An Phríomh-Oifig Staidrimh Central Statistics Office

Fás Aníos in Éirinn Growing Up in Ireland

An Roinn Leanaí, Comhionannais, Míchumais, Lánpháirtíochta agus Óige
Department of Children, Equality, Disability, Integration and Youth

Growing Up in Ireland

Cohort 08 at 13 years of age

Summary Guide for

Wave 6 of the Infant Cohort

August 2023

ESRI ECONOMIC & SOCIAL RESEARCH INSTITUTE

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Table of Contents

List of Tables

1 Introduction and Background

1.1 Introduction

This document provides a summary of the sixth wave of data collection with the Infant Cohort (also known as Cohort '08) at 13 years of age of the ***Growing Up in Ireland*** (***GUI***) study, as well as an overview of the microdata files (Research 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 years[1], 9 years and 13 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). 8,032 families participated in the fifth wave of the survey. Data collection took place between June 2017 and February 2018 when the study children were 9 years old.

The sixth wave of data collection (which is the focus of the current document) took place between September 2021 and June 2022, when the cohort were 13 years of age, resulting in a completed data file of 6,655 cases. Wave 6 also included a school-based phase of fieldwork.

This report describes the background, design, instruments and procedures used only in respect of Wave 6 of the Infant Cohort. Earlier waves of this cohort (and the Child Cohort '98) are the subjects of a parallel set of reports. The focus here is on the sample design and response rates in Wave 6 of the Infant Cohort, the nature and content of the questionnaires and other instrumentation, along with a broad overview of the datasets.

[1] On a postal basis

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 13 years old (Wave 6)*' released by the Department of Children, Disability, Equality, Integration and Youth.

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 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 21st century Ireland.

The study began in 2006 and is a Government-funded study of children being carried out jointly by the ESRI and Trinity College Dublin. It is managed by the Department of Children, Equality, Disability, Integration and Youth in association with the Central Statistics Office[2].

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 13 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 objectives[3]:

1. To describe the lives of Irish children, to establish what is typical and normal as well as what is atypical and problematic

[2] Since January 2023, Growing Up in Ireland is being carried out jointly by the Department of Children, Equality, Disability, Integration and Youth (DCEDIY) and the Central Statistics Office (CSO).

[3] *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. 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.[4]

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

[4] 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

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 Register[5] as having been born between 1st December 2007 and 30th June 2008. The target sample was selected over this seven-month 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 and subsequent waves of the study. 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 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,[6] 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

[5] 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.

[6] 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

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.

A total of 8,032 questionnaires were returned, amounting to 72.1% of the families interviewed at 9 months of age.

2.7 Sample Design at Wave 6 (13 years old)

A total of 9,723 children and their families were contacted for interview in Wave 6, when the children were 13 years old. This included families who were interviewed in Wave 5, when the children were aged 9 years and some families who were interviewed in previous waves but were not interviewed at wave 5. We did not contact families who had left Ireland or those who had asked not to be contacted again. Due to the on-going restrictions caused by Covid-19, interviewers carried out the interviews with families by

telephone, with a short follow up web-based self-complete survey.

A total of 6,655 questionnaire were returned, amounting to 59.8% of the families interviewed at 9 months of age.

2.8 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 6, 60% of the original Wave 1 sample participated in the interview.

Table 2.1 Participation rates Wave 1 to Wave 6

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 5**	**Wave 6**
Year	2008 – 2009	2010 – 2011	2013-2014	2015-2016	2017-2018	2021-2022
Child age	9 months	3 years	5 years	7-8 years	9 years	13 years
Wave participation	11,134	9,793	9,001	5,344	8,032	6,655
Participation rate		88.0%	80.8%	48.0%	72.1%	59.8%

2.9 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 6 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 13-year data consists in carrying forward the earlier weight (which controls for initial non-response and attrition up to the 9-year wave) and adjusting it for attrition between the 9-year and 13-year waves. The Study Team used the GROSS software, as in previous rounds of Growing Up in Ireland.[7] This has been used extensively by the Economic and Social 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 13-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 13-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 1[8]

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

[8] 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.

Most of these characteristics were measured at the 9-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.[9] The distribution of the child and family characteristics in the completed 13-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 13 years old. Given the fixed sample design, children who were living in Ireland at 13 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 13 years and who, accordingly, were no longer growing up in Ireland.

In preparing the Wave 6 dataset, four sets of weights were calculated. The first set should be applied in analysis based on the 6,655 families for whom there is a valid observation at 9 months and 13 years of age. The second set of weights should be used in analysis based on the smaller set of 6,056 families who participated at all five rounds of interviewer-led interviews. (Waves 1, 2, 3, 5 and 6). The third set should be used when analysing any of the primary caregiver sensitive questionnaire data, the fourth set should be used when analysing any of the young person sensitive questionnaire data.

Due to the split nature of data collection for this wave, i.e. interview administered telephone interview, followed by a self-complete web survey; not everyone who completed the telephone interview, also completed the self-complete web survey. In previous waves, as the interviewer was in the household while the respondent completed the self-complete questionnaire, this was not an issue. Given the importance of the data collected in the self-complete questionnaires, it was deemed necessary to calculate a separate weight to ensure the self-complete data was as representative of the population as possible.

Table 2.2 Weighting factors available for use with Infant Cohort at 13 years of age

Families participated at:	***No. of families***	***Weights Factors***
9 months and 13 years	6,655	WGT_13YRa
9 months, 3 years, 5 years, 9 years and 13 years	6,056	WGT_13YRb
9 months and 13 years PCG sensitive	3,963	WGT_13YRc
9 months and 13 years only YP sensitive	3,128	WGT_13YRd

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

The first set (subscript 'a') should be used when one is carrying out analysis on the most complete 13-year sample of 6,655 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 6,655 families. As noted above, population refers to the number of 13-year-olds who were resident in Ireland at 9 months of age and who continue to be resident in the country when they are 13 years old, accounting for those who no longer live in Ireland at 13 years of age or who have deceased since 9 months of age[10].

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 6,056 families who participated in all five interviewer-led rounds of the study, i.e. at 9 months, 3 years, 5 years, 9 years and 13 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 6,056 cases).

The third set (those with subscript 'c') perform exactly the same functions as the first set. The only difference is that they are applied to the smaller sample of 3,963 families who completed the Primary Caregiver Self-complete questionnaire in addition to the main questionnaire. 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 3,963 cases).

The fourth set of statistical adjustments (those with subscript 'd') perform exactly the same functions as the first set. The only difference is that they are applied to the smaller sample of 3,128 households where the 13 year old completed both the main questionnaire and the self-complete questionnaire. These weights should be applied when analysing data with self-complete questions. 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 3,128 cases).

Table 2.3 Combinations of data waves and statistical adjustment factors

	Combination of data waves	**Unweighted number of**	**Weighting Factor**	**Comments**
A	9 months in cross-section	11,134	WGT_9MTH	All families who participated in Wave 1
B	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)

[10] 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 13 years of age, it was decided to not include a grossing factor in this wave of the data.

C	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 9 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.
G	Standalone 13 years (matched 9 mth, and 13 years)	6,655	WGT_13Yra	Families who participated in Waves 1 and 6, 599 of whom did not participate in Waves 2, 3 or 5
H	Matched 9 months, 3 years, 5 years, 9 years & 13 years	6,056	WGT_13Yrb	Families who participated in all five Waves of interviewer-led interviewing to date.
I	Standalone 13 years (matched 9 mth, and 13 years) PCG Sensitive	3,963	WGT_13Yrc	Families who participated in Waves 1 and 6 who filled out a PCG self-complete questionnaire
J	Standalone 13 years (matched 9 mth, and 13 years) YP Sensitive	3,128	WGT_13Yrd	Families who participated in Waves 1 and 6 who filled out a young person self-complete questionnaire

For example, if one wishes to track the trajectories of a particular characteristics across *all five waves* one must use the matched data file containing 6,056 cases as in Row H in Table 2.3 However, if one is interested only in investigating differences between 9 months and 13 years, regardless of the intervening 3-year, 5-year or 9-year wave, one may use the larger sample of 6,655 cases (Row G in Table 2.3). 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 2.3 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 6 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 6 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 6 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.

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 13-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 and consultation with a group of 13-year olds who were members of Comhairle na nÓg[11], 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

The original plan for the pilot approved by the Steering Group and the Research Ethics Committee (REC) was for in-home fieldwork with the 13-year-olds and their parents was drawn up in Autumn 2019 and approved in March 2020, just before COVID-19 became a public health emergency. Following advice on the likely trajectory of the COVID-19 pandemic in 2021, the project Steering Group reached the decision to conduct the 2021/22 main fieldwork using a combination of telephone and web survey.

This sixth 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 by phone at home by a survey interviewer, with a web-based self-complete sensitive questionnaire. In the course of that interview, details were recorded on the school currently attended by the Study Child. The second component

[11] Comhairle na nÓg are child and youth councils in the 31 local authorities of the country, which give children and young people the opportunity to be involved in the development of local services and policies.

involved follow-up with relevant schools to complete school-based questionnaires with the School Principal.

Home-based fieldwork for this pilot phase of the project was carried out in July/August 2020. The sample used in this pilot consisted of 200 13-year-olds (including 5 twin pairs) in 195 families who had participated in the pilot at 9 years of age. Full details of the structure and design, along with results and recommendations from the pilot phase, are available in the Cohort '08 at 13 Years Old Pilot Report currently available on the Growing Up in Ireland website.

4 Survey Instruments

4.1 Introduction

This section provides a general overview of the instruments used in the Infant Cohort at age 13. The household was contacted by the Survey Interviewer by telephone and guided through the main questionnaires. Respondents were then given a unique randomly generated identifier to allow them individual access to the sensitive self-complete questionnaire.

4.2 Household-based instruments

Questionnaires with the Primary and Secondary Caregivers were completed over the telephone. The Survey Interviewer filled out the questionnaire on a laptop (Computer Assisted Telephone Interviewing; CATI). A common Self-complete Questionnaire was completed by Primary and Secondary caregivers on a web-based self-completion (Computer Assisted Web Interview; CAWI) basis. The respondent was provided with a unique id number to gain access to their self-complete questionnaire. Table 4.1 provides an outline of the different Wave 6 household based instruments, divided into sections according to topic. For more detailed information on all the questionnaires and instruments used at Wave 6, see www.growingup.gov.ie/questionnaires/.

In addition to their own main questionnaire and self-complete sensitive questionnaire, children took a cognitive test under the guidance of the telephone interviewer.

Table 4.1 Household based instruments used at Wave 6

Respondent	Mode of Completion	Summary of content
Primary Caregiver	CATI	**Main Questionnaire**
		Section A. Household composition
		Section Z. Covid-19 Experiences
		Section B. 13-Year Old's Health and Disabilities
		Section C. Primary Caregiver's Health
		Section D. 13-Year Old's Emotional Health and Well-

		Section E. Education and School
		Section F. Internet and Screen Time
		Section G. Family Relationships and Context
		Section H. Housing and Socio-demographic
		Section J. About You (the PCG)
		Section K. Neighbourhood/ Community
Secondary Caregiver	CATI	**Main Questionnaire**
		Section Z. Covid-19 Experiences
		Section C. Secondary Caregiver's Health
		Section F. Internet and Screen Time
		Section G. Family Relationships and Context
		Section H. Housing and Socio-demographic
		Section J. About You (the SCG)
Study Child	CATI	**Main Questionnaire**
		Activities and time at home during COVID-19 pandemic related-restrictions
		Activities
		Internet and screen time
		School and education
		Parental supervision and discipline; pocket money
		Physical activities; chores, food; and self-care
		Things the Young Person can have or do
		Feelings and how the Young Person sees themselves
		Siblings and friends
Primary & Secondary Caregiver	CAWI	**Self-complete Questionnaire**
		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
		Currently pregnant? Only asked if female
		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
Study Child	CAWI	**Self-Complete Questionnaire**
		Relationships and sexual education
		Sources of information on sex/relationships
		Sexual Orientation; puberty

		Antisocial behaviour
		Mental Health
		Bullying
		Cigarettes, alcohol and other substances
		Parenting Style
Primary & Secondary Caregiver	Measured by Respondent	Physical Measurements (height and weight)
Study Child	Measured by Respondent	Physical Measurements (height and weight)
Study Child	Administered under guidance of Interviewer	Cognitive Assessment

4.3 School-based instruments

A school-based questionnaire was completed by the Principal of the school the Study Child attends. It was completed on paper and returned to the GUI study team by post. The name and address of the school currently attended by the Study Child were collected during the household interview with the Primary Caregiver. The Principal was asked to fill in the school questionnaire in relation to the current situation. At the time of school-based fieldwork some schools had enrolled a number of Ukrainian pupils which led to additional English Language support needs in the school.

School-based fieldwork began after the Christmas break in January 2022. As Cohort '08 were dispersed around the country, and there was a high likelihood of there being at least one ***Growing Up in Ireland*** participant in each school, Principal Questionnaires were sent to every school in the country. Post-fieldwork, individual participants were matched to the data collected from the corresponding Principal Questionnaire for their school. Each pupil/participant attending the same school had the same school-level information appended to their data row.

An introductory letter explaining the project was sent to the school principal along with the paper questionnaire and a postage-paid envelope for its return. Following the mailing of the questionnaires to the schools, repeated telephone call-backs were made over a period of eight weeks to remind and encourage schools to complete and return them.

Table 4.2 School based instruments used at Wave 6

Respondent	Summary of content
Principal	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
	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.

5 Fieldwork and Implementation

5.1 Introduction

This chapter briefly outlines the fieldwork procedures at Wave 6. 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 Telephone Interview system (CATI)
4. Fieldwork procedures
5. Instruction and practice in the administration of the direct child assessments
6. Child protection guidelines and incident reporting
7. Ethics
8. Summary of other documentation used in the administration of the survey

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. Covid-19 restrictions were still in place during the fieldwork for the sixth wave of the infant cohort. The interviewer phoned each household to arrange an interview. At that phone call, interviewers asked to speak to the person listed at Wave 5 as the Study Child's Primary Caregiver. 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 sheet[12] which had already been posted to the family, and answering any queries the parent had, the interviewer asked the Primary Caregiver to verbally consent to being interviewed. Only after securing consent did the interviewer undertake any work with the family (interviews or tests).

6 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.

6.1.1 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.

[12] A copy of the information sheet provided to participants can be found at www.growingup.gov.ie

6.1.2 Retrospective tracing procedures

The telephone-based nature of the survey in context of Covid-19, which excluded the possibility of face-to-face fieldwork, meant that retrospective tracing procedures as carried out for previous waves were not possible. This probably contributed to the lower than usual response rate.

6.2 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.

7 Structure and Content of the Data File

7.1 Introduction

This section outlines the structure of the Research 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.

7.2 Anonymised (AMF) and Research (RMF) Microdata Files

Two data files are available for researchers: the Research 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. all 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.

7.3 Structure of the data file

Both the Research 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.

7.4 Variable naming

All variables for Wave 6 of the Infant Cohort are 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.

7.4.1 Naming of Home-based Primary and Secondary Caregiver Variables

For the home-based Primary and Secondary Caregiver questions, the prefix 'b' is followed by '6' to indicate the sixth 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 6 will have the variable name '**b6pcC1**'
- Question F8(a) from the Secondary Caregiver Main Questionnaire will have the variable name '**b6scF8a**'
- Question 12 from the Primary Caregiver Self-Complete Questionnaire will be named '**b6pcS12**'

7.4.2 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. This is followed by '6'

to indicate the sixth wave of data collection, and the question number.

Examples:

- Question 'F1' from the Child Main Questionnaire at Wave 6 will have the variable name '**bcq6F1**'
- Question 18 from the Child Self-Complete Questionnaire will have the variable name '**bcs6q18**'

7.4.3 Naming of School-based Variables

For the school-based questions, the initial 'b' prefix is followed by '6' (to indicate the sixth wave of data collection), and 'P' for Principal.

Examples:

- Question 7 on the Principal Questionnaire will be named '**b6p_7**'

7.4.4 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.

7.5 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 Principal questionnaire appear after the other questionnaires. For data protection purposes, all variables on the self-complete questionnaires can only be accessed through the RMF.

Table 7.1 Ordering of variables in the data file

Order	Questionnaire/Section	Variable prefix
1st	Household Grid	p1xxW6, p2xxW6
2nd	Primary Caregiver Main Questionnaire	B6pc
3rd	Primary Caregiver Self-complete Questionnaire (on RMF	B6pcs
4th	Secondary Caregiver Main Questionnaire	B6sc
5th	Secondary Caregiver Self-complete Questionnaire (on RMF	B6scs
6th	Child Main Questionnaire	B6q5

7th	Child Self-Complete Questionnaire (on RMF only)	Bcs6
8th	Physical Measurements	Bpc6, Bsc6
9th	Scale Scores	B6_
10th	Derived Variables	B6_
11th	Cognitive test	B6_
12th	Principal Questionnaire	B6_p

7.6 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**'.

7.7 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 'P1sexW6' where 'W6' indicates Wave 6 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 6 some information on household composition which was captured at Wave 5 was fed forward to the household grid at Wave 6. 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 labelled 'P1xxW6' etc included any corrections made at Wave 6. To ensure confidentiality, only the respondent who identified as the Primary Caregiver at Wave 5 could review the forward-fed information[13]. If the respondent identified as the Primary Caregiver at Wave 5 was no longer resident in the household at Wave 6, the person identifying as the Primary Caregiver at Wave 6 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 31

[13] 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.

onwards are new additions at Wave 3, and; individuals with an original line number from 51 onwards are new additions at Wave 5; individuals with an original line number from 61 onwards are new additions at Wave 6.

Whether or not the Primary Caregiver and Secondary Caregiver roles at Wave 6 are being filled by the same individual as in Wave 5 is indicated by the derived variables '**PCGstatw6'** and '**SCGstatw6'**.[14]

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 (**b6_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-guided household grid and that given in the Primary Caregiver Self-complete questionnaire (self-completed on CAWI). 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.

7.8 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 5 had swapped roles between waves (flagged by the variables '**PCGstatw6'** and '**SCGstatw6'**). Note that more detailed information on the inter-wave swapping of roles is provided in the *RMF*.

7.9 Twins

A data record exists for each child included in the sample. All non-singleton children (those with twins, triplets, etc.) are coded as '**b6_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

[14] Note this information will be unavailable for families who did not complete at Wave 5

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.

7.10 Weighting variables

As discussed in Section 2.8 above, in line with best practice in sample surveys, the data have been re-weighted or statistically adjusted to ensure that the sample is representative of the population[15] 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 weighting factors should be used in significance testing and data modelling.

The data file contains four weighting factors. The first is **WGT_13YRa** which is based on the families who participated at Wave 1 and Wave 6, but not necessarily Wave 2, 3 or 5. The weighting factor incorporates the structural adjustment of the completed sample to the population, whilst maintaining the total completed sample size of 6,655 cases.

The second weighting factor is **WGT_13YRb** and this relates to families who participated at all five main waves (Waves 1, 2, 3, 5 and 6) in the study to date - the reduced sample of 6,056 families. The weighting factor incorporates the structural adjustment of the completed sample to the population, whilst maintaining the completed sample size of 6,056 families who participated in all five main waves.

The third weighting factor is **WGT_13YRc** and this relates to Primary Caregivers who completed both the main primary caregiver questionnaire and the self-complete primary caregiver questionnaire. In previous waves the respondent was encouraged to fill out the self-complete questionnaire as soon as the interviewer was finished with the main questionnaire. For the most part, respondents filled out the self-complete while the interviewer was still in the house (perhaps while the interviewer was interviewing another member of the family). Covid-19 restrictions were still in place during fieldwork for the sixth wave of the survey so interviewers administered the main questionnaire over the phone. When they had completed the main survey the interviewer provided the respondent with a unique id number to access a self-complete e-questionnaire at a later time. The response rate was reduced as a result.

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

The fourth weighting factor is **WGT_13YRd** and this relates to Study Children who completed both the main 13-year old questionnaire and the self-complete questionnaire. Similar to the Primary Caregiver, the 13-year old was guided through the main questionnaire over the phone and given a unique id number to access the self-complete questionnaire at a later time.

The variables ***xxwave1***, ***xxwave2, xxwave3, xxwave4, xxwave5*** and ***xxwave6*** indicate if the case has data for each of the waves. A value of one indicates participation at the relevant wave. In the 13 year data file **xxwave1** and **xxwave6** are equal to 1 for all cases, as all cases in this file have completed both Wave 1 and Wave 6. Frequencies of these indicator variables are outlined below.

Table 7.2 Frequency distribution of 'xxwave' variables

Variable Name	Value	*N*
xxwave1	1	6655
	0	0
xxwave2	1	6478
	0	177
xxwave3	1	6414
	0	241
xxwave4	1	4413
	0	2242
xxwave5	1	6353
	0	302
xxwave6	1	6655
	0	0

7.11 Derived Variables

In addition to some of the derived variables mentioned above (e.g. '**b6_partner**', '**PCGstatw6**' and '**SCGstatw6**'), 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.

7.11.1 Household type (b6_hhtype4, b6_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. **'B6_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

'B6_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.

7.11.2 Equivalised Household Income (b6_equivinc, b6_eIncDec, b6_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 equivalised 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 (**b6_equivinc**) is available in the *RMF* only. In the *AMF*, equivalised household income is given in deciles (**b6_eIncDec**) and quintiles (**b6_eIncQuin**).

7.11.3 Household Class (b6_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 home[16]. On this basis, a social class classification was generated for both Primary and Secondary Caregiver. The classification used was that adopted by the Central Statistics Office (CSO) with seven categories, derived from the census of population, 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.

7.11.4 Household location (b6_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 (**b6region8**: South-East, Dublin, etc) as well as a three-fold classification (**b6region3**: Dublin, BMW, Rest) . There is also a two-fold classification (**b6_region**) which is an urban rural classification.

Unlike in previous waves, no question was asked directly to infer whether the locality could be perceived as rural or urban. This was excluded due to questionnaire space concerns. Instead **b6_region** was derived from the response in wave 5 and carried over when PCG had not changed address.

[16] Current occupation if economically active; previous occupation if retired or unemployed.

7.12 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.

7.12.1 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 D2. The SDQ comprises five subscales, a total difficulties score and, an impact score. The subscales and their corresponding variable names are listed in Table 7.3.

Table 7.3 Subscales of the Strengths and Difficulties Questionnaire and their relevant variable names

Subscale	Primary Caregiver Variable Name
Emotional	B6_SDQemotional
Conduct	B6_SDQconduct
Hyperactivity	B6_SDQhyper
Peer problems	B6_SDQpeerprobs
Prosocial	B6_SDQprosocial
Total difficulties	B6_SDQtotaldiffs

7.12.2 The Pianta Scale - Child Parent Relationship Scale (CPRS) (Pianta, 1992)

It is completed by both the Primary Caregiver and the Secondary Caregiver. Previously, ***Growing Up in Ireland*** has used both the *closeness* and *conflict* subscales from the measure; however, only the conflict subscale was included at the current wave because of the need to reduce the overall length of the survey and the closeness scale tending towards ceiling effects. The conflict subscale has eight statements to which participants respond on a five-point scale from 'definitely does not apply' to 'definitely applies' [to my relationship with the child]. The relevant scale appears as question G1 on the Primary Caregiver Main Questionnaire and as question G1 on the Secondary Caregiver Main Questionnaire. The measure produces a *Conflicts* subscale (**b6pc_conflict, b6sc_conflict**).

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

Question 28 on the 13 year old Sensitive Questionnaire were taken from self-report measures of parenting style which were used in LSAC. All parents have different patterns of interaction with their children. In this case, parenting style refers to the degree of warmth exhibited and patterns of control that parents tend to use when interacting with their children, such as when responding to bad behaviour.

7.12.4 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 10. The subscale is the six-item Parental Stressors sub-scale (**b6pc_stress**, **b6sc_stress**).

7.12.5 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 22). For both respondents, a total score was obtained which is a sum of the raw scores (**b6pc_CEStotal; b6sc_CEStotal).** Also included in the file are two variables (**b6pc_CESD; b6sc_CESD**) which categorise respondents into 'depressed' or 'not depressed'.

7.12.6 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 8 & 9, **b6pc_DAS**; **b6sc_DAS**).

7.13 Physical Measurements

7.13.1 Height & Weight

The heights of the Primary and Secondary Caregivers were fed-forward from the previous waves. Height for adults had been recorded in centimetres using a standard measuring stick (Leicester portable height measure) in previous waves. At the end of the CATI interview Primary and Secondary Caregivers were told that they would be asked to record their weight in the self-complete CAWI interview, and if possible they should weigh themselves before they start. The 13-year old and their PCG were told that the 13-year-old's height and weight would be recorded in the child Self-Complete questionnaire, and again they should measure themselves, with parental assistance if possible, before they start. All were given the option to enter their measurements in metric or imperial, in addition to being asked a follow up question if their measurements were an estimate or have been measured in the last month. 13-year-olds were also asked if their parent/guardian or another adult had helped them with the measurements.

All measurements were recorded on the laptop during the course of the CATI interview.

The heights and weights recorded were edited to remove clearly implausible values. The Wave 6 measurements (which include the forward-fed height values where available) can be found in the following variables:

- Primary Caregiver Height (**bpc6cms**)
- Secondary Caregiver Height (**bsc6cms**)
- Study Child Height (**b6kidcms**)
- Primary Caregiver Weight (**bpc6kgs**)
- Secondary Caregiver Weight (**bsc6kgs**)
- Study Child Weight (**b6kidkgs**)

7.13.2 Body Mass Index (BMI)

BMI scores were derived from the height and weight measurements reported for the Primary Caregiver (**bpc6bmi**), Secondary Caregiver (**bsc6bmi**) and Study Child (**b6kidbmi**)[17]. Categorised variables are also provided, which categorise Primary and Secondary Caregivers into underweight, healthy, overweight, obese (**bpc6bmi_cat**, **bsc6bmi_cat**) and Study Child into non-overweight, overweight, obese (**b6kidbmi_cat).**

7.14 Cognitive Assessments

A cognitive test assessing verbal semantic fluency (also referred to as an 'animal naming test') was administered by the interviewer to the Young Person over the phone after they had completed the Main Questionnaire. The test involved naming as many animals as possible in one minute; the interviewer timed the test and recorded the 13-year-old's score.

The total score on the selective attention test is contained in the variable **bcq6_cog.**

7.15 Coding & Editing

The CATI questionnaires administered in ***Growing Up in Ireland*** consisted mainly of closed questions[18]. The program included extensive range and cross-variable consistency checks (both hard and soft)[19]. 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

[17] On the AMF, BMI scores are derived from the original height and weight measurements before top and bottom coding.

[18] Almost all CAWI questions were closed.

[19] 'Hard' edit consistency checks in a CATI 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.

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.

7.16 Forward-feed from Wave 5

As discussed in Section 6.7 above, some variables were fed forward from previous waves to reduce interview time at Wave 6. Adult height was also forward-fed as noted earlier. A summary of all other variables that were fed forward at Wave 6 is provided in the Table 7.4 below.

Table 7.4 Details on variables forward-fed from previous waves (excl. household grid and adult height)

Variable name	**Variable description**	**Rules**
B6_region	Two-fold region – Urban/Rural	Taken from previous waves if the household has not moved address since the last wave
bpcJ7, bscJ7	Ireland as country of birth	Asked if missing from previous wave or new respondent
bpcJ9, bscJ9 *[RMF only]*	Length of time living in Ireland	Asked if missing from previous wave or new respondent
bpcJ10, bscJ10	Ethnicity	Asked if missing from previous wave or new respondent
B6pJ1_ed, b6sJ1_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

7.17 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 CATI interview but chosen to not complete the self-complete CAWI 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**: 'b6_pcgmain', 'b6_scgmain', 'b6_pcgsens', 'b6_scgsens', 'b6_prin'.** 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 CATI), 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. It is not possible from the data file to distinguish between a missing value that is a 'refusal' to answer or a 'don't know' or a skipped question (either because of the routing of the questionnaire or because the respondent skipped over the question). In the case of the sensitive questionnaires, respondents were able to skip over questions they did not want to answer. No distinction is made on the data file between the different types of unanswered variables.

A programming error in Blaise meant that not everyone who should have answered H23 on the Primary Caregiver Main questionnaire were routed in to answer these questions. These variables have been excluded from the data files to avoid under-reporting.

8 Short and Proxy Questionnaires

In order to facilitate as many young people as possible, the Study Team introduced new instruments and protocols to allow those unable to complete a full interview to take part. If the 13-year-old was unable to complete the full main questionnaire for any reason, they were offered the Short Questionnaire. Ideally, this Short Questionnaire would be completed by the 13-year-old themselves, although they were welcome to complete it with the help of an advocate if needed (e.g., a carer, sibling, or parent). If the 13-year-old was unable to complete the Short Questionnaire, even with the help of an advocate, the parent was invited to complete a Proxy Questionnaire on their behalf (with consent from both the child and parent). The Short and Proxy Questionnaires were not used as a means of encouraging reluctant participants who were otherwise able to complete a full interview.

The order of preference for completing the Young Person Questionnaire was as follows:

1. 13-year-old completes the full main questionnaires with the interviewer
2. 13-year-old completes the shorter main questionnaire with the interviewer
3. 13-year-old completes the shorter main questionnaire with the help of an advocate
4. Parent completes the proxy interview on behalf of the 13-year-old

The 13-year-old was invited to self-complete the Young Person Sensitive Questionnaire if able to do so. Shortened, assisted, or proxy questionnaires were not offered for sensitive items.

Both the Short and Proxy Questionnaires were made up of questions taken from the 13-year-old Main Questionnaire. The variable **b6_ypmain** on the RMF data can be used to identify questionnaires that were filled out by proxy or a shortened version of the questionnaire was used. It is up to the researcher's discretion whether to include data from shortened or proxy questionnaires in their analysis.

8.1.1 Short Questionnaire Structure

The short questionnaire contained eight sections:

A. Preliminaries
B. Activities and time at home during Covid-19 restrictions
C. Activities
D. Internet and screen time
E. School and education
F. Pocket money
G. Physical activities and chores
K. Siblings and friends

8.1.1.1 Section A - Preliminaries

Section A included a single item for the Young Person to confirm that they had read the information sheet, discussed participating with a parent/guardian, and agree to take part in the survey.

8.1.1.2 Section B - Activities and time at home during Covid-19 restrictions

Section B comprised of a single item from the ***Growing Up in Ireland*** Special COVID Survey addressing experiences of home learning during the COVID-19 restrictions.

SECTION B	**Construct**
	Learning experiences during school closures

8.1.1.3 Section C - Activities

Section C addressed the 13-year-old's participation in sports and/or cultural activities.

SECTION C	Construct
	Participation in team sports, cultural clubs, or other community groups
	Participation in informal sports, cultural, or social activities

8.1.1.4 Section D - Internet and screen time

Section D covered the Young Person's internet usage and screen time.

SECTION D	Construct
	Internet accessible devices
	Online profiles

8.1.1.5 Section E - School and education

Section E included items on the Young Person's school and education.

SECTION E	Construct
	Class
	Subjects taken
	Transition to secondary school
	Perception and experience of school

8.1.1.6 Section F - Pocket money

Section F asked about pocket money.

SECTION F	Construct
	Pocket money

8.1.1.7 Section G - Physical activities and chores

Section G included items on physical activity and chores.

SECTION G	Construct
	Physical activity
	Chores

8.1.1.8 Section K - Siblings and friends

Section K included items on siblings, friends, informal sources of support, and future career aspirations.

SECTION K	Construct
	Relationship with siblings
	Friendship network

8.1.2 Proxy Interview Structure

The proxy interview contained six sections:

A. Preliminaries
C. Activities
D. Internet and screen time
E. School and education
F. Pocket money
G. Physical activities and chores

8.1.2.1 Section A - Preliminaries

Section A included a single item the Young Person to confirm that they have read the information sheet, discussed participating with a parent/guardian, and agree to take part in the survey.

8.1.2.2 Section C - Activities

Section C addressed the 13-year-old's participation in sports and/or cultural activities.

SECTION C	Construct
	Participation in team sports, cultural clubs, or other community groups
	Participation in informal sports, cultural, or social activities

8.1.2.3 Section D - Internet and screen time

Section D covered the Young Person's internet usage and screen time.

SECTION D	Construct
	Internet accessible devices
	Online profiles

8.1.2.4 Section E - School and education

Section E included items on the Young Person's school and education.

SECTION E	Construct
	Class
	Subjects taken
	Transition to secondary school
	Perception and experience of school

8.1.2.5 Section F - Pocket money

Section F asked about pocket money.

SECTION F	Construct
	Pocket money

8.1.2.6 Section G - Physical activities and chores

Section G included items on physical activity and chores.

SECTION G	Construct
	Physical activity
	Chores

9 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 6 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.

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 Questionnaire	PCG	Peer	.55
		Emotional	.74
		Hyperactivity	.75
		Conduct	.55
		Total Difficulties	.72
		Prosocial	.65
Pianta Child-Parent Relationship Scale	PCG	Conflict	.84
	SCG	Conflict	.82
Dyadic Adjustment Scale	PCG		.75
	SCG		.71
Parental Satisfaction Scale	PCG		.62
	SCG		.61
Parental Stressor Scale	PCG		.76
	SCG		.73
Parenting Style Inventory	13-year old		.84
Parental Control	13-year old		.75
Rosenberg Self-Esteem	13-year old		.82
Mental Health Inventory	13-year old		.78
Short Mood and Feelings	13-year old		.92
CES-D Depression Scale	PCG		.83
	SCG		.81