Centre for Longitudinal Studies, Institute of Education

Teaching students quantitative methods using resources from the British Birth Cohorts

Assessment of Cognitive Development through Childhood

CognitiveAssessmentVariables.doc: this document includes:

a) A list of all the variables in the *n716reading.sav, n716maths.sav* and the *b516reading.sav SPSS* data files. (Original variable names from the full dataset and documentation of a particular survey are also included.)

b) *n716reading.sav* contains 14 variables in total: identifier, sex, the three reading assessment scores, standardised assessment scores at each age, exact age of the child at the assessment, estimated reading age for each age based on performance and real age.

c) *n716maths.sav* contains 14 variables in total: identifier, sex, the three mathematics assessment scores, standardised assessment scores at each age, exact age of the child at the assessment, estimated mathematics age for each age based on performance and real age.

d) *b516reading.sav* contains 14 variables in total: identifier, sex, the three reading assessment scores, standardised assessment scores at each age, exact age of the child at the assessment, estimated reading age for each age based on performance and real age.

e) SPSS syntax for constructing the estimated reading and mathematics ages for cohort members, and BCS70 age 10 reading score.

NCDS Reading Assessment data file: *'n716reading.sav'*

The data

This longitudinal dataset includes information from the 1958 National Child Development Study (NCDS) for looking at cognitive development, or more specifically reading, over childhood. The data file contains 14 variables, based around summary reading scores derived from performance of NCDS cohort members in assessments at age 7, 11 and 16 (1965, 1969 and 1974). The dataset only includes those who had 'age at test' information at age 7.

There are 14,983 cohort members (individuals), 7,689 (51.3%) are men and 7,294 (48.7%) are women. To enable the information included here to be cross referenced with the original questionnaire documentation, original variable names are included in the variable label. Table 1 contains a list of the variables in the *n716reading.sav* dataset. In 1965 when cohort members were age 7, of the 14,983 with 'age at test' information, 14,872 completed the Southgate Reading Test[1], 12,827 completed the NfER constructed Reading Comprehension Test when age 11[2] and 10,621 completed the same NfER Reading Comprehension Test when they were age 16 in 1974.

This file can be merged with the *n042KeyVariables.sav* which includes socio-economic characteristics of the cohort members' families at the time the assessments were completed and also some characteristics of the cohort members themselves in adulthood such as employment status and highest qualification at age 42.

[1] Southgate, V (1962) Southgate Group Reading Tests: Manual of Instructions. University of London Press

[2] The assessment used at age 11 and age 16 has not been published.

NCDS Mathematics Assessment data file: *'n716maths.sav'*

The data

This longitudinal dataset includes information from the 1958 National Child Development Study (NCDS) for looking at cognitive development, or more specifically mathematics, over childhood. The data file contains 14 variables, based around summary mathematics scores derived from performance of NCDS cohort members in assessments at age 7, 11 and 16 (1965, 1969 and 1974). The dataset only includes those who had 'age at test' information at age 7.

There are 14,983 cohort members (individuals), 7,689 (51.3%) are men and 7,294 (48.7%) are women. To enable the information included here to be cross referenced with the original questionnaire documentation, original variable names are included in the variable label. Table 2 contains a list of the variables in the *n716maths.sav* dataset. In 1965 when cohort members were age 7, of the 14,983 with 'age at test' information, 14,853 completed the Problem Arithmetic Test[3], 12,822 completed the NfER constructed Arithmetic/Mathematics Test[4] when age 11 and 10,566 completed the University of Manchester constructed Mathematics Test when they were age 16 in 1974.

This file can be merged with the *n042KeyVariables.sav* which includes socio-economic characteristics of the cohort members' families at the time the assessments were completed and also some characteristics of the cohort members themselves in adulthood such as employment status and highest qualification at age 42.

[3] Pringle, M.K, Butler, N, and Davie, R (1966) 11,000 Seven Year Olds. Longman, in association with National Children's Bureau

[4] The assessments at age 11 and age 16 have not been published.

BCS70 Reading Assessment data file: *'b516reading.sav'*

The data

This longitudinal dataset includes information from the 1970 British Cohort Study (BCS70) for looking at cognitive development, or more specifically reading, over childhood. The data file contains 15 variables, based around summary reading scores derived from performance of BCS70 cohort members' in assessments at age 5, 10 and 16 (1975, 1980 and 1986). The dataset only includes those who had 'age at test' information at age 5.

There are 12818 cohort members in the data file in total, 6631 (51.7%) are male and 6187 are female. To enable the information included here to be cross referenced with the original questionnaire documentation, original variable names are included in the variable label. Table 3 contains a list of the variables in the *b516reading.sav* dataset. In 1975 when cohort members were age 5, of the 12,818 with 'age at test' information, 10,098 completed the English Picture Vocabulary Test (EPVT)[5], 9775 completed the Shortened Edinburgh Reading Test[6] when age 10 and 4720 completed the Vocabulary Test[7] when they were age 16 in 1986.

This file can be merged with the *b034KeyVariables.sav* which includes socio-economic characteristics of the cohort members' families at the time the assessments were completed and also some characteristics of the cohort members themselves in adulthood such as employment status and highest qualification at age 34.

[5] Brimer, M. A., & Dunn, L. M. (1962). English Picture Vocabulary Test. Bristol: Education Evaluation Enterprises.
[6] Hodder and Stoughton (1979). The Shortened Edinburgh Reading Test. Kent: Hodder and Stoughton Educational Ltd
[7] APU Vocabulary Test (multiple choice format, 1986). Kent: Hodder and Stoughton Educational Ltd

Table 1: Variables in 'n716reading.sav' data file

Name	Label
serial	unique serial number of cohort member
n0sex	ncds age 0: sex of cohort member (n622)
n7read	ncds age 7: raw Southgate Reading Test score (n92)
n7age	ncds age 7: age of cohort member when sat assessments (derived from n16 n17)
zb7read	ncds age 7: standardised Southgate Reading Test score (n7read)
n7readage*	ncds age 7: estimated reading age at age 7 (derived from raw Southgate Reading Test score and age at testing)
n11read	ncds age 11: raw Reading Comprehension Test score(n923)
n11age	ncds age 11: age of cohort member when sat assessments (derived from n910 n911)
zn11read	ncds age 11: standardised Reading Comprehension Test score (n11read)
n11readage*	ncds age 11: estimated reading age at age 11 (derived from raw Reading Comprehension Test score and age at testing)
n16read	ncds age 16: raw Reading Comprehension Test score (n2928)
n16age	ncds age 16: age of cohort member when sat assessments (derived from n2925 n2927)
zn16read	ncds age 16: standardised Reading Comprehension Test score (n16read)
n16readage*	ncds age 16: estimated reading age at age 16 (derived from raw Reading Comprehension Test score and age at testing)

* for further information on how these variables were constructed see PLEWIS, I. (1996) Statistical methods for understanding cognitive growth: A review, a synthesis and an application. *British Journal of Mathematical and Statistical Psychology*, 49, 25-42

Table 2: Variables in 'n716maths.sav' data file

Name	Label
serial	unique serial number of cohort member
n0sex	ncds age 0: sex of cohort member (n622)
n7math	ncds age 7: raw Problem Arithmetic Test score (n90)
n7age	ncds age 7: age of cohort member when sat assessments (derived from n16, n17)
zb7math	ncds age 7: standardised Problem Arithmetic Test score (n7math)
n7mathage*	ncds age 7: estimated reading age at age 7 (derived from raw Problem Arithmetic Test score and age at testing)
n11math	ncds age 11: raw Arithmetic/Mathematics Test score(n923)
n11age	ncds age 11: age of cohort member when sat assessments (derived from n910 n911)
zn11math	ncds age 11: standardised Arithmetic/Mathematics Test score (n11math)
n11mathage*	ncds age 11: estimated reading age at age 11 (derived from raw Arithmetic/Mathematics Test score and age at testing)
n16math	ncds age 16: raw Mathematics Test score (n2930)
n16age	ncds age 16: age of cohort member when sat assessments (derived from n2925 n2927)
zn16math	ncds age 16: standardised Mathematics Test score (n16math)
n16mathage*	ncds age 16: estimated reading age at age 16 (derived from raw Mathematics Test score and age at testing)

* for further information on how these variables were constructed see PLEWIS, I. (1996) Statistical methods for understanding cognitive growth: A review, a synthesis and an application. *British Journal of Mathematical and Statistical Psychology*, 49, 25-42

Table 3: Variables in 'b516reading.sav' data file

Name	Label
key	unique case identifier of cohort member
bsex	bcs70 age 0 - 34: sex of cohort member (derived from variables in childhood a0255 d003 sex10)
b5read	bcs70 age 5: raw English Picture Vocabulary Test score (EPVT) (f117)
b5age	bcs70 age 5: age of cohort member when sat assessments (derived from f112)
zb5read	bcs70 age 5: standardised English Picture Vocabulary Test score (EPVT) (b5read)
b5readage*	bcs70 age 5: estimated reading age at age 5 (derived from zb5read and average age)
b10read	bcs70 age 10: raw Edinburgh Reading Test score (i3017 to i3069)
b10age	bcs70 age 10: age of cohort member when sat assessments (derived from i2503, i2503m, i2503y)
zb10read	bcs70 age 10: standardised Edinburgh Reading Test score (b10read)
b10readage*	bcs70 age 10: estimated reading age at age 10 (derived from raw Edinburgh Reading Test score and age at testing)
b16read	bcs70 age 16: raw Vocabulary Test score (cv01 to cv075)
b16age	bcs70 age 16: age of cohort member when sat assessments (derived from fdoc_mt fdoc_yr)
zb16read	bcs70 age 16: standardised Vocabulary Test score (b16read)
b16readage*	bcs70 age 16: estimated reading age at age 16 (derived from raw Vocabulary Test score and age at testing)

* for further information on how these variables were constructed see PLEWIS, I. (1996) Statistical methods for understanding cognitive growth: A review, a synthesis and an application. *British Journal of Mathematical and Statistical Psychology*, 49, 25-42

SPSS syntax for constructing estimated reading and mathematics ages in NCDS

```
COMPUTE n7readage = zn7read+7.
COMPUTE n7mathage = zn7maths+7.
EXECUTE .

COMPUTE n11readage = (1.57*zn11read) + 11.
COMPUTE n11mathage = (1.57*zn11maths) + 11.
EXECUTE .

COMPUTE n16readage = (2.29*zn16read) + 16.
COMPUTE n16mathage = (2.29*zn16maths) + 16.
EXECUTE .
```

SPSS syntax for constructing estimated reading ages in BCS70

```
COMPUTE b5readage = zb5read + 5.
COMPUTE b10readage = (1.57*zb10read) + 10.
COMPUTE b16readage = (2.29*zb16read) + 16.
EXECUTE .
```

SPSS syntax for constructing summary Reading score at age 10

BCS70 variables age 10 (1980)

```
THE SHORTENED EDINBURGH READING TEST.

FREQUENCIES
  VARIABLES=i3003 i3004 i3005 i3006 i3007 i3008 i3009 i3010 i3011 i3012 i3013
  i3014 i3015 i3016 i3017 i3018 i3019 i3020 i3021 i3022 i3023 i3024 i3025
  i3026 i3027 i3028 i3029 i3030 i3031 i3032 i3033 i3034 i3036 i3037 i3038
  i3039 i3040 i3041 i3042 i3043 i3044 i3045 i3046 i3047 i3048 i3049 i3050
  i3051 i3052 i3053 i3054 i3055 i3056 i3057 i3058 i3059 i3060 i3061 i3062
  i3063 i3064 i3065 i3066 i3067 i3068 i3069.


** RECODE ALL QUESTIONS WHEN FIRST ANSWER WAS THE CORRECT ANSWER.

RECODE
    i3004  i3005  i3012  i3015  i3017  i3018  i3019 i3020 i3021 i3024 i3031 i3036 i3037 i3038
    i3039  i3040  i3031  i3041  i3042  i3043  i3047 i3056 i3065 i3066
  (1=1) (2 THRU HIGHEST=0) (MISSING=COPY)  .

** RECODE ALL QUESTIONS WHEN SECOND ANSWER WAS THE CORRECT ANSWER.

RECODE
    i3006  i3013  i3029  i3033  i3034  i3045 i3049 i3052 i3058 i3063 i3067
(1=0) (2=1) (3 THRU HIGHEST=0) (MISSING=COPY) .
```

```
** RECODE ALL QUESTIONS WHEN THIRD ANSWER WAS THE CORRECT ANSWER.

RECODE
  i3003  i3007 i3008  i3016  i3025 i3030 i3046 i3048 i3050 i3057 i3061 i3069
(1=0) (2=0) (3=1) (4 THRU HIGHEST=0) (MISSING=COPY) .

** RECODE ALL QUESTIONS WHEN FOURTH ANSWER WAS THE CORRECT ANSWER.

RECODE
   i3011  i3014 i3022  i3044 i3051 i3055 i3060 i3062 i3064 i3068
 (1 THRU 3=0)  (4=1) (5 THRU HIGHEST=0) (MISSING=COPY) .

** RECODE ALL QUESTIONS WHEN FIFTH ANSWER WAS THE CORRECT ANSWER.

RECODE
   i3027  i3032 i3059
 (1 THRU 4=0)  (5=1) (6 THRU HIGHEST=0) (MISSING=COPY) .

** RECODE ALL QUESTIONS WHEN SIXTH ANSWER WAS THE CORRECT ANSWER .

RECODE
   i3010   i3023  i3028
 (1 THRU 5=0)  (6=1)  (7=0) (MISSING=COPY) .

** RECODE ALL QUESTIONS WHEN SEVENTH ANSWER WAS THE CORRECT ANSWER .

RECODE
  i3009 i3026
 (1 THRU 6=0)  (7=1) (8 THRU HIGHEST=0)  (MISSING=COPY) .

** TWO QUESTIONS CODED DIFFERENTLY: COHORT MEMBERS HAD TO PUT 5
SENTENCES IN THE CORRECT ORDER .

RECODE
  i3053
 (35214=1) (ELSE=0) (MISSING=COPY).

RECODE
  i3054
 (41532=1) (ELSE=0) (MISSING=COPY).


FREQUENCIES
  VARIABLES=i3003 i3004 i3005 i3006 i3007 i3008 i3009 i3010 i3011 i3012 i3013
  i3014 i3015 i3016 i3017 i3018 i3019 i3020 i3021 i3022 i3023 i3024 i3025  i3026 i3027 i3028
i3029 i3030 i3031 i3032 i3033 i3034 i3036 i3037 i3038  i3039 i3040 i3041 i3042 i3043 i3044
i3045 i3046 i3047 i3048 i3049 i3050 i3051 i3052 i3053 i3054 i3055 i3056 i3057 i3058 i3059
i3060 i3061 i3062 i3063 i3064 i3065 i3066 i3067 i3068 i3069  .

** REMEMBER THAT VALUE LABELS ARE NOT CORRECT ON RECODED VARIABLES NOW.

DO IF (SUM(i3003,i3004,i3005,i3006,i3007,i3008,i3009,i3010,i3011,i3012,i3013,
i3014,i3015,i3016,i3017,i3018,i3019,i3020,i3021,i3022,i3023,i3024,i3025,i3026,
i3027,i3028,i3029,i3030,i3031,i3032,i3033,i3034,i3036,i3037,i3038,i3039, i3040,
i3041,i3042,i3043,i3044,i3045,i3046,i3047,i3048,i3049,i3050,i3051,i3052,i3053,
```

```
i3054,i3055,i3056,i3057,i3058,i3059,i3060,i3061,i3062,i3063,i3064,i3065,i3066,i3067,
i3068,i3069) >= 1) .
COMPUTE b10read =  SUM(i3003,i3004,i3005,i3006,i3007,i3008,i3009,i3010,i3011,i3012,i3013,
i3014,i3015,i3016,i3017,i3018,i3019,i3020,i3021,i3022,i3023,i3024,i3025,i3026,i3027,i3028,
i3029,i3030,i3031,i3032,i3033,i3034,i3036,i3037,i3038,i3039,i3040,i3041,i3042,i3043,i3044,
i3045,i3046,i3047,i3048,i3049,i3050,i3051,i3052,i3053,i3054,i3055,i3056,i3057,i3058,i3059,
i3060,i3061,i3062,i3063,i3064,i3065,i3066,i3067,i3068,i3069)).
END IF.
RECODE B10READ (SYSMIS = -1).
MISSING VALUES B10READ (-1).
EXECUTE .

VARIABLE LABELS B10READ 'bcs70 age 10: raw Edinburgh Reading Test score
(i3017 to i3069)'.

FREQUENCIES
   VARIABLES=b10read .
```