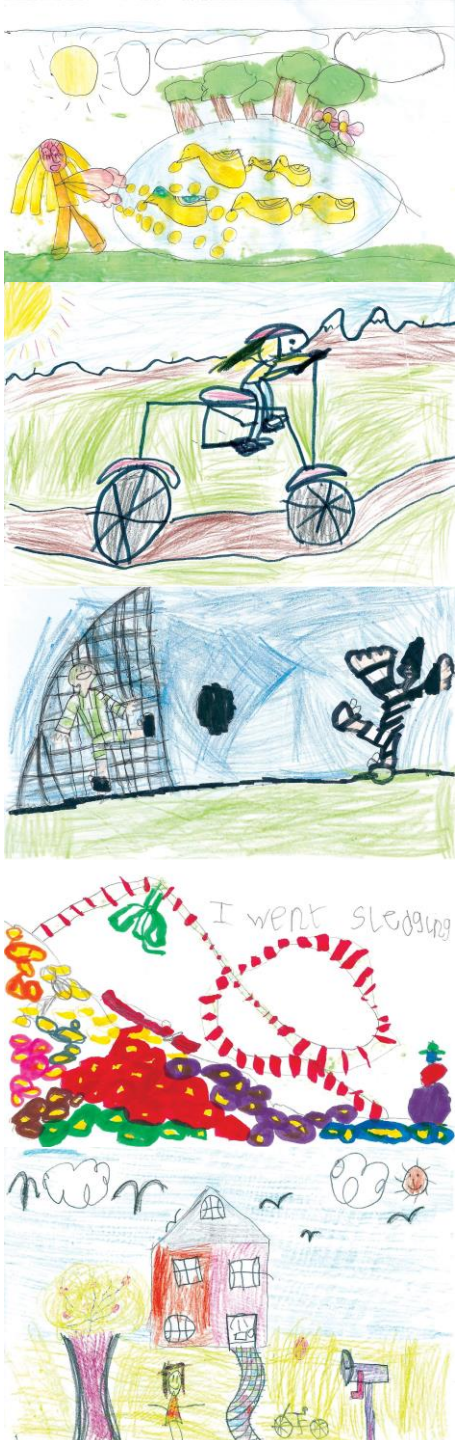
ScotCen
Social Research that works for society

# Growing Up in Scotland

Survey design and methodology:
sampling, response and weighting

# Overview

- Study design
  - Research objectives
  - Study design overview
  - Sample design
  - Data collection
- Response rates
- Weighting
  - Overview
  - Weighting for GUS



# 1. Study design

# The purpose of GUS

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

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

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


ScotCen
Social Research

# Study design: overview

- National sample capable of analysis by urban/rural, deprived/non-deprived and other sub-groups of interest
- Sample drawn from Child Benefit Records
  - Good coverage
  - Some limitations
- Three cohorts:
  - Birth cohort 1: 5217 children, born 2004/05, aged 10.5 months at the 1st interview
  - Child cohort: 2859 children, born 2002/03, aged 34.5 months at the 1st interview
  - Birth cohort 2: 6127 children, born 2010/11 aged 10.5 months at the 1st interview



ScotCen Social Research

# Ages and stages

| Child's age | Cohort/Year of data collection | | |
|---|---|---|---|
| | Child cohort | Birth cohort 1 | Birth cohort 2 |
| **10 months** | | 2005/06 | 2010/11 |
| **Age 2** | | 2006/07 | |
| **Age 3** | 2005/06 | 2007/08 | 2013 |
| **Age 4** | 2006/07 | 2008/09 | |
| **Age 5** | 2007/08 | 2009/10 | 2015 |
| **Age 6** | 2008/09 | 2010/11 | |
| **Age 8** | | 2012/13 | |
| **Primary 6 (Age 10)** | | 2014/15 | |



ScotCen Social Research

# Cross-sectional time-series

| Child's age | Child cohort (CC) | Birth cohort 1 (BC1) | Birth cohort 2 (BC2) |
|---|---|---|---|
| 10 months | | 2005/06 | 2010/11 |
| Age 2 | | 2006/07 | |
| Age 3 | 2005/06 | 2007/08 | 2013 |
| Age 4 | 2006/07 | 2008/09 | |
| Age 5 | 2007/08 | 2009/10 | 2015 |
| Age 6 | 2008/09 | 2010/11 | |
| Age 8 | | 2012/13 | |
| Primary 6 (Age 10) | | 2014/15 | |



ScotCen Social Research

# Cross-sectional time-series

| Child's age | Child cohort (CC) | Birth cohort 1 (BC1) | Birth cohort 2 (BC2) |
|---|---|---|---|
| 10 months | | 2005/06 | 2010/11 |
| Age 2 | | 2006/07 | |
| Age 3 | 2005/06 | 2007/08 | 2013 |
| Age 4 | 2006/07 | 2008/09 | |
| Age 5 | 2007/08 | 2009/10 | 2015 |
| Age 6 | 2008/09 | 2010/11 | |
| Age 8 | | 2012/13 | |
| Primary 6 (Age 10) | | 2014/15 | |



ScotCen Social Research

# Longitudinal data

| Child's age | Child cohort (CC) | Birth cohort 1 (BC1) | Birth cohort 2 (BC2) |
|---|---|---|---|
| 10 months | | 2005/06 | 2010/11 |
| Age 2 | | 2006/07 | |
| Age 3 | 2005/06 | 2007/08 | 2013 |
| Age 4 | 2006/07 | 2008/09 | |
| Age 5 | 2007/08 | 2009/10 | 2015 |
| Age 6 | 2008/09 | 2010/11 | |
| Age 8 | | 2012/13 | |
| Primary 6 (Age 10) | | 2014/15 | |

# Sample Design (1)

- **Random/Probability sample**
  - Every person in sample frame has a known (and non-zero) probability of selection
  - Statistical theory applies
    - You can calculate error
    - You can estimate non-response bias
  - Key concepts
    - **Sampling error** - extent to which estimates based on random samples vary from true value in the population
    - **Confidence interval** - estimate of the range in which actual value in the population will fall (+/-)
    - **Confidence level** - how confident you are about your estimates, e.g. 40% (+/- 3% at the 95% confidence level)

# Sample Design (2)

**Clustering - area level sampling**

- Areas made up by aggregating Data Zones
- Data Zones merged into larger areas
- Each merged area of similar ‘size’ (i.e. same average no. of births per year)
- List of areas sorted by Local Authority and then by SIMD
- 160 areas selected at random, a further 39 added later (130 for BC1)

**Individual level**

- Within each area, ALL children who met the date of birth criteria were selected (for child cohort, 3/5 of children selected)
- Sampling undertaken monthly
- Multiple child households

# Data collection

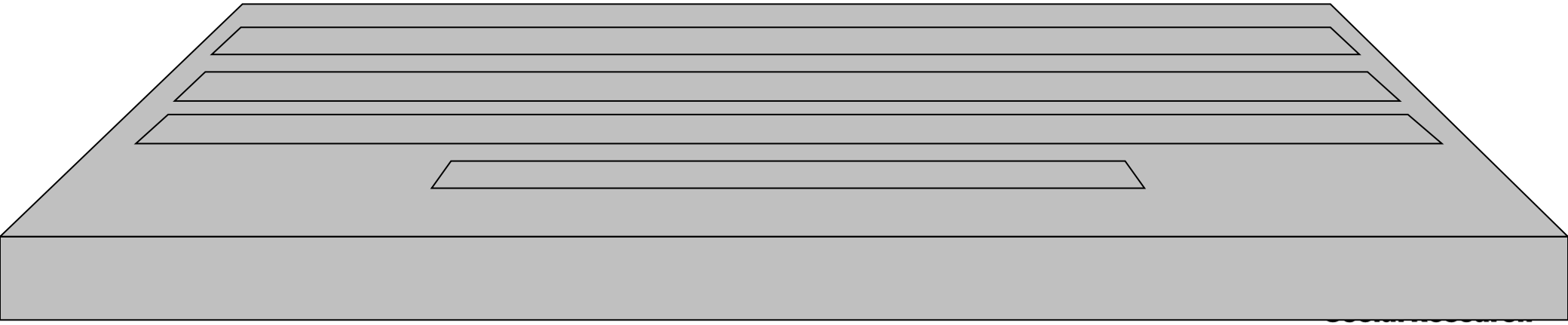
- Face-to-face CAPI interview with self-complete (CASI) section – just over 60 minutes
- Respondent to be child's 'main carer' but aimed to get mother as far as possible (and did so in upwards of 95% of cases at all sweeps)
- At subsequent sweeps aim is to interview, where possible, respondent from previous sweep
- Timing of fieldwork:
  - Monthly 'waves'
  - Target interview dates



ScotCen Social Research

# The CAPI Method

- Questions and answers held on a laptop computer
- Interviewers follow a set ‘script’ with question text and answers restricted
- Pre-program complex routing – so that only certain respondents are asked particular questions
- Build in logic and range checks – to ensure the data recorded makes sense
- Once an interview is completed, the data is downloaded to our Ops Dept in Brentwood ready for editing and analysis.



GUS Sweep 7

Forms Answer Navigate Options Help

SHOWCARD F2

How often do you look at books with Bart or read stories with him?

- 1. Every day / most days
- 2. Once or twice a week
- 3. Once a fortnight
- 4. Once every 1 or 2 months
- 5. Once every 3 or 4 months
- 6. Once every 6 months
- 7. Once a year or less often
- 8. Varies too much to say
- 9. Never

| | | | | | |
|---|---|---|---|---|---|
| Intro | 1 | Cont | ActsIntro | 1 | Cont |
| Pabs1 | 1 | SAgree | SActiv5 | | |
| Pabs2 | 2 | Agree | SActiv8 | | |
| Pabs3 | 2 | Agree | SActiv4 | | |

42/140 | P7100 | 9999901 | QStyles.SActiv5 | 03/09/2010 | 11:10:43

File Edit View Favorites Tools Help

Favorites | Growing Up in Scotland | Page | Safety | Tools

# Growing Up in Scotland

ScotCen Social Research

Welcome to Growing Up in Scotland. We're interested in hearing from you about family life in Scotland.

This questionnaire is voluntary and should take about 10 minutes to complete.

Contact us

Find out more

FAQs

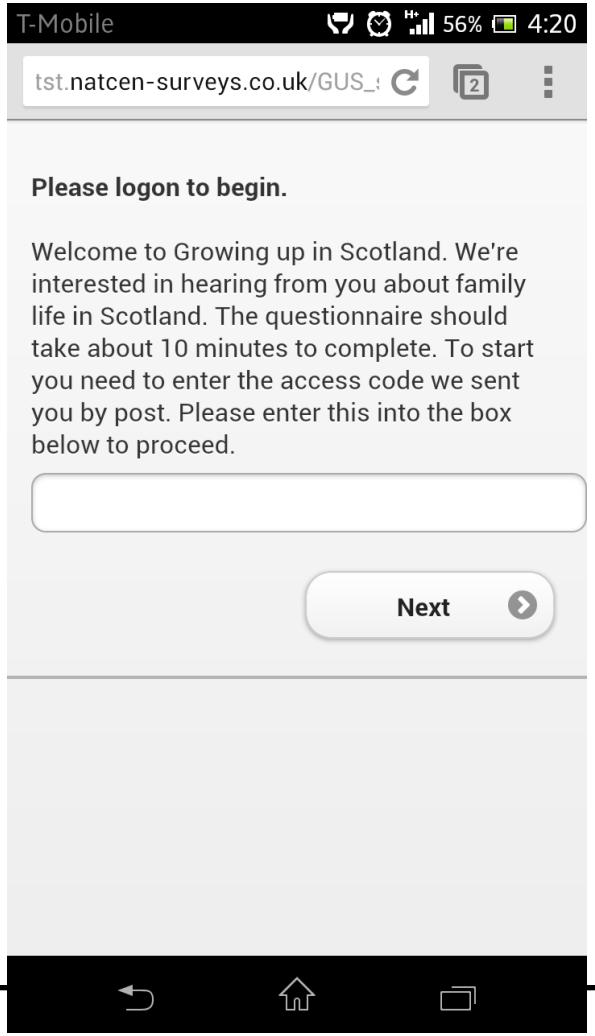
**Start Survey**

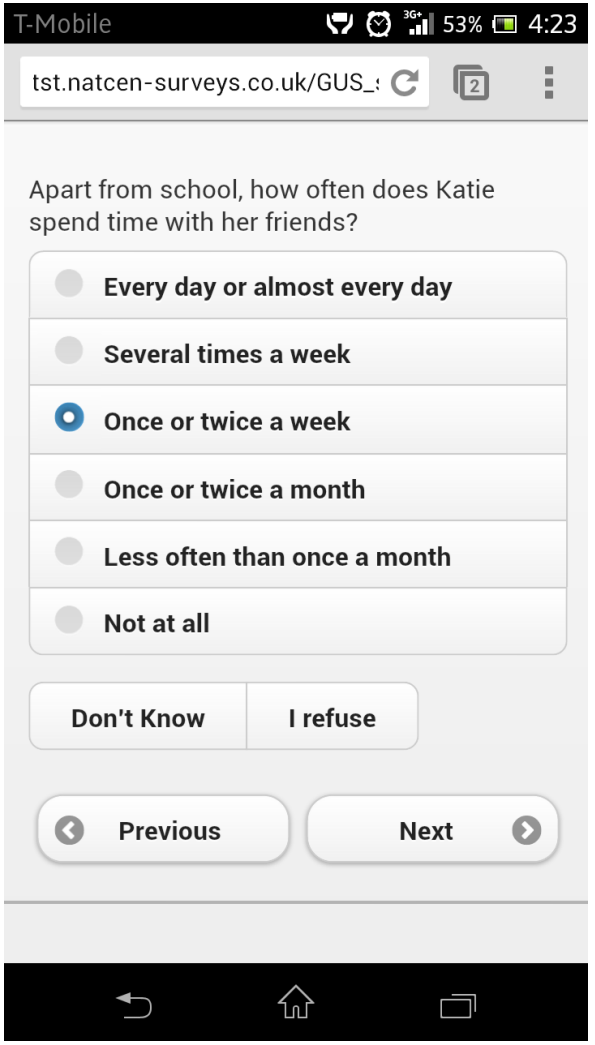




start | 4 Microsoft ... | Microsoft Word | Microsoft Excel | 2 Microsoft ... | Growing Up in... | 99% | 16:29

# Smartphone questionnaire






ScotCen Social Research

# Response rates

# 2.

# Response and attrition rates

| | No. cases achieved | Response rate | |
|---|---|---|---|
| | | As % of issued | As % of sw1 achieved |
| Sweep 1 – BC1 | 5217 | 80% | 100% |
| Sweep 1 – BC2 | 6127 | 65% | 100% |
| Sweep 2 | 4512 | 88% | 86% |
| Sweep 3 | 4193 | 90% | 80% |
| Sweep 4 | 3994 | 91% | 77% |
| Sweep 5 | 3833 | 92% | 74% |
| Sweep 6 | 3657 | 87% | 70% |
| Sweep 7 | 3453 | 83% | 66% |

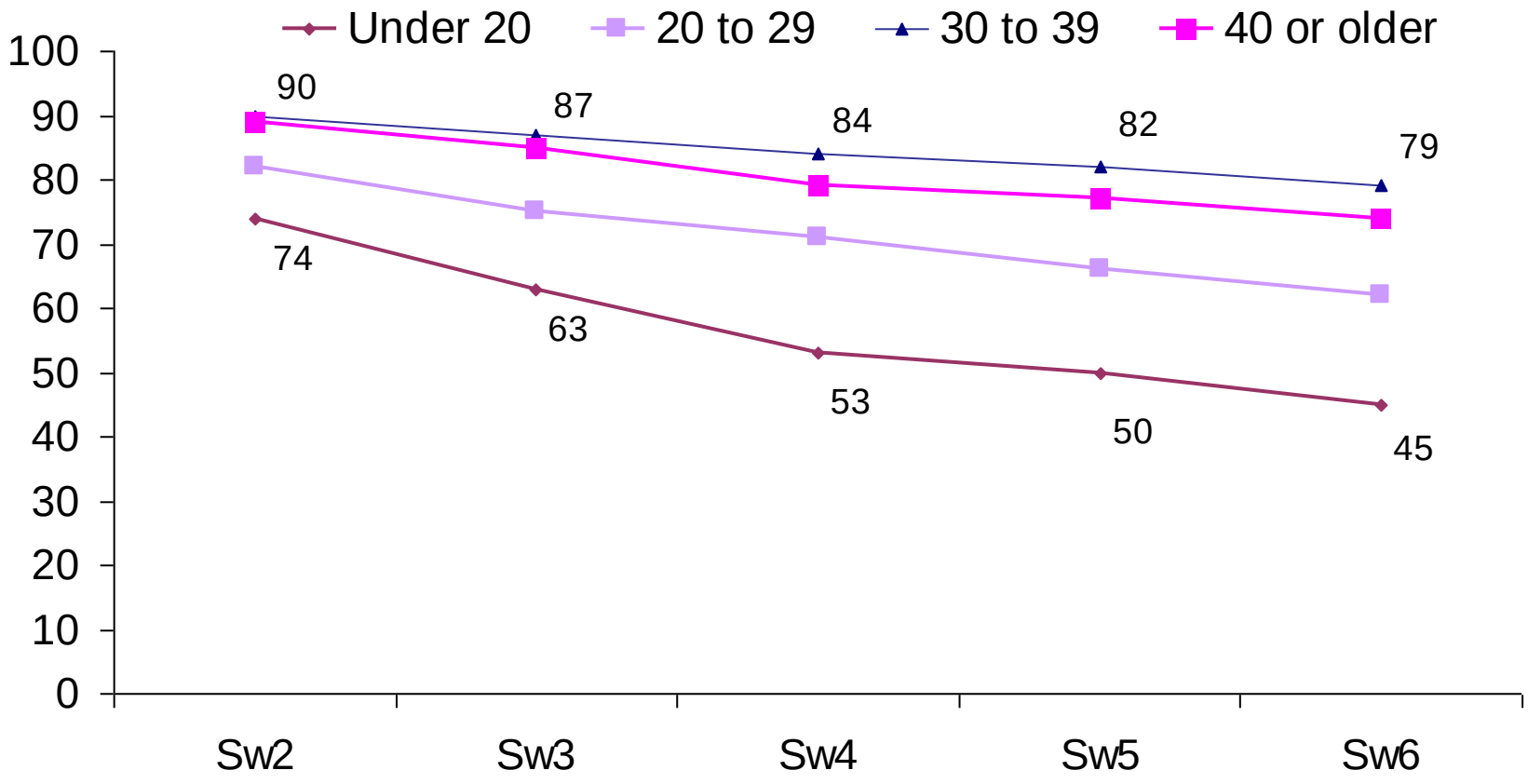
# Non-response

- Why is this an issue?
- After sweep 1, survey data and area-level variables are used to model non-response
- Factors affecting non-response are similar at each sweep.
- Analysis indicates that non-response more likely amongst
  - Lower income families
  - Lone parents
  - Families living in more deprived areas
  - Mothers who had not breastfed
  - Parents who did not attend parent and child groups
  - Younger mothers



ScotCen Social Research

# % of sw1 achieved who responded at subsequent sweeps by maternal age at child's birth





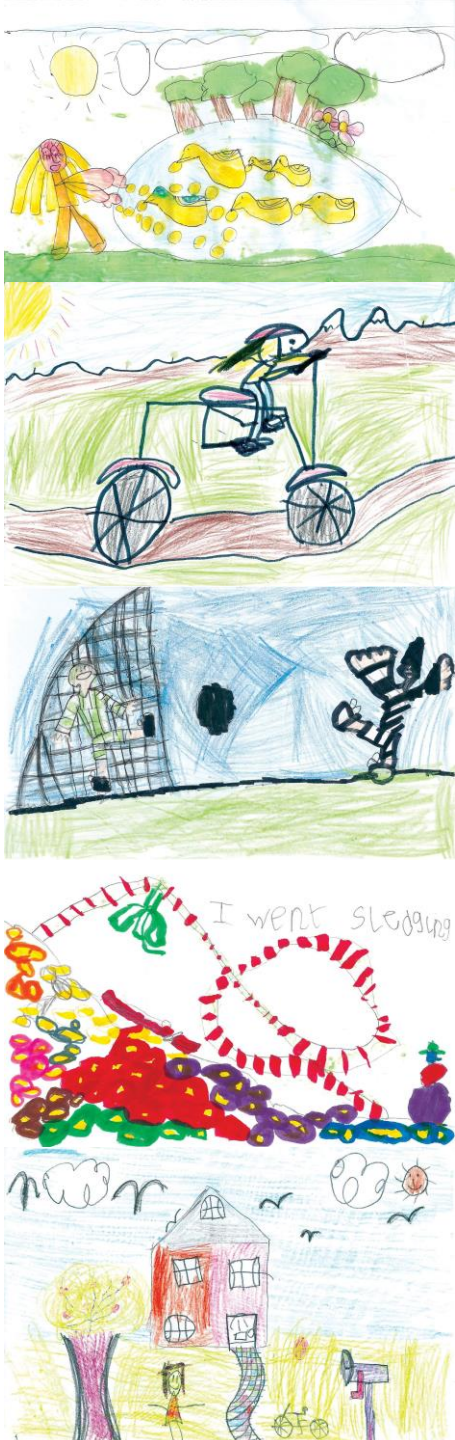

ScotCen Social Research

# Weighting

# 3.

# Weighting overview

- Why do we need weights?
  - To make the achieved sample look as much like the population as possible
    - Selection weights – correcting for unequal selection probabilities
    - Non-response weights – to correct for any bias in achieved sample
- Advantages:
  - Correct for selection and non-response bias
  - Allow inferences about national population, not the sample
- Disadvantages
  - Reduce sample efficiency

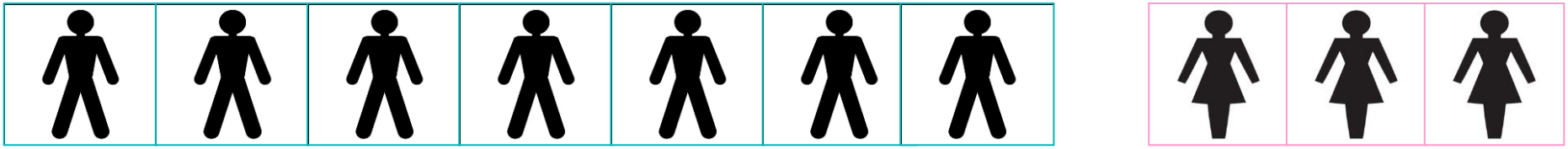


# How weights work

Unweighted sample

70% male

30% female



*x 0.71 =*

Weighted sample

*x 1.67 =*

50% male

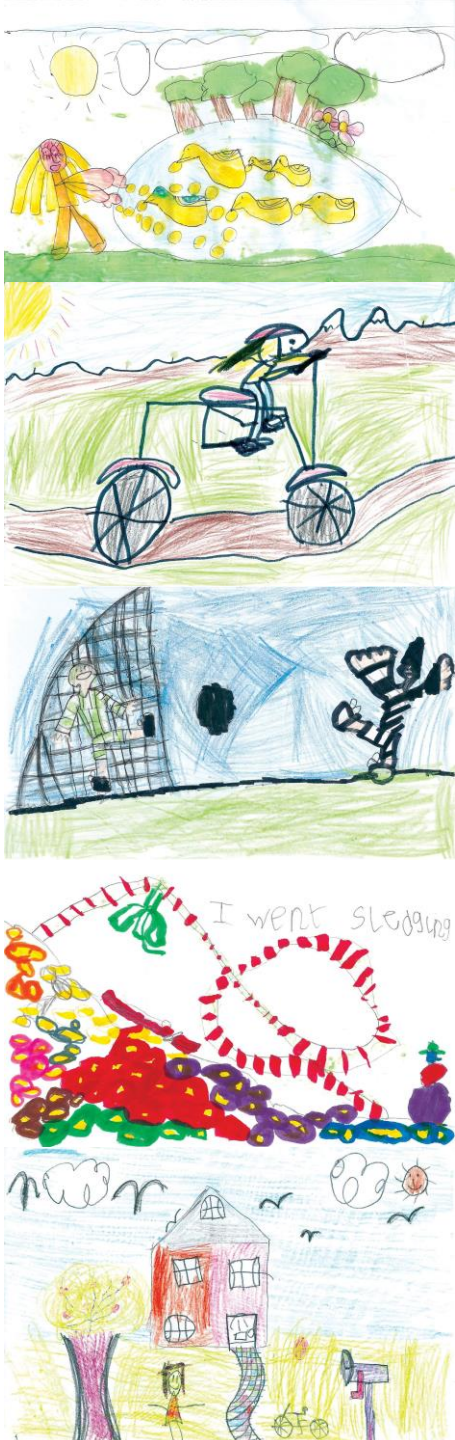
50% female



ScotCen Social Research

# Weighting overview

- Why do we need weights?
  - To make the achieved sample look as much like the population as possible
    - Selection weights – correcting for unequal selection probabilities
    - Non-response weights – to correct for any bias in achieved sample
- Advantages:
  - Correct for selection and non-response bias
  - Allow inferences about national population, not the sample
- Disadvantages
  - Reduce sample efficiency



# Sample efficiency (longitudinal)

| Cohort | Actual sample size | Effective sample size | Sample efficiency | 95% CI for an estimate of … | | |
|---|---|---|---|---|---|---|
| | | | | 10% | 30% | 50% |
| **BC1** | | | | | | |
| Sweep 1 | 5217 | 5061 | 97% | 0.8% | 1.3% | 1.4% |
| Sweep 2 | 4512 | 4294 | 95% | 0.9% | 1.4% | 1.5% |
| Sweep 3 | 4120 | 3829 | 93% | 1.0% | 1.5% | 1.6% |
| Sweep 4 | 3844 | 3484 | 91% | 1.0% | 1.5% | 1.7% |
| Sweep 5 | 3621 | 3221 | 89% | 1.0% | 1.6% | 1.7% |
| Sweep 6 | 3657 | 3262 | 89% | 1.2% | 1.8% | 1.9% |
| **BC2** | | | | | | |
| Sweep 1 | 6127 | 5818 | 95% | 0.8% | 1.2% | 1.3% |

# GUS weights

- Sweep 1 – all cohorts
  - Single weight corrects for selection and non-response bias
- Sweep 2:
  - Two weights:
    - Main interview weight
    - Partner weight
  - Each correct for non-response at sweep 2
  - The main interview weight includes the weight from sweep 1, the partner weight includes the sw2 main interview weight
- Sweeps 3 to 6
  - Two weights because two ‘samples’:
    - Those who responded at all sweeps
    - Those who responded at the individual sweep but missed an intervening sweep



ScotCen Social Research