



Leading education  
and social research  
Institute of Education  
University of London

# National Child Development Study

Sample of essays written at age 11 archived to  
accompany biographical interviews at age 50

Jane Elliott

User guide to the data (First Edition)

May 2013



**Centre for Longitudinal Studies**

Following lives from birth through the adult years  
[www.cls.ioe.ac.uk](http://www.cls.ioe.ac.uk)

CLS is an ESRC Resource Centre based at the Institute of Education, London



First published in 2013 by the  
Centre for Longitudinal Studies  
Institute of Education, University of London  
20 Bedford Way  
London WC1H 0AL  
[www.cls.ioe.ac.uk](http://www.cls.ioe.ac.uk)  
© Centre for Longitudinal Studies

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills. The views expressed in this work are those of the author(s) (amend as necessary) and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the author(s).

---

This document is available in alternative formats.  
Please contact the Centre for Longitudinal Studies.  
tel: +44 (0)20 7612 6875  
email: [clsfeedback@ioe.ac.uk](mailto:clsfeedback@ioe.ac.uk)

## Table of Contents

ACKNOWLEDGEMENTS.....	iii
PREFACE .....	iv
NCDS Essays Deposit: Elements .....	iv
INTRODUCTION.....	1
<i>NCDS</i> .....	1
BACKGROUND TO THE ORIGINAL DATA COLLECTION .....	2
<i>The context and timing of essay completion</i> .....	2
<i>Previous research using the essays</i> .....	3
THE SAMPLE OF ESSAYS.....	3
<i>Table 1: Essay Status</i> .....	3
<i>Table 2: Essay Status by coding at Age 11</i> .....	4
SUBSET OF QUANTITATIVE DATA.....	5
<i>Variables</i> .....	5
<i>Identifiers</i> .....	7
Appendix 1: Syntax for Derived Variables.....	8
<i>FAMBAK</i> .....	8
<i>ABIL11</i> .....	9
BIBLIOGRAPHY.....	10

## **ACKNOWLEDGEMENTS**

We would like to thank a number of people for their contribution to this deposit of a sample of essays written in 1969. Firstly, we would like to thank all the cohort members and their teachers, who generously gave their time to participate in this project and without whom this research resource would not now be available.

We wish to acknowledge the support for this deposit of our principal funders: the Economic and Social Research Council.

## **PREFACE**

This document has been prepared to accompany the deposit, with the UK Data Archive at the University of Essex, of a sub sample of the essays written at the age of eleven years by cohort members from the National Child Development Study (NCDS), a continuing, multidisciplinary, national, longitudinal study.

The elements of the deposit, to which reference will be made throughout this document, are identified below. Users are advised that they will need to consult all elements of the documentation to gain a full understanding of the data.

### **NCDS Essays Deposit: Elements**

<b>Title</b>	<b>Format</b>
NCDS Sample of 220 essays	Text
NCDS Sub-set of data to accompany 220 essays	SPSS
NCDS Sample of 220 essays: Guide to the Dataset	PDF

It is important to note that an additional sample of approximately 500 transcribed essays is deposited as study number 5790. Documentation for this separate sample of essays from a stratified sample is available with the deposit and a CLS working paper describing the key themes within the essays is also available from the CLS website (Elliott and Morrow, 2007).

The focus of this deposit is the essays written by those cohort members interviewed in 2009/2010 as part of a qualitative sub-study of the cohort.

## INTRODUCTION

This document has been prepared to accompany the deposit, with the UK Data Archive at the University of Essex, of a sub-sample of essays written by the members of the National Child Development Study (NCDS) when they were eleven years old in 1969. It is important to note that an additional sample of approximately 500 transcribed essays is deposited as study number 5790. Documentation for this separate sample of essays from a stratified sample is available with the deposit and a CLS working paper describing the key themes within the essays is also available from the CLS website (Elliott and Morrow, 2007).

The focus of this new deposit is the essays written by those cohort members interviewed in 2009/2010 as part of a qualitative sub-study of the cohort. The intention is to provide this specific sample of essays so that further qualitative material is available to researchers, collected during childhood, from those who took part in the biographical interview study.

## NCDS

NCDS started life as the Perinatal Mortality Survey, which was designed to examine the social and obstetric factors associated with stillbirth and infant mortality. In the first survey, data were collected about the births and families of 17,638 babies born in Great Britain during one week in March 1958. Since then, there have been eight surveys gathering information from respondents living in England, Scotland and Wales<sup>1</sup>, in order to monitor their health, education, social and economic circumstances. These surveys were carried out in 1965 (age seven), 1969 (age eleven), 1974 (age sixteen), 1981 (age 23), 1991 (age 33), 1999/2000 (age 42), 2004/2005 (age 46) and 2008/2009 (age 50). As part of the 1991 survey, information was additionally collected on the children of one in three cohort members; this included assessments of the behaviour and cognitive development of around 5,000 co-resident children. There have also been surveys of sub-samples of the cohort, the most recent occurring in 1995 (age 37), when a 10% representative sub-sample was assessed for difficulties with basic skills. Finally, during 2002-2004, 9,340 NCDS cohort members participated in a bio-medical survey, carried out by qualified nurses; the bio-medical survey did not cover in detail any of the topics included in the 2004/2005 survey.

Data for NCDS have so far been collected from a number of different sources; the midwife present at birth, the cohort members' parents, the head and class teachers, school health service personnel, the cohort members themselves, their spouses, cohabitantes and children, and the 1971 and 1981 censuses. Data has also been collected using a variety of methods; paper and electronic questionnaires and self-completion questionnaires, clinical records, medical examinations, physical measurements, ability tests, and educational assessments.<sup>2</sup>

---

<sup>1</sup> Including the Channel Islands, Isle of Man and other offshore islands.

<sup>2</sup> For further information on NCDS sweeps see: Power, C. and Elliott, J. (2005) Cohort Profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*, 2005, Information can also be found on the CLS website <http://www.cls.ioe.ac.uk>.

## **BACKGROUND TO THE ORIGINAL DATA COLLECTION**

When the children of the NCDS were eleven years old they were asked to complete assessments of reading, arithmetic and general ability and to answer a number of questions about their interests outside school, the school subjects they enjoyed most, and what they thought they were most likely to do when they left secondary school. In addition, they were asked to write an essay about what they thought their life would be like at age 25. The instructions given were as follows:

‘Imagine you are now 25 years old. Write about the life you are leading, your interests, your home life and your work at the age of 25. (You have 30 minutes to do this).’

Of the 14,757 children who participated in the age 11 sweep of the NCDS (representing 90.8% of the target sample of 16,253 (Plewis *et al.* 2004), a total of 13,669 (92.6%) completed an essay about their imagined life at age 25.

### **The context and timing of essay completion**

Approximately 90% of the essays were written in April, May or June 1969 i.e. very soon after the cohort members’ eleventh birthdays. A very few essays were written in January February and March of 1969, and a handful of essays were written as late as November 1970. By the end of the school year, in July 1969, over 96% of the essays had been completed with fewer than 1% of essays written in 1970. This means that at the time they wrote the essays almost all of the cohort members were in their final year of primary school.

Of those children who wrote essays, 47.5% were at a state junior school. A further 45.6% were attending a combined junior and infant school. Only 0.8% of the children were at secondary school, and just 3.9% were attending an independent school. The remaining children were either at a school for all age groups (1.1%) or at a special school (1%).

The survey of NCDS cohort members at age 11 also collected a great deal of information about the school the child was attending and this provides additional contextual information about the environment in which the essay was written. In the sample as a whole (14,106 children) 82% were attending a school that had a male head teacher with just 18% of schools having a female head teacher. In addition it was reported that 55% of children had a male class teacher and 45% had a female class teacher. It is also interesting to note that the mean class size was 34.3 and the median class size was 36. In other words 50% of these eleven-year-old children were in classes with 36 or more pupils.

Copies of the original essays (i.e. in the child’s handwriting) have been stored on microfiche and archived at the Centre for Longitudinal Studies. Some preliminary coding of the occupational aspirations of cohort members was carried out and archived together with the other quantitative data collected at age 11. In 2007 a stratified sub-sample of 560 essays were transcribed, anonymised and deposited at the data archive (SN 5790). Some analysis of this subset has been undertaken focussing on gender within the essays (Elliott, 2010).

## Previous research using the essays

Discussion with the researchers who were responsible for the NCDS in 1969 (Professors Ron Davie and Peter Wedge) suggests that there were no specific or immediate plans to analyse the essays written by the cohort members at age 11. At the time, the priority was to analyse the rich quantitative data collected in this sweep of the study. In the late 1960s it was hoped that the NCDS would continue to follow the cohort members and study their development through into adulthood. With this in mind, the qualitative information about future aspirations, collected in the essays written at age 11, was seen as having potential interest in the future for comparison with the actual trajectories of individual cohort members.

During the 1970s some analysis of a sub-sample of the essays was carried out which focussed on the 'syntactic maturity' of the children in the study. A random sample of 521 of the total of 13669 essays were coded for composition length and for 'Mean terminal-unit length' (MTUL). This is not equivalent to sentence length but rather refers to a main clause with attached subordinate clauses and is therefore the shortest unit which it is grammatically allowable to punctuate as a sentence. This measure had been found to correlate strongly with age in previous research and was used as a simple indicator of development in syntactic control (Richardson *et al.* 1976).

## THE SAMPLE OF ESSAYS

The transcribed and anonymised essays that comprise this deposit were all written by individuals who were interviewed as part of a sub-study that focused on a selected sub-sample of 220 cohort members. The subsample was extracted from cohort members living in four areas of Great Britain: Scotland, the North West, the South East and Wales. Further details of the sample selected for the qualitative study are provided in a CLS working paper (Elliott, et al 2010) and also in the user guide that accompanies the qualitative interviews.. Of the 220 essays written by cohort members who had been interviewed. Table 1 below shows the breakdown of transcribed essays; 153 were fully transcribed, 26 were transcribed but a few words were illegible and 41 were missing an essay either because the questionnaire has been lost or the essay was not still attached. In total the deposit therefore includes 179 transcribed essays - corresponding to 81% of the sub-sample who were interviewed at age 50. The associated file with a subset of quantitative data indicates which essays are missing or totally illegible, and includes data for all 220 cases.

**Table 1: Essay Status**

	Frequency
Valid	
1 Coded	153
3 Partially illegible	26
4 Questionnaire – no essay	29
5 Missing questionnaire	12
Total	220



Two essays previously un-coded at age 11 were found and are included, 23 previously coded were not able to be found amongst the questionnaires currently available at CLS.

**Table 2: Essay Status by coding at Age 11**

	Coded	Partially illegible	Questionnaire - no essay	No questionnaire	Total
-1 NA	2	0	16	2	20
0 Professional etc	22	3	1	1	27
1 Teachers etc	13	2	1	1	17
2 Nurse	10	2	1	0	13
3 Other non-manual, SCII	15	3	2	1	21
4 Typists, clerical	24	0	3	4	31
5 Shop assistants	9	0	1	0	10
6 Junior non-man	2	1	1	0	4
7 Personal service	8	0	0	1	9
9 Skilled manual	16	6	2	1	25
10 Semi-skilled man	5	1	0	0	6
11 Unskilled manual	2	1	0	0	3
12 Self employed	1	1	0	0	2
14 Housewife	3	1	1	0	5
15 Forces	6	1	0	0	7
16 Sportsmen, women	8	2	0	1	11
17 Students	2	0	0	0	2
18 Other work-n.e.c	4	2	0	0	6
19 No mention work	1	0	0	0	1
Total	153	26	29	12	220

## SUBSET OF QUANTITATIVE DATA

The quantitative data to accompany this sample of essays from NCDS has been supplied to the UK Data Archive in the form of an SPSS dataset which holds 25 variables for 220 cases.

### Variables

Variables included on the dataset are identified below. Users are referred to the documentation associated with the original age 11 survey of cohort members in 1969 for a more complete description of the variables extracted from the original age 11 data. Note that only a small subset of variables from the age 11 data, that are considered to be of direct relevance to the essay task, are included here. Note that variable n958 provides a code for the main occupation mentioned by the child in their essay written at age 11.

Variable name	Variable label	Origin
Project ID	Project identifier to link to interview	Original age 11 data
ESSPRES	Essay status	Availability & quality of essay for transcription
n622c	0-3D Sex of child	Original age 11 data
n802	2S Standard region	Original age 11 data
n810	2S Date on Schools Questionnaire-month	Original age 11 data
n811	2S Date on Schools Questionnaire-year	Original age 11 data
n812	2P Sex of school head or principal	Original age 11 data
n813	2S Sex of childs class teacher	Original age 11 data
n910	2T Date on Test Booklet-month	Original age 11 data
n911	2T Date on Test Booklet-year	Original age 11 data
n914	2T Verbal score on general ability test	Original age 11 data
n917	2T Non verbal score on gen ability test	Original age 11 data

n920	2T Total score on general ability test	Original age 11 data
n923	2T Reading comprehension test score	Original age 11 data
n926	2T Mathematics test score	Original age 11 data
n929	2T Copying designs test score	Original age 11 data
n1122	2P Child's mother figure	Original age 11 data
n1123	2P Why child away from own-adoptive mum	Original age 11 data
n1127	2P Child's father figure	Original age 11 data
n1128	2P Why child away from own-adoptive dad	Original age 11 data
n1171	2P Social Class of father or male head	Original age 11 data
n1685	2PD Social class father,male head-groupd	Original age 11 data
n958	2T Pupil's occupation at 25	Original age 11 data
abil11	General ability age 11	Derived variable
fambak	Family background variable used for essays analysis	Derived variable

## Identifiers

ProjectID is the unique identified used to label these essays to allow researchers to match them with the interview transcripts.

**NB:** Users can help improve the quality of the data by reporting data problems they encounter via the CLS website (<http://www.cls.ioe.ac.uk/>).

Users who **register** on the CLS website at ([www.cls.ioe.ac.uk/register](http://www.cls.ioe.ac.uk/register)) will receive email updates on data updates and deposits, as well CLS news, events and publications.

## Appendix 1: Syntax for Derived Variables

### FAMBAK

The variable FAMBAK is used to indicate the family background of the child *within* the subsample of essays. The three main variables used to define the variable 'fambak' were n1122, n1685 and n1127. Frequency tables for all these variables are provided below. Note that the variable FAMBAK will have missing values for a number of different family types. For example, where there is no natural mother or where the natural mother is living with the cohort member's stepfather.

```
*** only create the family background variable if child has a
*** natural mother
do if n1122=1.
if n1685=3 and n1127 ge 5 fambak=3.
if n1685=1 and n1127=1 fambak=1.
if n1685=2 and n1127=1 fambak=2.
end if.
variable labels fambak 'Family background variable used for
essays analysis'.
value labels fambak 1 'Nat Mum & Dad non-manual'
                   2 'Nat Mum & Dad manual'
                   3 'Nat Mum & no father figure'.
```

## ABIL11

The variable Abil11 indicates whether each child is in the top 25%, middle 50% or bottom 25% of the ability distribution as measured using the variable N920 'Total Score on General Ability test'.

The syntax used to create the variable is shown in the box below. Note that the 'Ntiles' subcommand is used to obtain the quartiles for the whole distribution of the variable N920 across the whole sample at age 11.

```
***examining the ability scores

freq vars n920
  /format=notable
  /NTILES= 4
  /STATISTICS=STDDEV MEAN MEDIAN
  /histogram
  /ORDER= ANALYSIS .

*** creating a variable indicating whether child is in the top
*** 25% bottom 25% or middle 50% using n920 Total score on
*** general ability test.
recode n920 (0 thru 30=1) (31 thru 55=2) (56 thru hi=3)
  (else=copy) into abil11.
missing values abil11 (-1).
freq vars=abil11.
crosstabs tables= n622c by abil11
  /cel=row cou.

value labels abil11 1 'low' 2 'medium' 3 'high'.
variable labels abil11 'General ability age 11'.
```

## BIBLIOGRAPHY

ELLIOTT, J. and MORROW, V. (2007) [\*Imagining the Future: Preliminary analysis of NCDS essays written by NCDS children at age 11.\*](#) CLS Working Paper 2007/1. London: Centre for Longitudinal Studies

Elliott, J. (2010). Imagining a Gendered Future: Children's Essays from the National Child Development Study in 1969. [\*Sociology\*](#), 44(6): 1073-1090.

ELLIOTT, J, MILES, A, PARSONS, S and SAVAGE, M. (2010) [\*The design and content of the 'Social participation' study: A qualitative sub-study conducted as part of the age 50 \(2008\) sweep of the National Child Development Study.\*](#) CLS Working Paper 2010/3. London: Centre for Longitudinal Studies.

RICHARDSON, K, CALNAN, M, ESSEN, J and LAMBERT, L. (1976) The linguistic maturity of 11 year olds. *Journal of Child Language*, 3(1), 95-115





Centre for Longitudinal Studies  
Institute of Education  
20 Bedford Way  
London WC1H 0AL  
Tel: 020 7612 6860  
Fax: 020 7612 6880  
Email: [clsfeedback@ioe.ac.uk](mailto:clsfeedback@ioe.ac.uk)  
Web: [www.cls.ioe.ac.uk](http://www.cls.ioe.ac.uk)