CLS

CLS Cohort Studies

Working Paper 2010/1

NCDS Cognitive Assessments at Age 50: Initial Results

Matthew Brown
Brian Dodgeon

Centre for Longitudinal Studies

March 2010

NCDS Cognitive Assessments at Age 50: Initial Results

Matthew Brown
Institute of Education, University of London

&

Brian Dodgeon
Institute of Education, University of London

March 2010

First published in March 2010 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk

© Centre for Longitudinal Studies

ISBN: 978-1-906929-13-8

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 authors and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the authors.

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

CONTENTS

Acknowledgements 4

Introduction 5

1. Structure of working paper 5

2. Assessments 6
   2.1 Word list recall 6
   2.2 Animal naming 6
   2.3 Letter cancellation 7
   2.4 Delayed word list recall 7
   2.5 Comparability with other studies 7
   2.6 Findings from other studies 8

3. Results 11
   3.1 Immediate word-list recall 12
   3.2 Word-list recall (Delayed) 13
   3.3 Animal naming 13
   3.4 Letter cancellation 14
   3.5 Correlations between tests 16

4 Variations in performance by key demographics and other factors 17
   4.1 Sex 17
   4.2 Socio-economic class 17
   4.3 Alcohol consumption 18
   4.4 Smoking 19
   4.5 Depression 19
   4.6 Qualifications 20
   4.7 General health 20
   4.8 Exercise 21

5. The importance of ‘interview’ factors 22
   5.1 Timing of interview 22
   5.2 Whether word-list read by computer or interviewer 24
   5.3 Word list 25
   5.4 Presence of others 27
   5.5 Month of interview 28
   5.6 Any other factors 29

6. Regression models 30
   6.1 Memory (Immediate and delayed) 31
   6.2 Animal naming 33
   6.3 Letter cancellation 34

7. Conclusions 37

References 38

ACKNOWLEDGEMENTS

Thanks are due to Professor Ian Deary of the Centre for Cognitive Ageing and Cognitive Epidemiology at the University of Edinburgh, who gave us detailed comments on the initial draft which proved very useful in structuring the final version, and giving us a good sense of how these initial results fit with the existing literature. Professor Jane Elliott of the Centre for Longitudinal Studies was also extremely helpful in formulating the initial ideas for the Working Paper, and giving constructive comments on the earlier drafts.

Introduction

The changing demographic profile of the British population, with a higher proportion of individuals being in the older age groups, has led to a particular policy interest in healthy ageing, and of particular concern is the growing number of individuals with cognitive impairment. It is recognized that poor cognitive function is linked with a heightened risk of experiencing dementia, disability and consequently institutionalization and mortality (Korten et al, 1999; Matthews & Denning, 2002) and the number of people with cognitive impairment is expected to increase from 461,000 to 765,000 between 1998 and 2031. By 2031 the costs of long-term care for older people with cognitive impairment are expected to rise to £16.7 billion in England alone (Comas-Herrera et al, 2007).

The National Child Development Study (NCDS) 2008/9 follow-up included a module comprised of a series of cognitive assessments which sought to measure memory, verbal fluency, perception and attention at age 50. These assessments are the first cognitive assessments conducted since cohort members were age 16; meaning that for the first time since adolescence analyses can be performed which use cognitive function as an outcome variable rather than a predictor. The rich life-history information which has been collected over the life-time of the study will allow the determinants of cognitive ability at age 50 to be examined; for example researchers will be able to investigate the impact of a full range of life-time health behaviours such as exercise, diet, smoking and alcohol consumption, all of which are potentially modifiable at the individual level. Understanding the impact of these behaviours on cognitive function is vital if levels of dementia in the general population are to be reduced in the future.

The intention is that cognitive assessments will be included in subsequent sweeps of the study in order that the data collected at age 50 can serve as a baseline, against which changes in cognitive ability which occur with ageing can be measured.

1. Structure of working paper

Section 1 describes the four assessments included in the NCDS 2008/9 survey. These assessments are widely used in longitudinal studies. Section 1 also provides information on the extent to which the assessments conducted as part of NCDS are comparable with the data collected by other studies and then provides some examples of other research which has been conducted using these assessments.

Participation in the assessments was of course voluntary, so Section 2 begins by providing details about participation rates. Section 2 then goes on to summarise the scores which were achieved in the assessments and examines correlations between the assessments. The relationships between performance in the tests and key socio-demographic factors are explored in Section 3, and in Section 4 we examine the impact of ‘interview’ factors on test performance such as the time of day at which the tests were administered. Finally, Section 5 presents a series of regression models which examine the relative importance of socio-demographic factors and 'interview' factors in terms of their impact on performance in the cognitive assessments.

2. Assessments

The cognitive assessment module is comprised of four tests as described below. The names and labels of the relevant variables from the deposited data are provided for reference:

2.1 Word list recall

Word list recall is a test of verbal learning and recall where participants are required to learn a list of 10 common words. The CAPI (Computer Assisted Personal Interviewing) program randomly selects one of four lists of words which are presented to the respondent by the computer using a recorded voice. In cases where the computer voice is not audible the list is read aloud by the interviewer, who is asked to imitate the pace and clarity of the recorded voice, reading the words at approximately 2-second intervals.

Once the list has been read out, cohort members have up to two minutes to recall as many as they can. Interviewers make a note of each word correctly recalled and enter the total into the CAPI program.

(The word list allocated to the participant was recorded so that it can be ensured that a different list will be used in the next follow-up in which the tests are included).

2.2 Animal naming

Animal naming is a test of verbal fluency which measures how quickly participants can think of words from a particular category. The participant is asked to name as many different animals as possible within one minute. Interviewers make a note of each named animal and enter the total into the CAPI program. Repetitions, named animals (e.g. Bambi) and redundancies (e.g. white cow, brown cow) are excluded from the total score. Successful participants will typically categorize animals into groups (pets, farm animals, fish etc) and move to a new category once no more animals can be thought of from a particular category; this requires organisation, abstraction and mental flexibility. This test has been widely used, and the present version was taken from the cognitive assessment section of the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) (Roth et al., 1986).

2.3 Letter cancellation

Letter cancellation is a test of attention, mental speed and visual scanning. The participant is given a page of random letters of the alphabet, set out in rows and columns, and is asked to cross out as many “Ps” and “Ws” as possible within one minute. An example is given at the top of the page to show the respondent how to cross out the letters. The page contains 26 rows and 30 columns and there are 65 target letters in all. Respondents are instructed to work across each row from left-to-right as if they were reading a page and they are asked to perform the task as quickly and accurately as possible. When the allotted time is over the respondent is asked to underline the last letter that their eye has reached. The total number of letters searched provides a measure of speed of processing. The number of target letters missed (P and W) up to the letter reached provides a measure of accuracy.

2.4 Delayed word list recall

Delayed word list recall is a test of delayed memory which asks the participant to recall as many words as they can from the original list presented to them during the first word-recall task. The word lists are not repeated and participants have again two minutes to recall as many words as they can. Interviewers make a note of each word correctly recalled and enter the total into the CAPI program. The delayed word-list recall task is completed approximately five minutes after the immediate word-list test (after the animal naming and letter cancellation tests have been conducted).

2.5 Comparability with other studies

When conducting the cognitive assessments interviewers working on the 2008-9 follow-up followed exactly the same procedures as were employed by interviewers working on the English Longitudinal Study of Ageing (ELSA) (http://www.ifs.org.uk/elsa/) and as such the data collected by the two studies will be directly comparable.

Word-list recall exercises and the letter cancellation task have also been included in the 1946 cohort study (the National Survey of Health and Development) (http://www.nshd.mrc.ac.uk/)[1]. There are however a number of small differences in protocol which will have an impact on the comparability of results:

- In the 1946 cohort study the word list recall exercise asks participants to recall 15 words whereas ELSA and NCDS ask respondents to recall 10 words.

- In the 1946 cohort study word list recall exercise, the words are shown to participants in a flip book (at intervals of two seconds) whereas in ELSA and

[1] The letter cancellation task was developed for the 1946 birth cohort study

NCDS the word lists are stored as sound files on the interviewer's laptop and 'read' to the respondent by the computer (unless the participant was unable to hear well or there was a technical problem with the computer in which case the words are read by interviewer). A person's ability to recall words which they have read may differ from their ability to recall words which they have heard spoken.

- In the 1946 study, participants are asked to recall the words on 3 occasions whereas ELSA/NCDS participants are only asked to recall the words twice. On the 1946 study, once the word-list recall task has been completed for the first time it is immediately repeated a second time whereas in ELSA and NCDS the task is only completed once. Each of the studies then include a delayed word-list recall exercise where the words are not repeated but the 1946 cohort study participants will be at an advantage as they will have had an extra opportunity to commit the words to memory.

- 1946 cohort members are given one task between the original word list recall exercise and the delayed word-list recall exercise (the letter-cancellation task) whereas ELSA/NCDS participants are given two tasks (the letter-cancellation exercise and the animal naming exercise).

The word-list recall exercises (immediate and delayed) were also included in the 1993, 1995, 1996, 1998, 2000, 2002 and 2004 sweeps of the Health and Retirement Study (HRS) (http://hrsonline.isr.umich.edu/) in the USA. The protocols followed by interviewers working on the HRS were exactly the same as those working on NCDS and ELSA meaning the data collected by each of the studies will be comparable. The only difference is that between the immediate word-list recall exercise and the delayed word-list recall exercise the HRS has included 5 minutes of questioning rather than additional cognitive assessments as included in ELSA/NCDS.

2.6 Findings from other studies

The four cognitive assessments described above have been included in each of the three waves of ELSA that have been conducted so far (2002, 2004 and 2006). Broad discussion of the impact of ageing and other socio-demographic factors on cognitive function can be found in chapters of the two Institute of Fiscal Studies (IFS) reports: Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing (Marmot et al, 2002) and Retirement, health and relationships of the older population in England: The 2004 English Longitudinal Study of Ageing (Banks et al, 2006).

The cognitive assessment scores achieved by the NCDS cohort are broadly comparable to the scores achieved by ELSA participants of a similar age. The 2002 ELSA report mentioned above provides figures for those aged 50-54; scores for this age group are typically slightly lower than those achieved by the NCDS cohort; however the ELSA study did show that performance in the tests is strongly related to age so this could be explained by ELSA's greater age range, although it could also be indicative of improvements in cognitive function over time (Llewellyn and Matthews, 2009). Interestingly, ELSA showed that women performed significantly better than men on the memory tests; a finding replicated amongst the NCDS cohort and common with findings from many other studies (Huppert and Whittington, 1993; Portin et al., 1995; Maitland et al., 2000), but the ELSA study also showed that men performed significantly better on the executive function tests (animal naming and letter cancellation). This was not the case amongst the NCDS cohort , although it is generally acknowledged that although performance on the animal naming test declines with age and is positively associated with education, the association with gender is unclear (Acevedo et al., 2000; Mathuranath et al., 2003).

Research based on the data collected in the cognitive assessments conducted as part of ELSA, HRS and the 1946 cohort study has investigated the links between cognitive function and a full range of life-time health related behaviours and other socio-economic factors. Using data collected by the 1946 cohort study, Richards et al. (2003) showed that, independent of sex, socioeconomic status, previous (adolescent) cognitive ability, and a range of health indicators, smoking is associated with faster declines in verbal memory and with slower visual search speeds . The effect of smoking on cognitive function was particularly severe for those individuals who smoked more than 20 cigarettes per day suggesting that smokers who survive into later life may therefore be at risk of clinically significant cognitive declines. The effects of cigarette smoke on cognitive function are not restricted to smokers themselves; Llewellyn et al (2009) showed using ELSA data that passive smoking is also associated with increased odds of cognitive impairment.

The link between alcohol consumption and cognitive function is less clear cut; using ELSA data Lang et al. (2007) found that moderate alcohol consumption in older adults is associated with better cognition (and well-being) than abstinence and using 1946 cohort study data Richards et al. (2005) showed that alcohol consumption was associated with a slower memory decline from 43 to 53 year olds in men. However, over the same age interval, alcohol consumption was associated with a more rapid decline in visual search speed in women.

Physical health has been shown to be strongly related to cognitive function. Using data from the HRS, Blaum et al. (2002) demonstrated the links between low cognitive performance and a range of chronic diseases and conditions and Richards et al. (2005) found, using 1946 cohort study data, evidence of a positive relationship between lung capacity and slower decline in psychomotor speed between 43 and 53 (as measured by the letter cancellation test). Richards et al. (2002) used data from the 1946 cohort study to investigate the relationship between height and weight and cognitive ability over the life-course. Birthweight was shown to be positively associated with cognition up to age 26, and with the likelihood of obtaining advanced educational qualifications. Height is positively associated with cognition at all ages, and also with educational attainment. Weight is not associated with cognition at ages 8 and 15, but is negatively associated with verbal ability at age 26, with verbal memory at age 43, and with educational attainment. These effects were independent of each other, and of family background.

Researchers have also investigated the relationship between cognitive function and mental-health and well-being. Gonzales et al. (2008) found that symptoms of depression were significantly associated with immediate word-list recall scores suggesting that memory decline may be a long-term feature associated with depression among the older population. Ertel et al. (2009), also using HRS data, provide evidence that social integration protects against memory loss; declines in memory (as measured by the immediate and delayed word-list exercises) are twice as rapid amongst the least socially integrated (as measured by marital status, volunteer activity and contact with children, parents and neighbours) as the most socially integrated.

Hatch et al. (2007) investigated the relationship between educational attainment and cognitive function in late-middle life using the 1946 cohort study. They found that educational attainment by early adulthood was positively associated with all measures of cognitive ability in late middle-life and that adult education is associated with better memory and better verbal fluency, but there is no association between adult education and mental speed or concentration.

The longitudinal nature of these studies also clearly makes possible the investigation of the impact of childhood experiences and circumstances on cognitive function in later life. Luo et al. (2005) use HRS data to highlight the relationship between lower childhood socio-economic status and poorer cognitive function in adulthood. Richards and Wadsworth (2004) showed using 1946 cohort study data that early adverse circumstances are strongly associated with lower cognitive ability in childhood and adolescence, and were detectable on measures of verbal ability, memory, and speed and concentration in midlife. However, these long term effects are mostly explained by the effects of adversity on childhood or adolescent cognitive ability or by differences in educational attainment and adult social class. An exception is the effect of poor material home conditions on visual search speed at 53 years, which was maintained after controlling for adolescent ability, as well as further controls for educational attainment, adult social class, physical growth, cigarette smoking, and affective state. They found no evidence of more rapid decline in memory and psychomotor function across middle age in those exposed to early adversity.

Finally, it is also possible to use the cognitive assessment data to conduct research of a more methodological nature. For example, Knauper et al. (1997) used HRS data to investigate the relationship between cognitive ability and an individual's ability to answer 'difficult' survey questions. It was shown that older respondents lower in cognitive ability are more affected by question difficulty than older respondents higher in cognitive ability. Respondents lower in cognitive ability gave more 'don't know' responses to difficult than to easy questions, while the differences between the difficult and easy questions were not as apparent for respondents with higher cognitive abilities. Knauper et al. argue that it must be recognized that this effect may result in biased survey findings as responses to difficult questions from respondents with lower cognitive ability would be underrepresented.

3. Results

In total, interviews were conducted with 9,790 cohort members, of whom 9,649 (99%) agreed to participate in the cognitive assessments. All 9,649 completed both the immediate word list recall test and the animal naming test but the letter cancellation test was successfully scored for 9,442[2] and the delayed memory test was completed by 9,592[3].

Histograms showing the distribution of scores achieved in each test are provided below in Sections 3.1 to 3.4.

---

[2] A small number of cohort members were not physically capable of participating in the letter cancellation test because of a physical handicap. Additionally, as this test was completed by the respondent in a paper booklet it was not scored immediately by the interviewer like the other three tests. Instead interviewers posted the paper booklets to the fieldwork contractor's Operations Centre where the test were scored by a specially trained team. A small number of booklets were either not completed properly, not returned by the interviewer or got lost in the post meaning they were unable to be scored.

[3] It was discovered post-fieldwork that a small number of interviewers did not complete the delayed word-list recall test with any of the cohort members that they interviewed.

3.1 Immediate word-list recall

Figure 3.1 : Word-list recall (Immediate)

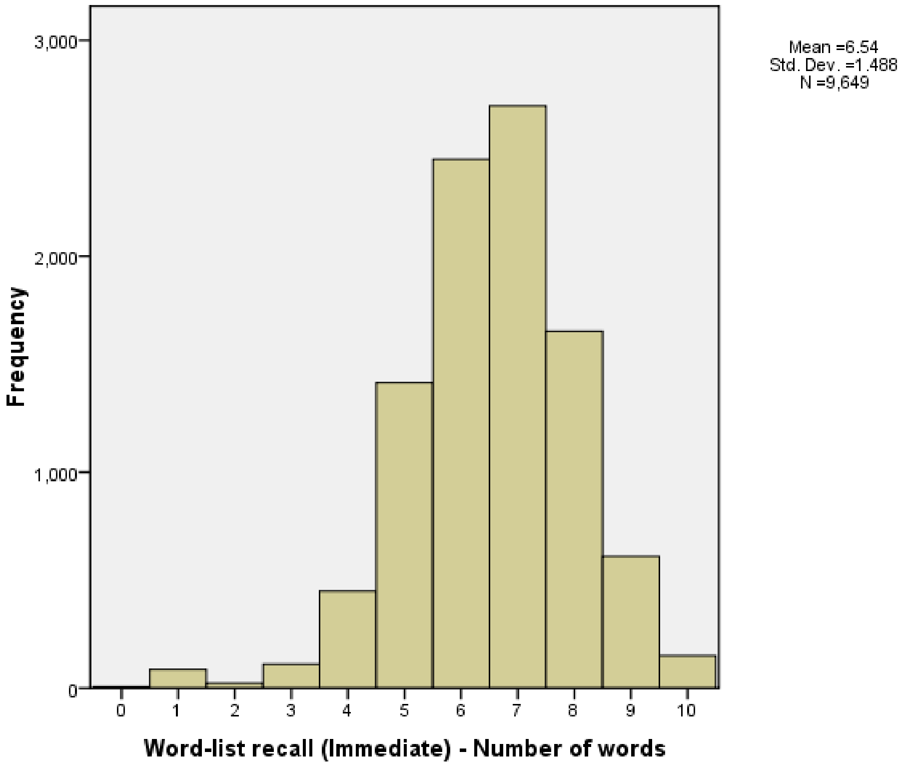


Figure 3.1 shows the distribution of scores on the immediate word-list recall exercise. The mean score achieved was 6.54 (out of 10) with a standard deviation of just under 1.5. Just under 84 per cent achieved a score of between 5 and 8. The distribution of scores is slightly skewed to the top end with around 8 per cent achieving scores at the top-end of the distribution (scores of nine or ten) and just over 2 per cent achieved scores at the very bottom of the distribution (scores of 0 to 3).

3.2 Word-list recall (Delayed)

Figure 3.2: Word-list recall (Delayed)

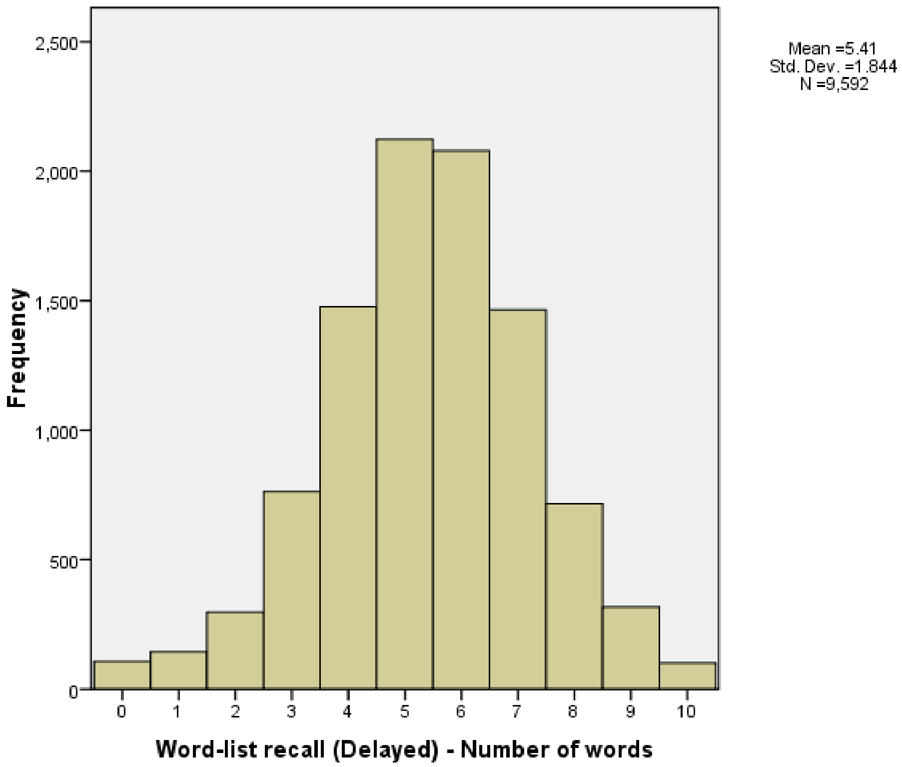
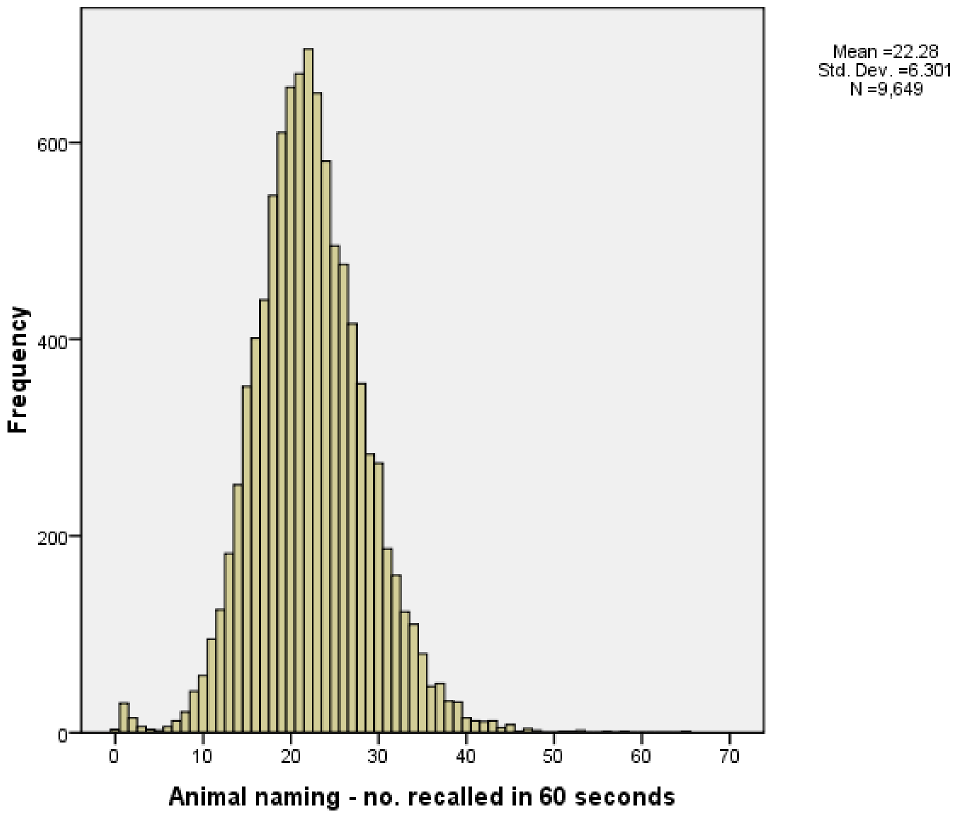


Figure 3.2 shows the distribution of scores on the delayed word-list recall task. The mean score achieved was 5.4 with a standard deviation of just over 1.8 (so a slightly larger spread than was observed in relation to the immediate recall task). Scores achieved in the delayed word-list recall exercise were fairly normally distributed with around three quarters (73%) achieving scores of between 4 and 7.

3.3 Animal naming

Figure 3.3 shows the distribution of scores on the animal naming task. Scores ranged between 0 and 65 with the mean number of animals recalled being 22.3 with a standard deviation of 6.3. Scores are reasonably normally distributed albeit with a small spike on scores of 1 and 2 (which may well result from interviewer data-entry errors) and a fairly long tail at the top of the distribution. Almost 89 per cent were able to recall between 11 and 30 animals.

Figure 3.3: Animal naming



3.4 Letter Cancellation

Figures 3.4 and 3.5 show the distribution of scores on the letter cancellation task in terms of speed and accuracy. The 'speed' score is measured by the total number of letters scanned. The grid for this task is comprised of 26 rows of 30 letters giving a maximum 'speed' score of 780 letters (a score which was achieved by 8 participants). The lowest score achieved was 84 (i.e. just short of three rows). The mean speed score was 334.1 (i.e. just over 11 rows) with a standard deviation of just under 89. Participants were instructed to underline the letter their eye had reached after 60 seconds. However, the distribution is clearly characterized by a series of spikes on particular values, in particular 919 cohort members achieved a speed score of 313. These spikes correspond with Ps and Ws (and in some instances with the last letter on a particular row) which suggests that in many cases the instruction was not adhered to correctly.

The accuracy score is measured by the number of Ps and Ws that were scanned but missed. A participant's 'accuracy' score is therefore likely to be closely related to their 'speed' score (as measured by the total number of letters scanned) as scanning a greater number of letters will of course increase the number of Ps and Ws that could potentially be missed.

Figure 3.4: Letter Cancellation (Speed)

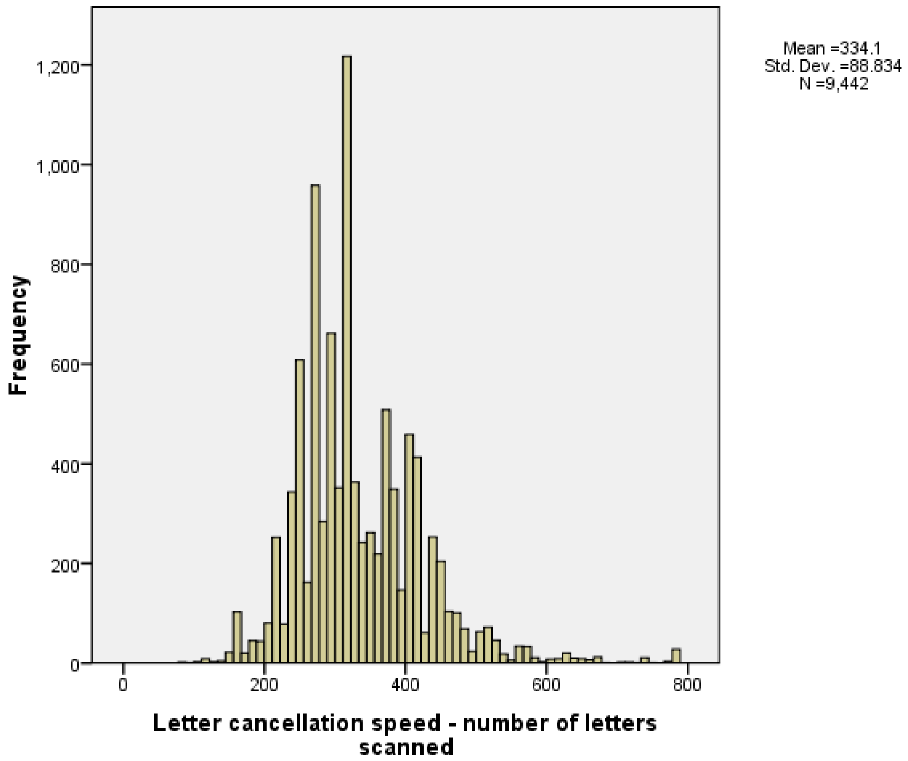
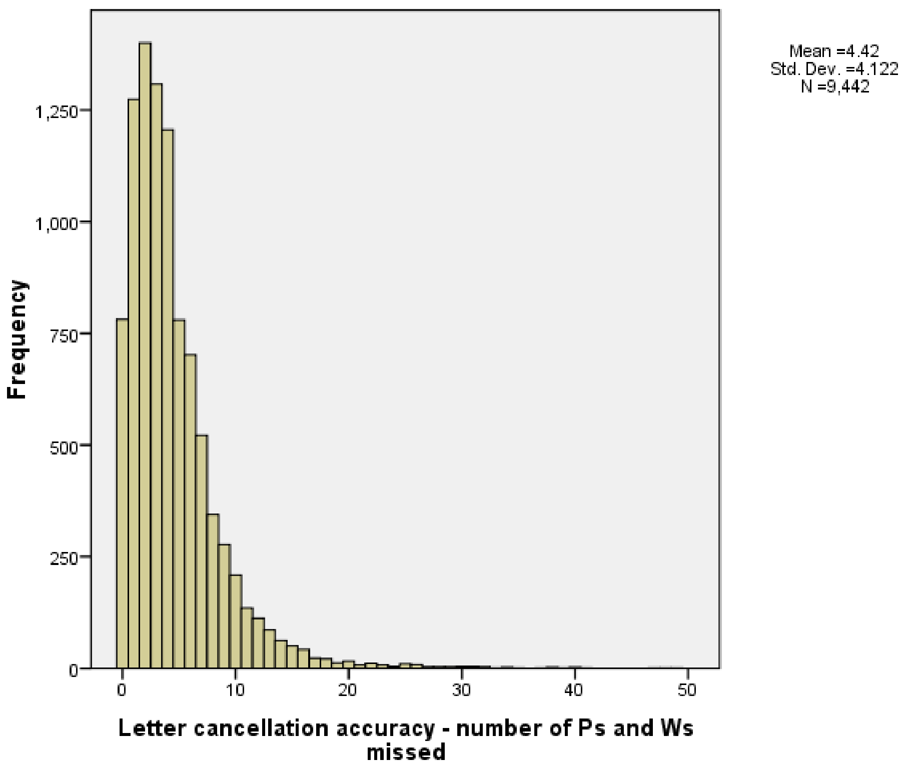


Figure 3.5: Letter Cancellation (Accuracy)



3.5 Correlations between Tests

As expected, there was a significant correlation between the results of the two word list recall tests, but only a small positive correlation between these recall tests and the animal naming task.

There was virtually no correlation at all between any of these three tests and the letter cancellation speed or accuracy measures, confirming that the word recall, animal naming and letter cancellation tests are measuring fundamentally different cognitive skills.

Note that Letter Cancellation Accuracy is measured negatively (i.e. higher scores mean more mistakes), resulting in negative correlations with Word Recall and Animal Naming, albeit small. Nevertheless, speed by accuracy is still a positive correlation, as noted in the last paragraph of section 3.1.

Table 3.5: Correlations between tests

Tests compared	**Correlation Coefficient (Pearson's R)**	**Significance**	**N**
Immediate Word List Recall by Delayed Word List Recall	0.65	0.000	9592
Immediate Word List Recall by Animal Naming	0.29	0.000	9649
Immediate Word List Recall by Letter Cancellation Speed	0.08	0.000	9442
Immediate Word List Recall by Letter Cancellation Accuracy	-0.09	0.000	9442
Delayed Word List Recall by Animal Naming	0.29	0.000	9592
Delayed Word List Recall by Letter Cancellation Speed	0.29	0.000	9385
Delayed Word List Recall by Letter Cancellation Accuracy	0.08	0.000	9385
Animal Naming by Letter Cancellation Speed	-0.14	0.000	9442
Animal Naming by Letter Cancellation Accuracy	-0.06	0.000	9442
Letter Cancellation Speed by Letter Cancellation Accuracy	0.59	0.000	9442

4. Variations in performance by key demographics and other factors

In this section we break the results down by some of the key variables which previous studies have found to be associated with cognition.

4.1 Sex

Women performed significantly better on the two word list recall tests ($p < 0.01$), but there was no significant gender difference on the animal naming task. Women were quicker than men at letter cancellation, but made more mistakes.

Table 4.1: Mean scores by sex

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Male	Mean	6.4	5.2	22.3	321	4.3
	StDev	1.5	1.8	6.4	86	4.0
	N	4746	4725	4746	4635	4635
Female	Mean	6.7	5.6	22.3	347	4.5
	StDev	1.5	1.8	6.2	90	4.2
	N	4903	4867	4903	4807	4807

4.2 Socio-Economic Class

Breaking down the results by socio-economic class (NS-SEC), we see the expected smooth class gradient, except that those in routine or semi-routine occupations performed slightly better on word recall and speed of letter cancellation than the 'smaller employers/lower supervisory/technical' class. But their animal naming and letter cancellation accuracy were not quite as good.

Table 4.2: Mean Scores by Socio-Economic Class (NS-SEC)

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Higher Professional/ Managerial	Mean	7.0	5.9	24.1	338	3.9
	StDev	1.4	1.8	6.1	85	3.5
	N	1160	1153	1160	1138	1138
Lower Managerial/ Intermediate	Mean	6.7	5.6	22.8	338	4.3
	StDev	1.4	1.8	6.3	89	4.0
	N	4182	4156	4182	4110	4110
Smaller employers/ Lower supervisory/ Technical	Mean	6.3	5.1	21.6	326	4.6
	StDev	1.4	1.8	6.1	87	4.3
	N	1816	1807	1816	1779	1779
Routine/ Semi-routine Occupations	Mean	6.4	5.3	21.4	336	4.6
	StDev	1.4	1.8	5.9	89	4.2
	N	1010	1004	1010	992	992
Not working	Mean	6.2	5.0	20.9	328	4.8
	StDev	1.7	2.0	6.3	92	4.5
	N	1461	1453	1461	1403	1403

4.3 Alcohol Consumption

Somewhat surprisingly, those consuming no alcohol do not perform as well on any of the five tests as those who drink moderately, and even those who are alcohol dependent perform as well as, or better, than non-drinkers on the first three tests.

Table 4.3: Mean Scores by Alcohol Consumption

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Does not drink alcohol	Mean	6.2	4.9	20.6	334	4.7
	StDev	1.5	2.0	6.5	98	4.4
	N	598	592	598	584	584
Unproblematic drinking	Mean	6.6	5.5	22.4	335	4.4
	StDev	1.5	1.8	6.3	88	4.0
	N	7288	7248	7288	7150	7150
Harmful or hazardous drinking	Mean	6.5	5.3	22.5	329	4.5
	StDev	1.5	1.8	6.2	87	4.1
	N	1340	1330	1340	1313	1313
Alcohol dependent	Mean	6.2	5.1	21.8	330	5.1
	StDev	1.6	2.0	6.5	99	5.0
	N	312	311	312	299	299

4.4 Smoking

There is little difference in performance between the two categories of non-smokers and occasional smokers, but regular smokers perform worse on all five tests.

Table 4.4: Mean Scores by Smoking Behaviour

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Never smoked	Mean	6.6	5.5	22.7	338	4.4
	StDev	1.5	1.9	6.5	88	4.0
	N	4481	4448	4481	4397	4397
Ex-Smoker	Mean	6.6	5.5	22.6	333	4.2
	StDev	1.5	1.8	6.0	89	3.9
	N	2990	2972	2990	2933	2933
Occasional smoker	Mean	6.5	5.5	22.2	335	4.4
	StDev	1.5	1.8	6.3	86	3.8
	N	317	316	317	310	310
Regular smoker	Mean	6.2	5.1	20.9	327	4.9
	StDev	1.5	1.9	6.2	92	4.8
	N	1860	1855	1860	1801	1801

4.5 Depression

There is a clear relationship between depression and cognitive ability, with those displaying signs of depression (score of 4 or more on the 9-point Malaise Scale) performing significantly worse on all tests except letter cancellation speed.

Table 4.5: Mean Scores by Depression (Malaise scale)

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Not depressed (Low malaise score)	Mean	6.6	5.5	22.5	334	4.3
	StDev	1.5	1.8	6.2	88	4.0
	N	8172	8124	8172	8013	8013
Depressed (High malaise score)	Mean	6.2	5.0	21.0	335	5.0
	StDev	1.5	1.9	6.3	94	4.8
	N	1395	1386	1395	1360	1360

4.6 Qualifications

We see the expected gradient, except that those with diplomas tended to perform worse than those with AS/A level or equivalent.

Table 4.6: Mean Scores by Highest Qualification

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
No qualifications	Mean	6.0	4.7	19.9	322	5.1
	StDev	1.4	1.7	5.6	88	4.8
	N	1384	1381	1384	1357	1357
CSE or equivalent	Mean	6.2	5.0	20.7	323	4.3
	StDev	1.4	1.7	5.8	85	3.9
	N	1218	1211	1218	1200	1200
GCSE or equivalent	Mean	6.6	5.4	22.2	334	4.2
	StDev	1.4	1.7	5.9	89	3.8
	N	2969	2945	2969	2911	2911
AS/A level or equivalent	Mean	7.1	6.1	24.2	340	4.0
	StDev	1.4	1.8	6.6	86	3.9
	N	767	763	767	751	751
Diploma	Mean	6.7	5.6	23.5	345	4.4
	StDev	1.5	1.8	5.5	97	4.0
	N	382	381	382	373	373
Degree or higher degree	Mean	7.2	6.2	25.4	349	4.2
	StDev	1.5	1.8	6.3	89	4.0
	N	1622	1610	1622	1593	1593

4.7 General Health

As self-reported health becomes increasingly poor, so does performance in the memory tests and the animal naming test. Speed and accuracy on the letter cancellation test would also seem to be affected by poor health but in terms of accuracy there is little difference between those reporting their health to be ‘excellent’, ‘very good’ or ‘good’ and those with rating their health as ‘fair’ have higher speed scores on average than those rating their health as ‘excellent’.

Table 4.7: Mean Scores by self-reported general health

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
Excellent	Mean	6.8	5.7	23.6	343	4.4
	StDev	1.4	1.8	6.5	88	4.3
	N	1877	1868	1877	1835	1835
Very good	Mean	6.7	5.6	22.5	337	4.3
	StDev	1.5	1.8	6.2	90	4.0
	N	3191	3161	3191	3136	3136
Good	Mean	6.5	5.3	22.1	332	4.3
	StDev	1.5	1.8	6.2	87	4.0
	N	2823	2815	2823	2772	2772
Fair	Mean	6.2	5.1	21.1	326	4.8
	StDev	1.5	1.9	6.2	88	4.3
	N	1228	1223	1228	1188	1188
Poor	Mean	6.0	4.7	20.1	320	4.9
	StDev	1.6	2.0	5.8	93	4.8
	N	527	522	527	508	508

4.8 Exercise

Those who exercised at least once a month performed significantly better on all tests than those who did not.

Table 4.8: Mean Scores by Amount of Exercise

		Immediate Word List Recall	Delayed Word List Recall	Animal Naming	Letter Cancell. Speed	Letter Cancell. Accuracy
At least once a month	Mean	6.6	5.4	22.5	336	4.4
	StDev	1.5	1.9	6.4	89	4.1
	N	7433	7433	7433	7285	7285
No exercise	Mean	6.2	5.1	21.0	327	4.5
	StDev	1.6	1.9	6.5	87	4.1
	N	2215	2215	2215	2156	2156

5. The importance of 'interview' factors

The previous section highlighted the impact of various socio-economic characteristics on cognitive function at age 50. In this section we investigate the impact of 'interview' factors such as the time at which cognitive assessments were conducted, or whether in the case of the verbal memory tests there are any differences in scores achieved by those cases who were 'read' the word-lists by the computer and those who were read the lists by the interviewer. These variables are often not provided to data analysts, and where provided are often ignored, but could potentially have a significant impact on test performance.

5.1 Timing of interview

Table 5.1 compares for each test the mean scores achieved by those interviewed in the morning, afternoon and evening.

Table 5.1: Mean scores by time of day

		Morning (Before 12:00)	Afternoon (12:00 to 17:59)	Evening (18:00 or later)
Immediate word list recall	Mean	6.67	6.49	6.50
	Std. Deviation	1.471	1.522	1.447
	Base:	2764	4058	2827
Delayed word list recall	Mean	5.53	5.36	5.36
	Std. Deviation	1.852	1.859	1.809
	Base:	2742	4040	2810
Animal naming	Mean	22.24	22.15	22.51
	Std. Deviation	6.272	6.220	6.440
	Base:	2764	4058	2827
Letter Cancellation (Speed)	Mean	338.81	334.72	328.59
	Std. Deviation	90.501	90.079	85.043
	Base:	2709	3967	2766
Letter Cancellation (Accuracy)	Mean	4.44	4.48	4.30
	Std. Deviation	4.235	4.032	4.139
	Base:	2709	3967	2766

In the two memory tests, the mean scores achieved by those interviewed in the morning were significantly higher than those who were interviewed in the afternoon or the evening ($p<0.01$), perhaps suggesting that as one becomes increasingly tired throughout the day one's performance in these tests begins to decline (although there was no significant difference between those tested in the afternoon and those tested in the evening).

On the animal naming test the highest scores were, however, achieved by those interviewed in the evening (although the difference in scores between those interviewed in the evening and those interviewed in the morning was not significant).

Letter cancellation speed scores were also affected by the time of day with those performing the test in the evening scanning significantly fewer letters than those performing the test in the morning or the afternoon, but time of day did not have a significant impact on accuracy scores.

Table 5.2 compares the mean scores achieved in each test by those interviewed on weekdays and those interviewed at weekends.

Table 5.2: Mean scores achieved on weekdays and at weekends

		Weekday	Weekend
Immediate word list recall	Mean	6.53	6.66
	Std. Deviation	1.489	1.471
	Base:	8719	930
Delayed word list recall	Mean	5.40	5.51
	Std. Deviation	1.843	1.852
	Base:	8668	924
Animal naming	Mean	22.27	22.36
	Std. Deviation	6.320	6.131
	Base:	8719	930
Letter Cancellation (Speed)	Mean	333.96	335.41
	Std. Deviation	88.553	91.460
	Base:	8532	910
Letter Cancellation (Accuracy)	Mean	4.41	4.45
	Std. Deviation	4.124	4.114
	Base:	8532	910

The mean score achieved in the immediate word-list recall task by those interviewed at weekends was significantly higher ($p<0.01$) than the mean score achieved by those interviewed during the on a week-day (6.66 compared with 6.53). There was also a difference in performance in the delayed memory test was of a slightly smaller magnitude (5.51 at weekends compared with 5.40 on week-days) and was only marginally significant ($p<0.1$). Nevertheless, this potentially suggests that weekends could be the optimum time for conducting memory tests; perhaps respondents are feeling more relaxed and perhaps able to focus more without the distractions of the working week? If this were to be true, then had more interviews taken place at the weekend the overall mean scores on these tests could potentially have been significantly higher. It would, however, at this stage be unreasonable to jump to this conclusion as we will need to control for characteristics which might be associated with cognitive function (health, socio-economic status, education level etc) as these factors might be associated with the likelihood of being interviewed at the weekend (and regression models are used to do exactly this in Section 6).

There were no significant differences in performance in the animal-naming task or the letter cancellation task (speed or accuracy).

It was shown above that mean scores in the two memory tests were highest for those interviewed in the morning and at the weekend. In Table 5.3 the variables indicating time of day and whether the interview took place during the week or at the weekend are combined:

Table 5.3: Means scores by day of week and time of day

		Weekday Morning	Weekday Afternoon	Weekday Evening	Weekend Morning	Weekend Afternoon	Weekend Evening
Immediate word list recall	Mean	6.64	6.48	6.50	6.81	6.51	6.79
	Std. Deviation	1.495	1.515	1.446	1.342	1.577	1.672
	Base:	2295	3611	2813	469	447	14
Delayed word list recall	Mean	5.51	5.36	5.36	5.59	5.41	5.86
	Std. Deviation	1.857	1.859	1.806	1.826	1.863	2.282
	Base:	2277	3595	2796	465	445	14

Scores achieved on the immediate word list recall test by those interviewed on a weekend morning are significantly higher than scores achieved at any other time of the week (other than weekend evenings when only a handful of interviews took place) ($p<0.05$). The scores achieved in the delayed word list recall test by those interviewed on a weekend morning were not significantly higher than scores achieved by those interviewed on a weekday morning or scores achieved by those interviewed at any other time in the weekend, but their scores were higher than scores achieved by those interviewed on weekday afternoons and evenings ($p<0.05$).

5.2 Whether word-list read by computer or interviewer

In order to ensure the tests were as standardised as possible the list of words for the immediate word-list recall test were recorded as sound files to be 'read' by the computer. In most cases this did happen, but in advance of the test interviewers played a test message in order that respondents could confirm they could comfortably here the computer's 'voice'; in the event of a technical problem or a hearing problem on the part of the respondent which resulted in the computer voice not being able to be heard, interviewers were instructed to read the word-list themselves at two second intervals.

In (98%) of cases the word-lists were read by the computer, but it is important to establish whether in the small proportion of cases where it was necessary for interviewers to read the word-lists there was any significant impact on test performance. Table 5.4 compares the mean scores on the two memory tests of those where the word-list was read by the interviewers with those where the word-list was read by the interviewer.

Table 5.4: Means scores by whether word-list read by computer or interviewer

		Word-list read by computer	Word-list read by interviewer
Immediate word list recall	Mean	6.55	6.24
	Std. Deviation	1.489	1.416
	Base:	9436	213
Delayed word list recall	Mean	5.42	4.89
	Std. Deviation	1.841	1.900
	Base:	9379	213

The mean scores achieved in both tests by those where the word-list was read to them by the interviewer were significantly lower ($p<0.01$) than the mean scores achieved by those where the list was read to them by the computer. These differences suggests that as might have been expected, the interviewer's voice seems not to have been as clear as the recorded word-lists played by the computer program or perhaps the interviewers read the words too quickly, slowly or at an irregular pace and this had an effect. It seems possible that the small number of respondents who were read the words by the interviewer may have achieved a higher score if they had been able to hear the recording.

5.3 Word list

Respondents were randomly allocated to one of four word-lists. However, when the immediate world list recall scores achieved by the respondents allocated to the four lists are compared there are some striking differences. The highest mean scores were achieved by those allocated to lists A and B (6.60) and the lowest mean score was achieved by those allocated to list D (6.44). The scores achieved by those allocated to list D were significantly lower than the scores achieved by those allocated to each other list (although the difference between the scores achieved by those allocated to list D and those allocated to list C was only moderately significant – $p < 0.1$).

In the immediate memory test there was no difference in mean scores between those allocated to lists A and B but when the delayed memory test was taken the mean score achieved by those allocated to list A was significantly lower ($p <0.05$) than those allocated to list B. It remained the case that those allocated to list D achieved the lowest mean score, but the mean score was only significantly lower than those allocated to list B ($p<0.01$).

Table 5.5: Mean scores by randomly allocated word-list

		List A	List B	List C	List D
Immediate word list recall	Mean	6.60	6.60	6.52	6.44
	Std. Deviation	1.547	1.473	1.483	1.442
	Base:	2381	2445	2397	2426
Delayed word list recall	Mean	5.39	5.51	5.41	5.33
	Std. Deviation	1.915	1.834	1.812	1.808
	Base:	2366	2435	2387	2404

The fact that those allocated to list D performed significantly more poorly on the immediate word-list recall task than those allocated to each other list ($p<0.05$) and significantly more poorly than those allocated to list B on the delayed word-list recall task perhaps suggests that this list was more difficult to commit to memory. If this were the case this could have implications not just when assessing the scores achieved this time around, but also when making comparisons with scores achieved the next time that the tests are included; for example, if a respondent was allocated list A in the 2008-9 follow-up, but in the subsequent follow-up was allocated to list D then a slightly lower score next time in the subsequent follow-up may not necessarily be the result of a decline in cognitive function, but could simply be a result of the fact that memorising the words in list D is a more difficult task.

However, the same word-lists have been employed in the first three waves of the ELSA survey and comparing the mean scores achieved in the immediate word-list recall exercise in each wave by those allocated to each of the four lists shows no evidence that scores achieved by those allocated to list D have been consistently lower than those allocated to other lists (Table X). In Wave 1 it was those allocated to list D that achieved the highest mean score (5.59 which was significantly higher than the mean score achieved by those allocated to each other list – $p < 0.05$); at Wave 2 the highest mean score was achieved by those allocated to list A (5.77 which was significantly higher than the mean score achieved by those allocated to lists B and C – $p < 0.05$ but not significantly higher than those allocated to list D) and in Wave 3 there were no significant differences between the word-lists.

Table 5.6 : Immediate memory test scores by randomly allocated list (ELSA Waves 1-3).

		List A	List B	List C	List D
Wave 1	Mean	5.50	5.42	5.33	5.59
	Std. Deviation	1.81	1.81	1.77	1.70
	Base	2879	2920	2934	2980
Wave 2	Mean	5.77	5.58	5.56	5.72
	Std. Deviation	1.84	1.83	1.83	1.72
	Base	2387	2282	2293	2274
Wave 3	Mean	5.83	5.73	5.77	5.82
	Std. Deviation	1.93	1.86	1.77	1.74
	Base	2317	2323	2412	2425

5.4 Presence of others

At the end of the cognitive assessments, module interviewers were asked to record whether there was anyone else in the room other than the interviewer and the respondent during the time when the assessments were being conducted. Table 5.7 below shows that in just under nine in ten cases (89%) the cognitive assessments were conducted in a room where nobody else was present (other than the interviewer). In cases where someone else was present it was most common that this person was the respondent's spouse or partner.

Table 5.7: Presence of others during cognitive assessments

	N	%
Nobody else present	8584	89.0
Respondent's spouse or partner	714	7.4
Other household member – adult	117	1.2
Other household member – child	171	1.8
Other person – not a household member	158	1.6
Base:	9649	

Table 5.8 below compares the mean scores for each assessment achieved by those who completed them alone and those who completed them with someone else in the room. In each case those who completed the tests alone performed significantly better than those who completed the tests in the presence of others ($p<0.01$ for immediate and delayed word-list recall, animal naming and letter cancellation speed, $p<0.05$ for letter cancellation accuracy).

Table 5.8: Mean scores by presence of others

		Nobody else in room (other than interviewer)	Others in room
Immediate word list recall	Mean	6.60	6.09
	Std. Deviation	1.478	1.493
	Base:	8584	1065
Delayed word list recall	Mean	5.48	4.87
	Std. Deviation	1.821	1.937
	Base:	8534	1058
Animal naming	Mean	22.43	21.01
	Std. Deviation	6.307	6.109
	Base:	8584	1065
Letter Cancellation (Speed)	Mean	335.23	324.86
	Std. Deviation	88.709	89.357
	Base:	8413	1029
Letter Cancellation (Accuracy)	Mean	4.39	4.68
	Std. Deviation	4.106	4.245
	Base:	8413	1029

5.5 Month of interview

Conducting cognitive assessments will for some interviewers have been an entirely new experience. Interviewers were given an opportunity to practice during their briefing sessions where a number of interviewers commented that they felt they may have difficulty keeping up with the respondent, particularly during the animal naming task. These interviewers were given assurance by more experienced interviewers and researchers that they would ‘get used to it’. In light of this it might be hypothesised that scores on the animal naming task might improve over the course of the fieldwork period as interviewers became more adept at noting down each animal mentioned by the respondent. Table 5.9 below shows the mean scores achieved in the animal naming task in each month of fieldwork and shows that there was no evidence of trend of scores increasing from month to month.

Table 5.9: Animal Naming Test Scores by Month of Interview

Month of interview	N	Mean	Std. Dev
August 2008	294	22.16	6.394
September 2008	1702	22.49	6.458
October 2008	1999	22.26	6.391
November 2008	1766	22.52	6.098
December 2008	1046	22.11	6.159
January 2009	1489	22.37	6.428
February 2009	824	22.11	6.021
March 2009	402	21.26	6.389
April 2009	116	21.16	5.897
May 2009	11	21.45	9.015

5.6 Any other factors

Finally, on completion of the cognitive assessment module interviewers were requested to record details of any factors which may have 'impaired the respondent's performance on one or more of the tests'. In seven per cent of cases interviewers recorded that there was at least one factor which impaired the performance of the respondent which as Table 5.10 below indicates were most commonly interruptions such as telephone calls and visitors or noisy environments.

Table 5.10: Factors affecting performance in cognitive assessments

Factor	N	%
Interruption or distraction e.g. phone call or visitor	118	1.22%
Noisy environment	102	1.06%
Deaf or hard hearing	57	0.59%
Too tired	49	0.51%
Illness or physical impairment	46	0.48%
Impaired concentration	45	0.47%
Blind or poor eyesight	44	0.46%
Very nervous or anxious	29	0.30%
Other mental impairment	18	0.19%
Problems with laptop	14	0.15%
Other answer	267	2.77%
Any factor affecting performance	688	7.13%

Table 5.11 below compares the mean scores achieved in situations where the interviewer did not report that any factors impaired upon the respondent's performance during the tests with those where at least one factor was recorded. Those where the interviewer recorded that at least one factor impaired performance achieved significantly lower mean scores on both memory tests and the animal naming task ($p<0.01$). They also achieved significantly lower accuracy scores on the

letter cancellation task although the speed at which the test was completed did not seem to be affected.

Researchers making use of the data from the cognitive assessment module will need to make decisions about whether to include the respondents flagged up by the interviewers as having factors impacting upon their test performance in their analyses. Clearly the exclusion of such cases would have a considerable impact upon overall average scores.

Table 5.11 Factors affecting performance in cognitive assessments

		At least one factor impaired performance	Nothing impaired performance
Immediate word list recall	Mean	5.74	6.60
	Std. Deviation	1.727	1.450
	Base:	688	8960
Delayed word list recall	Mean	4.45	5.48
	Std. Deviation	2.035	1.807
	Base:	688	8903
Animal naming	Mean	20.59	22.41
	Std. Deviation	6.411	6.275
	Base:	688	8960
Letter Cancellation (Speed)	Mean	333.22	334.17
	Std. Deviation	105.084	87.549
	Base:	637	8804
Letter Cancellation (Accuracy)	Mean	5.65	4.33
	Std. Deviation	5.430	3.997
	Base:	637	8804

6. Regression models

In this section we will, for each test, use regression models to examine the associations between performance in the cognitive assessments and a range of variables which previous research has linked with cognitive function (sex, educational qualifications, general health, childhood cognitive function (as measured by the general ability test at age 11), social class, excessive alcohol consumption, smoking and depression). In each case, we will then add the interview factors described in Section 4 to establish whether these factors can be seen to be having an impact on performance once the socio-economic and health behaviours described above are controlled for and whether the inclusion of these variables in the models increases the amount of variance in performance in the various tests that we are able to explain.

Characteristics of those included in the regression models (all participants in the cognitive assessments) are summarised in Table 6.1.

6.1 Memory (Immediate and Delayed)

Model 1 (see below) includes all the socio-economic and health-related factors listed above, but does not include any of the 'interview factors'. The R^2 for the immediate memory test model is 0.133, and for the delayed memory test is slightly higher at 0.15, so in neither case do the models explain a huge amount of the variation in performance. However, it is clear that many of the variables included have a very significant impact on scores achieved in these tests.

Controlling for other factors the respondent's sex is shown to be very significant on both tests with males performing significantly more poorly: particularly on the delayed memory test.

Educational qualifications become significant once a GCSE or equivalent is achieved. Gaining a diploma is associated with higher scores on both tests: the co-efficient is smaller than the coefficient associated with gaining both a GCSE and an A-Level or equivalent.

The general health coefficients are all negative when compared to the 'excellent' reference category but on the immediate memory test the coefficients do not increase in magnitude in a linear fashion as health becomes poorer, e.g. the negative coefficient associated with poor health is smaller than the negative coefficient associated with good health, although neither are significant. The only significant association is between 'fair' health and poorer scores ($p<0.05$) although the coefficient is fairly small.

In both tests, controlling for other factors it is clear that it is childhood cognitive function (as measured by the general ability test) that has the largest impact on performance in the two memory tests at age 50. In both cases this variable has the largest coefficient in the model and is highly significant.

Social class is associated with better performance on the immediate memory test but only at the top end of the social class spectrum (although again the coefficients are small). Social class is not however associated with better performance on the delayed test once other factors are controlled for.

Table 6.1 – Participant characteristics

		N	%
Sex	Male	4746	49.2
	Female	4903	50.8
	Total	9649	
General Health (Age 50)	Excellent	1877	19.5
	Very good	3191	33.1
	Good	2823	29.3
	Fair	1228	12.7
	Poor	527	5.5
	Total	9646	
Social Class (NS-SEC) (Age 50)	Higher professional /	1160	12.0
	Lower managerial /	4182	43.4
	Smaller employers /	1816	18.9
	Routine / Semi-	1010	10.5
	Not-working	1461	15.2
	Total	9629	
Alcohol Consumption (AUDIT) (Age 50)	Does not drink	598	6.3
	Unproblematic	7288	76.4
	Harmful or hazardous	1340	14.0
	Alcohol dependent	312	3.3
	Total	9,538	
Smoking behaviour (Age 50)	Never smoked	4481	46.4
	Ex-smoker	2990	31.0
	Occasional smoker	317	3.3
	Regular smoker	1860	19.3
	Total	9648	
Depression (Malaise) (Age 50)	Low malaise (no	8172	14.6
	High malaise	1395	85.4
	Total	9567	
Highest Qualification (Age 46)	No qualifications	1384	16.6
	CSE or equivalent	1218	14.6
	GCSE or equivalent	2969	35.6
	AS/A-Level or	767	9.2
	Diploma	382	4.6
	Degree or higher	1622	19.4
	Total	8342	
Childhood Cognitive Function (General Ability Test) (Age 11)	Scored between 0 and 80. Minimum score = 0, Maximum score = 79. Mean score = 45.41. Standard Deviation = 15.44	8,369	

Alcohol dependency is not shown to be significantly associated with immediate memory but interestingly both unproblematic levels of alcohol consumption and alcohol consumption defined as ‘hazardous’ or ‘harmful’ are both associated with higher scores on the immediate memory test. Alcohol problems do not however seem to be associated with performance on the delayed memory test.

Depression is significantly associated with poorer scores on both tests (albeit with small coefficients) and current smoking is significantly associated with poorer performance on the immediate memory test but not with the delayed memory test.

The addition of the ‘interview factors’ to the models (Model 2 below) moderately increases the R^2 values for both tests but only moderately (to 0.154 for the immediate test and to .165 to the delayed test) so the majority of variation in performance is still unexplained.

In Section 5 we showed that those taking the tests at weekends and those taking the tests in a room with no others present (other than the interviewer) achieved significantly higher scores. However, once the socio-economic and health behaviour factors (along with the other interview factors) are controlled for these effects become insignificant. However the time of day at which the tests are completed remains significant, with the association between completing the tests in the evening and poor scores being particularly noticeable, especially for the immediate memory test where the negative coefficient is only slightly lower than the negative coefficient associated with being ‘male’ and considerably larger than the negative coefficients associated with poor health, smoking and depression which perhaps suggests that this variable should be used by researchers as a control when analysing the memory test scores.

The positive effect of the word-lists being read by the computer remains significant for both tests and the negative effect of being allocated to Word List D remains significant for the immediate test (but not for the delayed).

The largest coefficient amongst the interview factors is that associated with ‘other factors’ which may have impaired performance which as mentioned earlier suggests that researchers analysing these scores should consider carefully whether to include cases where this applies.

6.2 Animal naming

Model 1 (which includes socio-economic and health-behaviour factors only) has an R^2 value very similar to that found for the immediate memory test (0.139 for animal naming compared with 0.133 for immediate memory) but there are some considerable differences in the associations between the covariates and performance in this test. As reported in Section 3, unlike the memory tests there is no significant gender difference on performance in the animal naming task. As in the memory tests, education becomes significant once a GCSE has been achieved and again the positive effect of achieving a diploma is smaller than the effect of achieving a GCSE or an A-Level. The positive coefficient associated with achieving a degree or higher qualification is twice that of achieving an A-Level and the effect of achieving a degree on performance in this test is larger than the effect on performance in the memory tests.

Poorer general health is associated with poorer performance on this test although the coefficients are fairly small.

As in the memory tests the most important factor in predicting performance on the animal naming test is childhood cognitive function. The positive effect of higher levels of childhood cognition on performance in this test is exactly the same as that found on the delayed memory test.

Social class, alcohol consumption, smoking and depression are generally not found to have a significant impact on performance in this test (although being an ex-smoker somewhat curiously seems to predict better performance than never smoking).

The addition of the interview factors to Model 2 does not increase the amount of variance explained by the model. It is only 'other factors' which might have impaired performance that have any effect and even these have a very moderate impact (considerably smaller than was found in relation to the memory tests).

6.3 Letter Cancellation

As mentioned previously, performance in the letter cancellation task is scored for both accuracy (as measured by the number of target letters scanned but missed) and speed (the total number of letters scanned).

Controlling for other factors, men were likely to scan a significantly smaller number of letters and perhaps as a result were likely to complete the test significantly more accurately (lower scores on the accuracy variable equate to a more accurate performance). In terms of speed on this test the sex effect was considerably larger than any other effect in the model.

Education only had a significant effect once a degree had been achieved and on both speed and accuracy the effect was positive. However the positive coefficients associated with gaining a degree are considerably smaller than on any of the other tests.

General health generally had very little effect although there is a modest negative effect of poor health on letter cancellation speed.

Childhood cognitive function is once again the most important predictor in terms of accuracy although the positive effect of increased childhood cognitive function is considerably smaller than that observed in the memory tests and the animal naming task. There is a significant positive association between childhood cognition and letter cancellation speed although the coefficient is smaller again.

Social class, alcohol consumption, smoking and depression are not found to be significantly related to accuracy or speed.

As with the animal naming task, the addition of the 'interview factors' to Model 2 does not increase the amount of variation explained by the model. The existence of 'other factors' which might have impaired performance has a modest effect on accuracy but no effect on speed.

Model 1: Socio-economic and health-related behaviour factors only.

	Immediate Memory Test		Delayed Memory Test		Animal Naming		Letter Cancellation (Accuracy)		Letter Cancellation (Speed)	
	Beta	**Sig.**	**Beta**	**Sig.**	**Beta**	**Sig.**	**Beta**	**Sig.**	**Beta**	**Sig.**
Sex (Ref = Female)										
Male	-.081	.000	-.120	.000	.019	.119	-.051	.000	-.149	.000
Highest qualification (Ref = No qualifications)										
CSE or	.012	.400	.023	.100	.013	.345	-.042	.006	.004	.768
GCSE or	.064	.000	.068	.000	.068	.000	-.028	.120	.020	.277
AS/A-Level or equivalent	.102	.000	.103	.000	.099	.000	.000	.993	.025	.109
Diploma	.038	.003	.035	.006	.059	.000	.007	.614	.024	.073
Degree or higher degree	.150	.000	.167	.000	.198	.000	.045	.017	.070	.000
General health (Ref = Excellent)										
Very good	-.013	.379	.006	.680	-.029	.052	-.018	.260	-.017	.288
Good	-.027	.069	-.021	.154	-.032	.034	-.026	.098	-.030	.059
Fair	-.036	.010	-.030	.031	-.032	.023	-.005	.721	-.028	.058
Poor	-.024	.065	-.039	.003	-.038	.005	-.015	.291	-.036	.011
Cognitive Function (Age 11)										
General Ability	.221	.000	.232	.000	.232	.000	-.177	.000	.058	.000
Social class (NS-SEC) (Ref = not working)										
Higher managerial / professional	.036	.024	.013	.403	-.001	.931	-.002	.928	.019	.265
Lower managerial / intermediate	.038	.040	.020	.286	.029	.116	.020	.317	.034	.084
Small employers /lower supervisory and technical	.005	.770	-.009	.566	.015	.356	.040	.024	.023	.182
Semi-routine / routine	.019	.192	.010	.503	.012	.403	.011	.487	.011	.465
Alcohol problem (Ref=no alcohol)										
Unproblematic drinking	.046	.028	.038	.071	.017	.430	-.014	.532	-.040	.071
Hazardous	.047	.016	.029	.141	.017	.387	.007	.758	-.018	.404
Alcohol	.008	.576	.005	.701	.004	.787	.015	.292	-.016	.265
Smoking (Ref = never smoked)										
Ex-smoker	.018	.126	.026	.030	.024	.043	-.012	.337	.003	.818
Occasional	-.004	.735	.008	.459	-.001	.952	-.003	.794	.005	.656
Current smoker	-.028	.024	.003	.808	-.012	.334	.015	.258	-.012	.361
Depression (Malaise) (Ref: no signs of depression)										
Signs of	-.034	.004	-.039	.001	-.014	.244	.033	.010	.010	.429
Model Fit										
F	48.704	<0.001	56.018	<0.001	51.086	<0.001	11.179	<0.001	12.724	<0.001
R^2	0.133		0.15		0.139		0.034		0.039	

*All coefficients are standardised

Model 2: Socio-economic factors and interview factors

	Beta	Sig.	Beta	Sig.	Beta	Sig.	Beta	Sig.	Beta	Sig.
Sex (Ref = Female)										
Male	-.076	.000	-.113	.000	.021	.091	-.049	.000	-.145	.000
Highest qualification (Ref = No qualifications)										
CSE or equivalent	.008	.580	.021	.139	.012	.384	-.041	.007	.004	.777
GCSE or	.060	.000	.061	.000	.067	.000	-.027	.138	.022	.220
AS/A-Level or equivalent	.097	.000	.102	.000	.098	.000	-.001	.948	.024	.119
Diploma	.035	.005	.033	.009	.058	.000	.007	.625	.024	.079
Degree or higher degree	.143	.000	.161	.000	.198	.000	.044	.018	.071	.000
General health (Ref = Excellent)										
Very good	-.012	.422	.008	.574	-.029	.052	-.016	.320	-.015	.345
Good	-.025	.094	-.017	.241	-.029	.051	-.026	.104	-.027	.082
Fair	-.030	.030	-.023	.088	-.030	.030	-.006	.678	-.026	.081
Poor	-.020	.127	-.031	.017	-.036	.007	-.015	.274	-.035	.014
Cognitive Function (Age 11)										
General Ability	.211	.000	.226	.000	.231	.000	-.172	.000	.058	.000
Social class (NS-SEC) (Ref = not working)										
Higher managerial / professional	.046	.005	.019	.229	.001	.958	-.003	.846	.027	.117
Lower managerial / intermediate	.049	.010	.031	.095	.030	.111	.015	.461	.043	.033
Small employers / lower supervisory and technical	.012	.478	-.002	.882	.017	.322	.036	.045	.029	.106
Semi-routine / routine	.023	.107	.014	.326	.012	.395	.008	.592	.014	.363
Alcohol problem (Ref=no alcohol)										
Unproblematic drinking	.047	.023	.035	.087	.016	.441	-.016	.466	-.042	.059
Hazardous	.047	.016	.028	.149	.017	.388	.003	.894	-.021	.325
Alcohol dependant	.008	.534	.008	.557	.004	.754	.013	.376	-.017	.236
Smoking (Ref = never smoked)										
Ex-smoker	.018	.125	.024	.041	.024	.044	-.013	.295	.002	.890
Occasional	-.006	.611	.007	.531	-.001	.958	-.004	.767	.004	.707
Current smoker	-.027	.028	.003	.794	-.011	.357	.015	.252	-.012	.362
Depression (Malaise) (Ref: no signs of depression)										
Signs of	-.031	.008	-.038	.001	-.013	.277	.029	.023	.009	.468
Time of day (Ref: Morning)										
Afternoon	-.049	.000	-.027	.040	.015	.249	.020	.159	.002	.866
Evening	-.069	.000	-.047	.001	-.001	.935	.010	.489	-.034	.025
Day of week (Ref: Week day)										
Weekend	.011	.320	.000	.978	-.015	.185	.008	.518	-.005	.707
Presence of others (Ref: No others present										
Others present	-.015	.168	-.015	.170	-.010	.372	-.010	.418	-.003	.824
Other factors impaired performance (Ref: No)										
Yes	-.110	.000	-.104	.000	-.032	.005	.050	.000	-.007	.574
Word-list (Ref: List A)										
List B	-.007	.597	.024	.077	-	-	-	-	-	-
List C	-.022	.110	.009	.491	-	-	-	-	-	-
List D	-.032	.017	-.008	.532	-	-	-	-	-	-
Whether word-lists read by computer (Ref: No)										
Yes	.032	.003	.038	.001	-	-	-	-	-	-
Model Fit										
F	41.648	<0.001	45.404	<0.001	42.527	<0.001	9.719	<0.001	10.774	<0.001
R^2	.154		0.165		0.140		0.036		0.040	

*All coefficients are standardised

7. Conclusions

This Working Paper is in the nature of a 'baseline' paper, which introduces the age 50 cognitive tests (the first undertaken by the whole cohort since childhood), and provides initial general analyses linking socio-economic and health factors to these cognitive outcomes at age 50, as well as looking at the effect of 'interview' factors on test results.

The intention is that it may be a useful starting point for research going in a number of different directions: e.g. identifying age-related cognitive decline and linking to factors during the lifecourse such as specific illnesses, accidents, traumas, depression or health behaviours; linking to results of earlier cognitive tests at ages 7, 11 and 16, to produce groupings of lifelong trajectories; or linking to genetic and biomedical risk factors identifiable from the NCDS 2002 Biomedical Survey, etc.

In this paper we have only shown the most direct associations between possible predictor variables and cognitive outcomes. The associations with educational levels and socio-economic class are as expected, and the findings on smoking, depression and exercise replicate what is found in the literature. Regression can have the undesired effect of eliminating some interesting variables that could be included if more sophisticated methods were used. There is much scope for exploring the effect of mediating variables, through the use of structural equations, latent class analysis or other types of modeling.

References

Acevedo, A., Loewenstein, D. A., Barker, W. W., Harwood, D. G., Luis, C., Bravo, M., et al. (2000). Category fluency test: Normative data for English- and Spanish-speaking elderly. *Journal of the International Neuropsychological Society*, 6, pp. 760–769.

Banks, J., Breeze, E., Lessof, C., & Nazroo, J. (2006) Retirement, health and relationships of the older population in England: THE 2004 ENGLISH LONGITUDINAL STUDY OF AGEING (Wave 2), London: Institute for Fiscal Studies.

Blaum, C., Ofstedal, M.B. & Liang, J. (2002) Low Cognitive Performance, Comorbid Disease, and Task Specific Disability: Findings from a nationally representative survey. *The Journals of Gerontology: Medical Sciences,* 57, pp. 523-531.

Comas-Herrera, A., Wittenberg, R., Pickard, L. & Knapp, M (2007) Cognitive impairment in older people: future demand for long-term care services and the associated costs, *International Journal of Geriatric Psychiatry*, 22, pp. 1027-45.

Ertel, K.A., Glymour, M.M. & Berkman, L.F. (2009) Effects of Social Integration on Preserving Memory Function in a Nationally Representative US Elderly Population. *American Journal of Public Health,* 98: pp. 1215-1220.

Gonzales, H.M., Bowen, M.E. & Fisher, G.G. (2008) Memory decline and depressive symptoms in a nationally representative sample of older adults: The Health and Retirement Study (1998-2004). *Dementia and Geriatric Cognitive Disorders.* 25, pp.266-71.

Hatch, S., Feinstein, L., Link, B., Wadsworth, M. & Richards, M. (2007) The continuing benefits of education: adult education and midlife cognitive ability in the British 1946 birth cohort. *Journal of Gerontology Series B,* 62, pp. 404-414.

Huppert, F.A. & Whittington, J.E. (1993) Changes in cognitive function in a population sample. In: Cox, B.D., Huppert, F.A. & Whichelow, M.J. (eds) *The Health and Lifestyle Survey: Seven Years On*. Aldershot: Dartmouth

Knauper, B., Belli, R.F., Hill, D.H., Herzog, A.R. (1997) Question Difficulty and Respondents' Cognitive Ability: The Impact on Data Quality. *Journal of Official Statistics.* 13 pp.181-199.

Korten, A. E., Jorm, A. F., Jiao, Z., Letenneur, L., Jacomb, A., & Henderson A. (1999) Health, cognitive, and psychosocial factors as predictors of mortality in an elderly community sample. *Journal of Epidemiology and Community Health*, 53, pp 83-88.

Lang, I. Wallace, R., Huppert, F. & Melzer, D. (2007) Moderate alcohol consumption in older adults is associated with better cognition and well-being than abstinence*. Age and Ageing, 36, pp. 256-261.*

Llewellyn, D., Lang, I., Langa, K., Naughton, F.& Matthews, F. (2009) Exposure to secondhand smoke and cognitive impairment in non-smokers: national cross sectional study with cotinine measurement. *British Medical Journal, 12, pp. 338-462.*

Llewellyn, D. & Matthews, F. (2009) Increasing Levels of Semantic Verbal Fluency in Elderly English Adults. *Aging, Neuropsychology and Cognition*, 25, pp. 1-13.

Luo, Y. & Waite, L.J. (2005) The Impact of Childhood and Adult SES on Physical, Mental, and Cognitive Well-Being in Later Life" *Journals of Gerontology Series B,* 60, pp. 93-101.

Maitland, S.B., Intrieri, R.C., Schaie, K.W. & Willis, S.L. (2000) Gender differences and changes in cognitive abilities across the adult life span. *Aging, Neuropsychology and Cognition*, 7, pp. 32–53.

Marmot, M., Banks, J., Blundell, R., Lessof, C. & Nazroo, J. (2002) *Health, wealth and lifestyles of the older population in England: THE 2002 ENGLISH LONGITUDINAL STUDY OF AGEING,* London: Institute for Fiscal Studies.

Matthews, F. & Denning, T. (2002) Prevalence of dementia in institutional care, *Lancet*, 360, pp. 225-226.

Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S. (2003) Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, 25, pp. 1057–1064.

Portin, R., Saarijarvi, S., Joukamaa, M. and Salokangas, R.K.R. (1995) Education, gender and cognitive performance in a 62-year old normal population: results from the Turva Project, *Psychological Medicine*, 25, pp. 1295–8.

Richards, M., Hardy, R., Kuh, D. & Wadsworth, M. (2002) Birthweight, postnatal growth and cognitive function in a national UK birth cohort. *International Journal of Epidemiology*, 31, pp. 342-8.

Richards, M., Hardy, R. & Wadsworth, M.E.J. (2005) Alcohol consumption and midlife cognitive function in the British 1946 Birth Cohort. *Alcohol and Alcoholism*, 40, pp. 112-17.

Richards, M., Jarvis, M.J., Thompson, N., Wadsworth, M.E.J. (2003).   Cigarette smoking and cognitive decline in midlife: evidence from a national birth cohort. *American Journal of Public Health*, 93, pp. 994-998.

Richards, M., Strachan, D., Hardy, R., Kuh, D., Wadsworth, M.E.J. (2005) Lung function and cognitive ability in a longitudinal birth cohort study. *Psychosomatic Medicine*, 67, pp. 02-08.

Richards, M. & Wadsworth M.E.J. (2004) Long-term effects of early adversity on cognitive function. *Archives of Diseases in Childhood*, 89, pp. 922-927.

Roth, M., Tym, E., Mountjoy, C.Q., Huppert, F.A., Hendrie, H., Verma, E and Goddard, R. (1986) CAMDEX. A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia, *The British Journal of Psychiatry,* 149, pp. 698-709.

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