National Child Development Study

Sample of essays (Age 11)

Deposited with the UK Data Archive

A Guide to the Dataset

October 2007

Jane Elliott

Centre for Longitudinal Studies

Institute of Education, University of London

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

NCDS Essays Project Team

The work of extracting, transcribing and documenting the essays was carried out by the following team based at CLS

CLS Team

Research design sample extraction and coding	Jane Elliott , Virginnia Morrow
Extracting sample of essays:	Denise Brown; Kevin Dodwell, George Andrew, Mary Ukah,
Transcription and coding of essays	Jenny Neuburger; Tashweka Anderson

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 and we would also like to acknowledge the support of the Nuffield Foundation who funded the transcription and coding of this set of essays.

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 essays written 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

Title	**Format**
NCDS Sample of essays	txt
NCDS Sub-set of data to accompany essays	SPSS
NCDS Sample of essays (age 11): Guide to the Dataset	PDF

CONTENTS

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 collected when cohort members were eleven years old in 1969.

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 seven surveys gathering information from respondents living in England, Scotland and Wales[1], 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) and 2004/2005 (age 46). 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. The 2004/2005 follow-up aimed to extend the data collection of the previous surveys. This latest wave of the NCDS was conducted for the first time as a telephone interview (CATI)[2] when the cohort members were 46 years of age.

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, cohabitees 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.[3]

[1] Including the Channel Islands, Isle of Man and other offshore islands.
[2] CATI stands for Computer Assisted Telephone Interview
[3] 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 given a short questionnaire to complete at school 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.

To date, these essays have not been fully coded or analysed for use with the other data collected in the study. 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, but this is the only aspect of the rich material contained in the essays that has been available to researchers.

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 (O'Donnell *et al*. 1967; Richardson *et al.* 1976*,* Richardson 1979). These analyses revealed that although the girls wrote significantly longer essays than the boys (with a mean of 228.8 words compared with 180.7 words) the slight difference in MTUL was not statistically significant. Richardson *et al.* suggested that this finding 'reflects the well-known 'conformism' of young girls relative to boys' in other words that social factors are a more likely explanation than differences in 'neurobiological maturation processes' (Richardson *et al*. 1976: 347). In a review of research on sex differences in specific abilities, Gipps and Murphy (1996) conclude that during the school years from 4-11 the two sexes perform very similarly on verbal tasks but from adolescence onwards there is evidence of female superiority in a variety of verbal abilities. The longer essays written by girls at age 11 could therefore be interpreted as evidence of girls' greater facility with language at this age. The research by Richardson et al also found that social class was significantly associated with composition length such that children from social class 1 produced essays that were on average about 25% longer than those from children from social class V, but once again there were no significant differences in MTUL suggesting no apparent differences in syntactic maturity between the social class groups.

THE SAMPLE OF ESSAYS

Rather than attempting to transcribe and code the whole sample of over 13,000 essays written by cohort members in 1969, a subsample of essays has been initially extracted for deposit. One approach would have been to sample these essays randomly, but in order to ensure sufficient numbers in subgroups to facilitate some simple comparisons to be made, a stratified sub-sample was extracted based on three key variables:

1) The gender of the cohort member
2) The social class and family background of the cohort member (see derived variable fambak) and
3) The ability of the cohort member (measured using a general ability test, similar to an IQ test, at age 11, see derived variable abil11)

The social class of the cohort member was classified according to the occupation of the father at age 11. This is recorded as variable N1685 ' **2PD Social class father,male head-groupd'**, in the set of data corresponding to the 1969 survey. Using this variable a variable with three categories was derived: children with fathers in non-manual occupations; children with fathers in manual occupations, and children living with lone mothers where there was no male head in the household 9see further details of the derived variable 'fambak' below).

At age 11 children in the NCDS completed a general ability test. Children's test results are recorded as variable n920 'Total score on general ability test' in the data set corresponding to the 1969 survey. These test scores are approximately normally distributed with a range of 0-80, a mean of 42.9, a median of 44 and an inter-quartile range of 31-56. For the purposes of constructing a stratified sub sample from the cohort the ability score was recoded to form three categories: children scoring in the lowest 25% (i.e. 31 or under); children scoring in the middle 50% (i.e. 32 -56) and children scoring in the top 25% (i.e. 57 or above).

The profile of this stratified sub-sample is provided in Table 1. The approach used to create this subsample, means that, although in the sample as a whole there is a statistically significant difference in the ability test scores of boys and girls (a mean of 41.8 for boys vs. 44.1 for girls) and in the ability test scores of those with fathers in non-manual, manual occupations and those in families with no male head (50.0 vs. 40.19 vs.39.46) in the stratified subsample of essays no such differences exist. In other words the sub-sample of essays has been drawn to reflect accurately the range of ability within the sample as a whole but to remove any relationship between ability and social class. This means, for example, that as can be seen from Table 1a the sampling fraction for children with *non-manual* fathers and low scores on the ability test is much higher (0.123) than the sampling fraction for children with *manual* fathers and low scores on the ability test (0.028). This reflects the fact that in the sample as a whole children from non-manual backgrounds tend to achieve higher scores on the ability test than those from manual backgrounds.

Table 1: Stratified sub-sample

		Gender of child						
		Male			Female			
		Social class of father			Social class of father			
		Non-manual	manual	No male head	Non-manual	manual	No male head	Total
General ability age 11	Low	30	30	10	30	30	10	140
	Medium	60	60	20	60	60	20	280
	High	30	30	10	30	30	10	140

	Total	120	120	40	120	120	40	560

Table 1a: Sampling fractions for the stratified sample

		Gender of child					
		Male			Female		
		Social class of father			Social class of father		
		Non-manual	manual	No male head	Non-manual	manual	No male head
General ability age 11	Low	0.123	0.028	0.102	0.190	0.034	0.122
	Medium	0.062	0.033	0.142	0.062	0.033	0.128
	High	0.040	0.045	0.227	0.036	0.043	0.141
	Total	0.061	0.034	0.141	0.061	0.035	0.129

Coding the essays

A draft coding frame was constructed and tested on a sub-sample of 60 essays. Some elements of the coding frame were shaped by the instructions given to the children to write about their 'interests, home life and work', while others emerged from recurrent topics introduced by the children themselves. The aim was to create a manageable coding frame with not too many 'variables' but which successfully captured most of the topics and themes that emerged repeatedly in the children's essays. The draft coding frame was subsequently modified slightly to reflect additional themes identified within the sub-sample of 50 essays, and further refined after discussions with our initial research assistants, (Jenny Neuberger and Tash Anderson). The sample of essays were then copy typed, each into a separate word document, as accurately as possible and retaining the children's original spellings, and coding of the main themes within the essays completed. Appendix A provides a summary list of variables and Appendix B provides the coding frame that was used. In the first stages of transcribing and coding the essays, research assistants checked each other's transcriptions and coding for accuracy, and to ensure that the coding frame was being interpreted consistently.

The coding frame was then used to record data on 28 variables for each essay. The aim was to use the coding frame to build up a database that would offer a detailed descriptive profile of the content of each essay, to identify (for example) how many of the essays included references to the cohort children's mother, father or siblings; how many discussed wanting to have children; and how many gave details about the skills they would use in their preferred occupation. Notes were also made of any exceptional essays. In effect, this process generated quantitative data that can be used both to provide an overall summary of the content of the essays, and also to explore any links between the characteristics of the children and the topics they chose to include in their essays. It is hoped this coding of the content of essays will help future researchers to identify sub-samples of essays that they would be interested in examining in more detail.

The achieved sample

A profile of the achieved sample of transcribed essays is provided in Table 2 below. Of the target sample of 560 essays, a total of 495 were transcribed. Of the 65 essays that it was not possible to transcribe, 11 were found to be completely illegible due to the quality of the microfiche copy; and 54 were missing from the microfiche archive held at CLS.

As can be seen from Table 2, the transcribed essays were distributed relatively evenly across the stratified groups and therefore the stratified sample does not appear to have been distorted by the relatively small number of essays that were not available to be transcribed. For example, the samples of 280 boys and 280 girls were reduced to 243 and 252 respectively, while the samples of children from families with non-manual fathers, manual fathers and lone mothers were reduced from 240, 240 and 80 to 208, 214, and 73 respectively.

Table 2: Profile of sample

		Gender of child						
		Male			Female			
		Social class of father			Social class of father			
		Non-manual	manual	No male head	Non-manual	manual	No male head	Total
General ability age 11	Low	28	24	9	23	27	10	121
	Medium	54	54	15	53	56	20	252
	High	24	26	9	26	27	10	122
	Total	106	104	33	102	110	40	495

SUBSET OF QUANTITATIVE DATA

The quantitative data to accompany the sample of essays from NCDS has been supplied to the UK Data Archive in the form of an SPSS dataset which holds 53 variables for 560 cases.

Variables

Variables included on the dataset are identified below. A fuller definition of the variables based on a qualitative coding of the transcribed essays is provided by the coding scheme used with the essays, which is reproduced as Appendix B 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. However, users can match the sample of essays to any of the data collected at age 11 using the serial number. Note that variable n958 provides a code for the main occupation mentioned by the child in their essay written at age 11. This variable is available for all 13669 cases who completed an essay.

Variable name	Variable label	Origin
serial	Serial number	Original age 11 data
Status	Availability of transcribed essay	Code based on essay
words	Number of words in essay	Code based on essay
mother	Is cohort member's mother mentioned?	Code based on essay
father	Is cohort member's father mentioned?	Code based on essay
siblings	Are cohort member's siblings mentioned?	Code based on essay
animals	Are animals or pets mentioned?	Code based on essay
othfam	Are other family members mentioned?	Code based on essay
friends	Are friends mentioned?	Code based on essay
livsit	Child's description of living situation	Code based on essay
Married	Child's statements about marriage	Code based on essay
kids	Number of children wanted	Code based on essay
occn	Number of occupations mentioned	Code based on essay
workch	Is work described as accommodating to children?	Code based on essay
Housewife	Being a housewife mentioned instead of occ	Code based on essay
partocc	Is partners occupation mentioned?	Code based on essay
occsk	Are skillsdiscussed in relation to occupation?	Code based on essay
hourswk	Are hours of work mentioned?	Code based on essay
occsoc	Are social elements of work mentioned?	Code based on essay
money	Are money, savings or earnings mentioned?	Code based on essay
leisure	Are leisure activities discussed?	Code based on essay
holidays	Are holidays discussed?	Code based on essay
housing	Nature of housing discussed?	Code based on essay
location	Where child is living at age 25	Code based on essay
domlab	Discussion of domestic labour	Code based on essay
congoods	Consumer goods mentioned (not car)	Code based on essay
car	Car specifically mentioned	Code based on essay
travel	Travel discussed apart from holidays	Code based on essay
football	Football discussed	Code based on essay
narrate	Does essay include narrative elements?	Code based on essay
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

NCDS cohort members have unique identifier, SERIAL, which appear on the datasets already lodged with the UK Data Archive. This identifier is also to be found on the new dataset and can be used to link the data longitudinally to earlier sweeps.

Data Dictionary from accompanying file

```
_ List of variables on the working file

Name (Position) Label

serial (1) Serial number
    Measurement Level: Nominal
    Column Width: 8  Alignment: Left
    Print Format: A7
    Write Format: A7
    Missing Values: ''

Status (2) Availability of transcribed essay
    Measurement Level: Nominal
    Column Width: 10  Alignment: Right
    Print Format: F15
    Write Format: F15

           Value    Label

              -1    awaiting coding
               1    coded
               2    illegible
               3    partially illegible
               4    missing essay
               5    missing fiche

words (3) Number of words in essay
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay

mother (4) Is cohort member's mother mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

_



father (5) Is cohort member's father mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9
```

```
           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

siblings (6) Are cohort member's siblings mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

animals (7) Are animals or pets mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

othfam (8) Are other family members mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

_



friends (9) Are friends mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes
```

livsit (10) Child's description of living situation
    Measurement Level: Nominal
    Column Width: 8 Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

Value		Label
-9	M	Missing or illegible essay
0		Not mentioned
1		Alone
2		Still living with parents
3		Married
4		Cohabiting
5		Friends
6		With other family members
7		Changes during essay
8		Other

Married (11) Child's statements about marriage
    Measurement Level: Nominal
    Column Width: 8 Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

Value		Label
-9	M	Missing or illegible essay
0		Not mentioned
1		Will get/be married
2		Doesn't want to be married
3		Not married
4		Other about marriage

_

kids (12) Number of children wanted
    Measurement Level: Nominal
    Column Width: 8 Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

Value		Label
-9	M	Missing or illegible essay
-1		Not mentioned
8		Vague
9		Other

occn (13) Number of occupations mentioned
    Measurement Level: Nominal
    Column Width: 8 Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

Value Label

```
              -9 M  Missing or illegible essay
               8    Vague
               9    Several mentioned as possibilities

workch (14) Is work described as accommodating to children?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

Housewife (15) Being a housewife mentioned instead of occ
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes
_



partocc (16) Is partners occupation mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Housewife
               2    Other occupation

occsk (17) Are skillsdiscussed in relation to occupation?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

hourswk (18) Are hours of work mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
```

```
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

occsoc (19) Are social elements of work mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes
_



money (20) Are money, savings or earnings mentioned?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

leisure (21) Are leisure activities discussed?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

holidays (22) Are holidays discussed?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
```

```
               1    Yes

housing (23) Nature of housing discussed?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    Not mentioned
               1    Flat
               2    Other house
_


               3    Rural countryside cottage or farm
               4    Other

location (24) Where child is living at age 25
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    Not mentioned
               1     London
               2    UK
               3    Abroad
               4    More than one location

domlab (25) Discussion of domestic labour
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    Not mentioned
               1    Childcare
               2    Cooking cleaning etc
               3    Childcare & cooking cleaning
               4    DIY or gardening
               5    Multiple tasks
               6    Sewing, knitting, needlework
               7    Other

_


congoods (26) Consumer goods mentioned (not car)
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
```

```
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

car (27) Car specifically mentioned
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    Not mentioned
               1    As transport
               2    Specific make

travel (28) Travel discussed apart from holidays
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes

football (29) Football discussed
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
               0    No
               1    Yes
_



narrate (30) Does essay include narrative elements?
    Measurement Level: Nominal
    Column Width: 8  Alignment: Right
    Print Format: F11
    Write Format: F11
    Missing Values: -9

           Value    Label

              -9 M  Missing or illegible essay
```

```
              0    No
              1    Some narrative clauses
              2    Extended narrative

n622c (31) 0-3D Sex of child
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  Not known
               1    Male
               2    Female

n802 (32) 2S Standard region
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  No educational
               1    North Western
               2    Northern
               3    East & W.Riding
               4    North Midlands
               5    Eastern
               6    London & S.East
               7    Southern
               8    South Western
               9    Midlands
              10    Wales
              11    Scotland

_



n810 (33) 2S Date on Schools Questionnaire-month
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

               1    January
               2    February
               3    March
               4    April
               5    May
               6    June
               7    July
               8    August
               9    September
              10    October
              11    November
```

```
              12    December

n811 (34) 2S Date on Schools Questionnaire-year
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              69    1969
              70    1970
              71    1971

n812 (35) 2P Sex of school head or principal
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

               1    Male
               2    Female

_



n813 (36) 2S Sex of childs class teacher
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

               1    Male
               2    Female

n910 (37) 2T Date on Test Booklet-month
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

               1    January
               2    February
               3    March
               4    April
               5    May
               6    June
               7    July
               8    August
               9    September
              10    October
              11    November
```

```
              12    December

_


n911 (38) 2T Date on Test Booklet-year
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              69    1969
              70    1970
              71    1971

n914 (39) 2T Verbal score on general ability test
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA

n917 (40) 2T Non verbal score on gen ability test
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA

n920 (41) 2T Total score on general ability test
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA

_


n923 (42) 2T Reading comprehension test score
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2
```

```
           Value    Label

              -1 M  NA

n926 (43) 2T Mathematics test score
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA

n929 (44) 2T Copying designs test score
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA

n1122 (45) 2P Child's mother figure
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  NA
               1    Natural mother
               2    Adoptive mother
               3    Stepma,inc cohab
               4    Foster mother
               5    Grandmother
               6    Elder sister
               7    No mother figure
               8    Other

_


n1123 (46) 2P Why child away from own-adoptive mum
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

           Value    Label

              -1 M  Own,adopt mum,NA
               9    Illigitimacy
              10    Broken marriage
              11    Death
              12    Other
```

n1127 (47) 2P Child's father figure
    Measurement Level: Ordinal
    Column Width: 8 Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

Value		Label
-1	M	NA
1		Natural father
2		Adoptive father
3		Steppa,inc cohab
4		Foster father
5		Grandfather
6		Elder brother
7		No father figure
8		Other

n1128 (48) 2P Why child away from own-adoptive dad
    Measurement Level: Ordinal
    Column Width: 8 Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

Value		Label
-1	M	Own,adopt dad,NA
9		Illigitimacy
10		Broken marriage
11		Death
12		Other

_

n1171 (49) 2P Social Class of father or male head
    Measurement Level: Ordinal
    Column Width: 8 Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

Value	Label
1	Social class I
2	Social class II
3	SC III non-man.
4	SC III manual
5	SC IV non-manual
6	SC IV manual
7	Social class V
8	Unclassifiable

n1685 (50) 2PD Social class father,male head-groupd
    Measurement Level: Ordinal
    Column Width: 8 Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

```
            Value    Label

               -1 M  NA,unclear
                1    Non manual
                2    Manual
                3    No male head

n958 (51) 2T Pupil's occupation at 25
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1, -2

            Value    Label

               -1 M  NA
                0    Professional etc
                1    Teachers etc
                2    Nurse
                3    Oth non-man,SCII
                4    Typists,clerical
                5    Shop assistants
                6    Junior non-man
                7    Personal service
                8    Foreman,manual
                9    Skilled manual
_


               10    Semi-skilled man
               11    Unskilled manual
               12    Self employed
               13    Farm workers
               14    Housewife
               15    Forces
               16    Sportsmen,women
               17    Students
               18    Other work-n.e.c
               19    No mention work
               20    Unclassifiable

abil11 (52) General ability age 11
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8.2
    Write Format: F8.2
    Missing Values: -1.00

            Value    Label

             1.00    low
             2.00    medium
             3.00    high

fambak (53) Family background variable used for essays analysis
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8.2
    Write Format: F8.2

            Value    Label
```

```
1.00    Nat Mum & Dad non-manual
2.00    Nat Mum & Dad manual
3.00    Nat Mum & no father figure
```

_

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) will receive email updates on data updates and deposits, as well CLS news, events and publications.

Syntax for Derived Variables

Fambak

The variable Fambak is used to indicate the family background of the child *within* the subsample of essays. Only those children still living with their natural mother were included in the sample of essays i.e. variable n1122=1. 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. The subsample of essays is selected from only those cases where the child is living with his or her natural mother and father (Fambak=1 or fambak=2) or where the child is living with his or her natural mother and there is no male head.

```
***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'.
```

n1122 2P Child's mother figure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Natural mother	13349	71.9	96.3	96.3
	2 Adoptive mother	133	.7	1.0	97.3
	3 Stepma,inc cohab	85	.5	.6	97.9
	4 Foster mother	39	.2	.3	98.2
	5 Grandmother	78	.4	.6	98.7
	6 Elder sister	20	.1	.1	98.9
	7 No mother figure	102	.5	.7	99.6
	8 Other	55	.3	.4	100.0
	Total	13861	74.7	100.0	
Missing	-1 NA	4697	25.3		
Total		18558	100.0		

n1127 2P Child's father figure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Natural father	12518	67.5	90.3	90.3
	2 Adoptive father	188	1.0	1.4	91.7
	3 Steppa,inc cohab	314	1.7	2.3	93.9
	4 Foster father	33	.2	.2	94.2
	5 Grandfather	73	.4	.5	94.7
	6 Elder brother	9	.0	.1	94.8
	7 No father figure	678	3.7	4.9	99.7
	8 Other	48	.3	.3	100.0
	Total	13861	74.7	100.0	
Missing	-1 NA	4697	25.3		
Total		18558	100.0		

n1685 2PD Social class father,male head-groupd

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Non manual	4559	24.6	33.5	33.5
	2 Manual	8332	44.9	61.3	94.8
	3 No male head	702	3.8	5.2	100.0
	Total	13593	73.2	100.0	
Missing	-1 NA,unclear	4965	26.8		
Total		18558	100.0		

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.

```
GET
  FILE='C:\Documents and Settings\Jane\My Documents\SPSS Analysis\Cohort data
files\ncds2.sav'.

****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'.
```

Appendix A: List of variables describing the content of the essays

serial	Serial number to link the essay to the quantitative record
Words	Number of words in the essay
Mother	Whether the cohort member's mother is mentioned
Father	Whether the cohort member's father is mentioned
Siblings	Whether the cohort member's siblings are mentioned
Othfam	Whether other family members are mentioned
Animals	Whether animals or pets are mentioned
Friends	Whether friends are mentioned
Livsit	Description of household at 25 (see coding scheme below for further details)
Married	Intentions for marriage by age 25 (see coding scheme below for further details)
Kids	Intentions for children by age 25 (see coding scheme below for further details)
Occn	Number of occupations mentioned
housewife	Whether cohort member explicitly mentions being a housewife
partocc	Partner's occupation mentioned
Occsk	Skills and activities described in relation to paid work (see coding scheme below for further details)
hourswk	Hours of work discussed (see coding scheme below for further details)
Occsoc	Social aspects of work mentioned (see coding scheme below for further details)
Money	Money, savings or earnings mentioned in the essay
Leisure	Leisure activities mentioned in the essay
Holidays	Holidays mentioned in the essay
Housing	Housing mentioned in the essay (see coding scheme below for further details)
Location	Does the child mention where they will be living at age 25? (see coding scheme below for further details)
domlab	Does the child mention domestic labour in the essay & if so what type (see coding scheme below for further details)
Car	Is a car mentioned in the essay (see coding scheme below for further details)
Travel	Is travel (apart from holidays) mentioned in the essay?
Football	Is football mentioned in the essay
Workch	Does the essay mention how paid work/career will need to accommodate responsibilities for children?
Narrate	Does the essay include narrative elements?
Extra	Is there some aspect of the essay that makes it particularly unusual or interesting?

Appendix B: Instructions for coders and coding frame for descriptive analysis of the essays

Instructions for coders

1) Type the essay as exactly as possible including spelling mistakes and aiming to mimic the original punctuation. If part of the essay is illegible write [illegible] to cover that section. There is no need to preserve the way the essay is set out in terms of words per line. Make sure that the serial number is typed at the top of the essay so that it can be matched with other data. It is essential that the serial number is correct and that no numbers are transposed so please double check this.
   Using the word count facility on word count the number of words in the essay and type this at the end of the essay.
2) Code the essay for the main themes it includes using the coding frame provided and then enter the codes onto the excel spreadsheet provided. If you are unsure about a code then leave the cell blank.

Coding frame (detailed) revised 17th May [Final version]

The majority of codes in the coding frame are dichotomous i.e. either the child has mentioned something in the essay or not. In all these cases 0 is used to indicate no mention and 1 to indicate that the topic has been mentioned.

Serial: This is the NCDS serial number, great accuracy is required

Words: number of words in essay (use word count in Word to obtain this)

Mother: whether cohort member's mother is mentioned (0= no mention,1= mother mentioned),
Note that if parents are mentioned this should be included as a reference to mother

Father: whether cohort member's father is mentioned (0,1),
Note that if parents are mentioned this should be included as a reference to father

Siblings: whether siblings mentioned (0,1)

Othfam: whether other family mentioned (0,1),
(e.g. 'I live with my Uncle')

Animals: whether pets or animals mentioned (0,1),

Friends, whether friends mentioned (0,1),
e.g. could be going out with friends, going on holiday with friends, living with friends

Livsit: this is a summary of the child's anticipated living situation at age 25 the following codes should be used:

not mentioned, 0
alone, 1
still living with parents, 2
married, 3
cohabiting, 4
friends, 5
other family, 6
changes during essay (e.g. gets married), 7
other, 8

Note that if the child describes a change during the essay for example being single and then getting married this should be included as code 7

married,

not mentioned, 0
will get/be married, 1
doesn't want to be married, 2

Not married 3
other about marriage, 4

Note that the main aim of this is to identify the children who expect to be married, those who categorically say no they don't ever want to marry and those who simply state they are not married.

kids, how many wanted e.g. -1=not mentioned, 0, 1,2, 3 etc 8=vague; 9=other (e.g. foster child),
Note that if children are not mentioned, this would be coded as -1. Children who say something fairly vague like 'I'll have some kids' should be coded as 8 to indicate that exact numbers of children are not specified.

Occn: Number of occupations specified 0,1,2 etc 8 = vague (i.e. I will have a job , but occ not specified)
9=several mentioned as possibilities
Note that several children seem to mention doing more than one job e.g. 'I'm going to be a footballer and an electrician'. This would be coded as 2. However if the child lists lots of different occupations that they might consider doing then code this as 9.

workch: work accommodating to care of children (0,1),
In some cases, girls in particular talk about working and then giving up to have children or having children and then working once the children are older, both of these should be captured by the workch code.

Housewife Whether child expects to care for home and family (instead of ,or as well as having an occupation) (0,1)

partocc, Is partner's occupation mentioned(include housewife as an occupation), e.g.
0 = not mentioned
1=housewife
2=other occupation
If the child mentions what they expect their husband or wife's occupation to be then code this as 2, except if the occupation mentioned is housewife, which would be coded as 1. e.g. 'I am a vet and so is my husband'

Occsk, skills discussed in relation to occupation (0,1),
Use this to indicate that the child has given some level of detail about what skills or activities are involved in the work they intend to do.

hourswk, Hours worked discussed (0,1)
if there is any mention of the timing of work then use this code e.g. 'I will work part-time' or 'I will work from 7am to 1pm on the farm'

Occsoc, Social elements of work mentioned (0,1),
If the child mentions social relations at work e.g. working with friends, getting told off by the boss, enjoying talking to other teachers etc the n code this as 1.

Money, money, savings, earnings specifically mentioned (0,1),
Use this if wages or other finance is specifically mentioned e.g. 'I will have saved £600' or 'I earn good wages'

Leisure, leisure activities discussed, (0,1)
This covers a very wide range of leisure activities e.g. going to the pub, watching TV, reading, swimming, going for day trips, visiting relatives.

holidays, holidays specifically mentioned, (0,1)
This should only be used for explicit holidays and not for days out or traveling for work.

housing, nature of housing discussed e.g. flat/house/quiet/in the country,
not mentioned: 0
flat: 1
other house: 2
rural/countryside/cottage in the country/farm: 3
other : 4
If the child mentions that they live in a house or a bungalow, for example, this would be coded 2

location, location mentioned
not mentioned: 0

London: 1
Other U.K: 2
overseas: 3
more than one location mentioned 4

e.g. If the child mentions that they live in Denton, this would be coded 2; If they say Spain, this would be 3.

domlab, domestic labour discussed whether done by self or other,
, not mentioned, 0
, child care/babysitting, 1
, cooking cleaning etc, 2
, child care and cooking/cleaning, 3
DIY and gardening, 4
Multiple tasks, 5
Sewing, knitting, needlework 6
Other 7

Note use this if any domestic labour is discussed e.g. my wife will look after the children, or we will have a maid to clear up for us.

Car: car specifically mentioned,
not mentioned, 0
car as means of transport, 1
specific type of car (as possession), 2

travel: travel (apart from holidays) e.g. for work (0,1),

football: is football discussed (0,1)
many of the boys either want to be footballers or talk about being footballers in their leisure time

Narrate: does essay include narrative elements,
no narrative, 0
some narrative clauses (this happened then this happened), 1
complex or extended narrative, 2

Note that many essays are pure descriptions with no sense of the passing of time or one thing happening after another – these should be coded as 0.
Other essays have some narrative elements e.g. relating what will happen in a typical day, one activity after another these use the code 1
In a few cases children have used the essay to tell an extended story about what will happen in their life these should use the full narrative code 2.

Extra: A very few essays (perhaps one in 20 stand out as being very different from the rest e.g. they have interpreted the task in an interesting and unusual way. Please use this code to identify essays that we may well want to look at again because they are so unusual.

Not Answered (-1)
Does Not Apply (-2)

7. Imagine that 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). N958

Do not write here.

Col. 51

Col. 52

Col. 53

Col. 54

Col. 55

Col. 56