





GROWING UP IN IRELAND A SUMMARY GUIDE TO COHORT '08 AT 9 YEARS (WAVE 5 OF THE INFANT COHORT AT 9 YEARS)

Amanda Quail, Caoimhe O'Reilly, Dorothy Watson,
Eoin McNamara, Desmond O'Mahony and Aisling Murray

Table of Contents

1		Introduction and Background	5
	1.1	Introduction	5
	1.2	Background	5
2		Sample Design	7
	2.1	Introduction	7
	2.2	Sample Design at Wave 1 (9 months old)	7
	2.3	Sample Design at Wave 2 (3 years old)	8
	2.4	Sample Design at Wave 3 (5 years old)	8
	2.5	Sample Design at Wave 4 (7/8 years old, the inter-wave postal phase)	8
	2.6	Sample Design at Wave 5 (9 years old)	8
	2.7	Response Rates	9
	2.8	Attrition and reweighting the data	9
3		Instrument Development & Piloting	13
	3.1	Introduction	13
	3.2	Instrument Development	13
	3.3	Piloting the Instruments	13
4		Survey Instruments	. 14
	4.1	Introduction	. 14
	4.2	Household-based instruments	. 14
	4.3	School-based instruments	. 16
5		Fieldwork and Implementation	. 18
	5.1	Introduction	. 18
	5.2	Interviewer Training	. 18
	5.3	Vetting	. 18
	5.4	Contacting a Household	. 18
	5.4.1	Follow-Up/Tracing Information	. 19
	5.4.2	Proactive tracing procedures	
	5.4.3	Retrospective tracing procedures	
_	5.5	Incidents	
6		Structure and Content of the Data File	
	6.1	Introduction	
	6.2	Anonymised (AMF) and Researcher (RMF) Microdata Files	
	6.3	Structure of the data file	
	6.4	Variable naming	
	6.4.1 6.4.2	Naming of Home-based Primary and Secondary Caregiver Variables Naming of Home-based Child Variables	
	6.4.2 6.4.3	Naming of School-based Variables	
	6.4.4	Naming of other variables	
	6.5	Variable order	
	6.6	Identification Codes	23
	6.7	The Household Grid	. 23
	6.8	The Main Respondent – Primary Caregiver	. 24
	6.9	Twins	. 25
	6.10	Weighting variables	. 25

	6.11	Derived Variables	26
	6.11.1	Household type (b5_hhtype4, b5_hhtype4_v2)	26
		Equivalised Household Income (b5_equivinc, b5_EIncDec, b5_EIncQuin)	
	6.11.3	Household Class (b5_hsdclass)	27
		Household location (b5_region)	
	6.12	Scaled Measures Used in the Study	28
	6.12.1	Strengths & Difficulties Questionnaire (SDQ; Goodman, 1997)	28
		The Pianta Scales - Child Parent Relationship Scale (CPRS) and Student-Teach	
		Relationship Scale (STRS) (Pianta, 1992)	
		$Parenting \ Style \ Measure \ (from \ the \ Longitudinal \ Study \ of \ Australian \ Children \ [LSAC]) \dots$	
		Parental Stress Scale (Berry and Jones, 1995)	
	6.12.5	Centre for Epidemiological Studies Depression Scale (CES-D) (Melchior, Huba, Brown	
		Reback, 1993)	
		Dyadic Adjustment Scale (DAS-4) (Sabourin et al., 2005)	
		FAST Alcohol Screening Test (Hodgson et al., 2002)	
	6.12.8	The Everyday Discrimination Scale short version (EDS) (Sternthal, Slopen, & William	
		2011)	
		Co-parenting scale (Feinberg, 2003)	
		Piers-Harris Scale (Piers, 1969)	
	6.13	Physical Measurements	
		Height & Weight	
		Body Mass Index (BMI)	
	6.14	Cognitive Assessments	
	6.15	Coding & Editing	31
	6.16	Forward-feed from Wave 3	32
	6.17	Missing data	33
7		Ethical Considerations	33
R	eference	25	34
Α	ppendix	: Indicative Cronbach's alphas for all scales used in the study	35

List of Tables

Table 2.1 Participation rates Wave 1 to Wave 5	9
Table 2.2 Weighting factors available for use with Infant Cohort at 9 years of age	11
Table 2.3 Combinations of data waves and statistical adjustment factors	11
Table 4.1 Household based instruments used at Wave 5	14
Table 4.2 School based instruments used at Wave 5	16
Table 6.1 Ordering of variables in the data file	23
Table 6.2 Frequency distribution of 'xxwave' variables	26
Table 6.3 Subscales of the Strengths and Difficulties Questionnaire and their relevant variable	ole names
	28
Table 6.4 Subscales of the Piers-Harris and their relevant variable names	30
Table 6.5 Details on variables forward-fed from previous waves (excl. household grid a	and adult
height)	32

1 Introduction and Background

1.1 Introduction

This document provides a summary of the fifth wave of data collection with the Infant Cohort (also known as Cohort '08) at 9 years of age of the *Growing Up in Ireland* (*GUI*) study, as well as an overview of the microdata files (Researcher and Anonymised) from that phase of the project.

Growing Up in Ireland - the National Longitudinal Study of Children – is the first project of its kind ever undertaken in Ireland and, as such, aims to explore the many and varied factors that contribute to or undermine the wellbeing of children currently living here. A two cohort longitudinal design was adopted and began with one cohort (the Infant Cohort or Cohort '08) of 11,134 infants (aged 9 months) and a second cohort (the Child Cohort or Cohort '98) of 8,568 9-year olds. Since the project is longitudinal in nature respondents in both cohorts are being interviewed on a number of occasions. The families of the infants were interviewed during Phase 1 of the GUI study when the children were 9 months, 3 years and subsequently 5 years of age, while the Child Cohort and their parents/guardians were interviewed at 9 and 13 years of age. Phase Two of GUI returns to the Infant Cohort at 7 years¹ and 9 years of age, and to the Child Cohort at 17/18 years and 20 years of age.

The 11,134 children representing the Infant Cohort were born between 1st December 2007 and the 30th June 2008 and data collection for the first wave, at age 9 months, took place between September 2008 and April 2009. 9,793 of these original families participated in the second wave of data collection, at age 3 years (Dec 2010 – July 2011), and 9,001 at age 5 years (Mar – Sep 2013). 5,344 completed a short postal survey at age 7/8 years (Spring 2016).

The fifth wave of data collection (which is the focus of the current document) took place between June 2017 and Feb 2018, when the cohort were 9 years of age, resulting in a completed data file of 8,032 cases. Wave 5 also included a school-based phase of fieldwork. The data from the home-based phase of fieldwork were archived in August 2019, with a second version of the data file including the school-based data released in October 2019.

This report describes the background, design, instruments and procedures used only in respect of Wave 5 of the Infant Cohort. Earlier waves of this cohort (and the Child Cohort) are the subjects of a parallel set of reports. The focus here is on the sample design and response rates in Wave 5 of the Infant Cohort, the nature and content of the questionnaires and other instrumentation, along with a broad overview of the datasets. For a more in-depth discussion of all these topics please see the report entitled 'Design, Instrumentation and Procedures for Cohort '08 of Growing Up in Ireland at 9 years old (Wave 5)' by Eoin McNamara, Desmond O'Mahony and Aisling Murray.

1.2 Background

The principal objective of *Growing Up in Ireland* is to increase our understanding of the determinants and drivers of children's wellbeing and its change and transformation over time, with a

¹ On a postal basis

view to improving our understanding of children's development across a range of domains. The study provides an evidence-informed input to policy formation and design of services for families and children in 21^{st} century Ireland.

Growing Up in Ireland is the national longitudinal study of children in Ireland. It is funded by the Department of Children and Youth Affairs, with a contribution from The Atlantic Philanthropies in Phase 2. The study is managed and overseen by the Department of Children and Youth Affairs in association with the Central Statistics Office. It is carried out by a consortium of researchers led by the Economic and Social Research Institute (ESRI) and Trinity College Dublin.

Growing Up in Ireland is designed to describe and analyse what it means to be a child or young person in Ireland today and to understand the factors associated with children's wellbeing, including those impacting on their physical health and development, social/emotional/behavioural wellbeing, and educational achievement/intellectual capacity. While current wellbeing is of immense importance, researchers are also cognisant of the future outcomes for the children and young people as they develop into adulthood. The longitudinal nature of the project allows one to record current data with a view to using them to assist in understanding future outcomes; in the case of the Infant Cohort, researchers are afforded the opportunity to track the same group of children from infancy through to nine years of age. By gathering comprehensive data on childhood development the Study will provide a statistical basis for policy formation and applied research across all aspects of a child's development – currently and into the future.

The Study has 9 over-arching objectives²:

- 1. To describe the lives of Irish children, to establish what is typical and normal as well as what is atypical and problematic
- 2. To chart the development of Irish children over time, to examine the progress and wellbeing of children at critical periods from birth to adulthood
- 3. To identify the key factors that, independently of others, most help or hinder children's development
- 4. To establish the effects of early child experiences on later life
- 5. To map dimensions of variation in children's lives
- 6. To identify the persistent adverse effects that lead to social disadvantage and exclusion, educational difficulties. ill health and deprivation
- 7. To obtain children's views and opinions on their lives
- 8. To provide a bank of data on the whole child
- 9. To provide evidence for the creation of effective and responsive policies and services for children and families

Full details on the underlying theoretical and conceptual framework can be found in Greene *et al.*, 2010.

² Request for Tenders (RFT) for Proposals to Undertake a National Longitudinal Study of Children in the Republic of Ireland, issued by the National Children's Office of the Department of Health and Children and the Department of Social and Family Affairs, December 2005, p.20.

2 Sample Design

2.1 Introduction

In order to provide the reader with an overview of the sampling procedures used in *Growing Up in Ireland (GUI)*, this section provides a brief outline of the sample design at the first wave of data collection with the Infant Cohort, as well as the subsequent waves. Response and attrition rates for the current wave of the study are then discussed. The process of statistically reweighting the data to ensure that they are fully representative of the population will also be outlined.

2.2 Sample Design at Wave 1 (9 months old)

Full details on the population, sampling frame and sample design for the Infant Cohort are given in the report entitled "Sample Design and Response in Wave 1 of the Infant Cohort (at 9 months) of *Growing Up in Ireland*": https://www.ucd.ie/t4cms/GUI-SampleDesignResponseInfants.pdf. This section presents a brief outline of the sampling at Wave 1, to provide the reader with a background to the sampling procedures used.

The Child Benefit register was used as the sampling frame for the first Wave of the Infant Cohort. Child Benefit is a universal monthly social welfare payment to families with children. Children should be registered with the appropriate authorities within 6 months of birth or becoming part of the family (e.g. through adoption), or of the family coming to reside in Ireland. This administrative database had some extremely attractive characteristics as a sampling frame. It contained a comprehensive up-to-date listing of eligible members of the relevant population; had a wide range of relevant characteristic variables of claimants (mostly mothers); and was already in an electronic form that could be accessed for sampling purposes.

There were just over 70,000 births in Ireland in 2007. The Wave 1 sample for the Infant Cohort study was selected from the 41,185 infants registered on the Child Benefit Register³ as having been born between 1st December 2007 and 30th June 2008. The target sample was selected over this sevenmonth period, with a view to carrying out fieldwork for Wave 1 when the children were 9 months of age, between September 2008 and March/April 2009. The sample was selected on a systematic basis, with a random start. Prior to selection, the sample was sorted by marital status of the claimant (usually the mother), county of residence and nationality of the claimant, as well as number of children in the payment or 'claim'. A simple systematic selection procedure based on a random start and constant sampling fraction was used. In total, 11,134 children were recruited onto the first wave of the study; representing a response rate of 65 per cent of all families approached and 69 per cent of valid contacts made in the course of the fieldwork.

The final completed Wave 1 sample of 11,134 infants and their families formed the target sample for Wave 2. The Study Child is the longitudinal focus of the study. We are interested throughout the study in tracking, interviewing, measuring and testing the child, regardless of changes in his/her family

³ Special permission was required to access the Child Benefit Register for sampling purposes and was possible only as the overall study is being conducted under the Statistics Act, 1993, which provides the legal basis of GUI.

composition, structure, location etc. In this respect the study is based on a pure, fixed panel of children who were nine months of age at the time of first interview.

2.3 Sample Design at Wave 2 (3 years old)

The Wave 2 target sample contained the 11,134 Study Children (and their families) who participated in the first round of interviewing. No additions were made to the sample since that time,⁴ with the only loss being through inter-wave non-response or attrition (including families who had moved away from Ireland between Waves 1 and Wave 2 or children who had deceased since the first round of interviewing). The longitudinal population at Wave 2, therefore, was the population of three year olds (and their families) who had been resident in Ireland at Wave 1 (when they were 9 months old) and who continued to be resident in Ireland at Wave 2 (at age 3 years). In total 9,973 families participated at Wave 2.

2.4 Sample Design at Wave 3 (5 years old)

The target sample at Wave 3 was made up of the 9,793 children and families who participated in Wave 2. In addition, it included most of those who participated at Wave 1 but who had refused or otherwise did not participate at Wave 2. Families who had moved abroad, moved within Ireland with no forwarding address, or who has said very definitely at Wave 2 that they did not wish to be contacted further about the study were not included in the Wave 3 sample. This resulted in a total Wave 3 sample of 10,586 families, of which, 9,001 completed a Wave 3 interview

2.5 Sample Design at Wave 4 (7/8 years old, the inter-wave postal phase)

For the fourth wave of the study, a single postal questionnaire was sent to the home with an accompanying letter and Information Sheet. The questionnaire was self-completed and returned by post by the Study Child's Primary Caregiver. Up to two reminders were sent by post and a sub-sample were followed up for reminders by telephone.

A total of 5,344 questionnaires were returned, amounting to 48% of the families interviewed at 9 months of age. However, this response rate does not take account the families who no longer lived in Ireland at the time of the survey, nor those whose letters were returned by the postal service as being unknown at the last address then available to the Study Team.

2.6 Sample Design at Wave 5 (9 years old)

A total of 10,052 children and their families were targeted in Wave 5, when the children were 9 years of age. This was made up of the families who had participated in the face-to-face interview in Wave 3 (when the Study Child was 5 years of age), as well as a small proportion of those who had not participated in Wave 3 but who had participated at one of the earlier rounds of the study.

⁴ Additions to membership of the Study Child's household between waves (in the form of new members residing in the household or being born into the household) are, of course, recorded on the household register in the relevant wave

2.7 Response Rates

As noted above, the Wave 1 sample was selected from the Child Benefit register and a total of 11,134 families participated at that stage of the study. These 11,134 respondent families made up the target sample for Wave 2. The target sample for subsequent waves included the 11,134 who participated at Wave 1 less those that moved outside of Ireland; or definitively opted out of the study ('hard refusals'); or, in a small number of cases, deceased between waves. Those who did not participate in any one wave due to time constraints etc. (so-called 'soft refusals') were included in the sample at the next wave.

Table 2.1 summarises the participation response rates (the proportion of the original Wave 1 cohort who participated at each wave) across each wave of the study. By Wave 5, 72.1% of the original Wave 1 sample participated in the face-to-face interview.

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Year	2008 –2009	2010 – 2011	2013-2014	2015-2016	2017-2018
Child age	9 months	3 years	5 years	7-8 years	9 years
Wave participation	11,134	9,793	9,001	5,344	8,032
Participation rate		88.0%	80.8%	48 N%	72 1%

Table 2.1 Participation rates Wave 1 to Wave 5

2.8 Attrition and reweighting the data

Non-response and inter-wave attrition are unavoidable in longitudinal surveys, regardless of tracking and conversion procedures employed. These become a problem where they are systematically related to family or other characteristics or with child outcomes. It is important to understand the levels and correlates of attrition and non-response to inform re-weighting procedures that statistically adjust the data for systematic non-response or attrition prior to analysis.

The *Design report* compared the distribution of the unweighted completed sample at Wave 5 to the weighted Wave 1 sample and found there was an under-representation of younger mothers and those in less advantaged circumstances (lower levels of education, income, lower social class, families where the parents are not in employment), one-parent families, renters of social housing (rented from the local authorities or Approved Housing Bodies), in the Dublin region, where the Study Child was never breastfed and where the Primary Caregiver smoked daily. There is also a slight under-representation of mothers born in Ireland (relative to those from outside of Ireland and the UK) and a very small under-representation of boys compared to girls.

The construction of the analysis weight for the 9-year data consists in carrying forward the earlier weight (which controls for initial non-response and attrition up to the 5-year wave) and adjusting it for attrition between the 5-year and 9-year waves. The Study Team used the GROSS software, as in previous rounds of *Growing Up in Ireland*.⁵ This has been used extensively by the Economic and Social

⁵ See, for example, Gomulka, J., 1992. "Grossing-Up Revisited", in R. Hancock and H. Sutherland (Eds.), Microsimulation Models for Public Policy Analysis: New Frontiers, STICERD, Occasional Paper 17, LSE. Gomulka, J., 1994. "Grossing Up: A Note on Calculating Household Weights from Family Composition

Research Institute (ESRI) since 1996. GROSS uses a minimum information-loss algorithm to fit a sample distribution of characteristics to population 'control totals'. An iterative procedure is used, allowing marginals of characteristics that are associated with one another to be fitted simultaneously.

The sample weights for the 9-year phase of the Infant Cohort were constructed by taking the weight from the previous wave as the initial weight, then calculating an adjustment factor for the carried-forward weight for each case so that the population distribution is replicated for the sample. The variables used to adjust for attrition and to generate the 9-year weights are those identified as being related to non-response:

- Age of Primary Caregiver at birth of the Study Child
- Educational attainment of Primary Caregiver
- Family structure / Primary Caregiver marital status (cohabiting or married)
- Family income quintile
- Family social class
- Work Status of Primary and Secondary Caregivers
- Where the Primary Caregiver was born (4 categories)
- Housing tenure (owner, Local Authority/Approved Housing Body renter, private renter, other)
- Primary Caregiver smoking (smokes daily)
- Primary Caregiver risk of depression (based on CESD)
- Study Child gender
- Whether the Study Child was ever breastfed
- Health / longstanding condition of Study Child in Wave 1⁶

Most of these characteristics were measured at the 5-year interview, apart from those which would not change over time (such as Study Child's gender and Primary Caregiver country of birth). The weights were truncated to avoid giving undue influence on results to individual cases (or a small number of cases) and to avoid excessively large sampling variances. The distribution of the child and family characteristics in the completed 9-year sample when these weights are applied are within one-half of a percentage point of the population distribution for all of the characteristics examined.

The longitudinal population is made up of children and their families who participated in the study at 9 months of age and who continued to live in Ireland when they were 9 years old. Given the fixed sample design, children who were living in Ireland at 9 years of age but who were not resident in the country at 9 months will not be included in this population. Equally, it does not include children who were resident in Ireland at 9 months of age but who had emigrated out of the country by 9 years and who, accordingly, were no longer growing up in Ireland.

Totals." University of Cambridge, Department of Economics, Microsimulation Unit Research Note MU/RN/4, March 1994.

⁶ Although Study Child health /longstanding condition (Wave 1) does not differ from the population distribution, it was included in the set of variables used to weight the data in order to ensure that its distribution was not distorted when adjustments to the weights were made for related characteristics such as Study Child gender and parental smoking.

⁷ The weights were truncated to one-fifth of the mean at the lower end and 5 times the mean at the higher end.

In preparing the Wave 5 dataset, two sets of weights were calculated. The first set should be applied in analysis based on the 8,032 families for whom there is a valid observation at 9 months and 9 years of age. The second set of weights should be used in analysis based on the smaller set of 7,507 families who participated at all four rounds of face-to-face interviews.

Table 2.2 Weighting factors available for use with Infant Cohort at 9 years of age

Families participated at:	No. of families	Weights & Grossing Factors
9 months and 9 years only	8,032	WGT_9YRa
9 months, 3 years, 5 years and 9 years	7,507	WGT_9YRb

The first set (subscript 'a') should be used when one is carrying out analysis on the most complete 9-year sample of 8,032 families. The 'weighting factor' adjusts the internal structure of the sample in line with the population, summing to the actual number of cases, i.e. to 8,032 families. As noted above, population refers to the number of 9-year-olds who were resident in Ireland at 9 months of age and who continue to be resident in the country when they are 9 years old, accounting for those who no longer live in Ireland at 9 years of age or who have deceased since 9 months of age⁸.

The second set of statistical adjustments (those with subscript 'b') perform exactly the same functions as the first set. The only difference is that they are applied to the slightly smaller sample of 7,507 families who participated in all four face-to-face rounds of the study, i.e. at 9 months, 3 years, 5 years and 9 years of age. As above, the 'weighting' factor adjusts the sample in line with the population structure and sums to the actual number of families (i.e. to 7,507 cases).

Table 2.3 Combinations of data waves and statistical adjustment factors

	Combination of data	Unweighted	Weighting	Comments
	waves	number of cases	Factor	Comments
Α	9 months in cross-section	11,134	WGT_9MTH	All families who participated in Wave 1
В	Standalone 3 years (matched 9 mth and 3 years)	9,793	WGT_3YR	All families who participated in Wave 2 (all of whom, by definition, also completed in Wave 1)
С	Standalone 5 years (matched 9 mth and 5 years)	9,001	WGT_5YRa	Families who participated in Waves 1 and 3, 289 of whom did not participate in Wave 2 (at age 3)
D	Matched 9 months, 3 years and 5 years	8,712	WGT_5Yrb	Families who participated in all three Waves of interviewing to date.
E	Standalone 9 years (matched 9 mth and 5 years)	8,032	WGT_9YRa	Families who participated in Waves 1 and 5, 525 of whom did not participate in either Wave 2 or Wave 3
F	Matched 9 months, 3 years, 5 years, and 9 years	7,507	WGT_9Yrb	Families who participated in all four Waves of face-to-face interviewing to date.

_

⁸ Note that earlier waves also included a grossing factor which maintained the same structural breakdown as the weighting factor but grossed the N of cases up to the population N. Given the difficulties in estimated the correct population N of children who were resident in Ireland at 9 months and continue to be resident at 9 years of age, it was decided to not include a grossing factor in this wave of the data.

For example, if one wishes to track the trajectories of a particular characteristics across *all four waves* one must use the matched data file containing 7,507 cases as in Row F in Table 2.8 However, if one is interested only in investigating differences between 9 months and 9 years, regardless of the intervening 3-year or 5-year wave, one may use the larger sample of 8,032 cases (Row E in Table 2.8). In sampling terms, it is generally preferable to use as large a sample as possible, to minimise standard errors and corresponding sampling fractions. Similarly, if one is investigating trends only in the 3-year component of the sample the 9,793 cases contained in Row B in Table 2.8 above is the appropriate file to use (with corresponding weighting factors). Finally, if one's focus is solely on the infants in the base-year at 9 months of age in cross-section the 11,134 cases in the file at Row A should be used, as this provides the largest sample for analysis.

3 Instrument Development & Piloting

3.1 Introduction

This chapter gives a brief outline of the consultative process of instrument development for Wave 5 of the Infant Cohort and provides a summary of the groups of experts who gave such valuable input during this process. An overview is given of the Pilot Phase of the Wave 5 data sweep, which consisted of two components: a household pilot and a school-based pilot.

3.2 Instrument Development

As at previous waves of the study, intensive consultation took place with various groups of experts in the development of the instruments and procedures used at Wave 5 of the Infant Cohort. These included, the Scientific Advisory Group (SAG), the International Advisors, the Study Child Consultative Process and the Stakeholder Groups.

A number of consultative meetings were held with the SAG, organized according to the thematic lines of the study: health and physical development; socio-emotional development and behaviour; educational and cognitive development; and social context, methodology and design.

Two international advisors with extensive experience on a number of similar longitudinal studies, including the National Child Development Study (NCDS), the Longitudinal Study of Australian Children (LSAC) and the National Longitudinal Study of Children and Youth (NLSCY) provided the study team with very experienced input at all levels and in respect of all topics and procedures, including the substance of the questions and scales, ethical issues around recording details on sensitive topics, and procedural issues on the implementation and administration of the questionnaires.

A two-part consultative process was conducted with 9-year-olds in preparation for the pilot and subsequent main phase of the project. The first part of the consultative process, focus groups in four schools, related to the main issues included in the questionnaires. The second focused on the best way in which to complete the child Self-Complete Questionnaires: on paper, on a laptop or on a tablet.

3.3 Piloting the Instruments

This fifth wave of the Infant Cohort study was preceded by a pilot phase which included two components. In the first instance the pilot sample of Study Children and their main caregivers were interviewed in the home on a face-to-face basis by a survey interviewer. In the course of that interview, details were recorded on the school currently attended by the Study Child. The second component involved follow-up with relevant schools to complete school-based questionnaires with the Principal and Study Child's teacher.

Home-based fieldwork for this pilot phase of the project was carried out in the Autumn/Winter of 2016. In total, 180 children and their families were included. Full details of the structure and design, along with results and recommendations from the pilot phase, are available in the Cohort '08 at 9 Years Old Pilot Report (GUI Study Team, in press).

4 Survey Instruments

4.1 Introduction

This section provides a general overview of the instruments used in the Infant Cohort at age nine. In the household, questionnaires were completed by the Primary Caregiver and Secondary Caregiver (where relevant), and physical measurements were taken. Direct physical and cognitive assessments of the Study Child were also carried out in the home. A number of school-based instruments were also administered where relevant; these will be further outlined in this section.

4.2 Household-based instruments

Questionnaires with the Primary and Secondary Caregivers were completed in the home using a laptop. The Primary Caregiver and Secondary Caregiver Main Questionnaires were administered by the interviewer on the laptop (Computer Assisted Personal Interviewing; CAPI). A common Self-complete Questionnaire was completed by Primary and Secondary caregivers on a self-completion (Computer Assisted Self-completion Interview; CASI) basis. Table 4.1 provides an outline of the different Wave 5 household based instruments, divided into sections according to topic. For more detailed information on the all questionnaires and instruments used at Wave 5, see www.growingup.ie/questionnaires/.

In addition to the survey questionnaires which were administered to the Primary and Secondary Caregivers, interviewers recorded the height and weight of the Study Child and the weights of Primary and Secondary Caregivers. Height measurements for adults were taken if this information was not available from previous waves (new respondent or missing). A medically approved mechanical SECA weighing scales was used for all weight measurements and a Leicester height stick was used for all height measurements.

Children undertook two standardised cognitive tests which were administered directly by the interviewer in the home. These tests were the Vocabulary test from the Drumcondra Primary Reading Test – Revised (DPRT-R) and the 'Map Mission' test of selective attention.

Table 4.1 Household based instruments used at Wave 5

Respondent	Mode of completion	Summary of content
		Main Questionnaire
		Section A. Household composition
		Section B. Child's sleep and relationships
		Section C. Child's physical health and development
	САРІ	Section D. Child's diet and exercise
Primary Caregiver		Section E. Parental health
		Section F. Child's play and activities
		Section G. Screen and internet use
		Section H. Child's emotional health and well-being
		Section I. Parenting and family context
		Section J. Child's education

Respondent Mode of completio		Summary of content		
		Section K. Peer relationships and bullying		
		Section L. Socio-demographics		
		Section M. About you (the PCG)		
		Section N. Neighbourhood/ Community		
		Main Questionnaire		
		Section B. Child's sleep and relationships		
		Section D. Child's diet and exercise		
		Section E. Parental health		
Secondary Caregiver	CAPI	Section F. Child's play and activities		
		Section I. Parenting and family context		
		Section L. Socio-demographics		
		Section M. About you (the PCG)		
		Section N. Neighbourhood/ Community		
		Self-complete Questionnaire		
		Details on persons who have left family since Wave 1 (PCG only)		
		Relationship to Study Child		
		Marital status		
		Quality of couple relationship		
		Quality of couple relationship (DAS-4 Scale)		
		Parenting style (LSAC scale)		
		Co-parenting relationship (CRS scale)		
		Parental stress (PSS)		
		Parental efficacy		
Primary & Secondary	CASI	Currently pregnant? Only asked if female		
Caregiver	<i>C, L</i>	Alcohol consumption		
		Smoking		
		Drug use		
		Depression & anxiety		
		Contact with the Criminal Justice System		
		Sharing of family chores/ child-rearing tasks		
		Attitude to corporal punishment		
		Perceived discrimination		
		Perceived employment security		
		Information on non-resident parent (if relevant)		
		Relationship with own parents at 9 years old		
Primary & Secondary Caregiver	Measured by interviewer	Physical Measurements (height and weight)		
Study Child	Measured by interviewer	Physical Measurements (height and weight)		
		Cognitive Assessment		
Study Child	Administered by interviewer	Drumcondra Primary Reading Test – Revised (DPRT-R)		
	by interviewer	'Map Mission' test of selective attention		

4.3 School-based instruments

Three school-based questionnaires were completed by the Study Child's school on paper and returned to the GUI study team by post – these were the Principal, Teacher-on-Self and Teacher-on-Pupil Questionnaires⁹. The name and address of the school currently attended by the Study Child were collected during the household interview with the Primary Caregiver. Signed consent was secured from the Primary Caregiver to approach the Study Child's teacher, in order to ask him/her to complete a detailed questionnaire about the child's engagement and performance in school. The school-based component of the 9 year (wave 5) survey adopted a multi-mode methodology based, in the first instance, on a postal approach to the school; in the second, on intensive telephone follow-up and in the third, on personal visit to the school by a survey interviewer.

The content of the three school-based questionnaires is outlined in Table 4.2. It should be noted that content of the Teacher-on-Pupil Questionnaire relates to one specific study child, as opposed to the Teacher-on-Self Questionnaire which asks about pupils on a general, school or class level.

Table 4.2 School based instruments used at Wave 5

Respondent	Summary of content
	Personal information
	Basic school information
	Ethos of the school
	Staff and classroom provision
	Year in which school was built and also year most recently refurbished.
	Adequacy of school facilities and resources
	Home-School Community Liaison Co-ordinator
	Free school meal
	Parents' association or council?
	Computer resources in the school
	School-community relationships
Principal	Extracurricular activities
-	School ethos
	Pupil population composition
	School attendance levels
	School catchment
	Emotional/behavioural problems and school supports
	Admission and streaming criteria
	Engagement with parents
	Pupil engagement with school
	Disciplinary policy in the school
	Bullying in the school
	Day-to-day problems and general environment in the school compared with other primary schools in the country.

_

⁹ In some cases, teachers will have completed more than one Teacher-on-Pupil questionnaire relating to different Study Children. This questionnaire is sometimes also referred to as Teacher-on-Child but it is the same questionnaire.

Respondent	Summary of content
	Teacher-on-Self
	Background characteristics of the teacher
	Basic characteristics of the class
	Subjects undertaken
	Teaching equipment
	Teaching methods
Teacher	Homework
reacties	Teacher's assessment of pupils
	Teacher control and input to decision-making in the classroom
	Teacher's perception of pupils
	Parental attendance at parent-teacher or school meetings
	Challenges for the teacher
	Perception of general school environment
	Job stress and job satisfaction
	Teacher-on-Pupil (Child)
	Characteristics of the Study Child
	Attending school in an appropriate state
	Completing homework
	Within-class grouping on the basis of reading/literacy and maths
Teacher	Study child's ability and attainment
reactiet	Strengths and Difficulties Questionnaire (SDQ)
	Parent – teacher meetings
	Study Child's disposition and attitudes to school
	Parent's engagement with the school and teacher
	The Pianta Student-Teacher Relationship Scale (STRS)
	Conditions that limit activities

5 Fieldwork and Implementation

5.1 Introduction

This chapter briefly outlines the fieldwork procedures at Wave 5. This includes the training and vetting of fieldworkers, protocols for making initial contact with a household, tracing methods, and incident reporting procedures.

5.2 Interviewer Training

Fieldwork was carried out by the ESRI's national panel of interviewers. All interviewers who worked on the home-based fieldwork received in-depth training prior to beginning work on the project. Training included the following modules:

- 1. Background and objectives of the study
- 2. Detailed review of the content of all questionnaires
- 3. Familiarisation with, and practice on, the Computer Assisted Personal Interview system (CAPI)
- 4. Fieldwork procedures
- 5. Adult and child measurements (height and weight)
- 6. Instruction and practice in the administration of the direct child assessments
- 7. Child protection guidelines and incident reporting
- Ethics
- 9. Summary of other documentation used in the administration of the survey

Additional training was provided to those interviewers who worked on the school phase of fieldwork.

5.3 Vetting

Growing Up in Ireland was carried out under the Statistics Act (1993). This is the same legislation as is used, for example, to carry out the Census of Population. Interviewers were appointed as 'Officers of Statistics' for the purposes of this project. This included a confidentiality clause on non-disclosure of information which was recorded in respect of a family or child to any unauthorised person, for any purpose.

In addition to being appointed Officers of Statistics, all interviewers and all other staff involved in the project were security vetted by An Garda Síochána (the Irish police force).

5.4 Contacting a Household

As in other waves of the study, the initial contact with the family at this wave was made via a letter from the Study Team. The interviewer subsequently made a personal visit to each household to arrange an interview. At that first visit, interviewers asked to speak to the person listed at Wave 3 as the Study Child's Primary Caregiver (Primary Caregiver details were not updated at the Wave 4 postal wave). If the person was still resident in the household, then s/he was asked to confirm that s/he was still the Primary

Caregiver. Having reminded the parent/guardian of the letter and information sheet¹⁰ which had already been posted to the family, and answering any queries the parent had, the interviewer asked the Primary Caregiver to sign two copies of the main consent form. One copy was returned by the interviewer to the field office, and the other was retained by the Primary Caregiver for his/her own records. Only after securing signed consent did the interviewer undertake any work with the family (interviews, tests or measurements).

An additional consent form was signed in relation to asking the Study Child's teacher to complete a questionnaire about him/her.

5.4.1 Follow-Up/Tracing Information

As discussed in Chapter Two, there are a number of variables associated with inter-wave attrition, some relating to the characteristics of the interview and others relating to characteristics of the individual respondent. The problem of attrition may be somewhat mitigated by implementing rigorous tracking procedures aimed at tracing respondents who, for example, change address between data sweeps. Lynn (2009) distinguishes between forward or proactive tracing methods, i.e. procedures put in place prior to the current phase of fieldwork; and retrospective tracing methods, i.e. procedures which are put in place after fieldwork, when it has been identified that the participant has changed address since the previous wave. Both proactive and retrospective tracing methods were implemented in the GUI study.

5.4.2 Proactive tracing procedures

A number of proactive procedures were adopted during data collection. These included recording contact information in respect of two of the respondent's close associates or family members (outside their own household) whom the Study Team could call if it was found in the subsequent wave that the respondent had moved since the previous interview. In addition, respondents were given a "change-of-address" postcard and asked to fill in their new contact details and return them to the Study Team in the event of them changing address between interview rounds.

5.4.3 Retrospective tracing procedures

Retrospective procedures were also adopted. When field interviewers identified that a family was no longer resident at the address known to the Study Team they attempted to obtain a new address from the current occupant or neighbours at the respondent's former address. In doing this the interviewer told the current occupier or neighbour that s/he wished to track the family who had previously participated in a survey, but did not divulge that it was the *Growing Up in Ireland* study or the nature or content of the project in question. In cases where a new address was successfully obtained in this manner, interviewers fed the new address back to Head Office where the family were reallocated to field staff (if in a different area from previous address).

Where a new address for a family who had moved could not be obtained by the interviewer, field support staff in Head Office accessed the alternative contact details provided for tracing purposes by the family in earlier waves of the study. These alternatives were contacted with a view to securing a

¹⁰ A copy of the information sheet provided to participants can be found at <u>www.growingup.ie</u>.

current address for the respondent Study Child.

A final (and extremely important) source of potential tracking in Wave 5 was the Child Benefit Register, which is maintained by the Department of Social Protection. In the course of the earlier interviews, respondents were asked to sign a consent form giving permission to track them using their Personal Public Service Number (PPSN) through the Child Benefit Register for tracing purposes associated with the study. This was implemented through the Department of Social Protection. To minimise the burden on the Department this approach was used as a final stage in tracing, only after field and other tracking procedures had been exhausted. Given the high quality of the contact information contained on the Child Benefit Register the success rate in securing alternative contact addresses was very high.

5.5 Incidents

A detailed *Growing Up in Ireland* Child Welfare and Protection protocol was developed by the Study Team. One aspect of this involved an incident report system. All incidents were immediately reported by interviewers to their Field Support Contact at Head Office and a detailed Incident Report Form was completed. Given that interviews often took place outside office hours, interviewers were provided with an emergency telephone number which could be used to contact the Study Team on a 24-hour, 7 day basis. Interviewers were instructed that in extreme circumstances, where a child or other vulnerable person was thought to be in immediate danger, they should use their own discretion and contact the Gardaí if necessary, without recourse to the Study Team.

6 Structure and Content of the Data File

6.1 Introduction

This section outlines the structure of the Researcher Microdata File (RMF) and Anonymised Microdata File (AMF) and provides a brief explanation of how the two data files differ in content. An overview is given of variable naming and ordering conventions and the reweighting process. Details are provided of the derived variables and those pertaining to the scaled measures used in the study, followed by the measurement variables, i.e. physical measurements and cognitive tests. Finally, the coding and editing process is outlined.

The Study Team would advise that the data are used in conjunction with the Questionnaire Documentation. This is probably the easiest way to get a broad overview of the topics included in the data file. Researchers should however note that there may be differences in value labels between the questionnaires and the data file, for the purposes of preparation and anonymisation. This is especially true for the AMF.

6.2 Anonymised (AMF) and Researcher (RMF) Microdata Files

Two data files are available for researchers: the Researcher Microdata File (RMF) and Anonymised Microdata File (AMF). The AMF is a publicly available anonymised dataset, while the RMF is a more detailed dataset, access to which is subject to appointment as an Officer of Statistics by the Central Statistics Office. Accordingly, some potentially disclosive variables which appear on the RMF have been removed from the AMF. Other variables have had their values banded into larger groups so that frequencies with low cell counts are not visible. In some instances this was achieved by either bottom or top coding (or both) of outlying cases. In others, continuous scores were grouped into categories. Information particularly likely to be sensitive in nature (i.e. the majority of the variables in the self-complete questionnaires) has been removed from the AMF. The user should therefore note that not every question from the questionnaires is included in the data file, particularly in the case of the AMF. A list of variables included in each data file is available via the accompanying summary data dictionary.

6.3 Structure of the data file

Both the Researcher Microdata File (RMF) and Anonymised Microdata File (AMF) are presented as a flat rectangular data file based on a simple concatenation of all questionnaires administered to respondents. The case-base is the Study Child. This means that the user does not have to be concerned about matching Primary and Secondary Caregiver questionnaires within household.

6.4 Variable naming

All variables for Wave 5 of the Infant Cohort of prefixed with a 'b' for 'birth cohort'; there are slight differences to the combination of preceding letters for the home-based versus school-based variables.

6.4.1 Naming of Home-based Primary and Secondary Caregiver Variables

For the home-based Primary and Secondary Caregiver questions, the prefix 'b' is followed by '5' to

indicate the fifth wave of data collection. This is followed by two letters which indicate the respondent ('pc' for Primary Caregiver, 'sc' for Secondary Caregiver) and the question number. In the case of the self-complete questionnaire, the question number is preceded by 'S'.

Examples:

- Question 'C1' from the Primary Caregiver Main Questionnaire at Wave 5 will have the variable name 'b5pcc1'
- Question B4(g) from the Secondary Caregiver Main Questionnaire will have the variable name 'b5scb4g'
- Question 12 from the Primary Caregiver Self-Complete Questionnaire will be named 'b5pcs12'

6.4.2 Naming of Home-based Child Variables

For the home-based Child questions, the prefix 'b' is followed either 'cq' for the Child Main questionnaire; 'cs' for the Child Self-Complete questionnaire or 'cph' for the Piers Harris questionnaire. This is followed by '5' to indicate the fifth wave of data collection, and the question number.

Examples:

- Question 13 from the Child Main Questionnaire at Wave 5 will have the variable name 'bcq5q13'
- Question 18 from the Child Self-Complete Questionnaire will have the variable name 'bcs5q18'
- Question 35 from the Child Piers Harris Questionnaire will have the variable name 'bcph5q35'

6.4.3 Naming of School-based Variables

For the school-based questions, the initial 'b' prefix is followed by '5' (to indicate the fifth wave of data collection), an underscore and subsequently letter(s) to indicate the questionnaire: 'P' for Principal, 'TC' for Teacher-on-Child (Pupil) and 'TS' for Teacher-on-Self.

Examples:

- Question 35 on the Principal Questionnaire will be named 'b5_p35'
- Question 4 on the Teacher-on-Pupil questionnaire will be named 'b5_tc4'
- Question 25 on the Teacher-on-Self questionnaire will be named 'b5_ts25'

6.4.4 Naming of other variables

Exceptions to the aforementioned variable naming conventions are variables from the household grid, derived variables and variables from the scaled measures, as well as direct measurements, i.e. physical measurements and cognitive tests.

6.5 Variable order

The first variables in the data file include the household identification code, details of family's

participation in subsequent waves and the weighting factors, all of which are detailed later in this chapter. Following these, blocks of variables appear in the dataset in the order listed in Table 6.1 (variable prefixes for blocks of variables are also shown). Note that derived variables appear at the end of the relevant block of variables, i.e. variables derived from the Primary and Secondary Caregiver (PCG and SCG) questionnaires appear after the other home-interview variables. Variables derived from the Teacher-on-Pupil questionnaire appear after the other Teacher-on-Pupil Variables.

Table 6.1 Ordering of variables in the data file

Order	Questionnaire/Section	Variable prefix
1 st	Household Grid	p1xxW5, p2xxW5
2 nd	Primary Caregiver Main Questionnaire	b5pc
3 rd	Primary Caregiver Self-complete Questionnaire	b5pcs
4 th	Secondary Caregiver Main Questionnaire	b5sc
5 th	Secondary Caregiver Self-complete Questionnaire	b5scs
6 th	Child Main Questionnaire	bcq5
7 th	Child Self-Complete Questionnaire	bcs5
8 th	Child Piers Harris Questionnaire	bcph5
9 th	Physical Measurements	bpc5, bsc5 or b5kid [study child]
10 th	Scale Scores	b5_
11 th	Derived Variables	b5_
12 th	Map/Attention test	b5_
13 th	Drumcondra test	b5_
14 th	Principal Questionnaire	b5_p
15 th	Teacher-on-Child Questionnaire (& relevant derived variables)	b5_tc
16 th	Teacher-on-Self Questionnaire	b5_ts

6.6 Identification Codes

Each household has a unique identification code, which is the same at all waves to enable matching of the data files where necessary. The sequence of identification codes runs from 300 to 1,113,400 and is indicated by the variable 'id'.

6.7 The Household Grid

The household grid contains the information on the members of the household, i.e. who lives in the household, his/her person number on the grid, gender, relationship to both the primary caregiver and the Study Child, age and principal economic status. For ease of reading, the household grid variables are prefixed with the person number. For example, the variable indicating the sex of the person on line 1 of the grid is 'P1sexW5' where 'W5' indicates Wave 5 data. Details were recorded such that the Primary Caregiver (usually the mother) was always on line 1, the Study Child was always on line 2, and the Secondary Caregiver (if relevant) was on line 3. The Study Child's twin or triplet etc was on lines 4, 5 as appropriate unless there was no Secondary Caregiver, in which case they were on lines 3, 4.

To save time in administering the interview at Wave 5 some information on household composition which was captured at Wave 3 was fed forward to the household grid at Wave 5. The Primary

Caregiver was asked to review this information and to correct any inaccuracies, either due to errors or changes which had taken place since the previous interview. New people could be added to the grid and others removed. The information represented by the variables labeled 'P1xxW5' etc. included any corrections made at Wave 5. To ensure confidentiality, only the respondent who identified as the Primary Caregiver at Wave 3 could review the forward-fed information¹¹. If the respondent identified as the Primary Caregiver at Wave 3 was no longer resident in the household at Wave 5, the person identifying as the Primary Caregiver at Wave 5 was asked to complete a new household grid, without any forward-fed information. On the *RMF* only, the original line number for the person at Wave 1 can be found in the variables named 'origlineP1' etc. Note that individuals with an original line number from 21 onwards are new additions to the grid at Wave 2; individuals with an original line number from 51 onwards are new additions at Wave 3, and; individuals with an original line number from 51 onwards are new additions at Wave 5. The variables named 'W5xstillp3' etc. indicate whether or not the person on that line number (e.g. line 3) at Wave 1 is still resident in the household.

Whether or not the Primary Caregiver and Secondary Caregiver roles at Wave 5 are being filled by the same individual as in Wave 3 is indicated by the derived variables 'PCGstatw5' and 'SCGstatw5'.¹²

As noted, where there is a Secondary Caregiver, s/he will be person 3 on the household grid. However, not all persons on line 3 of the household grid are Secondary Caregivers. For example, in a one-parent family the third person (if present) will be another household member (other than the Primary Caregiver or Study Child). A variable has been included in the database to indicate whether the Primary Caregiver has a partner (by definition the Secondary Caregiver) resident in the household (b5_partner).

Details obtained in the household grid, such as dates of birth, gender and relationships are very important in terms of calculating derived variables. Consequently, some editing of the information took place where it was clear from associated details that this was appropriate. There are, however, a few minor outstanding anomalies between the information given on the interviewer administered household grid and that given in the Primary Caregiver Self-complete questionnaire (self-completed on CASI). The reader should note that (for anonymisation purposes) exact dates of birth have been removed from the archived file and replaced with age in years.

6.8 The Main Respondent - Primary Caregiver

The Primary Caregiver was self-identified within the home as the person who provided most care to the Study Child and who knew most about him/her. In most cases, this was the child's mother. As noted above, in some cases the Primary and Secondary Caregiver from Wave 3 had swapped roles between waves (flagged by the variables 'PCGstatw5' and 'SCGstatw5'). Note that more detailed information on the inter-wave swapping of roles is provided in the *RMF*.

¹¹ This was done to meet the guarantees of confidentiality of information which were given to respondents in previous waves. At earlier interviews, respondents were told that no-one would have sight of the information which they provided in the course of their interview, including the information contained in the household grid.

¹² Note this information will be unavailable for families who did not complete at Wave 2

6.9 Twins

A data record exists for each child included in the sample. All non-singleton children (those with twins, triplets, etc.) are coded as 'b5_nonsingleton' in the file. In situations where there was a non-singleton in a family, a core questionnaire was administered to the Primary and Secondary Caregivers (where relevant) in the normal way to record the characteristics of the informant. These core questionnaires included details on, for example, the informant's health status and lifestyle, and socio-demographic characteristics. In addition, the Primary and Secondary Caregivers were asked to complete a questionnaire containing the relevant questions specific to each of the non-singleton Study Children – for example, in respect of the Primary and Secondary Caregiver's relationship with the child. Following the interview, a data record was constructed for each sampled non-singleton child to include the common questions from the Primary and Secondary Caregiver as well as the child-specific questions from the individual questionnaires.

6.10 Weighting variables

As discussed in Section 2.8 above, in line with best practice in sample surveys, the data have been reweighted or statistically adjusted to ensure that the sample is representative of the population¹³ from which it has been drawn. By doing this one ensures that the structure of the completed sample is in line with the structure of the population along key socio-demographic and other dimensions.

The data file contains two weighting factors. The first is **WGT_9YRa** which is based on the families who <u>participated at Wave 1 and Wave 5</u>, <u>but not necessarily Wave 2 or 3</u>. The weighting factor incorporates the structural adjustment of the completed sample to the population, whilst maintaining the total completed sample size of 8,032 cases.

The second weighting factor is **WGT_9YRb** and this relates to families who participated at all four main waves (Waves 1, 2, 3 and 5) in the study to date - the reduced sample of 7,507 families. The weighting factor incorporates the structural adjustment of the completed sample to the population, whilst maintaining the completed sample size of 7,507 families who participated in all four waves. The weighting factors should be used in significance testing and data modeling.

The variables **xxwave1**, **xxwave2**, **xxwave3**, **xxwave4** and **xxwave5** indicate if the case has data for each of the waves. A value of one indicates participation at the relevant wave. In the 9 year data file **xxwave1** and **xxwave5** are equal to 1 for all cases, as all cases in this file have completed both Wave 1 and Wave 5. Frequencies of these indicator variables are outlined below.

-

¹³ As noted in Chapter Two, given the fixed panel design of *Growing Up in Ireland* in the current context this is the population of 9-year-olds who were resident in Ireland at 9 months of age and who continued to live in the country at 9 years, adjusting for those who (with their families) had emigrated or deceased between 9 months and 9 years of age.

Table 6.2 Frequency distribution of 'xxwave' variables

Variable Name	Value	N
xxwave1	1	8032
	0	0
xxwave2	1	7769
	0	263
xxwave3	1	7700
	0	332
xxwave4	1	4985
	0	3047
xxwave5	1	8032
	0	0

6.11 Derived Variables

In addition to some of the derived variables mentioned above (e.g. 'b5_partner', 'PCGstatw5' and 'SCGstatw5'), a number of variables were derived to provide additional information on the circumstances of the household. These variables pertain to family composition, household income and household social class and are outlined below.

6.11.1 Household type (b5_hhtype4, b5_hhtype4_v2)

These fourfold variables give an indication of family composition. They are based on whether or not the Study Child is living in a one or two parent family as well as the number of children living in the household. 'B5_hhtype4' gives us a classification as follows:

- One-parent, one child under 18
- One-parent, two+ children under 18
- Two-parents, one child under 18
- Two-parents, two+ children under 18

'B5_hhtype4_v2' gives us a classification as follows:

- One-parent, max two children under 18
- One-parent, three+ children under 18
- Two-parents, max two children child under 18
- Two-parents, two+ children under 18

A child is defined solely in terms of age (under 18 years) and not in terms of relationship to the Study Child or others in the household.

6.11.2 Equivalised Household Income (b5_equivinc, b5_EIncDec, b5_EIncQuin)

In order to make meaningful comparisons between households on their income, household size and structure must be taken into account. This is done by creating an 'equivalised' income. In GUI, an equivalence scale was used to assign a "weight" to each household member. The equivalence scale

used assigned a weight of 1 to the first adult in the household,

0.66 to each subsequent adult (aged 14+ years living in the household) and 0.33 to each child (aged less than 14 years). The sum of these weights in each household gives the household's equivalent size – the size of the household in adult equivalents.

Disposable household income is recorded as total gross household income less statutory deductions of income tax and social insurance contributions. Household equivalised income is calculated as disposable household income divided by equivalised household size. This gives a measure of household disposable income which has been "equivalised" to account for the differences in size and composition of households in terms of the number of adults and/or children they contain.

The equivalised household income (**b5_equivinc**) is available in the *RMF* only. In the *AMF*, equivalised household income is given in deciles (**b3_EIncDec**) and quintiles (**b3_EIncQuin**).

6.11.3 Household Class (b5_hsdclass)

In the course of the survey, both caregivers (where relevant), were asked to provide details on their occupations from current or previous employment outside the home¹⁴. On this basis, a social class classification was generated for both Primary and Secondary Caregiver. The classification used was that adopted by the Irish Central Statistics Office (CSO) with seven categories as follows:

- 1. Professional workers
- 2. Managerial and technical
- 3. Non-manual
- 4. Skilled manual
- 5. Semi-skilled
- 6. Unskilled
- 7. All others gainfully occupied and unknown

The household's Social Class is then taken as the highest Social Class category of both partners in the household (as relevant). This standard procedure is referred to as the 'dominance criterion'. Households where both caregivers are currently economically inactive and have not held any previous employment in the past are classified as 'validly no social class', as they have no occupation code upon which to base their social class. Note that as past occupation is only taken into account if currently unemployed or retired and not for those now on home duties, the 'validly no social class' category may include households that had a valid social class classification in previous waves. Researchers conducting longitudinal analyses may wish to carry forward social class from previous waves for households in this category.

6.11.4 Household location (b5_region)

This variable is based on information provided by the Primary Caregiver in the course of the interview. There is an eight-fold classification on the RMF (**b5region8**: South-East, Dublin, etc) as well as a three-fold classification (**b5region3**: Dublin, BMW, Rest); and on the AMF it is reduced to an urban/rural

¹⁴ Current occupation if economically active; previous occupation if retired or unemployed.

dichotomy (b5_region).

6.12 Scaled Measures Used in the Study

A number of scaled measures were used in the *Growing Up in Ireland* study and scored according to protocols provided by the authors. These are briefly described below. An indication of the reliabilities of these scaled measures, as illustrated by Cronbach's alpha, are detailed in the appendix to this report.

6.12.1 Strengths & Difficulties Questionnaire (SDQ; Goodman, 1997)

The SDQ is a brief (25 item) behavioural screening questionnaire designed to assess emotional health and problem behaviours in children. The SDQ appears on the Primary Caregiver questionnaire as question H2 and on the Teacher-on-Pupil (Child) questionnaire as question 13. The SDQ comprises five subscales, a total difficulties score and, in the case of the parent SDQ, an impact score. The subscales and their corresponding variable names are listed in Table 6.3.

Table 6.3 Subscales of the Strengths and Difficulties Questionnaire and their relevant variable names

Subscale	Primary Caregiver Variable Name	Teacher-on-Pupil Variable Name	
Emotional	b5_SDQemotional	b5_TCsdqemotional	
Conduct	b5_SDQconduct	b5_TCsdqconduct	
Hyperactivity	b5_SDQhyper	b5_TCsdqhyper	
Peer problems	b5_SDQpeerprobs	b5_TCsdqpeerprobs	
Prosocial	b5_SDQprosocial	b5_TCsdqprososcial	
Total difficulties	b5_SDQtotaldiffs	b5_TCsdqtotaldiffs	

6.12.2 The Pianta Scales - Child Parent Relationship Scale (CPRS) and Student-Teacher Relationship Scale (STRS) (Pianta, 1992)

Each of these 15-item scales assess both the negative and positive aspects of the relationship between either parent and child or teacher and child. The relevant scale appears as question B4 on the Primary Caregiver Main Questionnaire, as question B4 on the Secondary Caregiver Main Questionnaire, and as question 18 on the Teacher-on-Pupil questionnaire. Both scales are very similar in content, with altered wording in the STRS (i.e. "this child" as opposed to "my child"). The measure produces a *Positive Aspects* subscale (b5pc_positive, b5sc_positive, b5_TCpositive) and a *Conflicts* subscale (b5pc_conflict, b5sc_conflict, b5_TCpositive).

6.12.3 Parenting Style Measure (from the Longitudinal Study of Australian Children [LSAC])

Questions 18 and 19 on the Primary Caregiver Sensitive Questionnaire (and question 18 and 19 on the Secondary Caregiver Sensitive Questionnaire) were taken from self-report measures of parenting style which were used in LSAC. These yield scores for three different parenting dimensions: *Warmth* (b5pc_warmth, b5sc_warmth), *Hostility* (b5pc_hostility, b5sc_hostility) and *Consistency* (b5pc_consistency, b5sc_consistency).

6.12.4 Parental Stress Scale (Berry and Jones, 1995)

One subscale of the Parental Stress Scale (Berry & Jones, 1995), which was designed to assess both positive and negative aspects of parenthood, appears on the Self-complete Questionnaire for both Primary and Secondary Caregivers as Question 21. The subscale is the six-item Parental Stressors sub-scale (**b5pc_stress**, **b5sc_stress**).

6.12.5 Centre for Epidemiological Studies Depression Scale (CES-D) (Melchior, Huba, Brown & Reback, 1993)

These eight questions provide a short self-report screening instrument for depression in the general population. Both Primary and Secondary Caregivers answered the CES-D as part of the self-complete questionnaires (question 39). For both respondents, a total score was obtained which is a sum of the raw scores (**b5pc_CEStotal**; **b5sc_CEStotal**). Also included in the file are two variables (**b5pc_CESD**; **b5sc_CESD**) which categorise respondents into 'depressed' or 'not depressed'.

6.12.6 Dyadic Adjustment Scale (DAS-4) (Sabourin et al., 2005)

The quality of the couple relationship was indexed using the short 4-item form of the Dyadic Adjustment Scale (DAS-4), which provides an assessment of dyadic satisfaction based on participants' self-report, and is used as a means of categorising marriages as either distressed or adjusted. Both Primary and Secondary Caregivers (where relevant) completed the DAS-4 in their sensitive questionnaire (questions 16 and 17, **b5pc_DAS**; **b5sc_DAS**).

6.12.7 FAST Alcohol Screening Test (Hodgson et al., 2002)

The FAST alcohol screening test was developed in the UK as a short screening tool for alcohol misuse. The scale comprises four items, however the test authors assert that 50% of people may be classified as 'hazardous' or 'not hazardous' drinkers using the answer to the first item "how often do you have EIGHT or more drinks on one occasion?" (six drinks for women). The items appear as S26-30 on the self-complete questionnaire for both the Primary and Secondary Caregiver. When these items are scored as 0 - 4, a person is classified as a 'hazardous' drinker if their total score is 3 or more. As anyone who answers S26a/b as having six or eight drinks on one occasion as weekly or more often is automatically classified as a hazardous drinker, not everyone will have a continuous score from 0 to 4. The classification is enclosed as **b5pc_fastclass** and **b5sc_fastclass**.

6.12.8 The Everyday Discrimination Scale short version (EDS) (Sternthal, Slopen, & Williams, 2011)

The Everyday Discrimination Scale (EDS) was used to measure discrimination. The original EDS was developed by Williams and colleagues (Williams, Yu, Jackson & Anderson, 1997) and a shortened version was developed by Sternhal and colleagues for the Chicago Community Adult Health Study. The scale consists of five items for example 'people act as if they are afraid of you' and has 6 response categories ranging from 1 (almost every day) to 6 (never). This scale is asked of both the primary and secondary caregiver in the supplementary questionnaire (question 44) and is **b5pceds** and **b5sceds** on the data file.

6.12.9 Co-parenting scale (Feinberg, 2003)

One subscale of the Co-parenting scale, the exposure to conflict subscale, was asked of both the primary and secondary caregiver in the sensitive questionnaire (question 20). The scale consists of five items for example 'Yell at each other within earshot of the child?' and has 7 response categories ranging from 0 (never) to 6 (Very often- several times a day). This scale was developed by Feinberg (2003) who asserted that co-parenting is central to family dynamics as well as overall quality of parenting. The variables on the data file are 'pcg_crsexposure' and 'scg_crsexposure;

6.12.10 Piers-Harris Scale (Piers, 1969)

The Piers-Harris scale is designed to measure a child's self-concept, with a higher score indicating a more positive self-concept. The version of the scale used in previous GUI waves was the 60 item self-concept scale. For this wave it was shortened to 31 items, by removing redundant items and retaining those with the best psychometric properties. Each of the 31 items has answer categories of yes or no (due to copy right issues it is not possible to share an example). This question appears on the child sensitive questionnaire (questions 25 – 55). For the total scale score and each of the six subscales there is a variable for the total score and a categorisation into 'Very low', 'Low', 'Low average', 'Average', 'High average', 'High' (see table below for variable names).

Note that for the shortened scale used in Wave 5, the mean values tended to skew slightly towards the upper third of each scale. This resulted in a clustering in the upper end of the range of categorised variables. Caution should be exercised in comparing the Piers Harris range scores in wave 5 to earlier waves.

Table 6.4 Subscales of the Piers-Harris and their relevant variable names

	Total Score	Categorisation
Total self-concept	W5PH_Totalscore	w5RangeTOT
Behavioural Adjustment	W5PH_Behaviour	w5RangeBEH
Intellectual and School Status	W5PH_Intellectual	w5RangelNT
Physical Appearance and Attributes	W5PH_Physical	w5RangePHY
Freedom from Anxiety	W5PH_Free_Anxiety	w5RangeFRE
Popularity	W5PH_Popularity	w5RangePOP
Happiness and Satisfaction	W5PH_Happiness	w5RangeHAP

6.13 Physical Measurements

6.13.1 Height & Weight

The heights of the Primary and Secondary Caregivers were fed-forward from the previous waves. Only in cases where this information was not available, or flagged for rechecking, were the heights of the Primary and Secondary Caregivers recorded at Wave 5. Weights for all Primary and Secondary Caregivers, and heights and weights of all study children were recorded (unless they were unable or unwilling to be measured).

All weights were recorded in kilograms using medically approved weighing scales: a flat mechanical

scale for adults (SECA 761). Height for both adults and children was recorded in centimetres using a standard measuring stick (Leicester portable height measure). All measurements were recorded on the laptop during the course of the CAPI interview.

The heights and weights recorded by the interviewer were edited to remove clearly implausible values. The Wave 5 measurements (which include the forward-fed height values where available) can be found in the following variables:

- Primary Caregiver Height (**bpc5cms**)
- Secondary Caregiver Height (bsc5cms)
- Study Child Height (b5kidcms)
- Primary Caregiver Weight (bpc5kgs)
- Secondary Caregiver Weight (bsc5kgs)
- Study Child Weight (b5kidkgs)

6.13.2 Body Mass Index (BMI)

BMI scores were derived from the height and weight measurements taken by the interviewer for the Primary Caregiver (**bpc5bmi**), Secondary Caregiver (**bsc5bmi**) and Study Child (**b5kidbmi**)¹⁵. Categorised variables are also provided, which categorise Primary and Secondary Caregivers into underweight, healthy, overweight, obese (**bpc5bmi_cat**, **bsc5bmi_cat**) and Study Child into non-overweight, overweight, obese (**b5kidbmi_cat**).

6.14 Cognitive Assessments

During the course of the home visit, interviewers administered two cognitive tests: an adaptation of the Drumcondra reading test which had previously been used with Child Cohort '98 at age 9 years, and a new test of selective attention (the 'map mission' test).

Variables in the data file for the Drumcondra reading test are:

- **b5_readclass -** Drumcondra Reading test class level sat Wave 5
- **b5_readatt** Drumcondra Reading test number of questions attempted Wave 5
- b5_readcorr Drumcondra Reading test number of correct answers Wave 5
- **b5_readpct -** Drumcondra Reading test percentage correct Wave 5
- b5_readingls Drumcondra Reading test Logit score Wave 5
- **b5_readinglsse** Drumcondra Reading test Logit score standard error Wave 5

The total score on the selective attention test is contained in the variable **b5** attentiontest.

6.15 Coding & Editing

The CAPI questionnaires administered in *Growing Up in Ireland* consisted mainly of closed questions¹⁶. The program included extensive range and cross-variable consistency checks (both hard

 $^{^{15}}$ On the AMF, BMI scores are derived from the original height and weight measurements before top and bottom coding.

¹⁶ Almost all CASI questions were closed.

and soft)¹⁷. This meant that much of the coding and data checking was effectively dealt with as the interview took place. However, in some cases open questions were needed to capture verbatim responses that would have been difficult to pre-code. Where relevant, these were coded into separate categorical variables after the interview was completed. Other questions did have a pre-defined coding frame but also had an 'other-specify' option for those responses which did not fit into any of the pre-coded categories - again answers were recorded on a verbatim basis by the interviewer. In this instance responses to these questions had to be recoded with additional categories. The newly coded responses for additional codes or variables appear in the *RMF* dataset. All verbatim text from the original responses has been removed as a safeguard to protecting respondent's identity. In terms of editing the data, regular checks were carried out on the data as they were returned from the field and inconsistencies dealt with.

The possibility of longitudinal inconsistencies arises with the collection of multiple waves of data, as well as cross-sectional inconsistencies within wave. For some key variables, such as social class, these were checked and edited to provide more consistency where appropriate. However, there remain some inconsistent cases where it was not possible to make a judgment on an appropriate edit.

6.16 Forward-feed from Wave 3

As discussed in Section 6.7 above, some variables were fed forward from previous waves to reduce interview time at Wave 5. Adult height was also forward-fed as noted earlier. A summary of all other variables that were fed forward at Wave 5 is provided in the Table 6.5 below.

Table 6.5 Details on variables forward-fed from previous waves (excl. household grid and adult height)

Variable name	Variable description	Rules	
b5pcm7, b5pcm8, b5pcm9, b5pcm10 b5scm7, b5scm8, b5scm9, b5scm10 [RMF only]	Literacy and numeracy	Asked if literacy or numeracy problems indicated at previous wave, missing or new respondent	
b5pcm13, b5scm13	Ireland as country of birth	Asked if missing from previous wave or new respondent	
b5pcm15, b5scm15 [RMF only]	Length of time living in Ireland	Asked if missing from previous wave or new respondent	
b5pcm1_ed, b5scm1_ed	Highest level of education	Asked to confirm if highest level of education from last wave was correct, or if it had changed. Asked if missing from previous wave or new respondent	

¹⁷ 'Hard' edit consistency checks in a CAPI program refer to cross-variable consistency checks which must be resolved by the interviewer in the field at the time of questionnaire administration. Until the inconsistency is resolved by the interviewer it will not be possible to continue administering the questionnaire. In contrast, a 'soft' edit consistency check is one which signals an apparent inconsistency, or extreme value from a respondent's answer to a question or set of questions. The extreme value may or may not be correct. If the interviewer administering the survey feels that it is a valid value, albeit extreme, s/he can suppress the soft edit check and continue with administering the survey.

6.17 Missing data

There were generally low levels of missing data throughout the interviews, however, missing data could arise in two ways. Firstly, respondents may have chosen to not answer an entire element of the interview. For example, the PCG may have completed the main CAPI interview but chosen to not complete the self-complete CASI interview or the SCG may have chosen to not complete the SCG interview at all. This type of missingness for each element of the interview is flagged by the following variables: 'b5_pcgmain', 'b5_scgmain', 'b5_pcgsens', 'b5_scgsens', 'b5_attention', 'b5_drum', 'b5_prin', 'b5_tos', and 'b5_toc'. Every variable within an uncompleted questionnaire will show as a system-missing value in the data file. Note that all cases in the data file have a completed PCG main interview.

The other type of missing data is individual variables within an otherwise complete interview. Respondents were given the option of choosing to skip over any individual question that they did not want to complete (in both CAPI and CASI), or they may have part-completed an interview but dropped out before the end of the interview, or they might simply have not known the answer to the question. This type of missing data shows as user-defined missing value in the data file and is labelled as 'not answered'. It is not possible from the data file to distinguish between a missing value that is a 'refusal' to answer rather than a 'don't know'.

7 Ethical Considerations

In undertaking research with families and children ethical considerations assumed primary importance. Procedures relating to child protection were informed by the Children First: National Guidance for the Protection and Welfare of Children (Department of Children and Youth Affairs, 2011) as well as the relevant Acts in Irish legislation. A number of acts are of particular relevance for this Study; they are the Data Protection Acts 1988, 2003, and 2018 (as well as the General Data Protection Regulation (GDPR) which was introduced after the end of Wave 5 fieldwork) and the Statistics Act, 1993. All interviewers, as well as other staff working on *Growing Up in Ireland*, were security vetted by An Garda Síochána (the Irish Police Service).

All work in Wave 5 of the Infant Cohort was carried out under ethical approval granted by a dedicated and independent Research Ethics Committee convened by the Department of Children and Youth Affairs, especially for *Growing Up in Ireland*. The Research Ethics Committee was very rigorous in its review and consideration of all the materials and procedures used in the project.

References

- Behr, A., Bellgard, E. and Rendtel, U. (2005). Extent and determinants of panel attrition in the European Community Household Panel. *European Sociological Review, 21* (5), 489-512.
- Berry, J.O. & Jones, W.H. (1995). The Parental Stress Scale: Initial Psychometric evidence. *Journal of Social and Personal Relationships, 12* (3), 463-472.
- Elliott, C.D., Smith, P. & McCulloch, K. (1997). *British Ability Scales Second Edition (BAS II): Technical Manual.* London: NFER-Nelson.
- Emlen, A. C., Koren, P. E., & Schultze, K. H. (2000). A Packet of Scales for Measuring Quality of Child Care from a Parent's Point of View. Portland University.
- Feinberg, M. E. (2003). The internal structure and ecological context of coparenting: A framework for research and intervention. *Parenting: Science and Practice, 3,* 95-131.
- Feinberg, M. E., Brown, L. D., & Kan, M. L. (2012). A multi-domain self-report measure of coparenting. *Parenting*, 12, 1-21
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology* and *Psychiatry, 38,* 581-586.
- Greene, S., Williams, J., Layte, R., Doyle, E., Harris, E., McCrory, C., Whelan, C. (2010). *Growing Up in Ireland, Background and Conceptual Framework.* Office of the Minister for Children and Youth Affairs.
- Gresham, F. M. & Elliott, S. N. (2008). *Social Skills Improvement System: Rating Scales Manual*. Minneapolis, MN: NCS Pearson, Inc.
- Hodgson, R.J., John, B., Alwyn, T., Hodgson, R.C., Waller, S., Thom, B. & Newcombe, R. (2002). Fast Screening for Alcohol Misuse: Manual for the FAST Alcohol Screening Test. London: Health Development Agency.
- Lynn, P. (2009). *Methodology of Longitudinal Surveys*. UK: John Wiley and Sons. McCane, R. A. & Widdowson, E. M. (2002). *The Composition of Foods (6th ed)*. London.
- Melchior, L., Huba, G., Brown, V., & Reback, C. (1993). A Short Depression Index for Women. *Educational and Psychological Measurement, 53,* 1117–1125
- Nicoletti, C. & Peracchi, F. (2005). Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel. *Journal of the Royal Statistical Society, 168* (4), 763-781.
- Pianta, R.C. (1992). Child-parent Relationship Scale. Unpublished measure, University of Virginia.
- Piers, E. V. (1969). Piers-Harris children's self-concept scale. Nashville, TN: Counselor Recordings and Tests.
- Sanson AV, Smart DF, Prior M, Oberklaid F & Pedlow R 1994. The structure of temperament from age 3 to 7 years: age, sex, and sociodemographic influences. *Merrill-Palmer Quarterly*, 40 (2), 233-252.
- Sabourin, S., Valois, P. & Lussier, Y (2005). Development and Validation of a brief version of the dyadic adjustment scale with a nonparametric item analysis model. *Psychological Assessment*, 17 (1), 15-27.
- Sternthal, M., Slopen, N., Williams, D. R. (2011). Racial Disparities in Health: How Much Does Stress *Really* Matter? *Du Bois Review, 8*(1), 95-113.
- Watson, N, & Wooden, M. (2009). Identifying Factors Affecting Longitudinal Survey Response. In P. Lynn (Ed.), Methodology of Longitudinal Surveys. Hoboken, NJ: John Wiley & Sons, Ltd.
- Williams, D. R., Yu, Y., Jackson, J. S., and Anderson, N. B. (1997) Racial Differences in Physical and Mental Health: Socioeconomic Status, Stress, and Discrimination. *Journal of Health Psychology*, 2(3), 335-351.
- Wrieden, W., Longbottom, P., & Barton, K. (2002). *Children's Food Portion Sizes: Estimation of Typical Portion Sizes for Children of Different Ages.* University of Dundee: Centre for Public Health Nutrition Research.

Appendix: Indicative Cronbach's alphas for all scales used in the study

Scale	Respondent	Subscale	α
Strength and Difficulties	PCG	Peer	.60
Questionnaire		Emotional	.69
		Hyperactivity	.80
		Conduct	.56
		Total Difficulties	.74
		Prosocial	.69
	Teacher	Peer	.69
		Emotional	.78
		Hyperactivity	.85
		Conduct	.73
		Total Difficulties	.87
		Prosocial	.80
Pianta Child-Parent	PCG	Positive	.67
Relationship Scale		Conflict	.79
	SCG	Positive	.69
		Conflict	.75
Pianta Student-Teacher	Teacher	Positive	.83
Relationship Scale		Conflict	.86
Parental Stress Scale	PCG		.78
	SCG		.77
Dyadic Adjustment Scale	PCG		.63
	SCG		.60
LSAC Parenting Measure	PCG	Warmth	.88
		Hostility	.68
		Consistency	.61
	SCG	Warmth	.88
		Hostility	.66
		Consistency	.61
Co-parenting Relationship	PCG		.71
Scale	SCG		.73
Parental Stressor Scale	PCG		.78
	SCG		.77
FAST Alcohol Consumption	PCG	Male	.64
Scale		Female	.45
	SCG	Male	.53
		Female	.74
CES-D Depression Scale	PCG		.85
	SCG		.81
Everyday Discrimination Scale	PCG	Majority group	.61
		Minority group	.78
	SCG	Majority group	.65
		Minority group	.83
Piers-Harris Scale (shortened)		Behaviour	.70
		Intellectual	.64
		Physical	.69
		Freedom from anxiety	.71
		Popularity	.65
		Happiness	.63
		Total	.80