dc - Number of adults in household
DcHGnmkd	Dc - Number of children in household
DcHGnmsb	Dc - Number of siblings in household
DcHGnp01	Dc - Number of natural parents in household
DcHGrsp01	Dc - Whether respondent is natural mother
DcHGrsp02	Dc - Whether respondent is natural father
DcHGnp02	Dc - Natural mother in household
DcHGnp03	Dc - Natural father in household
DcHGnp04	Dc - Respondent living with spouse/partner
DcMothID	Dc - Person number of mother
DcFathID	Dc - Person number of father
DcRespID	Dc - Person number of respondent
DcPartID	Dc - Person number of partner
DcRPAge	Dc - Respondent partners age
DcRPsex	Dc - Respondent partners sex

7.6 Childcare data

The childcare section of the CAPI questionnaire utilises feed-forward data. This technique allows information collected at the previous sweep(s) to be 'fed-forward' into the current sweep's CAPI questionnaire for the respondent to confirm or change rather than such information being

completely re-collected. This reduces respondent burden and allows for the saved time to be used elsewhere in the interview.

At sweep 3, for those cases where childcare had been used at the previous sweep, details of the previous sweep arrangements – including the provider name, provider type, the number of hours they looked after the child per week and the number of days over which those hours were spread – were fed-forward. The respondent could confirm whether all details were still correct, change the number of hours or days, or indicate that the arrangement was no longer being used. All respondents could also provide details of new arrangements which were in place at sweep 3 but had not been in place at the previous sweep. The multiple sets of information collected create a particularly complex data structure.

To make this complex picture more comprehensible, the childcare data can be usefully separated into three sections suitable for different types of analysis. The first is concerned with **continuity of provision** from sweep to sweep. The relevant variables include those which contain the details of the childcare arrangements of the previous sweep, and those which confirm whether or not the arrangement is still in place, and for those arrangements which have been ceased, the reasons why. These variables are detailed in Table 7.3.

Table 7.3 Childcare variables for exploring continuity of provision

Variable name	Description
MaCtya01/DbCtya01	Sw1 / Sw2 1st childcare provider type
MaCtma01/DbCtma01	Sw1 / Sw2 1st childcare provider - no of hours per week
MaCdya01/DbCdya01	Sw1 / Sw2 1st childcare provider - no of days per week
MaCtyb01/DbCtyb01	Sw1 / Sw2 2nd childcare provider type
MaCtmb01/DbCtmb01	Sw1 / Sw2 2nd childcare provider - no of hours per week
MaCdyb01/DbCdyb01	Sw1 / Sw2 2nd childcare provider - no of days per week
MaCtyc01/DbCtyc01	Sw1 / Sw2 3rd childcare provider type
MaCtmc01/DbCtmc01	Sw1 / Sw2 3rd childcare provider - no of hours per week
MaCdyc01/DbCdyc01	Sw1 / Sw2 3rd childcare provider - no of days per week
MaCtyd01/DbCtyd01	Sw1 / Sw2 4th childcare provider type
MaCtmd01/DbCtmd01	Sw1 / Sw2 4th childcare provider - no of hours per week
MaCdyd01/DbCdyd01	Sw1 / Sw2 4th childcare provider - no of days per week
MaCtye01/DbCtye01	Sw1 / Sw2 5th childcare provider type
MaCtme01/DbCtme01	Sw1 / Sw2 5th childcare provider - no of hours per week
MaCdye01/DbCdye01	Sw1 / Sw2 5th childcare provider - no of days per week
McCsta01	Mc Whether still using 1st provider from last sweep
McCcta01	Mc Previous 1st ccare provider - revised hrs at Sw3
McCcda01	Mc Previous 1st ccare provider - revised days at Sw3
McCrsa01	Mc - Why not using prev provider 1 at Sw3
McCstb01	Mc Whether still using 2nd provider from last sweep
McCctb01	Mc Previous 2nd ccare provider - revised hrs at Sw3
McCcdb01	Mc Previous 2nd ccare provider - revised days at Sw3
McCrsb01	Mc - Why not using prev provider 2 at Sw3
McCstc01	Mc Whether still using 3rd provider from last sweep
McCctc01	Mc Previous 3rd ccare provider - revised hrs at Sw3
McCcdc01	Mc Previous 3rd ccare provider - revised days at Sw3
McCrsc01	Mc - Why not using prev provider 3 at Sw3

McCstd01	Mc Whether still using 4th provider from last sweep
McCctd01	Mc Previous 4th ccare provider - revised hrs at Sw3
McCcdd01	Mc Previous 4th ccare provider - revised days at Sw3
McCrsd01	Mc - Why not using prev provider 4 at Sw3
McCste01	Mc Whether still using 5th provider from last sweep
McCcte01	Mc Previous 5th ccare provider - revised hrs at Sw3
McCcde01	Mc Previous 5th ccare provider - revised days at Sw3
McCrse01	Mc - Why not using prev provider 5 at Sw3
DcCstp01	Dc Whether any of the previous ccare arrgmts stopped
DcCstp02	Dc No of previous sweep providers stopped
DcCnpv01	Dc No of ccare provs from last sweep still being used
DcCapv01	Dc Whether resp still uses a previous ccare provider

The second section is concerned with the details of **new arrangements** which were in place at sweep 3.   These variables include details of the provider type, the number of hours and days per week they look after the child, the child's age when the arrangement commenced and the reasons given for using the provision. Details of the variables are listed in Table 7.4.

Table 7.4 Variables for exploring new childcare arrangements at sweep 3

Variable name	Description
McCany02	Mc If no ccare at last sweep whether using ccare at Sw3
McCany03	Mc If ccare at last sweep - any new prov at Sw3
McCtya01	Mc New provider 1 - type
McCtma01	Mc 1st new ccare provider - hours per week
McCdya01	Mc 1st new ccare provider - number of days per week
McCaga01	Mc Age (months) started new provider 1
McCwya01 – McCwya18	Mc Reasons for using 1st new provider
McCtyb01	Mc New provider 2 - type
McCtmb01	Mc 2nd new ccare provider - hours per week
McCdyb01	Mc 2nd new ccare provider - number of days per week
McCagb01	Mc Age (months) started new provider 2
McCwyb01 – McCwyb18	Mc Reasons for using 2nd new provider
McCtyc01	Mc New provider 3 - type
McCtmc01	Mc 3rd new ccare provider - hours per week
McCdyc01	Mc 3rd new ccare provider - number of days per week
McCagc01	Mc Age (months) started new provider 3
McCwyc01 – McCwyc18	Mc Reasons for using 3rd new provider
McCtyd01	Mc New provider 4 - type
McCtmd01	Mc 4th new ccare provider - hours per week
McCdyd01	Mc 4th new ccare provider - number of days per week
McCagd01	Mc Age (months) started new provider 4
McCwyd01 – McCwyd18	Mc Reasons for using 4th new provider
DcCnnp01	Dc No of new childcare arrangements at Sweep 3

Information from the first two sections was used to derive a set of variables forming the third section – **current arrangements.** These derived variables indicate - for all childcare arrangements in place at the time of the sweep 3 interview - the provider type, number of hours and days of the

arrangement, and whether or not it is a new arrangement at sweep 3. A range of summary variables indicating, for example, use of any childcare, total number of providers, total hours looked after by all providers and use of different provision are also included. These variables are detailed in Table 7.5.

Table 7.5 Variables for exploring current childcare arrangements at sweep 3

Variable name	Description
DcCtya01	Dc - Childcare prov A: provider type
DcCnwa	Dc - Provider A: new or existing
DcCtma01	Dc Provider A: No of hours per week
DcCdya01	Dc Provider A: No of days per week
DcCtyb01	Dc - Childcare prov B: provider type
DcCnwb	Dc - Provider B: new or existing
DcCtmb01	Dc Provider B: No of hours per week
DcCdyb01	Dc Provider B: No of days per week
DcCtyc01	Dc - Childcare prov C: provider type
DcCnwc	Dc - Provider C: new or existing
DcCtmc01	Dc Provider C: No of hours per week
DcCdyc01	Dc Provider C: No of days per week
DcCtyd01	Dc - Childcare prov D: provider type
DcCnwd	Dc - Provider D: new or existing
DcCtmd01	Dc Provider D: No of hours per week
DcCdyd01	Dc Provider D: No of days per week
DcCtye01	Dc - Childcare prov E: provider type
DcCnwe	Dc - Provider E: new or existing
DcCtme01	Dc Provider E: No of hours per week
DcCdye01	Dc Provider E: No of days per week
DcCany01	Dc Whether resp uses regular CCare at Sw3 (not including the excluded pre-school cases – see 7.6.1)
DcCtot01	Dc Number of ccare providers at Sw3 (not including the excluded pre-school cases – see 7.6.1)

Although not listed in Table 7.5, this section also covers variables associated with cost, availability, choice and preferences. Details of these questions and the corresponding variables are available in the sweep 3 questionnaire which accompanies this user guide.

7.6.1 Childcare and Pre-school arrangements

At the interviews for sweeps 2 and 3, children in the child cohort were aged between 3 and 5 years old. At this age, children in Scotland are eligible for funded pre-school places in private and education authority run nursery classes, nursery schools, and playgroups. It became clear on analysis of data from sweep 2 that a number of parents whose children were attending pre-school had not provided those pre-school details in the childcare section. The exclusion of these pre-school arrangements from the childcare data meant that data on the proportion of parents using childcare, the number of providers being used, the mix of provision and the total number of hours, was inaccurate in that it missed the pre-school arrangement.

To resolve this, a number of derived variables were created which incorporated information from the pre-school module and created a more accurate picture of current childcare use amongst parents.

As the childcare data at sweep 3 draws on feed-forward data from sweep 2, these amendments affect the childcare data at sweep 3. The individual details of the pre-school place at sweep 3 have been derived using the information from Sweep 2 variables *MbPRyn01* 'Mb Child attends pre-school at Sw2' and *MbPRty01* 'Mb Pre-school provider type' at Sweep 2 and incorporated in to 'current childcare arrangement' derived variables for those cases where the cohort child had not started Primary School at Sweep 3.

The final set of derived variables detailing *current* childcare arrangements for all cases at sweep 3 are listed in Table 7.6.

Table 7.6 Childcare variables including a correction for the excluded pre-school cases

Variable name	Description
DcCany02	Whether or not using childcare (including those who had excluded pre-school arrangements)
DcCtot02	Number of childcare providers being used at sw3 (including previously excluded pre-school arrgts)
DcCPrSpv	Sw3 Childcare provider E – derived provider type for those who did not provide pre-school details in childcare section
DcCPrSHr	No of hours looked after per week by provider F (excluded pre-school provider)
DcCPrSDy	No of days looked after per week by provider F (excluded pre-school provider)
DcCtmi01	Dc No of hrs child looked after by someone else (average week)
DcCtmi02	Dc No of hrs child looked after by someone else in an average week - BANDED
DcCday01	Dc Highest number of days per week in any one childcare arrangement
DcCtyp01	Dc Does respondent use grandparents for childcare?
DcCtyp02	Dc Does respondent use another relative for childcare?
DcCtyp03	Dc Does respondent use private creche/nursery for childcare?
DcCtyp04	Dc Does respondent use a childminder for childcare?
DcCtyp05	Dc Does respondent use a local authority playgroup for childcare?
DcCtyp06	Dc Does respondent use a local authority nursery for childcare?
DcCtyp07	Dc Does respondent use a private playgroup for childcare?
DcCtyp08	Dc Does respondent use a community/voluntary playgroup for childcare?
DcCtyp09	Dc Does respondent use an ex-spouse or partner for childcare?
DcCtyp10	Dc Does respondent use the childs older sibling for childcare?
DcCtyp11	Dc Does respondent use a friend or neighbour for childcare?
DcCtyp12	Dc Does respondent use a daily visiting nanny for childcare?
DcCtyp13	Dc Does respondent use a live-in nanny for childcare?
DcCtyp14	Dc Does respondent use a babysitter for childcare?
DcCtyp15	Dc Does respondent use a workplace creche or nursery for childcare?
DcCtyp16	Dc Does respondent use a family centre for childcare?
DcCtyp17	Dc Does respondent use a nursery class attached to a primary school for childcare?
DcCtyp18	Dc Does respondent use an agency carer?
DcCtyp19	Dc Does respondent use another type of childcare provider for childcare?

DcCtyp20	Dc Does respondent currently use OTHER INFORMAL childcare?
DcCtyp21	Dc Does respondent currently use NURSERY OR CRECHE for childcare?
DcCtyp22	Dc Does respondent currently use PLAYGROUP for childcare?
DcCtyp23	Dc Does respondent currently use OTHER PROVIDERS for childcare?
DcCtyp30	Dc Does respondent currently use informal childcare?
DcCtyp31	Dc Does respondent currently use formal childcare?
DcCtyp32	Dc Current use of formal and informal childcare

7.7 Indicators and summary variables

7.7.1 Socio-economic characteristics: National Statistics Socio-economic Classification (NS-SEC)

The National Statistics Socio-economic Classification (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, and 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/about-statistics/classifications/current/ns-sec/index.html

7.7.2 Socio-economic characteristics: Equivalised household annual income

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.

Table 7.7 Income equivalence scales for household members

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 (DcEqvinc) and quintiles of the scale (DcEqv5) are available in the datasets.

7.7.3   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 following categories:

Table 7.8   Scottish Government Urban Rural Classification

Classification	Description
1. Large Urban Areas	Settlements of over 125,000 people
2. Other Urban Areas	Settlements of 10,000 to 125,000 people
3. 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
4. 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
5. Accessible Rural	Settlements of less than 3,000 people and within 30 minutes drive of a settlement of 10,000 or more
6. 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 Scottish Government (2008) *Scottish Government Urban Rural Classification 2007 – 2008.* This document is available online at http://www.scotland.gov.uk/Publications/2008/07/29152642/0

7.7.4   Area-level variables: Scottish Index of Multiple Deprivation

The Scottish Index of Multiple Deprivation (SIMD) 2006 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 2006 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. The classificatory variable contained in the GUS Sweep 3 datasets is based on the 2006 version of SIMD. It should be noted that the analyses in the GUS Sweep 1 report are based on the 2004 version of SIMD as the 2006 version had not been published at the time the report was being written.

In the dataset, 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.

Further details on SIMD can be found on the Scottish Government Website http://www.scotland.gov.uk/Topics/Statistics/SIMD/Overview

7.7.5   Cognitive Assessments: British Ability Scales – Naming Vocabulary and Picture Similarities

Score variables in the dataset
The dataset provides the following scores for each assessment:

Raw score – the number of correct responses amongst the items administered. Note that because children take different sets of items, their raw scores cannot be compared directly.

Ability Score – Derived from the raw score 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. Note that DcPSAbSc/DcNVAbSc variables must be used for comparisons with Birth Cohort 2 (see below).

Normative scores – Derived from standard BAS tables and defined with reference to the standardisation sample used in developing the assessment. Both T-scores (with mean=50 and standard deviation=10) and their equivalent Percentiles are provided. Note that DcPSTSc/DcNVTSc variables must be used for comparisons with Birth Cohort 2 (see below).

Variable names and labels are detailed in table 7.9.

Table 7.9   BAS Summary Score Variables

Variable name	Label
DcPicRaw	Dc Picture Similarities Raw Score
DcPicSAS	Dc Picture Similarities Ability Score
DcPicSTS	Dc Picture Similarities T-Score
DcPicSPt	Dc Picture Similarities Percentile Equivalent
DcNamRaw	Dc Naming Vocabulary Raw Score
DcNamVAS	Dc Naming Vocabulary Ability Score
DcNamVTS	Dc Naming Vocabulary T-Score
DcNamPt	Dc Naming Vocabulary Percentile Equivalent
DcPSAbSc	Dc Picture Similarities Ability Score (for comparing with BC2)
DcPSTSc	Dc Picture Similarities T-Score (for comparing with BC2)
DcNVAbSc	Dc Naming Vocabulary Ability Score (for comparing with BC2)
DcNVTSc	Dc Naming Vocabulary T-Score (for comparing with BC2)

Comparison with BC2

Birth Cohort 2 children carried out Naming Vocabulary and Picture Similarities exercises when they were the same age (34 months). However, different editions of the assessments were used: For BC1, the 2nd edition assessment was used (BAS-II), whereas for BC2 the 3rd edition was used (BAS3). 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 original 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 two cohorts <u>MUST</u> use the adjusted BC1 scores (DcPSAbSc, DcPSTSc, DcNVAbSc, DcNVTSc).

Note also that because of the adjustments, it is not possible to convert differences in average cognitive ability scores to developmental age in months when using the adjusted scores.

General 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 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 was 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 7.10 contains details of the relevant variables where this information is recorded.

Table 7.10   Variables recording difficulties experienced during assessments

Variable name	Label
CAssPrb1	Mc - No difficulties experienced during assessments
CAssPrb2	Mc - Difficulties experienced because: Assessment was interrupted
CAssPrb3	Mc - Difficulties experienced because: Child was ill
CAssPrb4	Mc - Difficulties experienced because: Child was tired
CAssPrb5	Mc - Difficulties experienced because: Parent interfered
CAssPrb6	Mc - Difficulties experienced because: Something else

7.7.6   Child Development: Strengths and Difficulties Questionnaire

Parents in the child cohort completed the Strengths and Difficulties Questionnaire (SDQ). The SDQ is a brief behavioural screening questionnaire designed for use with 3-16 year olds. 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 7.11 with syntax illustrated in the derived variables documentation.

Table 7.11   Derived variables associated with the Strengths and Difficulties Questionnaire

Variable name	Description
DcDsdem1	Dc SDQ: Emotional symptoms score
DcDsdco1	Dc SDQ: Conduct problems score
DcDsdhy1	Dc SDQ: Hyper-activity or inattention score
DcDsdpr1	Dc SDQ: Peer problems score
DcDsdps1	Dc SDQ: Pro-social score
DcDsdto1	Dc SDQ: Total difficulties score

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

7.7.7 Child Development: Insecure Attachment

Selected items taken from the Temperament scale developed for the Child Surveys of the National Longitudinal Study of Youth (NLSY) were included as a measure of insecure attachment. The full NLSY scale contains a mixture of maternal-report items (for all children younger than seven years) and interviewer observations and different scales and items are used for children of different ages: the maternal scale "How My Infant Usually Acts" addresses the activity, predictability, fearfulness, positive affect, and friendliness of infants below age one; "How My Toddler Usually Acts" addresses the fearfulness, positive affect, and friendliness of one-year-olds; "How My Child Usually Acts" measures the compliance and attachment of two- and three-year-olds and additionally, the friendliness of children aged four through six.

Five items measuring insecure attachment, from the 'How My Toddler Usually Acts' scale, were included in sweep 3 of GUS. The variable names and labels are detailed in Table 7.11.

Table 7.11   Selected items measuring 'insecure attachment' originally developed for NLSY

Variable name	Description
McHatt01	Mc - child stays close to resp
McHatt02	Mc – child copies resp
McHatt03	Mc – child gets upset when resp leaves room
McHatt04	Mc – child gets impatient (is demanding)
McHatt05	Mc – child tries to help resp (is empathetic)

Further details on the NLSY Child survey can be found at:
http://www.bls.gov/nls/nlsy79ch.htm

For more information on the various temperament scales developed for use in the survey, consult the *NLSY79 Child & Young Adult Data Users Guide* which is available here:
http://www.nlsinfo.org/pub/usersvc/Child-Young-Adult/2004ChildYA-DataUsersGuide.pdf

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

At sweeps 1 and 3 of GUS, 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.

Table 7.12   Constituent and derived variables associated with the SF-12

Variable name	Description
McHpgn01	Mc - How is resp health in general
McHlmt01	Mc - Resp health limits moderate activities
McHlmt02	Mc - Resp health limits climbing stairs
McHlmt03	Mc - Resp health limited accomplishments past 4 wks
McHlmt04	Mc - Resp health limited reg activities past 4 wks
McHlmt05	Mc - Resp mental health limited accomplishments past 4 wks

McHlmt06	Mc - Resp mental health limited quality of accomplishments past 4 wks
McHlmt07	Mc - Resp physical pain limited normal work past 4 wks
McHpgn02	Mc - Time resp felt calm in past 4 wks
McHpgn03	Mc - Time resp felt energetic in past 4 wks
McHpgn04	Mc - Time resp felt down in past 4 wks
McHpgn05	Mc - Time resp health interfered socially in past 4 wks
DcSF12ph	Dc - Physical PCS - 12 Scale
DcSF12mn	Dc - Mental MCS - 12 Scale

7.8   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:

1. Those containing text
2. Those which contained a personal identifier (e.g. name/address)
3. 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 Scottish index of multiple deprivation summary variable.

7.9   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. .The derived variable specifications should be consulted for details.

8 Documentation

The documentation has been organised into the following sections:
- Survey materials containing interviewer and coding instructions.
- Data documentation containing the questionnaire with variable names added, the list of variables in the dataset (including derived variables), a separate list of derived variables with their SPSS

syntax and the show cards.

9 Related publications

Further information about GUS Sweep 3 is available in:

Bradshaw, P., Sharp, S, Webster, C. and Jamieson, L. (2009) *Growing Up in Scotland: Parenting and the Neighbourhood Context*, Edinburgh: Scottish Government
http://www.scotland.gov.uk/Publications/2009/03/13143448/11

Bradshaw, P. and Wasoff, F. (2009) *Growing Up in Scotland: Multiple Childcare Provision and its Effects on Child Outcomes*, Edinburgh: Scottish Government
http://www.scotland.gov.uk/Resource/Doc/263884/0079032.pdf

Bromley, C. (2009) *Growing Up in Scotland: the Impact of Children's Early Activities on Cognitive Development,* Edinburgh: Scottish Government
http://www.scotland.gov.uk/Resource/Doc/263956/0079071.pdf

Marryat, L., Reid, S. and Wasoff, F. (2009) *Growing Up in Scotland Sweep 3 Non-resident Parent Report,* Edinburgh: Scottish Government
http://www.scotland.gov.uk/Publications/2009/01/21085002/0

Marryat, L., Skafida, V. and Webster, C. (2009) *Growing Up in Scotland Sweep 3 Food and Activity Report,* Edinburgh: Scottish Government
http://www.scotland.gov.uk/Publications/2009/01/21085143/0

Other publications which include the use of GUS data include:

Bradshaw P, Cunningham-Burley S, Dobbie F, McGregor A, Marryat L, Ormston, R. and Wasoff F. (2008) *Growing Up in Scotland: Sweep 2 Overview Report,* Edinburgh: The Scottish Government

Anderson S, Bradshaw P, Cunningham-Burley S, Hayes F, Jamieson L, McGregor A, Marryat L and Wasoff F. (2007) *Growing Up in Scotland: Sweep 1 Overview Report,* Edinburgh:The Scottish Executive

Bradshaw, P. with Jamieson, L. and Wasoff, F. (2008) *Use of informal support by families with young children*, Edinburgh: Scottish Government

Bradshaw, P. and Martin, C. with Cunningham-Burley, S. (2008) *Exploring the experience and outcomes for advantaged and disadvantaged families* Edinburgh: Scottish Government

Jamieson, L. with Ormston, R. and Bradshaw, P. (2008) *Growing Up in Rural Scotland*, Edinburgh: Scottish Government

Skafida, V. (2008) "Breastfeeding in Scotland: The impact of advice for mothers", *Centre for Research on Families and Relationships, Briefing 36, February 2008,* Edinburgh: Centre for Research on Families and Relationships

Skafida, V. (2009) "The relative importance of social class and maternal education for breast-feeding initiation", *Public Health Nutrition,* First View Article, published 26 Feb 2009

Links to these reports and others, along with additional related information are available on the GUS website: http://www.growingupinscotland.org.uk/

10 Contact details

Contacts at the Scottish Centre for Social Research, 73 Lothian Road, Edinburgh, EH3 9AW

GUS Project Manager	Paul Bradshaw	0131 221 2564	p.bradshaw@scotcen.org.uk
GUS Data Manager	Joan Corbett	0131 221 2552	j.corbett@scotcen.org.uk
GUS Data Assistant	Mireille Ferrandon	0131 221 2578	m.ferrandon@scotcen.org.uk

Appendix A: Full non-response models

Table A1 Non-response model for birth cohort (Sample A)

	B	S.E.	Wald	df	Sig.	Exp(B)
Tenure			18.1	2	0.000	
Owner occupiers					(baseline)	
Rents HA/council	-0.239	0.152	2.5	1	0.115	0.79
Rents private	-0.728	0.173	17.7	1	0.000	0.48
Age of mother (grouped)			50.8	4	0.000	
<20					(baseline)	
20-24	0.061	0.173	0.1	1	0.726	1.06
25-29	0.694	0.192	13.0	1	0.000	2.00
30-34	0.977	0.202	23.3	1	0.000	2.66
35+	1.123	0.231	23.7	1	0.000	3.07
Ethnicity of mother			12.5	1	0.000	
White					(baseline)	
Other ethnic background	-0.818	0.231	12.5	1	0.000	0.44
Household employment			7.9	2	0.019	
At least one parent/carer in full-time employment					(baseline)	
At least one parent/carer in part-time employment	-0.059	0.135	0.2	1	0.660	0.94
No parent/carer working	-0.434	0.164	7.0	1	0.008	0.65
Whether child was breastfed			14.1	1	0.000	
Yes					(baseline)	
No	-0.443	0.118	14.1	1	0.000	0.64
Constant	2.324	0.208	124.3	1	0.000	10.21

Notes:
1. The response is 1 = sample A response to wave 3, 0 = sample A non-response.
2. Model is weighted by wave 2 birthweight
3. The model R^2 = 0.044 (Cox and Snells).
4. B is the estimate coefficient with standard error S.E.
5. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df. If the test is significant (sig < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model.
6. The Wald test for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.

Table A2 Non-response model for child cohort (Sample A)

	B	S.E.	Wald	df	Sig.	Exp(B)
Tenure			7.2	2	0.028	
Owner occupiers					(baseline)	
Rents HA/council	-0.079	0.210	0.1	1	0.708	0.92
Rents private	-0.595	0.238	6.2	1	0.013	0.55
Age of mother (grouped)			12.8	4	0.012	
<20					(baseline)	
20-24	0.072	0.265	0.1	1	0.786	1.07
25-29	0.318	0.280	1.3	1	0.256	1.38
30-34	0.636	0.292	4.7	1	0.030	1.89
35+	0.932	0.338	7.6	1	0.006	2.54
Family status			13.8	1	0.000	
Single parent					(baseline)	
Couple parent	0.681	0.183	13.8	1	0.000	1.98
Number of children in the household			13.6	3	0.003	
1					(baseline)	
2	-0.490	0.188	6.8	1	0.009	0.61
3	-0.263	0.251	1.1	1	0.295	0.77
4+	-1.027	0.306	11.3	1	0.001	0.36
Respondent NS-SEC (5 groups)			15.4	5	0.009	
Managerial and professional occupations					(baseline)	
Intermediate occupations	0.127	0.244	0.3	1	0.603	1.14
Small employers and own account workers	-0.406	0.329	1.5	1	0.216	0.67
Lower supervisory and technical occupations	-0.680	0.274	6.2	1	0.013	0.51
Semi-routine and routine occupations	0.139	0.218	0.4	1	0.524	1.15
Missing/never worked	-0.478	0.323	2.2	1	0.138	0.62
Constant	1.967	0.335	34.5	1	0.000	7.15

Notes:
1. The response is 1 = sample A response to wave 3, 0 = sample A non-response.
2. Model is weighted by wave 2 child weight
3. The model R^2 = 0.044 (Cox and Snells).
4. B is the estimate coefficient with standard error S.E.
5. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df. If the test is significant (sig < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model.
6. The Wald test for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.

Table A3 Distribution of sample A

	BIRTH COHORT			CHILD COHORT		
	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight
	%	%	%	%	%	%
Tenure						
Owner occupier	62.7	65.0	62.8	63.0	64.8	63.1
Rents HA/council	28.0	26.6	28.0	28.2	27.1	28.0
Rents private	9.3	8.4	9.2	8.8	8.1	8.8
Family status						
Lone parent	19.9	18.3	19.7	23.3	21.5	23.1
Couple parent	80.1	81.7	80.3	76.7	78.5	76.9
Whether child was mother's first-born						
First born	49.7	49.1	49.8	47.9	47.6	47.6
Other children	50.3	50.9	50.2	52.1	52.4	52.4
Mother's age						
<20	7.6	6.6	7.6	7.0	6.3	6.8
20-24	17.3	15.8	17.3	18.3	17.5	18.3
25-29	23.4	23.7	23.4	22.6	22.5	22.7
30-34	31.3	32.6	31.4	33.1	33.9	33.1
35+	20.3	21.3	20.3	19.0	19.8	19.0
Highest education level of respondent						
Degree or equivalent	26.7	28.1	27.0	27.1	28.2	27.5
Vocational qualification below degree	37.1	36.9	36.8	37.5	37.1	37.2
Higher Grade or equivalent	8.1	8.2	8.2	7.4	7.5	7.5
Standard Grade or equivalent	18.7	17.9	18.6	17.7	17.1	17.4
No Qualifications	9.4	8.9	9.3	10.4	10.0	10.3
Annual household income						
<£10,000	15.4	14.4	15.7	15.2	14.2	15.1
£10,000-£19,999	21.0	20.5	20.9	20.6	20.0	20.4
£20,000-£31,999	22.9	23.2	22.8	21.5	21.8	21.6
£32,000+	35.0	36.7	35.3	35.9	37.5	36.3
Missing	5.7	5.2	5.2	6.9	6.6	6.6

Table A3 Distribution of sample A (continued)

	BIRTH COHORT			CHILD COHORT		
	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight
Respondent NSSEC						
Managerial and professional occupations	34.6	36.1	34.8	34.5	35.5	34.5
Intermediate occupations	19.3	19.9	19.7	17.0	17.3	17.0
Small employers and own account workers	5.1	5.3	5.2	5.5	5.5	5.5
Lower supervisory and technical occupations	6.0	5.5	5.6	6.5	6.0	6.5
Semi-routine and routine occupations	30.4	29.0	30.0	32.1	31.8	32.0
Missing/never worked	4.6	4.2	4.7	4.4	4.0	4.4
Ethnicity of respondent						
White	96.1	96.4	96.2	96.2	96.3	96.3
Other ethnic background	3.9	3.6	3.8	3.8	3.7	3.7
Household employment						
At least one parent/carer in full-time employment	34.3	35.0	34.3	33.3	33.9	33.4
At least one parent/carer in part-time employment	49.0	49.9	49.0	49.4	50.4	49.8
No parent/carer working	16.8	15.1	16.7	17.3	15.6	16.8
Mother's employment status						
Childs mother working - full-time	14.7	15.0	14.6	15.5	15.8	15.5
Childs mother working - part-time	45.9	46.9	45.9	46.8	47.9	47.3
Childs mother not working	39.4	38.1	39.4	37.7	36.3	37.3
Number of children in the household						
1	41.1	40.8	41.2	26.2	26.5	26.2
2	38.0	38.3	37.9	49.9	49.7	50.0
3	15.5	15.9	15.8	18.3	18.5	18.2
4+	5.3	5.1	5.1	5.6	5.3	5.7

Table A3 Distribution of sample A (continued)

	BIRTH COHORT			CHILD COHORT		
	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight	Wave 2 weighted by W2 weight	Wave 3 weighted by W2 weight	Wave 3 weighted by W3 weight
Area Urban/rural indicator						
1 Large urban area	38.5	38.2	38.3	36.3	36.2	36.3
2 Other urban area	31.2	30.5	30.7	31.3	31.1	31.2
3 Accessible small town	9.6	9.9	10.0	10.9	10.8	10.9
4 Remote small town	0.9	1.0	1.0	1.4	1.4	1.4
5 Very remote small town	1.8	1.9	1.9	1.4	1.4	1.4
6 Accessible rural	13.6	14.1	13.8	14.1	14.3	14.1
7 Remote rural	2.1	2.2	2.1	2.6	2.6	2.6
8 Very remote rural	2.1	2.2	2.2	1.9	2.0	2.0
Use regular childcare?						
Yes	68.3	68.7	68.3	83.5	83.9	83.9
No	31.7	31.3	31.7	16.5	16.1	16.1
Area deprivation quintiles						
0.95 - 7.75 (least deprived)	18.4	19.4	18.7	19.7	20.3	19.8
7.75 - 13.56	19.2	19.8	19.3	20.7	21.0	20.8
13.56 - 21.04	19.5	19.8	19.6	19.6	19.9	19.8
21.05 - 33.70	18.2	17.8	18.1	17.4	17.0	17.2
33.71 -89.09 (most deprived)	24.7	23.3	24.3	22.5	21.8	22.4
Regularly attended toddler groups in the last year?						
Yes	39.0	40.3	39.5	42.4	43.5	43.1
No	61.0	59.7	60.5	57.6	56.5	56.9
Was child breastfed?						
Yes	60.1	61.8	60.1	59.5	60.1	59.3
No	39.9	38.2	39.9	40.5	39.9	40.7
Does child have a longstanding illness or disability?						
Yes	11.0	10.9	11.0	15.6	15.3	15.5
No	89.0	89.1	89.0	84.4	84.7	84.5
Respondent's self-reported general health						
Excellent	19.1	19.1	18.8	20.7	20.4	20.2
Very Good	41.1	41.4	41.2	37.9	38.2	37.9
Good	27.0	26.9	27.1	27.1	27.5	27.7
Fair	10.7	10.5	10.7	11.7	11.4	11.7
Poor	2.1	2.1	2.1	2.6	2.4	2.5