

Growing Up in Ireland Study Overview / Data Workshop

Dr Eoin McNamara Derek Nolan

Research and Evaluation Unit Department of Children, Equality, Disability, Integration and Youth

> **Fás Aníos in Éirinn** Growing Up in Ireland

Structure

1) Introduction to Growing Up in Ireland

- Study background
- Sample design
- Study design

2) Online Resources

Publications, questionnaires and technical documents

3) Accessing the Data

- Familiarising yourself with the data
- Applying for AMF / RMF
- Practical examples

3b) Intro to AMF Tables

- How to read tables
- What next?

1) Introduction to *Growing Up in Ireland*

- Study Background
- Sample Design
- Study Design

History of the Study

- Growing Up in Ireland (GUI) is the national longitudinal study of children and young people
- Established by the Irish Government in 2006
 - Calls for a study from 1980s-90s
 - Dearth of Irish data on key areas of child research
 - Ref: UK's National Child Development Study
- Original governance structure:
 - Led by the Economic & Social Research Institute and Trinity; data archived by the CSO
 - Funded by the Irish Government (Dept of Children)
- Revised structure (as of Jan 2023):
 - Now managed and overseen by the Central Statistics Office (CSO) and DCEDIY
 - Funded by the Irish Government (CSO & DCEDIY)

Objectives of GUI

- Provide evidence to create effective and responsive policies and services for children, young people and their families
- Study the lives of children/young people in Ireland
- Establish what is typical/normal, as well as what is atypical/problematic
- Identify the key factors that most help or hinder children's development
- Establish the effect of early child experiences on later life
- Obtain children's views and opinions on their lives

About GUI

- Growing Up in Ireland is a longitudinal study
 - Interview young people and families at successive waves
 - Waves often align with key milestones in the young person's life
- It utilises a fixed panel design
 - Same children with no additions in between waves (e.g. excludes recent immigrants)
- GUI was originally a dual cohort study
 - Cohort '98: recruited at age 9yrs
 - Cohort '08: recruited at age 9mths
- New birth cohort launched in Sept 2024 (Cohort '24)
 - Fieldwork underway (n~2,500)
 - Already planning for Wave 2 at 3yrs

Longitudinal Design

- Longitudinal design involves interviewing same people on numerous occasions
 - Tracks the progress of the same child, and their family, over an extended period
 - Differs from cross-sectional design (different respondents at each wave)
- WHY? Longitudinal design allows us to consider...

CAUSAL PROCESSES

- Effects of early childhood experiences on later development
- Change over time and dynamics of behaviour

EVALUATE POLICIES

- When and how it is best to intervene to support children/families
- Gauge the effectiveness of interventions and policies

International Examples of Cohort Studies

Centre for Longitudinal Studies (UK):

- National Child Development Study 1958
- British Cohort Study 1970
- Millennium Cohort Study 2001

Comparable international studies:

- Longitudinal Study of Australian Children 2004
- Growing Up in Scotland 2005
- Growing Up in New Zealand 2009
- The Irish Longitudinal Study of Aging (TILDA) 2009

Extensive list available here: https://en.wikipedia.org/wiki/Longitudinal_study

Three Cohorts of GUI: Timeline





What Data are Available?

Cohort '08

Wave	Age	Fieldwork	Archived
1	9 months	April 2009	Yes
2	3 years	July 2011	Yes
3	5 years	Sept 2013	Yes
4	7/8 years	Sept 2016	Yes
5	9 years	July 2018	Yes
	COVID Survey	Dec 2020	Yes
6	13 years	June 2022	Yes
7	17 years	2026	N/A

Cohort '98

Wave	Age	Fieldwork	Archived
1	9 years	June 2008	Yes
2	13 years	March 2012	Yes
3	17 years	July 2016	Yes
4	20 years	April 2019	Yes
	COVID Survey	Dec 2020	Yes
5	25 years	April 2024	2025

Examples of Policy Impact

DRCD

Civic engagement rural areas

Arts Council

strategic planning for children's cultural participation

DES

Review of career guidance

DCEDIY

Covid: schools re-opening ECCE take up
Early Learning impacts
Work life balance issues
BOBF priorities
ABC evaluation
Housing conditions & child outcomes
Parenting relationships
Migrant children
Discrimination

NCCA

Transitions to primary school

DoH

Medical card usage Population health planning Screen time

HRB

Pathways and drug use

DECCAE/ComReg

Mobile phones and learning

HSE

Young people's health behaviours Sex and sexual behaviour among young people

NDA

Parental educational expectations of children with disabilities

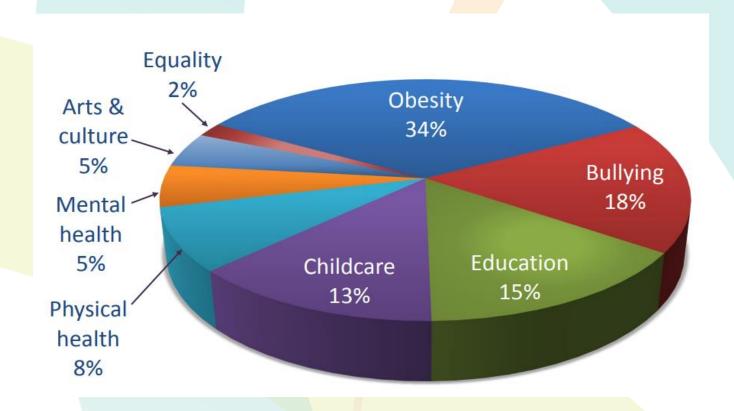
Institute of Public Health

Advising on upcoming legislation Vaping Gambling

NCSE

Estimating special needs prevalence Educational outcomes for students with SEN

GUI in Dáil Debates (100 times 2007- 18)



Examples of Research Impact

Trinity College

Digital media, screen time & mental health Ethnicity & breastfeeding rates Chronic illness and educational failure Folic acid & cleft palate in infants

UCC

Individual, family and environmental effects on physical activity Childhood obesity and the local food environment

Treoir

Unmarried and solo parent experiences

NUIG

Overweight, obesity and health Socioeconomic profile of childhood disability

Dental Hospital, Trinity

Dental problems across GUI waves

NCI

home learning environments and disadvantage

Maynooth University

Children & grandparents Effects of school age childcare on outcomes

HEA/Trinity gambling and participation in sport over time

ESRI

Attitudes to the Irish language Ante natal care pathways Access to GP services and GP fees Energy poverty and child health Primary to post primary transitions

Mary I Limerick

Home learning environment and cognitive development Gender and attitudes to maths Determinants of active commuting to school

DCU

Parent book reading at 9 months and vocabulary development at 3

Media Coverage

THE IRISH TIMES

Mon. Nov 29, 2021



Positive outcomes for children who have good relationship with father, study finds

Long working hours a barrier to fathers' involvement, ESRI research shows

Ø Mon. Nov 8, 2021, 01:43

Independent.ie

Business Sport Life Style Entertainment Travel

Sections

Education Health Courts Crime Centenaries

Opinion

School is a bigger influence on teen behaviour than the neighbourhood they live in, says ESRI report



Report highlights inequalities among 13year-olds

Updated / Wednesday, 17 Oct 2018 20:34









THE IRISH TIMES

Mon, Nov 29, 202

NEWS SPORT BUSINESS OPINION LIFE & STYLE CULTURE

Ireland > Irish News

Children say using internet better than playing with friends, study finds

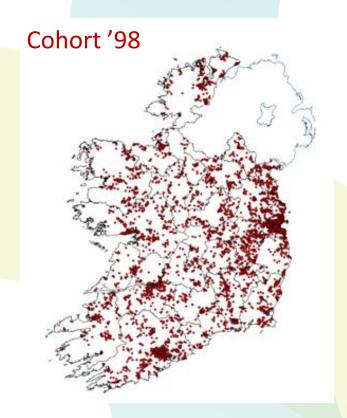
Growing Up in Ireland study finds most children enjoy online games more than playing with friends

Wed, Jun 16, 2021, 00:01 Updated: Wed, Jun 16, 2021, 00:14

1) Introduction to *Growing Up in Ireland*

- Study Background
- Sample Design
- Study Design

Samples Distribution





Sample Representativeness

1 in every 6 children from each cohort is a GUI participant

























Infant Cohort: 11,134 out of 75,173 babies born in 2008

Child Cohort: 8,568 out of 53,969 kids born in 1998

Sampling

Cohort '98	Cohort '08
56,500 9-year-olds in population	73,6 <mark>62 infants</mark> (less than one year old) in population
Random sample of 8,500 9-year-olds resident in Ireland	Random sample of 11,000 9-month-olds resident in Ireland
Represented 14% of all 9-year-olds	Represented 15% of all 9-month-olds
Two stage, clustered sample design:	Child Benefit Register used as sampling frame
Stratified random sample of Primary schools, 82.3% response rate	CBR is very representative source
Random sample of children within school	Sampled over 7-month period

Follow-up at Subsequent Waves

- Tracing information collected at Wave 1
 - PPSN
 - Family / friend contact details
- Initial contact from Head Office
 - Followed up by face-to-face visit from interviewer
 - If possible, same interviewer as Wave 1
- ~85-90% response rate at subsequent waves
- Fixed panel design
 - Considering booster samples

Sample sizes at each wave

Cohort '08

Wave	1	2	3	4	5	6
Sample	11,134	9,793	9,001	5,344*	8,032	6,723

Cohort '98

Wave	1	2	3	4	5
Sample	8,568	7,525	6,216	5,190	3,380

In between waves, the sample size can change for 4 main reasons:

- 1. child/family was unable to be reached
- 2. child/family decided not to participate at the current wave
- 3. child/family gave a "hard refusal" for current and future waves
- 4. child has emigrated / deceased

For latter cases, those children are removed from any future waves

Re-weighting the Sample

- Data can be re-weighted to account for differential attrition across waves
- Differential attrition leads to differences between structure of completed sample (at any wave) and wider population
 - According to key socio-demographic indicators
- Re-weighting ensures sample is representative of the population (*at time of recruitment, Wave 1)
- Re-weighting should be carried out prior to any analysis
- * Covered in more detail in final section of workshop

1) Introduction to *Growing Up in Ireland*

- Study Background
- Sample Design
- Study Design

Respondents

- Primary Caregiver (PCG)* / Secondary Caregiver (SCG)**
 - Interviews

- Physical Measurements

- Child / Young Person
 - Interviews
 - Physical measurements

- Cognitive tests
- Time-use diaries

- Principal / teacher
 - Postal questionnaires
- * **PCG** self-defined, person who provides most care to the child usually mother
- **SCG self-defined, resident spouse/partner of PCG usually father

Securing Informed Consent

 Information sheets, consent forms, assent forms sent in advance of home visit

- Independent consent/assent for parent(s) and child
- Signed consent (and assent) essential before any data collection undertaken

Fieldwork

- Interviews conducted face-to-face in the home, where possible
- Main interview administered by interviewer on Computer Assisted Personal Interview (CAPI) basis
- Sensitive interview self-administered on Computer Assisted Self Interview (CASI) basis
- Cognitive tests / physical measurements completed with interviewer
- Time-use / food frequency diaries left behind and posted to field office
- Principal/teacher questionnaires self-completed and posted

Fieldwork Adaptations

- COVID-19 pandemic: not possible to conduct in-person interviews for Cohort '08 at age 13
 - Interviews were administered online (CAWI)
- Questionnaires were shorter than originally planned
 - Some loss in cross-cohort and longitudinal consistency
 - Cognitive testing / physical measurements not possible
 - Impacted response rates
- Positive takeaways to revised mode too
 - Less paperwork, lower costs
- Mixed modes incorporated into subsequent study designs
 - In-person, telephone, online

Information Recorded - Cohort '08

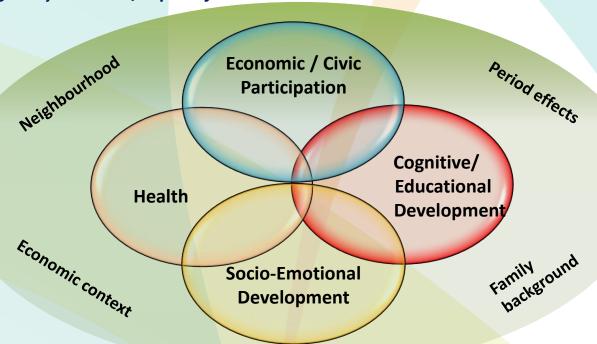
	PCG	SCG	Child	Cognitive Tests	Principal	Teacher	Physical Measures	Time Use
Wave 1: 9mth								
Wave 2: 3yr								
Wave 3: 5yr								
Wave 5: 9yr								
Wave 6: 13yr								

Information Recorded – Cohort '98

	PCG	SCG	Young Person	Cognitive Tests	Principal	Teacher	Physical Measures	Time Use
Wave 1: 9yr								
Wave 2: 13yr								
Wave 3: 17yr								
Wave 4: 20yr								
Wave 5: 25yr								

Multi-disciplinary Coverage





Questionnaire Topics

Socio-emotional Well-being	Physical Health	Education / Cog. Development	Economic & Civ. Participation	Family Background
Child's relationships	Pregnancy, birth, breastfeeding, etc.	Childcare arrangements	Young Person's economic status	Family social class
Child's lifestyle	Child's general health / longstanding conditions	Child's education / learning environment	Young Person's labour force experience	Family income
Child's socio-emotional development	Healthcare utilisation	Child's cognitive development	Young Person's Income	Sociodemographic information
Family context	Child's nutrition / diet	Child's attitudes to school	Political and community engagement	Parental education
Parenting	Child's physical activity levels/exercise	Educational performance	Religion and spirituality	Parental employment
Marital/Partner relationship	Physical measurements	Further/higher education	Confidence in state institutions	Family structure
Non-resident parent	Risky behaviours, parental health and lifestyle	School characteristics	Perceived discrimination	

Classificatory Variables

- Gender
- Household composition partner, # of children
- Parental employment status / education
- Family income
- Urban or rural

- * Included on data files
- * Useful for initial data analysis

Time-Use Diary

- Split into 15-minute slots for full day (96)
- List of activities tick box to indicate which activity was being undertaken in each time slot
- Completed by Study Child and/or Primary Caregiver
- Specified diary day to ensure an even spread of days

Activity AM	00.0 15 3			01.00 15 30			2.00 a 5 30			0 am 30 45	. 1	04.00 15 30		15	.00 aı 30 4	5	15.3	0 am 30 45	Нi	07.00 15 30	45	15	3.00 :		09.0 15 3		0.00 a 5 30		11.0 15 3	
 SLEEPING / RESTING (including time trying to get to sleep, trying to get up) PERSONAL CARE OR CETTING READY (showering, washing, dressing, brushing teeth or hair, doing make-up, getting changed or ready for school, for training, for going out or for going to bed) 	+		Ť		T	Ħ		Ⅱ	Ī	П	Ť	Ħ	Ī		T	Ī	Ť		t		↱	•	t			t		\pm	\pm	Ш
3. EATING (breakfast, lunch, dinner, tea)			_								_	\perp							┸			_ ⊢	•			ш	ш	ш	\perp	டட
4. TRAVELLING (to or from school or elsewhere)		ш	_								_								_				_	_			\perp	\perp	\perp	\Box
S. AT SCHOOL/COLLEGE																									_	-	\pm	\pm	\pm	-
6. AT WORK		Ш	\perp							ш	\perp								ш								ш	\Box	\perp	
7. DOING HOMEWORK OR STUDY																														
S. JUST HANGING AROUND WITH FRIENDS (outside or inside)		Ш									\perp								\mathbf{L}											
9. SPENDING TIME WITH FAMILY																														
10. PLAYING WITH OR EXERCISING A PET		П									\perp																	\Box	\perp	
11. AT THE GYM, PLAYING SPORT OR DOING PHYSICAL EXERCISE (training, matches)																														
12. ATTENDING A SPORTS EVENT		П									_																		\perp	
13. USING THE INTERNET / EMAILING (including social networking, browsing etc)																														
 PLAYING COMPUTER GAMES (e.g. Playstation, PSP, X-Box or Wii) 																													\perp	டட
15. TALKING ON THE PHONE OR TEXTING																											ш		\perp	
16. MUSIC LESSONS (OR PRACTICING MUSIC), DRAMA, CLASSES ETC																														
17. WATCHING TV, FILMS, VIDEOS OR DVDS																											ш		\perp	
18. LISTENING TO MUSIC																														
19. READING FOR PLEASURE OR INTEREST (NOT FOR SCHOOL/COLLEGE/STUDY)																											ш		\perp	
 HOUSEWORK (preparing food, tidying bedroom, feeding pets) 																														டட
21. HOBBIES AND OTHER LEISURE ACTIVITIES		LI									_																ш			டட
 OUT SHOPPING TO BUY THINGS (groceries, clothes etc). 			_						\perp		_								_			\Box				Ш	\Box	\Box	\perp	
23. GOING TO DISCOS OR BARS, ETC.																											ш			
24. GOING TO PARTY OR OTHER SOCIAL EVENT (in people's houses)			_						\perp		_	\Box							_		\perp					ш	\Box	\Box	\perp	\Box
25. OTHER 1 (SPECIFY)		\Box	┸					\perp		\Box	⊥	Ш						\Box	1	\Box			\perp		\perp		ш	\perp	丄	ட

2) Online Resources

Publications, questionnaires and technical documents

GUI Website

www.growingup.gov.ie

- First & best resource for study info & technical documents
 - Questionnaires
 - Design reports
 - Summary guides
 - Data dictionaries
 - Official GUI publications
 - Other publications
 - Conference archive
 - Data workshop archive



Questionnaires – growingup.gov.ie

- * Very useful resource for survey content
- All questionnaires available online
 - All waves, both cohorts
 - Respondents: PCG, SCG, child, teacher/principal
 - Questionnaires: main, self-complete (sensitive)
- Online questionnaires reflect interviews
 - Instructions to interviewers
 - Routing
 - Exact question & response category wording
 - * Minor redactions (if scales under copyright)

WAVE 6 OUESTIONNAIRES - AGE 13 YEARS

Parents/Guardian Questionnaires:

- · Primary Caregiver Main Questionnaire
- · Primary Caregiver Self-Complete Questionnaire
- · Secondary Caregiver Main Questionnaire
- · Secondary Caregiver Self-Complete Questionnaire
- · Primary Caregiver Twin Questionnaire

Young Person Questionnaires:

- · Young Person Main Questionnaire
- Young Person Self-Complete Questionnaire
- Young Person Short & Proxy Questionnaire

Teacher / Principal Questionnaires:

School Principal's Questionnaire

Questionnaires – growingup.gov.ie

L5. How many bedrooms do you have in your home? number of bedrooms [INTERVIEWER IF A STUDIO APARTMENT RECORD AS ZERO BEDROOMS]										
L6.Do you feel that your current accommodation ((excluding location) is suitable for your family's needs?									
Yes1	No									
L7. [CARD L7] Why is that? a. Too small b. Not a child-friendly layout c. Poor conditions in the home (damp, drafts, leaks et d. Other (specify)	tc)									
L8. [Card L8] Which of these descriptions BEST d [INTERVIEWER: IF RESPONDENT IS ON MATERNI RETURN TO, SHE SHOULD BE CODED AS 0] 0. Currently on maternity leave, but have a job to return to	4. Student full-time									

Technical Documents – growingup.gov.ie

Summary Data Dictionary:

- Short version of data dictionary
- Lists variable name and label
- Colour coded by questionnaire

9yr Name	9yr Label
id	Household ID code
WGT_9YRa	Weighting Factor - 9yr Full sample (Xsectional)
WGT_9YRb	Weighting Factor - 9yr Reduced sample (Longitudinal)
xxwave1	Family participated in Wave 1 - 9 months
xxwave2	Family participated in Wave 2 - 3 years
xxwave3	Family participated in Wave 3 - 5 years
xxwave4	Family participated in Wave 4 - 7/8 years
xxwave5	Family participated in Wave 5 - 9 years
PCGstatw5	PCG same as Wave 3
SCGstatw5	SCG if present is same as Wave 3
b5_partner	A1b. Do you have a spouse/partner who lives here with you in the household?
b5pcA4	A4. Total number of people in household - Wave 5
p1sexw5	Person 1 gender Wave 5 Grid (PCG)
p1yearsw5	Person 1 age Wave 5 Grid (PCG)
p2sexw5	Person 2 gender Wave 5 Grid (Study Child)
p2yearsw5	Person 2 age Wave 5 Grid (Study Child)
p3sexw5	Person 3 gender Wave 5 Grid
p3yearsw5	Person 3 age Wave 5 Grid

Data Dictionary:

- All info in summary data dictionary +...
- Value labels (answer cats.)

MMM5

		Value
Standard Attributes	Position	316
	Label	M5. Do you have any family living in this area
	Type	Numeric
	Measurement	Nominal
Valid Values	1	yes
	2	no
Missing Values	8	Refusal
	9	Dontknow

MS14

		Value
Standard Attributes	Position	317
	Label	S14. Current Marital Status
	Туре	Numeric
	Measurement	Nominal
Valid Values	1	Married and living with husband / wife
	2	Married and separated from husband / wif
	3	Divorced
	4	Widowed
	5	Never married
Missing Values	9	Dontknow

Technical Documents – growingup.gov.ie

Summary Guides:

- Background to the study
- Sample design
- Instrument development
- Fieldwork and implementation
- Structure / content of the datasets

Design Reports:

- Similar content to summary guide +...
- Conceptual framework
- Consultation process
- Ethical considerations
- Justification for questions
- Psychometric info on scales

Official Publications – growingup.gov.ie

Key Findings

- Produced after each wave of data collection.
- 3-4 domain-specific short reports, detailing a selection of key descriptive findings







Official Publications – growingup.gov.ie

Descriptive Reports

- Produced for each wave
- Wider scope than KFs
 - Some are theme-specific
- More in-depth analysis
 - correlations, regression, long. trends
- Not exhaustive
 - Many variables not included in analysis

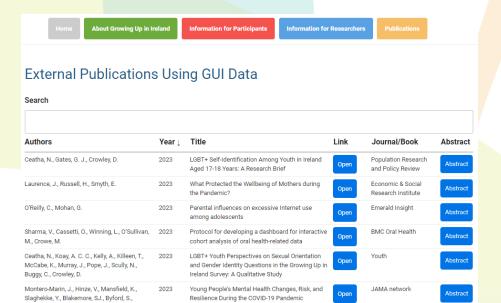




External Publications – growingup.gov.ie

An up-to-date list of External Publications

- Journal articles, reports
- * Must use GUI data





Exploring cumulative disadvantage in early school leaving and

planned post-school pathways among those identified with

special educational needs in Irish primary schools

Eamonn Carroll . Selina McCov. Georgiana Mihut

Conference Archive

- 15 Annual Research Conferences to date
- Showcase new research using GUI data
- 20-30 research presentations per conference
- Ref: Book of abstracts, conference programme
- Useful resource for researchers:
- What research has been conducted to date?
- What topics / interactions can I explore?



Growing Up in Ireland Annual Conference 8th Nov 2023

Room 1: Stratocaster A&B

9:00	Welcome - Laura McGarrigle, Assistant Secretary, DCEDIY										
9:10	Address by Minister Roderic O'Gorman TD Undate on Growing Up in Ireland by Ciara Pidoeon, DCEDIY, and CSO										
9:25	Update on Growing Up in Ireland by Ciara Pidgeon, DCEDIY, and CSO										
9.45	Five-minute room change										
	Room 1: Stratocaster A&B	Room 2: Alhambra	Room 3: Stratocaster C								
	Session A	Session B	Session C								
	Chair: Dr Anne Nolan, ESRI	Chair: Dr Eoin McNamara, DCEDIY	Chair: Dr Jan Skopek, TGD								
9:50	Volunteering among young adults in Ireland Emer Smyth	Identifying high-risk subgroups for self-harm in adolescents and young adults: a longitudinal latent class analysis of risk markers David McEvoy	Impact of bullying and parent-child conflict on self-concept: Analysis using secondary data from waves 1 and 2 of Growing up in Ireland Survey								
			Kayla O'Flaherty								
10:10	Caregiving among young adults: antecedents and outcomes Helen Russell	Young adult functional outcomes of childhood psychopathology Niamh Dooley	The relationship between victimisation, depressive symptoms and self-concept in 9-year-old children								
			Mary Bollard								
10.30	Coffee break										
	Room 1: Stratocaster A&B	Room 2: Alhambra	Room 3: Stratocaster C								
	Session D	Session E	Session F								
	Chair: Dr Ciara Reynolds, IPH	Chair: Dr Anna Visser, Dept. of An Taoiseach	Chair: Marian Brattman, Tusla								
11.00	The association between gambling and mental health outcomes for young people in Ireland	Digital use and digital inequality among Irish children from different ethnic backgrounds	A latent class analysis of mental health symptoms in primary school children: Exploring associations with								
	Gretta Mohan	Melissa Bohnert	school attendance problems Jane Sharpe								
11.20	Use of pornography and sexual health and wellbeing in young adulthood: Evidence from Growing	Narrowing English language achievements gaps by migration background and the role of school	Externalising behaviour among primary school children Emer Smyth								
	Up in Ireland Anne Nolan	Frances McGinnity	Linei Gilyti								
11.40	Associations between parental and child drinking behaviours Eoin McNamara	Lone parent benefit reform in Ireland: beyond the labour market effects Claire Keane	Disruptions and discontinuities in child development: The impact of the pandemic on children's psychological distress								
			Ross MacMillan								

ISSDA and CSO Websites

ISSDA:

- www.ucd.ie/issda/data/guichild/ Cohort '98
- www.ucd.ie/issda/data/guiinfant/ Cohort '08

CSO:

- https://www.cso.ie/en/surveys/householdsurveys/growingupinirelandsurvey/
- How to apply for data*
- Provide relevant technical documents and supplementary documentation

* Covered in detail in next section

3) Accessing the Data

- Familiarising yourself with the data
- Applying for AMF / RMF
- Practical examples

Familiarise Yourself with the Data

Suggested steps to familiarise yourself with the data...

- 1. Questionnaires general overview of the structure/content of all interviews
- **2. Summary data dictionaries** search for specific topics, keywords
- 3. Design reports find info on methodology and justification/validity of items
- **4. Publications / conference presentations explore findings to date, significant trends & interactions, potential analysis techniques**
- 5. AMF tables / AMF analysis

AMF Frequency Tables

- Provide introduction to data for policy-makers
 - Also useful for data users

- Tables displaying response frequencies (n, %)
 - Focus is on main questionnaires
 - Respondents: Young Person, Primary and Secondary Caregiver, Principal
- Introduce data and encourage further investigation
 - Policy-maker: commission research
 - Researchers: conduct further analysis



Survey Responses to Questionnaires Used with Cohort '08 at age 13

An Introduction to the Anonymised Microdata File (AMF) for Policy-Makers and Data Users

lune 2024







A Note on How to Read the Tables

No Answer

Many questions will display the row "No answer" as observed in the table below:

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	5702	85.7	89.5	89.5
	Female	672	10.1	10.5	100.0
	Total	6375	95.8	100.0	10111111111
No answer		280	4.2		
Total		6655	100.0		

This row covers respondents who refused to answer that question, didn't know how to answer or the question was not applicable. For the above question, it is likely that the majority of the "No answer" responses come from households with only a Primary Caregiver and a Young Person and no 3rd person.

Understanding the Different "Percent" Columns

Person 3 gender Wave 6 Grid							
		Frequency	Percent	Valid Percent	Cumulative Percent		
Valid	Male	5702	85.7	89.5	89.5		
	Female	672	10.1	10.5	100.0		
	Total	6375	95.8	100.0			
No answer Total		280	4.2				
		6655	100.0				

The percent column displays each response category as a percentage of the total sample; in this case, only 95.8% (6,375 participants) of the total sample (6,655 participants) answered the question and 4.2% (280 participants) did not. The valid percent column displays each response category as a percentage of the participants that responded to the question (i.e., as a percentage of the 6,375 participants who answered "male" or "female"). The cumulative percent column displays a running sum of the valid percent.

Example:

	E10. How do you feel about school in general?									
					Cumulative					
		Frequency	Percent	Valid Percent	Percent					
Valid	I like it very much	1313	19.7	21.3	21.3					
	I like it quite a bit	2238	33.6	36.4	57.7					
	I like it a bit	1929	29.0	31.3	89.1					
	I don't like it very much	506	7.6	8.2	97.3					
	I hate it	168	2.5	2.7	100.0					
	Total	6154	92.5	100.0						
	No answer	501	7.5							
	Total	6655	100.0							

E19. What is the highest qualification you expect to get by the time you finish your education?								
		Frequency	Percent	Valid Percent	Cumulative Percent			
Valid	Junior Cert	93	1.4	1.7	1.7			
	Leaving Cert	669	10.1	11.9	13.6			
	Certificate or Diploma (including PLC., apprenticeship)	777	11.7	13.9	27.5			
	Degree or higher degree	4067	61.1	72.5	100.0			
	Total	5606	84.2	100.0				
	No answer	1049	15.8					
	Total	6655	100.0					

Data File Characteristics

- Wide format data file, available in SPSS / Stata / SAS
 - All data relating to a child (and their family) contained in a single entry (one row)
- Not all questions will be on a data file (AMF vs RMF)
 - Sensitive information (from self-complete questionnaire)
 - Details that might compromise anonymity / confidentiality
- Variable labels are shortened version of questionnaire wording
 - Same for value labels
- Check value labels on the data file
 - may not exactly match questionnaire answer categories, if categories have been collapsed
 - e.g. Likert scale responses grouped together: satisfied + very satisfied

Data View - SPSS

D											
GUI Data_9YearCohort.sav [DataSet3] - IBM SPSS Statistics Data Editor											
<u>File</u> <u>E</u> dit	<u>V</u> iew <u>D</u> ata <u>T</u>	ransform <u>A</u> nalyz	re <u>G</u> raphs <u>U</u> tilition	es Extensions	<u>W</u> indow <u>H</u> elp						
	Annananti										
Visible: 850 of 850 Var ✓ ID ✓ Wgt 9yr ✓ Gross 9yr ✓ Partner ✓ Int type ✓ MMA2 ✓ MMA3 ✓ MMA4 ✓ MMA4 ✓ MMA5ap1 ✓ MMagep1 ✓ MMagep1 ✓ Mma5rm											
	Ø ID				Int_type	MMA2	MMA3	MMA4	🗞 mma5ap1		♣ mma5rm
1	1000	.30	1.97	0	2	2	1	2	2	32.00	
2	2000	1.85	12.23	0	2	2		2	2	45.00	
3	3000	1.08	7.10	0	2	2		2	2	29.00	
4	4000	.86	5.67	0	2	2		2	2	48.00	
5	5000	.89	5.86	0	2	2		2	2	33.00	
6	6000	.32	2.11	0	2	2		2	2	37.00	
7	7000	1.71	11.30	0	2	2		2	2	35.00	
8	8000	.56	3.71	0	2	2		2	2	50.00	
9	9000	.69	4.53	0	2	2		2	2	34.00	
10	10000	.81	5.34	0	2	2		2	2	35.00	
11	11000	.96	6.33	0	2	2		2	2	34.00	
12	12000	.69	4.56	0	2	2		2	2	36.00	
13	13000	1.94	12.79	0	2	2		2	2	30.00	
14	14000	2.40	15.79	0	2	2		2	2	48.00	
15	15000	1.93	12.71	0	2	2		2	2	28.00	
16	16000	.60	3.98	0	2	2		2	2	36.00	
17	17000	.59	3.90	0	2	2		2	2	32.00	
18	18000	.63	4.16	0	2	2		2	2	34.00	
19	19000	1.79	11.78	0	2	2		2	2		
					_	_		_	_	. 100	

Variable View - SPSS

ta GI	II Data 9VearCohort	cav [DataSet3]	- IRM SDSS S	tatistics Data	Editor						_ -0
	Edit View Bata				AL ME			ano			
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	0	Household ID	None	None	10	≣ Right		ゝ Input
2	Wgt_9yr	Numeric	8	2		None	None	10	■ Right		ゝ Input
3	Gross_9yr	Numeric	8	2		None	None	11	≣ Right		ゝ Input
4	Partner	Numeric	8	0	Partner in hou	{0, No partner}	None	10	■ Right		ゝ Input
5	Int_type	Numeric	8	0	Household inte	{1, Both caregivers i	None	10	■ Right		ゝ Input
6	MMA2	Numeric	1	0	A2. Record ge	{1, male}	8, 9	6	■ Right	Nominal	ゝ Input
7	MMA3	Numeric	2	0	A3.Resps relat	{1, biological parent}	98, 99	6	■ Right	Nominal	ゝ Input
8	MMA4	Numeric	2	0	A4.How many	{1, one}	98, 99	6	■ Right	🚜 Nominal	ゝ Input
9	mma5ap1	Numeric	1	0	Gender P1	{1, male}	8, 9	9	■ Right	🚜 Nominal	ゝ Input
10	MMagep1	Numeric	8	2	Age Person 1	{26.00, 26 years or I	None	10	■ Right		ゝ Input
11	mma5rmp1	Numeric	2	0	Relationship m	{1, Husband/wife}	98, 99	9	■ Right	& Nominal	ゝ Input
12	mma5rcp1	Numeric	2	0	Relationship St	{1, Husband/wife}	98, 99	10	■ Right	& Nominal	ゝ Input
13	mma5pesp1	Numeric	1	0	PES P1	{1, Pre-school}	8, 9	5	■ Right	🚜 Nominal	ゝ Input
14	mma5ap2	Numeric	1	0	Gender P2	{1, male}	8, 9	9	■ Right	Nominal	ゝ Input
15	MMagep2	Numeric	8	2	Age Person 2	None	None	10	■ Right		ゝ Input
16	mma5rmp2	Numeric	2	0	Relationship m	{1, Husband/wife}	98, 99	10	■ Right	Nominal	> Input
17	mma5rcp2	Numeric	2	0	Relationship St	{1, Husband/wife}	98, 99	10	■ Right	🚜 Nominal	ゝ Input
18	mma5pesp2	Numeric	1	0	PES P2	{1, Pre-school}	8, 9	6	■ Right	& Nominal	> Input
19	mma5ap3	Numeric	1	0	Gender P3	{1, male}	8, 9	9	≣ Right	& Nominal	ゝ Input
20	MMagep3	Numeric	8	2	Age Person 3	{50.00, 50 years and	None	10	≅ Right		ゝ Input
21	mma5rmp3	Numeric	2	0	Relationship m	{1, Husband/wife}	98, 99	10	≅ Right	& Nominal	ゝ Input
22	mma5rcp3	Numeric	2	0	Relationship St	{1, Husband/wife}	98, 99	10	≡ Right	& Nominal	ゝ Input

Classificatory Variables

- Core set of classificatory variables included in all data files
- Primarily socio-demographic indicators
- Useful for (initial) bivariate analysis
- Control for these in adjusted regression models

Variable Name (e.g. Wave 2)	Variable
p2sexW2	Child's gender
B2_hhtype4	Household type: 4-category, # of parents & children
B2_partner	PCG has spouse/partner living in the home
B2_EIncQuin	Equivalized household income quintile: deciles also available
B2_hsdclass	Family social class: professional, skilled manual, etc
B2region	Region: urban or rural

Scales

- Standardized measures (sets of questions) measuring an underlying concept
- Examples in GUI:
 - SDQ SMFQ Pianta scale
 - CES-D FAST / AUDIT
- Widely-established instruments
 - Allow for inter- and intra-cohort comparison
 - Used in comparable international studies too
- Tested for reliability, validity, replicability
 - Refer to Design Reports (scoring, psychometrics)
- AMF contains scale totals, RMF may include individual items (if not subject to copyright)
 - All derived variables (i.e. scale and subscale totals) found towards end of data file

3) Accessing the Data

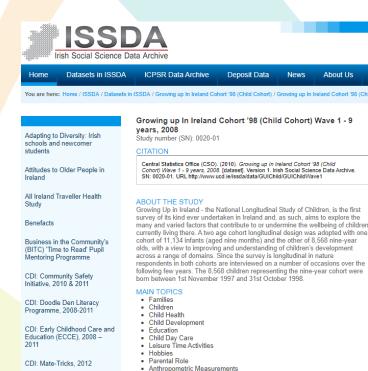
- Familiarising yourself with the data
- Applying for AMF / RMF
- Practical examples

Two Types of Data File

- Anonymised Microdata File (AMF)
- 2. Research Microdata File (RMF)
- Both files made available to researchers after each wave of data collection
 - AMF from ISSDA
 - RMF from CSO
- Data are confidential and anonymised
- Substantial overlap between both file types
 - AMF easier to access
 - RMF provides more detailed and sensitive data

Anonymised Microdata Files (AMF)

- The AMF is prepared and approved by the CSO
- Distributed by Irish Social Science Data Archive (ISSDA)
 - www.ucd.ie/issda/data/guichild/
 - www.ucd.ie/issda/data/guiinfant/
- Fewer 'sensitive' variables included
 - i.e. questions from self-complete questionnaires
- Identifying variables removed
- Top & bottom coding
- Collapsed categories



About Us

Anonymised Microdata Files (AMF)

How to Apply for AMFs

- Download the request form from ISSDA site
- Identify data files (study number, SN)
- 3. Complete all sections of the form:
 - Personal / institution details
 - Short description of intended use of the data
- 4. Sign the End User Licence, email completed application to issda@ucd.ie
- 5. Allow up to three working days for the application to be processed (hopefully quicker)
- 6. Receive a link to download data + separate password

ACCESS INFORMATION Accessing the data

To access the data, please complete a <u>ISSDA Data Request Form for Research Purposes</u>, sign it, and send it to ISSDA by <u>email</u>.

For teaching purposes, please complete the ISSDA Data Request Form for Teaching Purposes, and follow the procedures, as above. Teaching requests are approved on a once-off module/workshop basis. Subsequent occurrences of the module/workshop require a new teaching request form.

Data will be disseminated on receipt of a fully completed, signed form. Incomplete or unsigned forms will be returned to the data requester for completion.

AMF – Data Request Form

Please provide a short description (approx. 100 words) of your intended use of the dataset/s, including title and details of your intended research project*:							
Type of user *	□Academic Staff	☐ Post Doc	□PhD	□Masters			
	□Undergraduate	□Independent Researcher	☑Government/ Policy researcher				
	□Other [please spe	cify]					

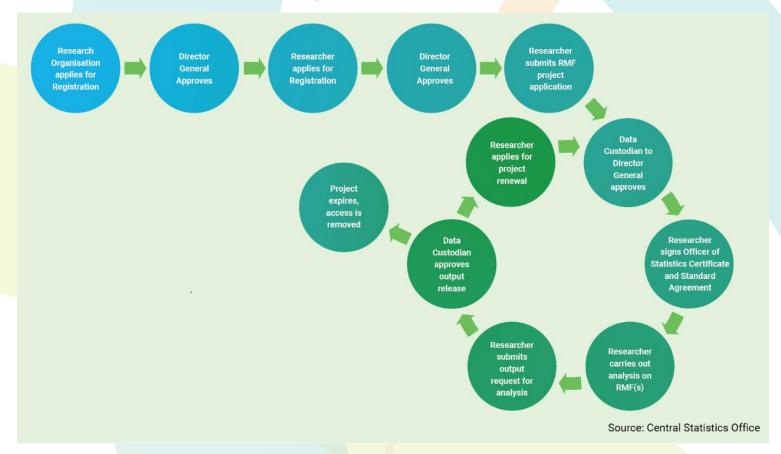
When to Apply for an RMF?

- From preliminary AMF analysis, it will be clear if the variables or level of detail you need are not included on the AMF
- If a variable / question / scale is present on question naire but not on the AMF, it will be available on the RMF
 - Majority of self-complete (sensitive) items are only available on RMF
 - Individual items from scales
- RMFs are only available through the CSO
- More detailed than AMF
 - Less collapsing of variable categories
- Substantially tighter controls / longer turnaround time

Research Microdata Files (RMF)

- Applicants (researchers) must be employed by, or formally related to, a registered research organisation
- If your organisation is not registered (or if you're unsure), contact the Researcher Coordination Unit (RCU) rcu@cso.ie
- All registered research org's will have a designated RMF contact they will need to countersign any RMF application form
- RMF access for students is restricted to those undertaking post-graduate work
 - Supervisor(s) must also apply and complete Officer of Statistics training
- Your organisation's RMF Contact who will set you up on ROSA with a profile, etc.
 - Researcher needs to complete their registration and Officer of Statistics training on ROSA
- On approval, ROSA will generate and issue the RMF Standard Agreement and Officer of Statistics cert accept/sign
- Data files will be made available on **remote desktop**; researchers can use SPSS, R and Stata

RMF Process Lifecycle



Finished Analysis

- Access to both AMF and RMF is project specific and time limited
 - You will need to re-apply for an extension
- Inform ISSDA/CSO when finished with AMF/RMF
- RMF access will be directly controlled by CSO who have an ethical obligation to monitor statistical outputs before releasing them for use
- No copies of the data should be retained by the researcher

3) Accessing the Data

- Familiarising yourself with the data
- Applying for AMF / RMF
- AMF Frequency Tables
- Practical examples

A Note on Syntax

*Examples hereafter use SPSS, but many similarities with other common statistical programs

There are two ways to conduct statistical analysis (and manage/manipulate data)

- 1) Using graphical interface (i.e. drop-down menus)
- 2) Using syntax
 - Syntax is a programming command language; write commands to run procedures

Key Advantages of Syntax:

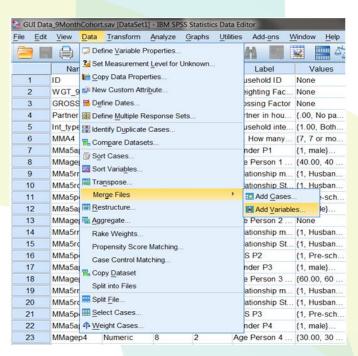
- Can be saved, providing a record of all analyses you've performed
- Repeatable and flexible can be edited, improved, fixed and re-run
- Retrace your steps, keep notes, and communicate your process

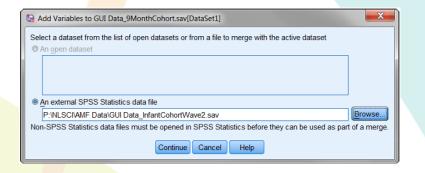
Matching Files

WHY?

If we want to conduct any longitudinal analysis i.e. compare results across waves

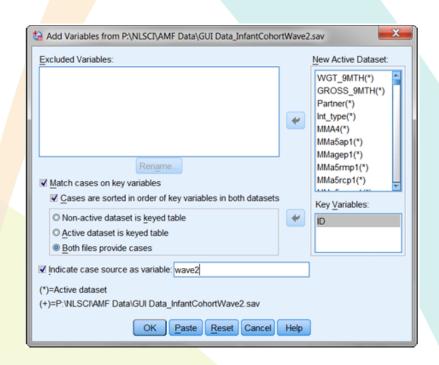
HOW? Open a data file (e.g. C'08 at Wave 1), then using drop-down menu...





Matching Files

- Match cases on key variables
 - Cases are sorted in order of key variables in both datasets
 - By default, sorted by ID
- Indicate case source as variable rename
 - New variable 'wave2', cases coded 1 if in Wave 2, 0 if not
- Repeat for subsequent waves...
- Save
- Conduct longitudinal analysis
- * Remember, file structure is wide
- * Still just one row per ID



Matching Files

Using Syntax...

```
FILE HANDLE wave1 name = 'P:\NLSCI\AMF Data\GUI Data_9MonthCohort.sav'.

FILE HANDLE wave2 name = 'P:\NLSCI\AMF Data\GUI Data_InfantCohortWave2.sav'.

FILE HANDLE wave3 name = 'P:\NLSCI\AMF Data\GUI Data_InfantCohortWave3.sav'.
```

FILE HANDLE merged name = 'P:\NLSCI\AMF Data\GUI Data_InfantCohortWave1Wave2Wave3.sav.sav'.

GET FILE wave1. SORT CASES by id.

MATCH FILES FILE = * / FILE = wave2 / IN wave2 / BY id / MAP. FREQUENCIES wave2.

MATCH FILES FILE = * / FILE = wave3 / IN wave3 / BY id / MAP. FREQUENCIES wave3.

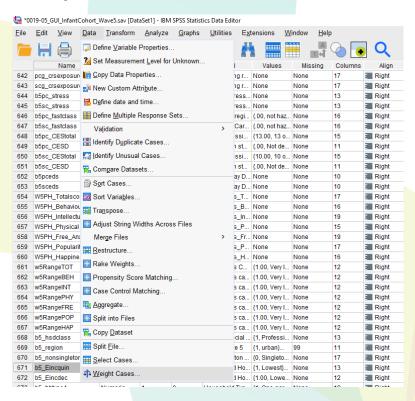
SAVE OUTFILE = merged.

- Data can be re-weighted to account for differential attrition across waves
 - Leads to differences between structure of completed sample (any wave) and wider population
 - Ensures sample is representative of the population (*at time of recruitment fixed panel design)
 - Based on key socio-demographic variables
- Re-weighting should be carried out prior to any analysis
 - Standard procedure / considered best practice in sample surveys
- System used for GUI is called GROSS (used at ESRI)
 - Minimum information loss algorithm which fits population marginals in a regression framework and adjusts the sample to ensure that it produces estimates which match known population parameters
 - Similar to CALMAR and ADJUST

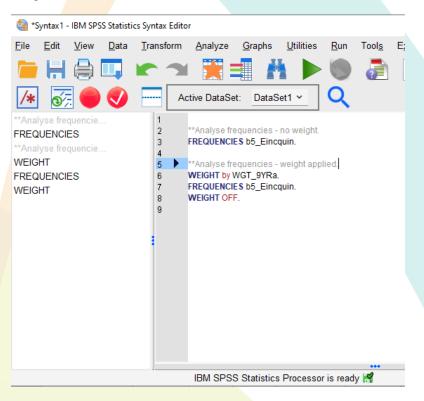
- Data file will usually* contain two weighting factors
 - Using an example of Cohort '08 at 9yr data (Wave 5)....
- 1. WGT_9YRA; based on children who participated at Wave 1 & Wave 5
 - For cross-sectional (one wave only) analysis
 - May have participated at Wave 2-4, but not relevant to our analysis
- **2. WGT_9YRB**; based on children who participated at all main waves (1,2,3,5)
 - For **longitudinal** analysis
 - A reduced sample = smaller N than using WGT_9YRA

^{*} older data files also have a grossing factor (e.g. GROSS_9YR), which calibrates to the population total of 60-70,000 children (varies by wave)

- Drop-down menu



- Syntax



- No weight

WEIGHT OFF. FREQUENCIES b5_Eincquin.

b5_Eincquin Equivalised Household Net Annual Income -Quintiles W5

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Lowest	1154	14.4	15.8	15.8
	2 2nd	1355	16.9	18.6	34.4
	3 3rd	1433	17.8	19.7	54.1
	4 4th	1646	20.5	22.6	76.6
	5 Highest	1704	21.2	23.4	100.0
	Total	7292	90.8	100.0	
Missing	System	740	9.2		
Total		8032	100.0		

- Weight applied

WEIGHT by WGT_9YRa. FREQUENCIES b5_Eincquin.

b5_Eincquin Equivalised Household Net Annual Income -Quintiles W5

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Lowest	1449	18.0	20.0	20.0
	2 2nd	1448	18.0	20.0	40.0
	3 3rd	1443	18.0	19.9	60.0
	4 4th	1447	18.0	20.0	80.0
	5 Highest	1450	18.1	20.0	100.0
	Total	7238	90.1	100.0	
Missing	System	794	9.9		
Total		8032	100.0		

Thank You

Questions?

Contact...
eoin.mcnamara@equality.gov.ie

Ownership and Use of the Data

- Data collected under the Statistics Act 1993
 - Clearly sets out the terms and conditions of use of the data recorded under it
 - Ensures that the integrity and confidentiality of the data is maintained
 - Data shall be used for statistical compilation and analysis only
- RMF data users must complete Officer of Statistics training and abide by the conditions of the RMF Standard Agreement
- Data are owned by the State and accessed under licence from the CSO
- No data which can be related to an identifiable person shall be disseminated, shown or communicated to any person or body

Example of 'Edge Cases', Cohort '98 at 17 years

- Data becomes increasingly complex as multiple patterns of participation become possible
- Parent and Young Person (YP) give consent individually. This may lead to fragmentary cases
- YP deemed to be most important source of information from Wave 3 onwards
- Only the 6,216 cases containing a YP interview included in the AMF/RMF (214 Parent interviews with no YP are not published)
- SES for 59 cases with no Parent interview can be taken from the previous wave

Wave 3		Interview from Parent		
		No	Yes	
Interview from	No	-	214	214
YP	Yes	59	6,157	6,216