CLS
Cohort
Studies

Investigating individual differences in memory and cognition in the National Child Development Study cohort members using a life course approach

Helen M. Knight
Matt Brown
Brian Dodgeon
Barbara Maughan
Martin Richards
Jane Elliott
Barbara Sahakian
Trevor Robbins

December 2010

CLS

Centre for Longitudinal Studies

Understanding Individual Behaviours Exploratory Networks

Investigating individual differences in memory and cognition in the National Child Development Study (NCDS) cohort members using a life course approach

Pilot Study Results Report

Knight, H.M.[1,2], Brown, M. [3], Dodgeon, B.[3], Maughan, B. [4], Richards, M [5], Elliott, J. [3], Sahakian, B. J.[1,6], Robbins, T.W,[1,2].

[1]Behavioural and Clinical Neuroscience Institute (BCNI), University of Cambridge, Downing Street Cambridge CB2 3EB, UK.

[2]Department of Experimental Psychology, University of Cambridge, Downing Street Cambridge CB2 3EB, UK.

[3]Centre for Longitudinal Studies, Institute of Education, 20 Bedford Way, London WC1H 0AL, UK.

[4]Social, Genetic and Developmental Psychiatry Centre (MRC), Institute of Psychiatry, DeCrespigny Park, London, SE5 8AF

[5]Centre for Family Research, Faculty of Politics, Psychology, Sociology and International Studies, University of Cambridge, Free School Lane, Cambridge CB2 3RF

[6] Department of Psychiatry, University of Cambridge School of Clinical Medicine, Addenbrookes Hospital Cambridge CB2 2QQ, UK.

December 2010

Contents

ACKNOWLEDGEMENTS

1. Introduction

The National Child Development Study (NCDS), also known as the 1958 birth cohort, is a continuing, multi-disciplinary longitudinal study of individuals all born in Great Britain in one week in March 1958. To date, data exist on over 9,000 cohort members generated in eight major surveys conducted throughout their life span and which now provide significant information on physical, biomedical, health, educational and economic factors as well as social development and inter-generational relationships (Power and Elliott, 2006). Importantly, the last survey administered was in 2008/9 when cohort members had reached the age of 50.

In recent decades, the changing demographic profile of the British population showing an increased life expectancy and decreased fecundity, has lead to an increase in economic pressure to provide for a more aged population (Bray, 2008; ONS, 2008). Consequently, cognitive epidemiology, the study of relationships between intelligence, cognitive health, disease specific outcomes and morbidity, has concurrently grown in importance. A central concept in this field is the idea that environmental and lifestyle factors may have a negative or positive influence on cognitive ability at different stages in life. For example, childhood trauma and experiencing stress or social exclusion in adulthood are factors reported to have a negative impact on cognition, whereas education, regular physical and mental activity, dietary interventions and social stimulation are variables thought to buffer cognitive decline in later adult life (Beddington et al., 2008).

A neurobiological explanation for this phenomenon is known as the cognitive reserve hypothesis. The cognitive reserve hypothesis proposes that an "individual's resistance to cognitive impairment that arises as a consequences of brain pathology caused by injury, disease or the normal aging process" and is thus related to the 'brain's reserve capacity to adapt to neuropathological change' (Stern, 2002; Barnett and Sahakian, 2008). Indeed, evidence supports the existence of risk and protective factors such as education, occupational attainment, exercise and social/mental stimulation, influencing the expression of certain neuropsychiatric conditions, for example, Alzheimer's disease (Barnett and Sahakian, 2008; Beddington et al., 2008).

Alzheimer's disease (AD) is a common, typically late-life, neurodegenerative disease characterized by progressive decline in memory, executive function and language (Waldemar et al., 2007). Critical to the treatment of AD is the early detection of pathological change in cognition indicative of a prodromal stage termed amnesic mild cognitive impairment (aMCI). Previous research has indicated that poor performance in a visuo-spatial associative learning task, (the Cambridge Automated Test Battery (CANTAB) Paired Associates Learning (PAL) task), accurately distinguished AD and questionable dementia (QD) patients (analogous to a diagnosis of MCI) from depressed and control subjects (Swainson et al., 2001). Moreover, using a second small QD population impaired performance in the PAL task in conjunction with performance in a second task, the Graded Naming Test (GNT), was highly predictive of a subsequent clinical diagnosis of AD (Blackwell et al., 2004). However, there is still a need for the development and testing of sensitive neuropsychological tests like

PAL, which might be used as a reliable marker of very early (before a diagnosis of MCI) decline and hence, in the future, could aid in the assessment of the efficacy of novel pharmacological agents by providing a tool by which performance can be assessed from very early to late stages of AD.

British cohort studies have been a valuable research resource for examining individual differences in cognitive trajectories during both normal aging and in later-life pathological states. For example, the Lothian Birth Cohort 1936 study used performance in the Moray House Test, a general intelligence test administered at the age of 11, and follow-up assessments of cognitive ability, biomedical factors and physical fitness at the age of 70 in over a thousand cohort members to examine relationships between cognitive change and key genotypic and phenotypic factors (Deary et al., 2007). Furthermore, in a large continuing longitudinal birth cohort study similar to the NCDS, but in which all members were born in 1946 (known as the National Study of Health and Development (NSHD) or 1946 cohort), measures of cognitive function in childhood and at the ages of 43 and 53 were used to investigate a range of life course determinants with the aim of modelling risk factor trajectories in relation to a later life mental and physical outcomes (Wadsworth et al., 2006). However, the sample size of the 1946 cohort (approx 3,000 plus for cognitive variables) is far exceeded by the NCDS cohort population of over 9,000 people.

Cognitive tests have also been administered to NCDS cohort members, both in their childhood and more recently, when they had reached mid-life aged 50. At the ages of 7, 11 and 16, individuals were given Maths and English tests and, in addition, at the age of 11 a general ability test (GAT). During the last NCDS8 survey, members completed a cognitive assessment battery, which included three neuropsychological tests: an animal naming test, a letter cancellation test (both assessing measures of executive function) and a word list recall task (assessing new learning and memory- a cognitive measure that is most consistent with deficits evident at the early stages of AD). Similar to the goals of the aforementioned cohort studies, cognitive capability measured in NCDS cohort members in early and later life provides the potential to track individual cognitive trajectories and in particular identify a group of individuals who show a marked decline in cognition analogous to early aMCI.

1.1 Aims of the study

The primary aim of the UIBEN pilot study was to administer a battery of computerised tasks designed to assess a variety of cognitive constructs as well as a short questionnaire to a sub-population of selected NCDS cohort members. This approach was intended to provide output in two ways. First, to provide an assessment of the feasibility and practicalities of performing similar computerised batteries of cognitive tests in a larger-scale study potentially conducted nationwide using the whole NCDS cohort. Discussion of findings relating to these issues is described in a second UIBEN working paper (Brown et al., 2010). Second, to generate data to test a working hypothesis; namely, that a group identified showing marked decline in cognitive function in mid-life will perform significantly worse than two matched control

groups when tested on the CANTAB visuo-spatial PAL task. Thus, evidence of substantial decline in performance of this test in a sample of individuals from the NCDS cohort might support idea that PAL performance could be used as a reliable marker of very early AD.

Ancillary objectives

There were five ancillary objectives to the pilot study:

1. To assess performance on all test/task outcome measures with regard to group (decline, and control groups) effects.

2. To assess performance on all task outcome measures with regard to gender effects.

3. To perform a factor analysis using the pilot study test indices to identify core psychological constructs.

4. To examine the NCDS survey data for potential differences between decline and control groups in terms of demographic and epidemiological variables consistent with the cognitive reserve hypothesis. This will be conducted for the whole cohort decline group and matching controls groups (~3,600 individuals) as well as the sub-groups who were tested on the UIBEN pilot study battery.

5. To investigate the responses given in the pilot study questionnaire with regards to items relating to health and mental activity routines. The findings relating to the fMRI questions are discussed in a third separate UIBEN technical report (Brown & Knight 2010).

2. Methods

2.1 Construction of groups based on childhood and adult cognitive ability

The first objective was to identify a group of NCDS cohort members who demonstrate cognitive decline in adulthood and to compare this Decline group with two control groups, one matched for childhood ability (control group 1- 'consistent achievers') and the other for mid-life cognitive performance (control group 2 - 'consistent non-achiever'). Preliminary analysis by means of latent class longitudinal growth curve modelling1 of the childhood Maths/English test scores at the ages of 7, 11 and 16, and the general ability test score (GAT) at the age of 11 alone, indicated

---

[1] LCGM is a semi-parametric statistical technique used when longitudinal data follow a pattern of change in which both the strength and the direction of the relationship between the independent and dependent variables differ across cases. The analysis identifies subgroups of individuals following a distinct pattern of change over time on the variable of interest.

that although cognitive ability was subject to fair amount of variation between ages 7 and 11, performance remained stable between the ages of 11 and 16 (data not shown). There was also a high correlation between the age 11 GAT results and those of the combined English/Maths tests at ages 11 and 16. Therefore, it is reasonable to assume that the GAT score is a representative measure of stable late-childhood cognitive ability. Groups were constructed by performing an ordinary least squares (OLS) regression using the age 11 General Ability Test as predictor variable (range 0-80), and the sum of the age 50 immediate & delayed word recall tests as the outcome variable (range 0-20). The sum of the immediate & delayed recall scores was chosen as the outcome measure owing to insufficient variation in scores (range 0-10) for both immediate, delayed alone and immediate minus delayed, for adequate numbers of people showing decline to be identified.

The Decline group members (whole cohort N = 1229) were identified as having a combined word recall score more than 1 standard deviation below that predicted from their GAT childhood scores. Individuals belonging to the first comparison control group (whole cohort N= 1330), control group 1, were identified as having similar scores at age 11 to those in the Decline group, but with age 50 recall scores consistent with that age 11 level of achievement (residuals on the OLS prediction only accepted between -0.2 and +0.2 SD), i.e. not having a marked drop in cognitive performance in mid-life: they are therefore also referred to as 'consistent achievers' in this report. Individuals belonging to the second comparison control group (whole cohort N= 1066), control group 2, were identified as having similar summed recall scores at age 50 to the Decline group, but an age 11 GAT score sufficiently low that they are consistent with that age 50 level of achievement, i.e. lower performance and no marked decline. For this reason the control group 2 is also referred to as 'consistently low achievers'.

One significant methodological problem in constructing groups using this linear regression approach is the 'floor effect' caused by the intercept being significantly above zero (in fact 8.53). So an individual with a score of zero at age 11 will nevertheless be predicted to score 8.5 out of 20 at age 50, and since the standard deviation is 3.15, a score of 5 or less at age 50 would put the person into the 'cognitive Decline' group (i.e. >1SD below predicted), even though their cognitive ability was evidently already low in childhood. A related problem occurs in trying to construct control group 2 (consistently low achievers) by 'matching' their scores at age 50 with those of the Decline group. Since around 10% of the Decline group scored 5 or less at age 50, it would be impossible to match these 10% of cases with individuals whose age 11 score was sufficiently low to represent 'no marked decline' (i.e. not more than 1 SD below predicted). Nevertheless the mean residuals of the control group 2, although not zero (as they should be ideally), were less than the residuals of the Decline group (-0.43 as opposed to -1.63).

2.2 Pilot study neuropsychological testing

2.2.1 Research participants

471 cohort members who live within a 55 mile radius of Cambridge city centre (the catchment area) had valid childhood GAT and word recall scores. Of these, 54 individuals were identified as fulfilling the Decline group criterion, and were subsequently individually matched with 54 individuals belonging to the control group 1 and 54 cohort members belonging to the control group 2.

Using survey data obtained in the last five years, selected cohort members were then further screened for health problems such as neurological conditions, head trauma, serious medical problems (e.g. poor vision in both eyes) or 'hard/illegal' drug dependent behaviour which could affect cognitive performance and exclusion criteria positive individuals were excluded from the study. Participants were tested between March and June 2010 when all cohort members were 52 years old. In total 16, Decline group members (9 male, 7 female) 14 control group 1 members (8 male, 6 female) and 12 control 2 group members (7 male, 5 female) were successfully tested and related data included in the statistical analysis of cognitive performance. Fewer women than men took part in the study (24 males, 18 females), but this gender difference is not statistically significant ($\chi^2 = 0.012$, $P = 0.994$).

As discussed in the previous section, owing to the 'floor effect', some cohort members who have poor cognitive performance in childhood will be predicted to have suffered some cognitive deterioration independent of a 'real' decline in mid-life ability. In the Cambridge pilot study subpopulation, 9 out of the 12 individuals in control group 2 have residuals lower than -0.5 SD and half (6 out of the 12 individuals) have residuals lower than -0.65 SD. Thus, although none are as low as -1 SD which would indicate an overlap in cognitive trajectories between the control group 2 and Decline group, these groups' cognitive trajectories are not as distinct as was desired.

2.2.2 Equipment and Procedure

Participants were all tested in a room at the Behavioural and Clinical Neuroscience Institute (BCNI), Cambridge, specifically designated for testing individuals on neuropsychological tasks. Informed consent was given by cohort members before the testing session. All computerised tasks were run on a Sahara Slate PC® i400 Series 12.1" touch-screen tablet PC and using CANTAB eclipse test software (Cambridge Cognition Ltd). Loud speakers were attached to the tablets and used during testing to ensure standardised quality and volume of sound.

The battery of tests took approximately 1.5 hours to administer and included five CANTAB tasks, the three cognitive tests performed in the NCDS8 sweep and a short questionnaire, details of which are provided below. Two versions of the battery, which included two different modes of the CANTAB AGN and CGT tasks, were randomly allocated to the participants prior to testing. A 27-item self-completion questionnaire, computerised and administered to participants on the touch-screen

Sahara slate tablet PC, was designed to assess health and health related issues as well as gathering information on how participants felt about the testing session. A script of the questionnaire is presented in Appendix 1. Participants were first given instructions on how to complete the questionnaire and then asked if they would like to self-complete with or without assistance from the researcher. Initial items focused on the current health status and mental activity routines followed by a series of questions concerning feedback about the set up of the study, neuropsychological tasks administered and willingness to participate in similar studies in the future. The final section of the questionnaire was designed to attain information concerning how participants might feel about taking part in fMRI studies and what sort of medical feedback they would wish to receive from such a scan. This part of the questionnaire involved the researcher briefly explaining fMRI technology and aims of fMRI research studies as well as discussion of ethical issues surrounding feedback from such studies.

The word list recall tests, animal naming verbal fluency test and letter cancellation test completed in the NCDS8 sweep were administered in the present battery using the same instructions and format, paper assessment booklet and recordings of the four word lists (Brown and Dodgeon, 2010). Each participant was allocated a word list prior to testing and was not presented with the same word list that they had previously experienced. A sound check was conducted before the word list was played to ensure that volume of sound was appropriate for each participant. The word lists and timer for the tests were available as buttons for the experimenter to press in a program loaded onto the desktop of the Sahara slate tablet PC. The NCDS8 survey assessments and the current study did, however, differ in two important respects. First, a third recall test of the word list was added to the battery and administered approximately 15 minutes after originally presentation of the word list. Second, participants were assessed in a professional testing environment and not at home, where the presence of other people and surrounding noise could have been a distracting factor.

Cambridge Automated Test Battery

The Cambridge Automated Test Battery (CANTAB) is a battery of computerized tests administered to participants with a touch-sensitive screen which automatically records and derives different performance indices within each test (Robbins et al., 1994). It was originally devised to adapt paradigms developed for testing of animal models of dementia and hence was initially used to assess cognitive function in elderly and demented subjects (Robbins et al., 1998). CANTAB batteries have now been extensively used to test clinical populations with a range of neuropsychological conditions including Alzheimer's disease, Parkinson's disease, depression, schizophrenia, and anorexia nervosa, as well as large numbers of non-clinical control volunteers of different age groups (Sahakian et al., 1988; Abas et al., 1990; Sahakian and Owen, 1992; Fowler et al., 2006). In the current study, the battery consisted of five CANTAB tests, specific details of which are provided below.

Visuo-spatial Paired Associates Learning task
The Visual spatial Paired Associates Learning (PAL) test (Sahakian et al., 1988) is a visuo-spatial associative learning test, which assesses visual memory and new learning. In this test, subjects were presented with six or eight white boxes, each of which opened up in a random order. The task was to look for coloured patterns in the boxes and to remember which pattern belonged in which box. In the current study, there were six levels of difficulty relating to stages in which 2, 3, 4, 5, 6, and 8 patterns were presented. Subjects had up to six chances to complete each stage and automatically moved from one stage to the next by correctly locating all of the patterns. If, however, a subject was unsuccessful after six attempts, the task was stopped at that point. Performance was scored using four main outcome measures: **Errors** (PAL total errors, PAL total errors at the 6 and 8 pattern stage and PAL mean errors to success); **Trials** (PAL total trials to success, PAL total errors at the 6 and 8 pattern stage and PAL mean errors to success); **First trial memory score** (the number of patterns correctly located after the first trial summed across the stages completed) and **Stages completed** (total stages completed and stages completed on first trial). For both the trial and error indices an adjustment is made for each stage not attempted due to previous failure.

Rapid Visual Information Processing task
The Rapid Visual Information Processing (RVIP) test (Wesnes and Warburton, 1984) assesses sustained attention capacity with a working memory component. It is also reported to be a sensitive measure of general information processing performance. Digits ranging from 2 to 9 appear at a rate of 100 digits per minute, one at a time on the screen. Subjects are instructed to press the response pad when they detect any one of three consecutive odd or even digit sequences, e.g. 2,4,6; 3,5,7). There are 9 main outcome measures: **total hits; total misses; total false alarms; total correct rejections; probability of a hit; probability of a false alarm; mean latency; RVP A'** (the signal detection measure of sensitivity to the target regardless of response tendency); and, **RVP B'** (the signal detection measure of the bias to make 'yes' response, i.e. false alarms).

The Graded Naming Test (GNT)
The GNT (McKenna and Warrington, 1980) assesses semantic and/or verbal memory. Subjects were presented with thirty line drawings and asked to name what each drawing represents. The researcher records correct and incorrect responses by pressing a button on a response pad. There is only one mean outcome measure, **total errors** made.

The Affective Go/No-go (AGN) task
The Affective Go/No-go (AGN) task (Murphy et al., 1999) assesses affective decision making and information processing biases for positive and negative stimuli. In this test, a series of words appear very rapidly on the screen. Half of these words are positive, or happy words and half are negative or sad words. Subjects are instructed to press the response pad as fast as they can as soon as they see a target valence word, e.g. happy word. Each word was presented for 300ms, followed by a 900 ms inter-stimulus interval within eight word blocks,

each containing 18 affective valenced and target word valence changed every two word blocks. The present study used two modes of the task, which differed only in the order of word target valences. Three main outcome measures were extracted from this task: ***mean correct latency/reactions times for correct responses*** (total mean correct latency; mean correct [shift or non-shift] mean correct (by positive or negative target type]*; **total commissions*** indicating total incorrect responses to a distracter stimulus (total commissions, total commissions in a shift and non shift blocks, total commissions for positive or negative target words); ***total omissions*** indicating total missed responses to targets (total omissions, total omissions in a shift and non shift blocks, total omissions for positive or negative target words).

Cambridge Gambling Task
The Cambridge Gambling Task (CGT) (Rogers et al., 1999) assesses decision-making and risk-taking behaviour outside a learning context. The subject is told that the computer has hidden a yellow token inside one of ten red or blue boxes aligned in a row at the top of the screen. The subject is instructed to choose whether they believe the token is hidden in a red or blue box and then to decide how many points (from an initial 100 points) they wish to gamble on being correct. The likelihood of each choice being correct is indicated on each trial by the ratio of red to blue boxes displayed and hence results in outcomes of more likely (9:1, 8:2, 7:3) or almost equally likely (6:4, 5:5) probability of winning. Sequences of trials were run in blocks under two conditions: an ascending condition in which the points that can be bet starts low and get progressive larger, e.g. 5%, 25%, 50% 75% up to 95%, or a descending condition in which the available points to bet start high and get progressively smaller. The CGT version used in the present study had a shortened interval time between increasing/decreasing bets of 2 seconds. The five principal outcome measures in this task are as follows: ***deliberation time*** (mean decision latency); ***quality of decision*** (the proportion of trials on which the subject chose to gamble on the more likely outcome); ***risk taking*** (the mean proportion of current points total that the subject stakes on each gamble test trial for which they had chosen the more likely outcome); ***risk adjustment*** (the degree to which a subject varies their risk taking in response to the ratio of red to blue boxes on each trial); ***CGT delay aversion*** *(*the difference between the risk-taking score in the descend and the ascend condition); ***CGT overall proportion bet*** (the mean proportion of the current points total that the subject chooses to risk on each trial regardless of the likely outcome).

2.2.3 Statistical analysis of task performance

Performance data was first examined for deviation from normality at P< 0.01 using the Shapiro-Wilks test. Data demonstrating a normal distribution were subjected to one-way analysis of variance (ANOVA) with group as a factor and post hoc comparisons of means between two groups with (Tukey's Honestly Significant Differences (HSD) test) and without (independent-sample T test) correction for multiple comparisons. If the distribution was deemed skewed, a non-parametric test,

Kruskal-Wallis H, was used to test for statistical significance between group means and the non-parametric Mann-Whitney U test for differences between the means of two groups (using the Games-Howell procedure when corrected for multiple comparisons). To decrease extreme skewness and stabilize variance, data pertaining to the outcome measures, letter cancellation omissions, PAL total errors adjusted, PAL mean trials to success, PAL total trials (adjusted), AGN total omissions positive and CGT deliberation time were either $\log_{10}$ or square root transformed. Analyses of covariance (ANCOVA) were used to investigate the relationship between task indices of group CANTAB task performance, gender and general ability test scores (GAT). All tests performed were two-tailed and hypothesis tests were assessed according to a type 1 error rate of 5%.

2.3 Factor analysis

Correlation matrices were first used to explore interrelationships between different test score outcome measures with the aim to identify the main test variables to be used in factor analysis. Principal component analysis was conducted using data from 11 key outcome variables. The method employed was a varimax rotation with Kaiser normalization performed using the SPSS package (Norusis, 1990).

2.4 Analysis of demographic, epidemiological and pilot study questionnaire variables

Key demographic and epidemiological variables were examined to assess whether differences between the groups exist and whether such differences support the idea that particular life factors may contribute to protection against cognitive decline. Initially, analyses were performed using data relating to the derived groups for the whole cohort, approximately 43% of all cases (Decline $N = 1229$; males 674, females 555; control group 1 $N = 1330$; males 635, females 695; control group 2 $N = 1066$; males 608, females 458). Analysis of the same variables was subsequently repeated using data from only the pilot study Cambridge subgroups as well as examining responses given in the pilot study questionnaire.

Statistical analysis was performed using the SPSS statistics package v17.0 (Norusis, 1990). Interval variable data demonstrating a normal distribution were subjected to one-way analysis of variance (ANOVA) with group as a factor. If the distribution was deemed skewed, the non- parametric test, Kruskal-Wallis H was used to test for statistical significance between the group means. Post hoc tests employed included independent-sample T test (parametric) and Mann-Whitney U test (non-parametric). Nominal data, described in the text as percentages and counts, was analyzed by means of chi square test or Fisher's exact test when counts/sample sizes were small. Pearson's Correlation Coefficient (*r)* was used to examine correlations between variables and to estimate effect size. All tests performed were two-tailed and statistical significance was considered at $P < 0.05$.

3. Results

3.1 Pilot Study neuropsychological task performance

3.1.1 Differences in cognitive performance between the groups

Table 3.1.1 presents group means and p values for the main outcome indices for childhood GAT, and pilot study repeated word recall, animal naming, letter cancellation tests. For both GAT and summed word recall performance there were significant group effects: GAT, (χ^2 = 13.216, df = 2, N = 42, p = 0.001); Summed immediate and delayed 1 recall [F (2, 39) = 6.26, p = 0.004]: Summed immediate and delayed 2 recall [F (2, 39) = 6.43, p = 0.004]; Summed delayed 1 and 2 recall [F (2, 39) = 3.28, p = 0.048]. As expected, and consistent with how the groups were selected, the control group 1 (CTL1) and Decline group (DEC) had significantly higher mean GAT scores than control group 2 (CTL2), DEC vs CTL2 U = 110.50, Z = -.062, p = 0.001; CTL1 vs CTL2 U = 101.50, Z = -3.113, p = 0.002. Similarly, and as would be predicted for the main word recall performance measure, the summed immediate and delayed 1 recall score, control group 1 had the highest mean which significantly differed from the Decline group, t(28) = -3.605, p < .001. However, the mean for control group 2 was higher than the Decline group (the difference showing a trend towards significance), but does not significantly differ from the control group 1 mean t(24) = 1.280, p < 0.213. This pattern of group performance was also evident for the summed immediate and delayed 2 recall score as well as summed delayed 1 and 2 recall score.

Comparison of the summed immediate and delayed recall scores attained from the NCDS8 age 50 survey with the current pilot study data showed that all three groups had a higher mean recall score in the second retest as detailed in Table 3.1.1. However, the difference between the initial and retest summed recall scores was found to be significant for only the Decline and control group 2 indicating that these groups showed a marked improvement in recall performance as compared to control group 1 (decline group t(15) = 2.514, p < 0.024; control group 1 t(13) = 0.809, p < 0.433; control group 2 t(11) = 2.746, p < 0.019).

A significant group effect was also shown for animal naming performance [F (2, 39) = 5.116, p = 0.01]. Again, as expected, control group 1 showed significantly better performance than control group 2 t(24) = 2.998, p < 0.006) and the Decline group performance was somewhere in-between. No group effects were evident for the letter cancellation outcome measures means as shown in Table 3.1.1.

Table 3.1.1 Group means for GAT, word recall test, PAL and RVIP CANTAB outcome measures.

TASK OUTCOME MEASURE	Mean Decline	Mean Control 1	Mean Control 2	Group difference statistic (F or χ^2)	Group (p value)	DEC vs CTL1 (p)	DEC vs CTL2 (p)	CTL1 vs CTL2 (p)
General ability score (GAT)	50.31 ± 3.500	50.43 ± 3.348	31.58 ± 3.499	χ^2 13.216	0.001**	0.95 (cor 1.000)	0.001 (cor 0.002)**	0.002 (cor 0.002)**
Recall imm & del 1 age 50	8.13 ± 0.44	12.29 ± 0.19	9.42 ± 0.23	43.532	0.000**	0.000**	0.025*	0.000**
Recall imm & del 1 pilot re-test	9.44 ± 0.61	12.79 ± 0.71	11.42 ± 0.80	6.264	0.004**	0.001**(cor 0.003)	0.055 (cor 0.129)	0.213
Animal naming	23.25 ± 0.98	25.43 ± 1.41	19.92 ± 1.10	5.116	0.011*	0.550	0.162	0.006 (cor 0.008)**
Letter cancellation omissions	4.56 ± 1.63	3.21 ± 0.68	4.17 ± 0.73	χ^2 0.774	0.679	0.850	0.426	0.466
Letter cancellation speed	308.94 ± 18.26	306.86 ± 20.12	310.08 ± 15.12	χ^2 0.189	0.910	0.967	0.728	0.680
Letter cancellation correct	19.56 ± 0.979	20.57 ± 1.36	19.67 ± 0.94	0.250	0.780	1.000	1.000	1.000
PAL total errors adjusted	34.06 ± 7.776	18.50 ± 2.936	30.58 ± 5.578	χ^2 3.762	0.152	0.119 (corr 0.174)	0.834	0.076 (cor 0.165)
PAL total errors adjusted log10	1.4080 ± 0.08014	1.1711 ± 0.0907	1.4098 ± 0.07735	2.686	0.081	0.059 (corr 0.110)	0.988	0.061 (cor 0.142)
PAL total errors 8 shapes (adj)	17.00 ± 3.490	8.57 ± 1.806	14.08 ± 3.539	χ^2 3.464	0.177	0.070 (corr 0.104)	0.780	0.212
PAL mean trials to success log10	0.400 ± 0.041	0.294 ± 0.032	0.407 ± 0.037	χ^2 5.458	0.065	0.055 (corr 0.112)	0.641	0.041 (cor 0.113)
PAL Total trials (adj)	15.31 ± 1.399	12.00 ± 0.792	15.25 ± 1.067	χ^2 5.889	0.053	0.052 (corr 0.120)	0.698	0.030 (cor 0.058)
PAL Total trials (adj) log10	1.163 ± 0.035	1.067 ± 0.031	1.171 ± 0.031	3.214	0.051	0.045 (corr 0.088)	0.860	0.021 (cor 0.083)
RVP total hits	16.38 ± 1.287	20.36 ± 1.333	14.00 ± 0.816	6.557	0.004**	0.041 (cor 0.056)*	0.162	0.001 (cor 0.003)**
RVP total misses	10.62 ± 1.287	6.64 ± 1.333	12.92 ± 0.811	6.421	0.004**	0.041 (cor 0.055)*	0.176	0.001 (cor 0.003)**
RVP total false alarms	1.56 ± 0.626	1.14 ± 0.312	2.08 ± 0.773	χ^2 0.679	0.712	0.935	0.507	0.563
RVP total correct rejections	249.56 ± 2.911	257.57 ± 2.910	244.25 ± 1.978	5.628	0.007**	0.063 (cor 0.096)	0.172	0.001 (cor 0.006)**
RVP probability of hit	0.6065 ± 0.048	0.7540 ± 0.049	0.5201 ± 0.030	6.490	0.004**	0.041 (cor 0.055)*	0.169	0.001 (cor 0.003)**
RVP probability of false alarm	0.0064 ± 0.003	0.0045 ± 0.001	0.0084 ± 0.003	χ^2 0.809	0.667	0.912	0.472	0.392
RVP A'	0.8960 ± 0.014	0.9364 ± 0.127	0.8747 ± 0.008	5.773	0.006**	0.047 (cor 0.064)*	0.250	0.001 (cor 0.006)**
RVP B"	0.9375 ± 0.027	0.9391 ± 0.025	0.9373 ± 0.022	χ^2 0.231	0.891	0.804	0.637	0.807
RVP mean latency	410.41 ± 21.555	404.94 ± 17.971	413.47 ± 19.45	χ^2 0.232	0.891	0.934	0.642	0.742

N = 42. Means are presented with standard error of mean (SEM). (Cor) denotes P value corrected for multiple comparisons. P values shown relate to two-tailed probability. * denotes statistical significance at P < 0.05 and ** denotes statistical significance at P < 0.01.

Paired Associate Learning

Four PAL outcome measures; total errors, total errors at 8 shapes, mean trials to success and total trials, show a trend in significance for a group effect (Table 3.1.1). For all these performance indices control group 1 (consistent achievers) showed better performance, i.e. need fewer trials to locate all the patterns correctly or make fewer errors, than the Decline and control group 2 (consistent non-achievers) groups. *Post-hoc* contrasts indicated that the difference in performance between the control groups or control group 1 and the decline group, reached either significance or a trend in significance for: total errors (CTL1 vs CTL2 $t(24) = -1.966$, $p < 0.061$; CTL1 vs DEC $t(28) = 1.965$, $p < 0.059$); mean trial to success (CTL1 vs CTL2 $U = 44.000$, $Z = -2.071$, $p = 0.041$, CTL1? vs DEC $U = 66.500$ $Z = -1.919$, $p = 0.055$); and total trials (CTL1 vs CTL2 $t(24) = -2.464$, $p < 0.021$, CTL1 vs DEC $t(28) = 2.094$, $p < 0.045$). There was, however, no difference in performance for these measures between control group 2 and the Decline group. A plot of values for each group illustrating this point is shown for total trials in Graph 3.1.1.

Interestingly, at the most difficult level of 8 patterns, the Decline group made more errors, that is, performed worse than both the control groups although the difference only showed a trend for significance between the control group 1 and the decline group (CTL1 vs DEC $U = 68.500$, $Z = -1.814$, $p = 0.070$; CTL2 vs DEC $U = 90.000$, $Z = -0.280$, $p = 0.780$; CTL1 vs CTL2 $U = 59.000$, $Z = -1.290$, $p = 0.212$) as shown in Graph 3.1.2.

As a guide to how much larger a sample might be required for such a result to attain statistical significance, an 84-person dataset was created by simply duplicating the results (i.e. Decline group = 32 participants; control group 1 = 28; control group 2 = 24). This, of course, produces exactly the same means, but the significance levels are as shown in Table 1, Appendix 2. With twice as many participants, every outcome variable shows group effect, which is now significant except for the adjusted total trials 8 shapes (the latter very close to significance).

Relationship between memory recall and PAL performance
Impairment in memory function is widely acknowledged to be a central feature of early Alzheimer's disease. In the present study, the main PAL outcome measure, total trials needed to successfully complete each stage, showed a significant negative correlation with immediate and delayed word recall ($r = -0.383$, $p = 0.012$) as indicated in Graph 3.1.3. However, as this relationship does not indicate a very high correlation, i.e. a value close to 1, this would suggest that much of the variation in PAL performance cannot be explained in terms of memory alone and hence supports the proposal that the PAL task is assessing additional cognitive measures such as spatial associative learning and attention.

Graph 3.1.1 Pal total trials (adjusted log10 transformed) as a function of groups
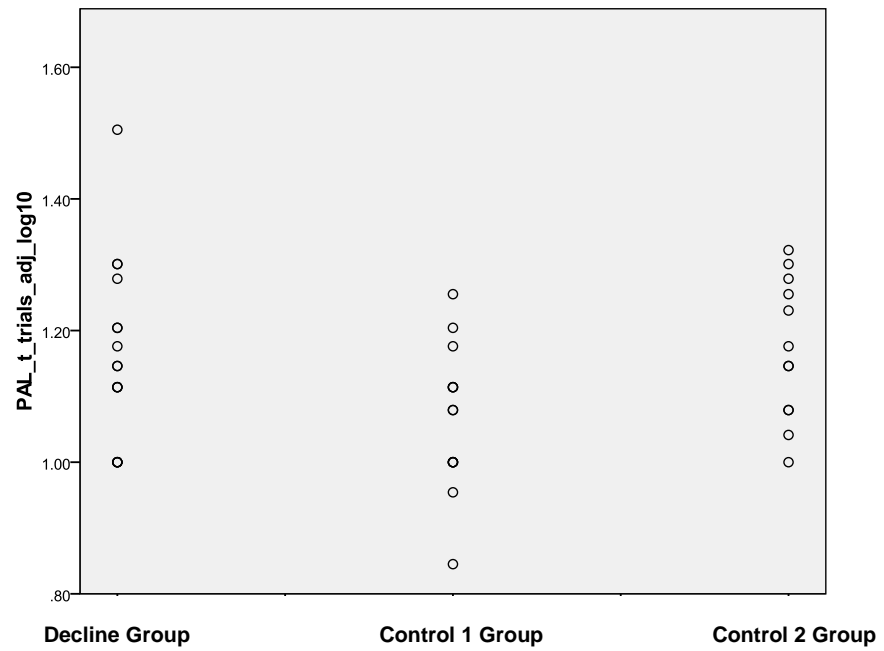
Graph 3.1.2 Pal total error 8 patterns as a function of groups

Rapid Visual Information Processing
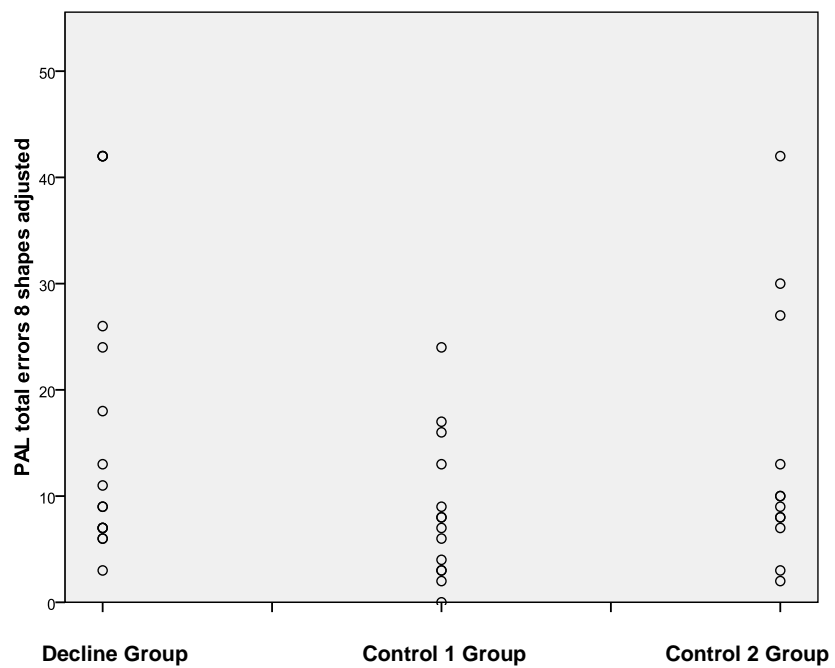
Five RVIP outcome measures, total hits, total misses, total correct rejections, probability of hit and RVP A', showed a significant group effect (total hits [$F_{(2, 39)}$ = 6.557, p = 0.004]; total misses [$F_{(2, 39)}$ = 6.421, p = 0.004]; total correct rejections [$F_{(2, 39)}$ = 5.628, p = 0.007]; probability of hit [$F_{(2, 39)}$ = 6.490, p = 0.004]; and RVP A' [$F_{(2, 39)}$ = 5.773, p = 0.006] (detailed in Table 3.1.1.) Again, these effects appear to be driven by better control group 1 and poorer Decline and control group 2 performance on these measures, that is, correctly responding to a target, correctly rejecting incorrect stimuli, and accurate signal detection independent of response tendency. Post-hoc contrasts indicated statistically significant or a trend towards statistically significant differences between the control groups and control group 1 and the Decline group (detailed in Table 3.1.2). However, even though the control group 2 had consistently lower means than the decline group for these variables, performance did not significantly differ between these groups.

All other indices of RVIP performance, e.g. total false alarms, probability of false alarm and RVP B' did not differ between the three groups including mean latency, the time taken to respond to correct stimuli.

Relationship between PAL and RVIP performance

In an attempt to separate out how much of the PAL performance of the respective groups was due to attention as opposed to memory, an OLS regression with RVIP 'A' as predictor, and PAL as outcome variable was performed. Saving the residuals from this regression defines a variable that can be viewed as reflecting the 'pure' spatial memory element of the PAL results. The resulting comparison of means is shown in Figure 3.1.4 (note the regression was performed using the *negative* of RVIP 'A', so both indicators are in the same direction). We see that, to the extent that this can be considered a measure of pure spatial memory, the Decline group performed markedly worse than control group 2, who themselves did better even than control group 1 (this will be a function of the relatively excellent performance of control group 1 in the RVIP 'A' – see Fig 3.1.4). However, the analysis does not approach significance (p = 0.364).

Summary of PAL and RVIP task performance
In summary, a similar pattern of differences in performance between the groups was shown for PAL and RVIP outcome measures. The Decline and control group 2 groups performed worse than the control group 1 (consistent achievers) making more errors in both PAL and RVIP tasks suggesting a deficit in memory, new learning and sustained attention as compared to control group 1 individuals. However, an OLS regression with RVIP 'A' as predictor, and PAL as outcome variable indicated that the decline group showed poorer mean performance than the control group 2 group in what could be assumed to be 'pure' spatial memory.

Graph 3.1.3 Correlation between recall test performance and main PAL outcome measure
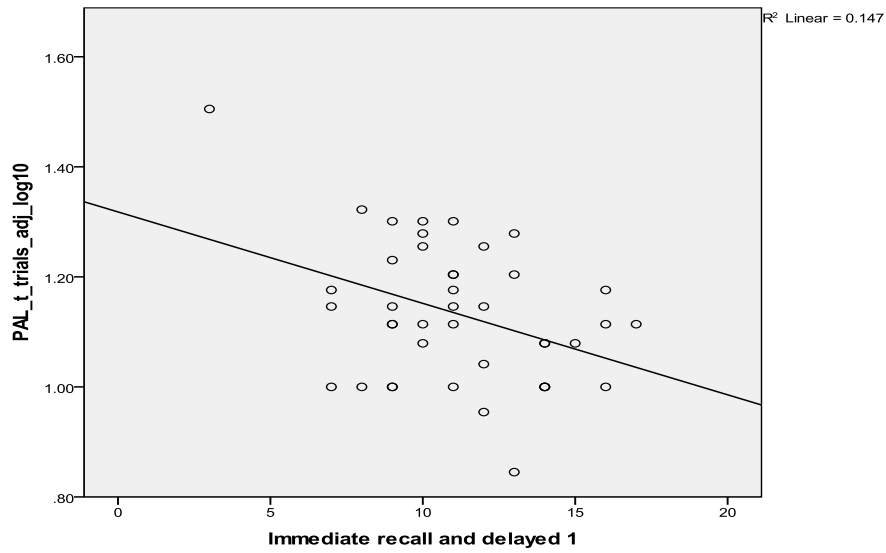
Graph 3.1.4 Residuals after regressing PAL t-err 8 shapes onto (-ve of) RVIP 'A' score

Table 3.1.2 Group means for AGN, GNT and CGT CANTAB outcome measures

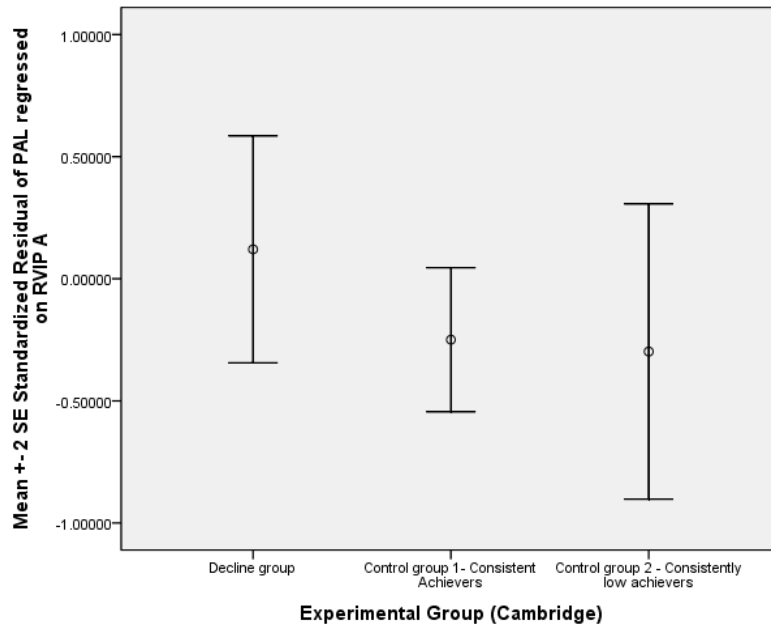| TASK OUTCOME MEASURE	Mean Decline	Mean Control 1	Mean Control 2	Group difference statistic	Group (P value)	DEC vs CTL1 (P)	DEC vs CTL2 (P)	CTL1 vs CTL2 (P)
AGN Mean correct latency	483.41± 17.66	498.55 ± 15.56	539.78 ± 24.13	2.246	0.119	1.000	0.056 (cor 0.106)	0.153
AGN Mean correct latency positive	480.52± 18.83	491.68± 14.83	533.63 ± 22.74	2.117	0.127	1.000	0.073 (cor 0.120)	0.125
AGN Mean correct latency negative	485.94± 17.94	505.37± 17.04	545.05± 25.57	2.221	0.122	1.000	0.060 (cor 0.104)	0.193
AGN Total commissions	9.19 ± 1.950	7.14 ± 1.540	8.67 ± 1.896	χ^2 0.667	0.716	0.505	0.926	0.453
AGN Total commissions positive	5.50 ± 1.049	4.36 ± 1.020	5.00 ± 1.052	χ^2 0.766	0.678	0.427	0.797	0.500
AGN Total commissions negative	3.69 ± 0.961	2.79 ± 0.631	3.67 ± 0.995	χ^2 0.118	0.943	0.782	0.906	0.754
AGN Total omissions	8.56 ± 3.108	5.14 ± 2.135	6.00 ± 1.624	χ^2 1.243	0.537	0.358	0.981	0.313
AGN Total omissions positive	5.69 ± 1.658	2.21± 0.806	3.00 ± 0.888	χ^2 3.974	0.137	0.061 (cor 0.127)	0.250	0.332
AGN Total omissions negative	2.88 ± 1.533	2.93 ± 1.424	3.00 ± 0.969	χ^2 2.786	0.248	0.525	0.086 (cor 0.997)	0.382
GNT Total errors	6.69 ± 0.978	7.50 ± 1.152	9.75 ± 0.760	2.375	0.106	0.593	0.027 (cor 0.095)*	0.129
CGT Quality of decision making	0.9447 ± 0.0267	0.9719 ± 0.020	0.9411 ± 0.0196	χ^2 6.364	0.042*	0.224	0.208	0.014 (cor 0.520)*
CGT Deliberation time log10	3.392 ± 0.044	3.296 ± 0.031	3.337 ± 0.033	χ^2 3.824	0.148	0.051 (cor 0.170)	0.404	0.328
CGT Risk taking	0.5490 ± 0.035	0.5921 ± 0.029	0.5553 ± 0.050	0.387	0.681	0.618	0.917	0.518
CGT Risk adjustment	1.236 ± 0.260	1.322± 0.164	1.113 ± 0.2605	0.069	0.934	0.789	0.919	0.688
CGT Delay aversion	0.1045 ± 0.032	0.0843 ± 0.045	0.0959 ± 0.0704	χ^2 0.325	0.850	0.618	0.642	0.980
CGT Overall proportion bet	0.5159 ± 0.0340	0.5472 ± 0.034	0.5184 ± 0.050	0.227	0.798	0.483	0.966	0.601

N = 42. Means are presented with standard error of mean (SEM). (Cor) denotes P value corrected for multiple comparisons. P values shown relate to two-tailed probability.

Graph 3.1.5 Mean ratio of positive to total omissions in the AGN task

17
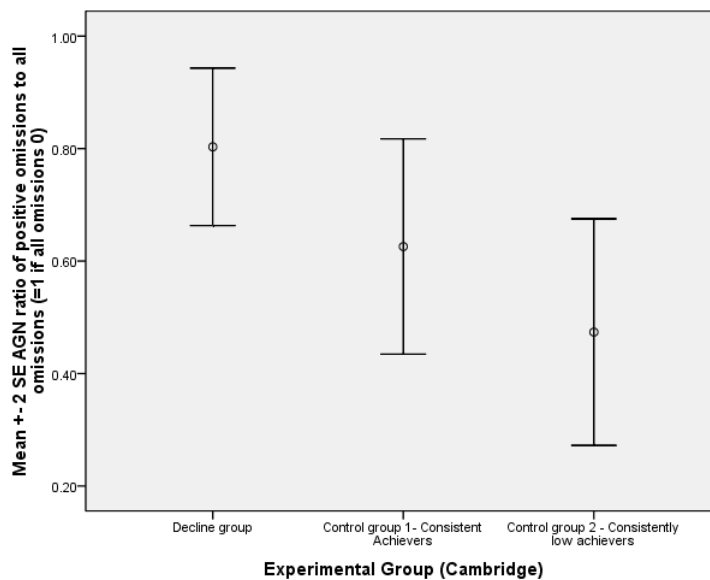
The Affective Go/No-go (AGN)Task

No significant group effects were shown for any of the *AGN* outcome measures as detailed in Table 3.1.2 The Decline group were fastest (had the shortest latencies) and control group 2 were the slowest for all of the correct latency conditions examined, e.g. total latency, positive and negative valence words, respectively. *Post-hoc* comparisons showed that there was a trend towards significance for all the mean latency measures between the Decline and control group 2: total correct latency (DEC vs CTL2 $t(26)$ = -1.932, $p < 0.064$); correct latency shift (DEC vs CTL2 $t(26)$ = -1.862, $p < 0.074$); correct latency non-shift (DEC vs CTL2 $t(26)$ = -1.943, $p < 0.063$); correct latency positive words only (DEC vs CTL2 $t(26)$ = -1.865, $p < 0.073$); correct latency negative words only (DEC vs CTL2 $t(26)$ = -1.966, $p < 0.060$).

As well as being faster than either of the control groups, the Decline group also made the most no responses to a target stimulus (omissions) under all condition assessed, except for negative valence words. Consequently, a variable was constructed which examined the ratio of positive omissions to total omissions (i.e. pos + neg). In the event of total omissions being zero, this ratio was set to 1. The results indicated that 80% of omissions made by the Decline group were for positive words as against 47% for control group 2 (Decline group mean ratio positive to total omissions = 0.8031 ± 0.07; CTL1 mean ratio positive to total omissions = 0.6259 ± 0.096; CTL2 mean ratio positive to total omissions = 0.4737 ± 0.1; (p=0.039), shown in Graph 3.1.5. The decline group thus made significantly more omissions for positive words. Furthermore, the difference in mean positive omissions showed a trend towards significance between the Decline and control group 1 (CTL1 vs DEC $U = 67.500$, $Z = -1.871$, $p = 0.061$). However, the Decline group made the fewest omissions for negative words (*post-hoc* contrasts CTL1 vs DEC $U = 96.000$, $Z = -0.699$, $p = 0.525$, DEC vs CTL2 $U = 60.500$, $Z = -1.717$, $p = 0.086$).

This would suggest that members of the Decline group process and respond to positive and negative valence words differently. More specifically it might suggest they have a deficit in responding to positive target words similar to total omissions but show no deficit and perform well when targeting negative words. In contrast, control group 1 members made the fewest incorrect responses to distracter stimuli (commissions) whereas performance between control group 2 and the Decline group were comparable. All *Post-hoc* comparisons assessing differences between any two groups in commission measures were found not to be statistically significant.

Graded Naming Task

Every participant attempted to name each object during the graded naming task. No significant group effect was shown for the number of objects incorrectly identified (total errors) [$F(2, 39) = 2.375$, $p = 0.106$]. However, the Decline group made the fewest errors and control group 2 the most errors with the difference in performance between these two groups reaching statistical significance (DEC vs CTL2 $t(26)$ = -2.339, $p < 0.027$) shown in Table 3.1.2. These results suggest that the Decline group does not have a deficit in semantic memory.

Cambridge Gambling Task

Only one CGT outcome measure, quality of decision making, showed a significant group effect (χ^2 = 6.364, df = 2, N = 42, p = 0.042) which appears to be driven by better quality of decision making in the control group 1 indicated in Table 3.1.2. Post-hoc comparisons showed that the difference between the control groups is significant (CTL1 vs CTL2 U = 119.000, Z = -2.460, p = 0.027), but not performance between the Decline group and either control groups. In contrast, the Decline group has the longest mean deliberation time, i.e. are the slowest to make a choice on which colour to bet, with control group 1 being the fastest and control group 2 having a mean deliberation time in-between. *Post-hoc* analysis indicated that the difference between the Decline group and control group 1 in deliberation time almost reaches statistical significance (CTL1 vs DEC U = 170.000, Z = -1.954, p = 0.051). The other indices of CGT performance did not differ between the three groups.

Summary of AGN, GNT & CGT task performance

In summary, the Decline group made the fastest responses but made the most omission errors for total correct and positive valence words in the AGN task. Omission errors for negative words were, however, comparable between the groups suggesting that Decline group individuals when asked to trade-off accuracy versus speed, process and respond to positive and negative valence words differently. In contrast, in the CGT task, a test in which participants were not informed that speed would be an assessed measure, the Decline group had the longest deliberation time. Control group 2 made significantly more errors in GNT task and showed the worst quality of decision making in the CGT task. As might be predicted, this would suggest that control group 2 individuals have impaired semantic memory (GNT performance) and lower IQ-related or general cognitive ability (CGT quality of decision making) as compared to the Decline and control group 1 individuals.

3.1.2 Effects of Gender

Table 3.1.3 presents male and female mean scores, SEM, *P*-value and *P*-value after loading for GAT as a covariate, for task outcome measures which showed a statistically significant or trend towards a statistically significant difference between the sexes. First of note, there would appear to be a clear gender difference in childhood GAT scores, with females performing better than males (χ^2 = 4.036, df =1, N = 42, p = 0.045). This result is consistent with findings for the whole cohort (mean males, N = 4165, 44.14 ± 0.25, mean females, N = 4301, 46.35 ± 0.23 [$F_{(1, 8464)}$ = 43.314, p < 0.001], and for the subpopulation (approx 43% of cohort) belonging to the whole cohort three groups (mean males, N = 1917, 40.80 ± 0.36, mean females, N = 1708, 42.82 ± 0.39 [$F_{(1, 3623)}$ = 14.425, p < 0.001].

Therefore, to explore further potential gender differences which could not be explained in terms of GAT differences, the outcome measures were analysed with GAT as a covariate. Consequently, only one task performance measure, CGT risk adjustment, remained highly significant before and after controlling for GAT. As shown in Graph 3.1.6 women showed less risk adjustment than men increasing their bets less as the ratio of red to blue boxes increases. This result that male participants

exhibited a greater modulation of risk-taking in response to probability of winning is consistent with previous reported findings performed using a large, normal (non-clinical) population which ranged in age between 17- 79 years old (Deakin et al., 2004).

Between group gender effects
Gender effects were also examined across the groups using a derived variable, group-gender, which re-grouped participants into 6 categories based on their gender and Decline, control group 1 and control group 2 membership, e.g. category 1 was made of all men belonging to the Decline group. Table 3.1.4 presents the mean of for each group, group-gender and *post-hoc* comparison *p*-values both with and without loading GAT as a covariate. As expected, there was significant group-gender effect for GAT (χ^2 = 17.662, df =5, N = 42, p = 0.003) with better performance by females most prominent for control group 1 (CTL1 males vs CTL1 females U = 7.000, Z = -2.197, p = 0.028).

Only four outcome measures, summed immediate & delayed 1 recall, summed immediate & delayed 2 recall, CGT deliberation time and CGT risk adjustment showed a significant effect after loading for GAT: summed immediate & delayed 1 recall [$F_{(5, 35)}$ = 3.756, p = 0.008]; summed immediate & delayed 2 recall [$F_{(5, 35)}$ = 3.721, p = 0.008]; CGT deliberation time (log10) [$F_{(5, 35)}$ = 4.284, p = 0.006]; CGT risk adjustment [$F_{(5, 35)}$ = 2.240, p = 0.040].

The group-gender differences in word recall performance again appears to be driven by a difference between males and females in control group 1. However, post-hoc comparisons between the groups do not reach significance (shown in Table 3.1.4). Similarly, men and women in control group 1 also differed in deliberation time during the CGT task with men being significantly faster than women in making a choice (p = 0.005). In contrast, the gender difference in risk adjustment performance discussed above was more evident in the control group 2 and the Decline group (DEC males vs DEC females p = 0.006; DEC males vs DEC females p = 0.044).

Comparisons of performance on all task measures also revealed a significant difference in letter cancellation omissions (accuracy) scores between Decline group males and females. Men in the Decline group made significantly more errors, i.e. were less accurate than women, although this effect shows only a trend in significance when loading for GAT suggesting that childhood ability could be a contributing factor. Finally, five RVIP outcome measures, total hits, total misses, probability of hit, probability of false alarm and RVP B' showed a significant group-gender effect before, but not after, controlling for GAT score. Once again, the difference between males and females on these measures was most marked in control group 1 with men making more correct response hits, fewer target misses, having a higher probability of a hit, a lower probability of a false alarm (see Table 3.1.4 for *post-hoc* comparisons).

Table 3.1.3 Male and female mean scores, SEM, *P*-value and *P*-value after loading for GAT

OUTCOME MEASURE	Mean Total Men (N=24)	Mean Total Women (N=18)	Gender difference statistic	Gender difference (*P* value)	ANCOVA *P* value minus GAT
General ability score	40.00 ± 3.004	50.33 ± 3.415	χ² 4.036	0.045*	
Word recall 2 delayed 1	4.46 ± 0.37	5.50 ± 0.36	3.833	0.057	0.164
Summed recall 1 & 2	10.42 ± 0.59	12.06 ± 0.65	3.424	0.072	0.153
Summed recall 1 & 3	10.33 ± 0.54	11.78 ± 0.61	3.083	0.087	0.178
Summed recall 2 & 3	8.83 ± 0.69	10.72 ± 0.67	3.717	0.061	0.177
AGN Mean correct latency shift	526.12± 17.40	480.18± 13.12	3.949	0.054	0.119
CGT Risk adjustment	1.573 ± 0.184	0.827 ± 0.148	9.01	0.005**	0.002**

Graph 3.1.6. CGT risk adjustment scores as a function of sex
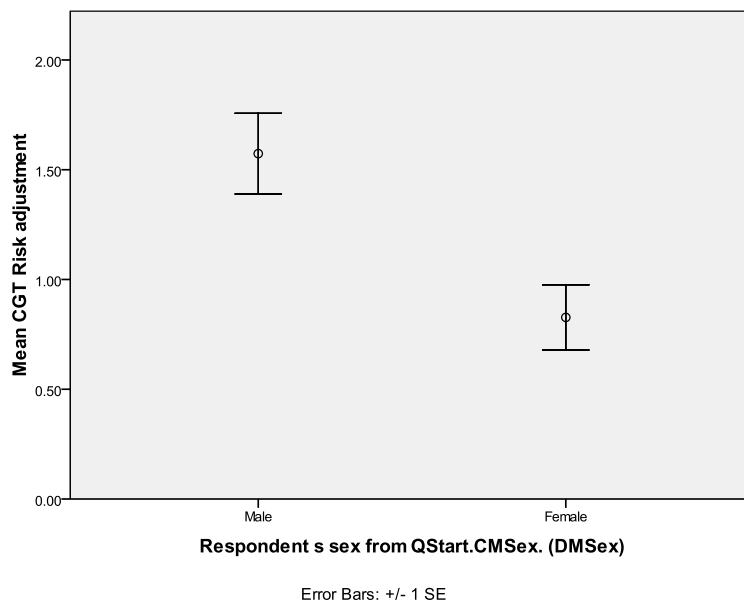

Error Bars: +/- 1 SE

21

Table 3.1.4: Gender effects for each group.

Measure	Decline Men (N=9)	Decline Women (N=7)	Cont 1 Men (N=8)	Mean Cont 1 Women (N=6)	Cont 2 Men (N=7)	Mean Cont 2 Women (N=5)	Gender/group statistic	Group/Gen (p value)	ANCOVA p minus GAT	Decline M vs W (p)	CTL 1 M vs W (P)	CTL 2 M vs W (p)	Decline M vs W (p) adj GAT	CTL1 M vs W (p) adj GAT	CTL2 M vs W (p) adj GAT
GAT	47.11 ± 4.797	54.43 ± 5.051	45.25 ± 4.165	57.33 ± 3.703	28.29 ± 4.196	36.20 ± 5.809	χ^2 17.662	0.003**		0.266	0.028	0.416			
Summed recall 1 & 2	8.89 ± 1.02	10.14 ± 0.40	11.75 ± 0.73	14.17 ± 1.20	10.86 ± 0.76	12.20 ± 1.20	6.264	0.004**	0.008**	0.321	0.093	0.435	0.514	0.224	0.548
Summed recall 1 & 3	8.89 ± 0.98	10.00 ± 0.49	11.63 ± 0.73	13.67 ± 1.02	10.71 ± 0.87	12.00 ± 1.23	6.434	0.004**	0.008**	0.369	0.120	0.396	0.544	0.287	0.531
Let can omissions	6.67 ± 2.630	1.86 ± 1.056	3.00 ± 0.107	3.50 ± 0.992	4.29 ± 0.747	4.00 ± 1.517	χ^2 4.161	0.526	0.315	0.049*	0.742	0.868	0.075	0.445	0.814
RVP total hits	16.11 ± 2.071	16.71 ± 1.443	21.00 ± 1.350	19.50 ± 2.668	13.86 ± 1.280	14.20 ± 0.970	2.535	0.046*	0.080	0.825	0.598	0.847	0.674	0.092	0.604
RVP total misses	10.89 ± 2.071	10.29 ± 1.443	6.00 ± 1.350	7.50 ± 2.668	13.00 ± 1.272	12.80 ± 0.970	2.482	0.050	0.082	0.825	0.598	0.910	0.679	0.095	0.568
RVP P Hit	0.5967 ± 0.077	0.6190 ± 0.053	0.7778 ± 0.050	0.7222 ± 0.099	0.5159 ± 0.047	0.5259 ± 0.036	2.509	0.048*	0.081	0.825	0.598	0.878	0.667	0.093	0.586
RVP PFA	0.0049 ± 0.004	0.0084 ± 0.003	0.0024 ± 0.002	0.0072 ± 0.001	0.0122 ± 0.005	0.0032 ± 0.002	χ^2 11.089	0.050	0.352	0.071	0.032*	0.152	0.285	0.161	0.164
CGT DT log10	3.4401 ± 0.070	3.3302 ± 0.043	3.2389 ± 0.020	3.3722 ± 0.054	3.3469 ± 0.050	3.3442 ± 0.022	χ^2 9.533	0.090	0.006**	0.315	0.028*	0.639	0.276	0.005*	0.763
CGT Risk adjustment	1.714 ± 0.348	0.620 ± 0.262	1.425 ± 0.220	1.185 ± 0.256	1.561 ± 0.405	0.687 ± 0.190	2.185	0.077	0.040*	0.032*	0.491	0.118	0.006**	0.347	0.044*

Table 3.1.4. Gender effects for each group. Group means, group interaction statistic and p values and *post-hoc* comparisons with and without loading for GAT are presented.

A summary of the main pilot study CANTAB task findings are presented in Box 3.11 below.

Box 3.11 Summary of main CANTAB task findings

- The Decline group and control group 2 performed significantly worse than control group 1 individuals in the PAL and RVIP CANTAB tasks. Performance in these tasks, which assess working memory, spatial associative learning and attention, was comparable between the Decline and control group 2 individuals.

- The Decline group made the fastest responses, but made the most omission errors for total correct and positive valence words in the AGN task.

- Omission errors for negative words in the AGN task were, however, similar between the groups suggesting that Decline group process and respond to positive and negative valence words differently.

- The Decline group showed the longest deliberation time in the CGT task.

- Control group 2 individuals made significantly more errors in GNT task (semantic memory) and showed the worst quality of decision making in the CGT task.

- The Decline group showed no deficit in the GNT task indicating that semantic memory is not impaired.

- Male participants exhibited a greater modulation of risk-taking in response to probability of winning in the CGT.

3.2 Factor analysis

Factor analysis of eleven key neuropsychological task outcome measures

Correlation matrices were first used to explore interrelationships between different test score outcome measures with the aim to identify the main test variables to be used in factor analysis. Consequently, eleven variables representing the key outcome indices for all tests were selected and interrelationships re-examined (shown in the correlation matrix, Appendix 3).

A summary of loadings for eleven key variables following principal component analysis together with their correlations with GAT is presented in Table 3.2.1 ($N =$ 42). A two-factor solution was derived with eigenvalues of 4.290 and 2.511 which represents 61.8% of the variance. The first factor accounts for co-variation in increased deliberation time, letter cancellation omissions, AGN total omissions, PAL total trials to success, and RVP mean latency; and, decreased quality of decision making, summed recall scores, RVP A' and AGN mean correct latency (the relationships are diagrammatically represented in Figure 3.2.1). Thus it would appear that factor 1 consistently shows a correlation between 'poor' performances across these tasks with the exception of AGN mean latency in which increases in speed of correct AGN responses relate to poor performance in the other variables. In contrast, the second factor accounts for decreases in the number of animals named (animal naming task), and RVP A' and increases in total errors made during the graded naming test, mean RVP latency and AGN correct latency.

Factor 2 was found to have a significant negative correlation with childhood GAT ($r =$ -0.692, $p = 0.000$) while factor 1 showed no correlation with GAT ($r = -0.132$, $p =0.406$) as detailed in Table 3.2.2. Moreover, when examining these factors as variables which may differ between the groups, i.e. decline and control groups, a significant difference was found for factor 2 [$F(2, 39) = 5.674$, $p = 0.007$] but not for factor 1 ($\chi^2 = 4.443$, df = 2, $N = 42$, , $p = 0.108$). However, after loading GAT as a covariate (hence to eliminate the possibility that differences between the groups is being driven mostly by childhood GAT), there is no longer a statistical difference between the groups for factor 2, but the difference between the groups for factor 1 now shows a trend towards significance. Of interest, the Decline group has the highest mean for factor 1 and the group mean differs significantly from the mean of control group 2. Therefore, a possible interpretation of the factor structure is that factor 2 represents general learning and semantic memory ability whereas factor 1 represents attentional processing and learning which may be pertinent as a marker/measure of decline relevant to mild cognitive impairment in mid-life.

Table 3.2.1: Summary of loadings for eleven key task outcome measures on factors 1 and 2 following factor analysis.

CGT DT; CGT deliberation time; CGT QDM, CGT quality of decision making; LC omis, letter cancelation omissions; AGN omis, AGN omissions; AGN cor lat, AGN mean correct latency; PAL t trials, PAL total trial; recall, summed immediate and delayed 1 recall; Animal nam, Animal naming score; GNT err, GNT total errors; RVP A', RVP A prime; RVP lat, RVP mean latency.

	CGT DT	CGT QDM	LC omis	AGN omis	AGN cor lat	PAL t trials	recall	Animal nam	GNT err	RVP A'	RVP Lat
Factor 1	0.820	0.782	-0.782	0.774	-0.734	0.706	-0.454	-0.148		-0.522	0.503
Factor 2	0.355		0.291	0.257	0.467	0.368	-0.244	-0.749	0.710	-0.655	0.571

Figure 3.2.1 Component plot for eleven main task outcome measures following factor analysis.

Component Plot in Rotated Space

Table 3.2.3: Correlations with childhood GAT and differences between the groups with and without loading for GAT for factors 1 & 2.

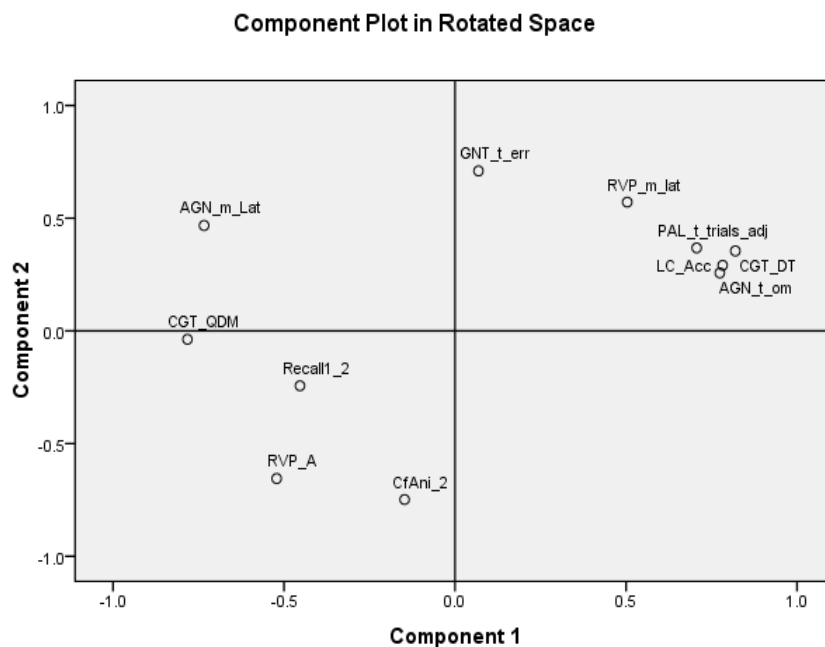|	Childhood GAT (r)	Difference groups	GAT effect (F)	Dif groups after adj for GAT	Mean DEC (adj GAT)	Mean CTL1 (adj GAT)	Mean CTL2 (adj GAT)	DEC vs CTL 1 (adj GAT p)	DEC vs CTL 2 (adj GAT p)	CTL1 vs CTL2 (adj GAT p)
Factor 1	-0.132 (p) 0.406	χ^2 4.443 (p) 0.108	F 2.857 (p) 0.099	F 2.883 (p) 0.068	0.469 ± 0.248	-0.119 ± 0.265	-0.487 ± 0.320	0.102	0.032*	0.408
Factor 2	-0.692 (p) 0.000	F 5.674 (p) 0.007	F 20.654 (p) 0.000	F 0.751 (p) 0.479	0.060 ± 0.190	-0.205 ± 0.203	0.160 ± 0.246	0.331	0.764	0.286

Table 3.2.3. Correlations with childhood GAT and differences between the groups with and without loading for GAT for factors 1 & 2.

3.3 Analysis of demographic, epidemiological and pilot study questionnaire variables

3.3.1 Whole cohort groups

Survey variables examined for differences between the whole cohort groups (~ 43% of NCDS cohort) included socio-economic status, educational qualifications, obstetric factors, physical and mental health & quality of life or wellbeing, alcohol-drinking behaviour, body mass index and childhood adversity factors. Tables 3.3.1.1 and 3.3.1.2 summarise the main results.

Socio-economic status

Using a simplified 5 category scale derived from the NCDS8 survey variable n8nssec, the three groups show a significantly difference in socio-economic status, (χ^2 = 75.121, df = 8, N = 3616, p = 0.000). As predicted the consistent achievers and Decline group had a significant higher percentage of professional/managerial individuals [DEC vs CTL2 (χ^2 = 33.289, df = 1, N = 177, p = 0.0001); CTL1 vs CTL2 χ^2 = 38.516, df = 1, N = 195, p = 0.0001)] than control group 2, but no difference in percentage of individuals between control group 1 and the Decline group. At the other end of the scale, control group 2 and the Decline group have a significantly higher percentage of individuals who are not working as compared to control group 1, [CLT1 vs DEC (χ^2 = 15.136, df = 1, N = 419, p = 0.0001); CTL1 vs CTL2 (χ^2 = 13.383, df = 1, N = 385, p = 0.0003)] .

Looking more closely at reasons for 'not working', there is no difference between the groups for the explanatory answer 'looking after family', but a significant difference in numbers between the control groups (control group 2 higher) for 'unemployed and actively seeking work' [CTL1 vs CTL2 (χ^2 = 4.667, df =1, N = 2289, p = 0.031)]. There is also a very clear difference in groups in percentage of individuals not working owing to temporary or permanent sickness/disability with control group 2 and the Decline group having significantly more individuals off work owing to ill health ([CLT1 vs DEC (χ^2 = 27.444, df = 1, N = 2552, p = 0.0001); CTL1 vs CTL2 (χ^2 = 14.674, df = 1, N = 2289, p = 0.0001)].

Table 3.3.1.1 Summary of socio-economic status, educational qualifications, obstetric factors and BMI variables as a function of whole cohort group status.

Factor	Decline *N*, Mean or %	Control 1 *N*, Mean or %	Control 2 *N*, Mean or %	Group difference statistic	Group (*p*)	Effect size (*r*)
Socio-economic Class						
NS-SEC NCDS8	(*N*= 1225)	(*N*= 1327)	(*N*= 1064)	χ² 75.121	0.000	0.049
Higher professional/managerial	10.8%	11.3%	4.2%			
Lower managerial / intermediate	39.1%	44.5%	38.3%			
Small employers / Lower supervisory/technica	20.2%	19.7%	24.2%			
Routine / Semi-routine occupations	10.4%	10.9%	14.4%			
Not-working	19.4%	13.6%	19.2%			
Not-working category	(*N*= 221)	(*N*= 156)	(*N*= 189)	χ² 37.437	0.001	0.051
Unemployed and seeking work	3.3%	2.3%	3.8%			
Retired	0.7%	0.8%	0.4%			
Looking after home/family	5.0%	4.8%	6.1%			
Sick/disabled	9.0%	3.9%	7.9%			
Highest Educational qualification						
NS-SEC NCDS8 (derived)	(*N*= 1229)	(*N*= 1330)	(*N*= 1066)	χ² 234.64	0.000	-0.151
Higher degree	2.3%	3.7%	0.7%			
Degree	13.1%	15.9%	4.2%			
Diploma	4.2%	5.0%	3.5%			
2 A' levels, Scottish highers	5.6%	7.3%	3.5%			
1 A' level, 2 AS	0.6%	0.5%	0.3%			
GCSE, Scottish standards	34.3%	36.2%	29.5%			
Other Scottish qualifications	15.3%	13.6%	21.4%			
No qualifications	24.6%	17.7%	37.0%			
Qualifications gained over the last 10yrs	(*N*= 1171)	(*N*= 1278)	(*N*= 1023)			
Yes	26.4%	34.3%	28.4%	χ² 19.582	0.000	0.021
Obstetric factors	(*N*= 1095-1175)	(*N*= 1164-1270)	(*N*= 882-993)			
Birth weight in ounces	117.490 ± 0.542	118.153 ± 0.524	115.388 ± 0.590	F 6.496	0.002	
Premature births	4.4%	3.6%	5.1%	χ² 2.752	0.253	-0.012
Smoking during pregnancy	31.4%	32.0%	39.5%	χ² 18.718	0.000	0.066
Body mass index	(*N*= 1166)	(*N*= 1277)	(*N*= 1023)			
BMI mean	27.066 ± 0.161	27.419 ± 0.127	27.889 ± 0.171	F 6.005	0.002	
BMI categorised (obese only)	23.4%	24.1%	28.7%			

Educational qualifications

The variable, highest educational qualification attainment, was derived from the NCDS5 (1991) survey variable 'highest academic qualification' which was then updated by information acquired in subsequent 2000, 2004 and 2008 surveys. The three groups were shown to significantly differ from each other in highest educational qualification attainment, (x^2 = 234.640, df =14, N = 3625, p = 0.000). As expected, control group 1 has the highest percentages of individuals with degrees and higher degrees, control group 2 the lowest and the Decline group a percentage in the middle. All differences between each group for this category reaches statistical significance [CTL vs DEC (x^2 = 7.059, df = 1, N = 2559, p = 0.007); DEC vs CTL2 (x^2 = 65.857, df =1, N = 2295, p = 0.000); CTL1 vs CTL2 (x^2 = 92.997, df =1, N = 2396, p = 0.000)].

This pattern is reversed when examining the percentage of individuals with no qualifications with control group 2 having significantly more 'no qualification' individuals then the Decline group whilst the Control 1 group has the lowest percentage overall [CTL vs DEC (x^2 = 717.529, df = 1, N = 2559, p = 0.000); DEC vs CTL2 (x^2 = 21.660, df = 1, N = 2295, p = 0.000); CTL1 vs CTL2 (x^2 = 64.360, df = 1, N = 2396, p = 0.000)]. Similar results, i.e. a group effect and between group category-related comparisons, were obtained when examining a second qualification derived variable which includes all vocational qualifications as well as academic qualifications attained (results not shown).

An additional qualification variable was derived to estimate the percentage of individuals who attained an academic or vocational qualification between 2000 and 2008/9. A significant difference in percentage that attained a qualification over the last ten years is shown for the groups (x^2 = 19.582, df = 2, N = 3472, p = 0.000). Interestingly, although control group 1 has the highest percentage as might be expected, there is no difference between the Decline and control group 2 as shown in Graph 3.3.1.1 This may suggest that the behaviour of the Decline group is already changing.

Birth weight

A significant difference between the groups was found for birth weight [$F_{(2, 3381)}$ = 6.496, p = 0.002]. Comparisons between individual groups indicated that control group 2 had a significantly lower mean birth weight as compared to the other two groups [CTL1 vs DEC $t_{(2402)}$ = -0.884, p < 0.377; DEC vs CTL2 $t_{(2139)}$ = 2.611, p < 0.009; CTL1 vs CTL2 $t_{(2221)}$ = 3.506, p < 0.000] as shown in Graph 3.3.1.2. This finding is consistent with previous reported research indicating a positive association between birth weight and cognitive ability in childhood and later in life (Richards et al., 2001, 2002).

Graph 3.3.1.1 Percentage of individuals who attained an educational qualification over the last ten years.
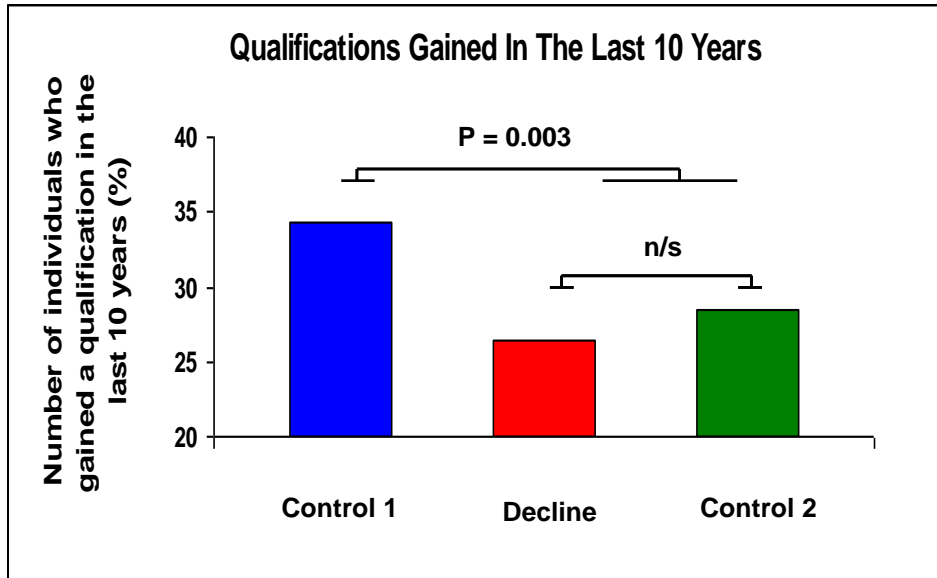
Graph 3.3.1.2 Mean birth weight (in ounces) as a function of group status.
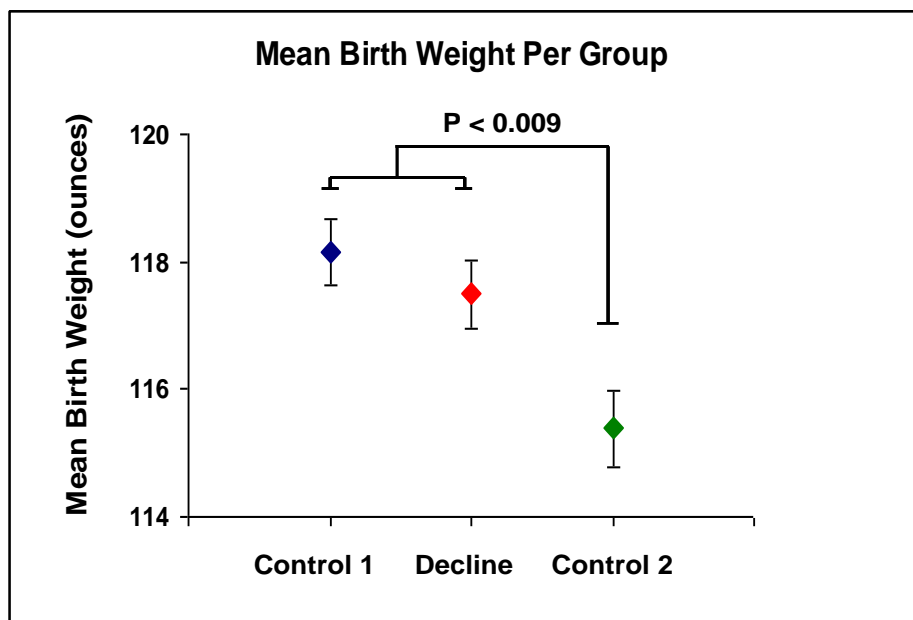
Prematurity

A derived variable which denotes whether an individual was born earlier than 37 weeks (gestation age) was used to assess differences between the groups in prematurity and whether there exists a correlation between prematurity and low birth weight. The three groups do not differ in percentage of premature births (χ^2 = 2.752, df = 2, N = 3141, p = 0.253). However, being born premature did show a significant positive correlation with birth weight (r = 0.272, p = 0.000).

Smoking during pregnancy

The percentage of mothers who smoked after the fourth month of pregnancy does significantly differ between the three groups (χ^2 = 18.963, df = 2, N = 3438, p = 0.000). Control group 2 has the highest percentage of individuals whose mother's smoked with the difference in percentages reaching significance when comparing this group with both the Decline group and control group 1 [DEC vs CTL2 (χ^2 = 21.660, df =1, N = 2295, p = 0.000); CTL1 vs CTL2 (χ^2 = 64.360, df =1, N = 2396, p = 0.000)]. Furthermore, smoking after the fourth month of pregnancy shows a significant negative correlation with birth weight (r = -0.150, p =0.000) suggesting that smoking later in pregnancy could be a potential contributing factor to lower birth weight.

Body Mass Index (BMI)

A significant difference between the three groups was found for mean BMI estimated at the age of 50 [$F_{(2, 3463)}$ = 6.005, p = 0.002]. The Decline group had the lowest mean BMI and control group 2 the highest with both the Decline group and control group 1 significantly differing in means from control group 2 [CTL1 vs DEC $t(2441)$ = -1.560, p < 0.119; DEC vs CTL $t(2187)$ = -3.492, p < 0.000; CTL1 vs CTL2 $t(2298)$ = -2.015, p < 0.044]. Body Mass Index (BMI) was also analysed by binning the age 50 BMI value into four categories: underweight, normal, overweight, obese. No group effect was evident when examining these categories (χ^2 = 10.635, df = 1, N = 3466, p = 0.100) However, a significantly higher percentage of control group 2 belong to the category of 'obese' as compared to the Decline group and control group 1 [CTL vs DEC (χ^2 = 0.169, df = 1, N = 2443 p = 0.624) ns; DEC vs CTL2 (χ^2 = 8.052, df = 1, N = 2189 p = 0.005); CTL1 vs CTL2 (χ^2= 6.274, df =1, N = 2300, p = 0.012)].

Exercise

A difference in the amount of exercise undertaken was explored by analysing the percentage of individuals who self reported undertaking regular physical exercise at least once a month at the age of 50. The groups show an overall difference in percentage of exercise self reported (χ^2 = 13.744, df = 2, N = 3625 p = 0.001) with control group 1 having a higher percentage than both the Decline group and control group 2 [CTL vs DEC (χ^2 = 7.051, df = 1, N = 2559, p = 0.008) ns; DEC vs CTL2 (χ^2

= 0.969, df =1, N = 2295, p = 0.325 ns); CTL1 vs CTL2 (χ^2 = 12.721, df = 1, N = 2396, p = 0.000)]. This result is consistent with data for the whole of the cohort population which shows that individuals at the age of 50 who exercised at least once a month performed significantly better on all NCDS8 cognitive tests including word recall than those who did not (Brown and Dodgeon, 2010). However, variables relating to amount of physical exercise undertaken during childhood and more specifically at the at the ages of 11 and 16, do not show differences in frequency of exercise between the three groups, although control group 2 consistently has the lowest percentage of individuals reporting undertaking frequent ('daily'/ 'sometimes') exercise (results not shown). This would suggest a lack of relationship between amount of physical exercise in early life and cognitive decline later in mid-life.

Medical/health conditions

The percentage of individuals registered disabled or report long-term disability/health problems was found to differ between the three groups (χ^2= 25.834, df = 4, N = 3066, p < 0.001). A significantly higher percentage of individuals in control group 1 reported no disability or long-term health issues as compared to the Decline and control 2 group, with the Decline group having the highest percentage of all groups [CTL vs DEC (χ^2 = 16.537, df = 1, N = 2200, p = 0.000); DEC vs CTL2 (χ^2 = 2.699, df =1, N = 1895, p = 0.1004) ns; CTL1 vs CTL2 (χ^2 = 64.795, df = 1, N = 2037, p = 0.029)].

Table 3.3.1.2 Summary of percentages and means relating to exercise and health variables as a function of whole cohort group status.

Factor	Decline N, Mean or %	Control 1 N, Mean or %	Control 2 N, Mean or %	Group difference statistic	Group (p)	Effect size (r)
Exercise	(N= 1229)	(N= 1330)	(N= 1066)			
NCDS8 regular exercise	73.2%	77.7%	71.4%	χ^2 13.744	**0.001**	0.014
Physical health						
Registered disabled or long-term health issues	(N= 1029)	(N= 1171)	(N= 866)	χ^2 25.834	**0.000**	0.034
Registered disabled	5.2%	2.2%	4.6%			
Long-term illness/disability	41.6%	36.0%	38.5%			
No long-term health problems	53.2%	61.7%	56.9%			
SF-36 survey physical health measures	(N= 1073-1113)	(N= 1213-1239)	(N= 929-962)			
Physical functioning	83.060 ± 0.760	87.316 ± 0.572	82.141 ± 0.862	χ^2 12.500	**0.002**	
Role-limitations due to physical health	79.963 ± 1.084	83.711 ± 0.918	79.952 ± 1.148	χ^2 6.855	**0.032**	
Pain score	75.126 ± 0.787	78.502 ± 0.675	72.633 ± 0.891	χ^2 21.359	**0.000**	
General health	65.613 ± 0.701	69.656 ± 0.617	65.178 ± 0.742	χ^2 28.465	**0.000**	
Diabetes	(21) 1.9%	(21) 1.7%	(17) 1.8%	χ^2 3.785	0.706	
Epilepsy	(18) 1.5%	(4) 0.3%	(13) 1.2%	χ^2 10.067	**0.007**	
High blood pressure	(203) 16.5%	(187) 14.1%	(194) 18.2%	χ^2 7.728	**0.021**	
Mental health						
Malaise inventory (high score)	(218) 18.0%	(184) 13.9%	(206) 19.6%	χ^2 14.571	**0.001**	
	(N= 1055-1078)	(N= 1195-1216)	(N= 911-936)			
SF-36 survey mental health measures						
Emotional well-being score	74.303 ± 0.575	75.692 ± 0.514	72.512 ± 0.634	F 7.675	**0.000**	
Role-limitations due to emotional problems	83.862 ± 1.005	86.683 ± 0.846	83.118 ± 1.092	F 3.891	**0.021**	
Social functioning score	54.244 ± 0.552	55.339 ± 0.492	54.617 ± 0.603	F 1.118	0.327	
Energy/fatigue score	59.472 ± 0.676	61.228 ± 0.591	60.173 ± 0.381	F 2.338	0.097	
CASP-12- Quality of life	25.638 ± 0.185	26.380 ± 0.164	25.187 ± 0.191	F 11.449	**0.000**	
WEMWB survey	48.486 ± 0.264	49.473 ± 0.238	48.177 ± 0.273	F 7.366	**0.001**	

The SF-36 survey is a multi-purpose, short-form health survey with 8 outcome measures, which are said cluster into two groups, one relating to physical health and the second to mental health issues. All four SF-36 physical health outcome variables derived from NCDS8 survey data (physical functioning, role-limitations due to physical health, pain score and general health) show differences between the three groups in mean outcome scores (presented in Table 3.3.2). Moreover, for all these measures, and consistent with the results above relating to long-standing illnesses, control group 1 has the highest mean, i.e. indicating the best self-reported physical health, and which significantly differs from the other two groups.

Looking at individual diseases and conditions can be problematic because numbers may be very small. Nevertheless for disorders that have a relatively prevalence in the general population, it is clear that certain conditions, such as epilepsy and high blood pressure, show the same pattern described above, but not for other common disorders such as diabetes (see Table 3.3.1.2 for further details).

Mental health variables and quality of life/mental well being survey scores

The 2008/9 survey also included a nine item version of the Malaise Inventory, a survey designed to assess mental health issues. A Malaise Inventory score of 4 or higher is considered to be a sign that an individual is experiencing symptoms associated with depression. The three groups significantly differ in the percentage of individuals with either low or high binned scores, (χ^2 = 14.571, df = 2, N = 3583, p = 0.001). More specifically, the Decline and control group 2 have similar percentages of individuals with a high score, which is significantly higher than the mean for control group 1 [CTL vs DEC (χ^2 = 7.550, df = 1, N = 2530 p = 0.006); DEC vs CTL2 (χ^2 = 0.786, df = 1, N = 2263, p = 0.375) ns; CTL1 vs CTL2 (χ^2 = 13.082, df = 1, N = 2373, p = 0.000)] – suggesting these two groups have a higher incidence of depression. The present results are in support of previous data indicating that elevated depressive symptoms are associated with significantly lower word list recall scores in individuals over 50 and also associated with declines in episodic learning and memory over time (Gonzalez et al., 2008).

However, when examining the four SF-36 survey mental health outcome measures, only two of the four (emotional well-being and role-limitations due to emotional problems) show significant differences in mean scores between the three groups (presented in Table 3.3.2). Again, control group 1 has a higher mean score suggesting better self-reported emotional well-being than the other two groups, (emotional well-being [CTL1 vs DEC t(2288) = -1.808, p < 0.071; DEC vs CTL2 t(2006) = 2.097, p < 0.036; CTL1 vs CTL2 t(2148) = 3.896, p < 0.000]; role-limitations due to emotional problems [CTL1 vs DEC t(2284) = -2.147, p < 0.032; DEC vs CTL2 t(1997) = 0.502, p < 0.616 ns; CTL1 vs CTL2 t(2139) = 2.582, p < 0.010]. Furthermore, two additional surveys, measuring mental wellbeing and quality of life, were administered during the 2008/9 sweep. The mean total scores for both the CSAP-12 and Warwick Edinburgh Mental Well-Being Scale showed same significant pattern of differences between the groups (see Table 3.3.2), that is, control group 1 has a significantly higher mean than the Decline group and control group 2 (CASP-12

[CTL1 vs DEC t(2248) = -3.017, p < 0.003; DEC vs CTL2 t(1969) = 1.692, p < 0.091; CTL1 vs CTL2 t(2109) = 4.755, p < 0.000]; WEMWBS [CTL1 vs DEC t(2251) = -2.844, p < 0.004; DEC vs CTL2 t(1965) = 0.812, p < 0.417 ns; CTL1 vs CTL2 t(2106) = 3.663, p < 0.000].

Alcohol consumption behaviour

The 2008/9 AUDIT 'Alcohol Use Disorders Identification Test' scores were binned into 3 categories: unproblematic drinking, harmful drinking and alcohol dependency indicated in Table 3.3.1.3. A significant difference between the three groups across the categories is apparent [(χ^2 = 9.550, df = 4, N = 3310, p = 0.049)]. The most notable difference is in the category 'alcohol dependency' in which the Decline group has a higher percentage of individuals than either control group. However, these differences do not reach statistical significance [CTL vs DEC (χ^2 = 3.115, df = 1, N = 2353, p = 0.078) trend; DEC vs CTL2 (χ^2 = 0.556, df =1, N = 2064, p = 0.0.4557 ns); CTL1 vs CTL2 (χ^2 = 0.885, df = 1, N = 2203, p = 0.347 ns)].

Childhood adversity questions

A 16-item survey of childhood adversity measures adapted from the Australian Path Through Life Study (Rosenman and Rodgers, 2004, 2006) was administered to cohort members during the biomedical survey. Items in this survey range from questions on quality of relationships with parents to different types of child abuse. Total derived score ranges from 0 to 16, although no cohort member reported a total of 16 and information concerning individual items was not made available for analysis. As shown in Table 3.3.1.3, there is statistically significant difference in the percentages of individuals from each group reporting total number of childhood adversities experienced (χ^2 = 47.433, df = 28, N = 2975, p = 0.012). Breaking this down further and starting with no reported adversities, the groups show a difference in the total number of individuals reporting having experienced no adversities with control group 1 having a significantly higher percentage than the Decline group and control group 2 [CTL vs DEC (χ^2 = 3.987, df = 1, N = 2146, p = 0.046); DEC vs CTL2 (χ^2 = 0.000, df = 1, N = 1841, p = 0.999 ns); CTL1 vs CTL2 χ^2 (1, N = 1963, = 3.560, p = 0.059)].

At the other end of scale, there is also a significant difference in percentages of individuals between the groups who report experiencing 8 or more adversities. Again, the percentages between the Decline and control 2 are similar and are significantly higher than then the control 1 group as shown in graph 3.3.3. CTL1 vs DEC (χ^2 = 5.773, df = 1, N = 2146 p = 0.0163); DEC vs CTL2 (χ^2 = 1.573, df =1, N = 1841, p = 0.2097 ns); CTL1 vs CTL2 (χ^2 = 12.924, df = 1, N = 1963, p = 0.0003)].

Table 3.3.1.3: Summary of percentages and p values relating to alcohol consumption behaviour and childhood adversity variables as a function of whole cohort group status.

Factor	Decline *N, %*	Control 1 *N, %*	Control 2 *N, %*	Group difference statistic	Group (*p*)	Effect size (*r*)
Alcohol drinking behaviour						
AUDIT score age 50 (categorized)	(*N*= 1107)	(*N*= 1246)	(*N*= 957)	χ^2 9.550	**0.049**	-0.021
Unproblematic	79.3%	82.4%	80.4%			
Harmful	15.5%	14.8%	15.9%			
Dependency	5.1%	2.8%	3.8%			
Childhood adversities						
Path Through Life Scale	(*N*= 1012)	(*N*= 1134)	(*N*= 829)	χ^2 47.433	**0.012**	0.028
No adversities	42.10%	46.40%	42.10%			
0-7 adversities	96.20%	98.00%	95.00%			
8-16 adversities	3.80%	2.00%	5.00%			

Graph 3.3.1.3 Percentage of individuals who self-report experiencing 8 or more adversities during childhood.
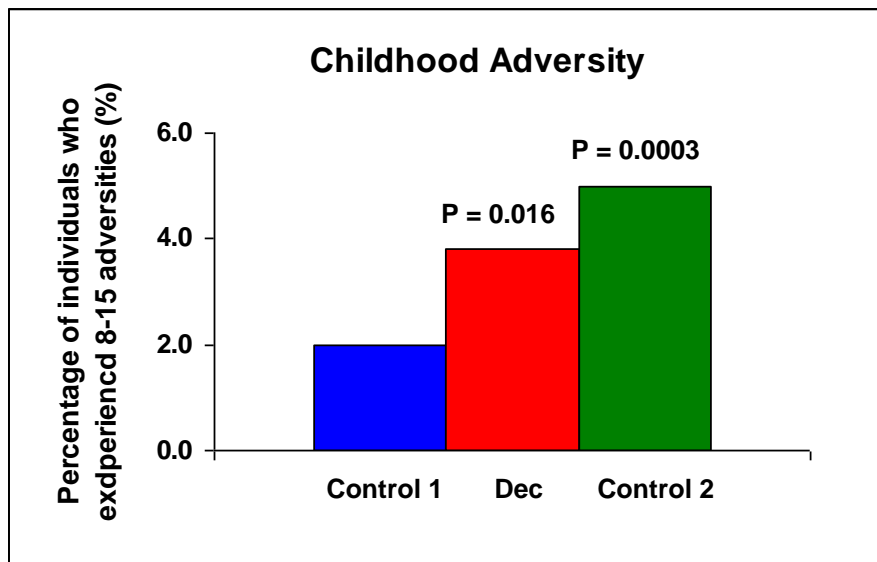
3.3.2 Cambridge Pilot study groups

Tables 3.3.2.1 and 3.3.2.2 present the results of analysis of the same variables as described in the previous section, but examined for differences between the Cambridge area pilot study groups ($N = 42$). Exploring data relating to such small numbers is difficult to interpret especially when variable outcomes involve division into categories. Nevertheless, three observations are worthy of note. First, similar to the findings for the whole cohort groups, the pilot study control group 2 has the lowest mean birth weight [CamCTL1 vs CamDEC $t(27) = 1.382$, $p < 0.178$ ns; CamDEC vs CamCTL2 $t(25) = 2.332$, $p < 0.028$; CamCTL1 vs CamCTL2 $t(24) = 0.922$, $p < 0.366$ ns].

Second, there is a significant difference between the three pilot study groups in percentage of individuals registered disabled or who have a long-term health problem ($\chi^2 = 9.447$, df = 2, $N = 37$, $p = 0.009$). However, in contrast to findings for the whole cohort groups which indicated control group 1 to have the lowest percentage of individual registered disabled or with long-term conditions, in the pilot study control group 1 has the highest percentage and the control group 2 by far the lowest [CamCTL1 vs CamDEC $p = 0.424$ ns); CamDEC vs CamCTL2 $p = 0.041$); CamCTL1 vs CamCTL2, $p = 0.004$) Fisher's exact test]. This might suggest that the Cambridge pilot study population does not represent the whole cohort group population in at least with respect to medical health.

Third, four individuals belonging to the pilot study Cambridge Decline group have high Malaise scores whereas only one from control group 1 and none from the control group 2 also have scores that are classified as high. However, the responses given during the pilot study questionnaire to a question asking about whether cohort members have consulted their GP or a psychiatrist about feeling depressed or anxious in the past 12 months, indicated 4 individuals, 1 in the Decline group, 1 in the control group 1 and 2 in the control group 2, had indeed answered yes. As only one individual had both a high Malaise score and said yes to the recent questionnaire item, it is possible that depressive symptoms in the remaining 8 cohort members (who either had a high Malaise score or said they had visited a GP or psychiatrist about feeling depressed) may have fluctuated over the last two years.

Table 3.3.2.1 Summary of socio-economic status, educational qualifications, obstetric factors and BMI variables as a function of pilot study group status.

Factor	Decline *N*, Mean or %	Control 1 *N*, Mean or %	Control 2 *N*, Mean or %	Group difference statistic	Group (*p*)
Socio-economic Class					
NS-SEC NCDS8	(*N*= 16)	(*N*= 14)	(*N*= 12)	χ^2 10.938	0.205
Higher professional/managerial	0.0%	(2) 14.3%	(2)16.7%		
Lower managerial / intermediate	(8) 50%	(7) 50%	(6) 50%		
Small employers / Lower supervisory/technical	(3) 18.8%	0.0%	(4) 33.3%		
Routine / Semi-routine occupations	(2) 12.5%	(2) 14.3%	0.0%		
Not-working	(3) 18.8%	(3) 21.4%	0.0%		
Highest Educational qualification					
NS-SEC NCDS8 (derived)	(*N*= 16)	(*N*= 14)	(*N*= 12)	χ^2 7.921	0.791
Higher degree	(2) 12.5%	(1) 7.1%	(1) 8.3%		
Degree	(2) 12.5%	(3) 21.4%	0.0%		
Diploma	0.0%	(2) 14.3%	(1) 8.3%		
2 A' levels, Scottish highers	(1) 6.3%	(1) 7.1%	(1) 8.3%		
GCSE, Scottish standards	(5) 31.3%	(5) 35.7%	(2) 25.0%		
Other Scottish qualifications	(2) 12.5%	(1) 7.1%	(2) 16.7%		
No qualifications	(4) 25.0%	(1) 7.1%	(4) 33.3%		
Qualifications gained over the last 10yrs					
Yes	(3) 23.1%	(6) 46.2%	(6) 50.0%	χ^2 2.262	0.323
Obstetric factors					
Birth weight in ounces	121.93 ± 0.4.225	113.29 ± 4.631	107.17 ± 4.714	F 2.681	0.081
Smoking during pregnancy	(5) 31.5%	(4) 26.6%	(8) 66.7%	χ^2 4.805	0.090
Body mass index	(*N*= 14)	(*N*= 13)	(*N*= 12)		
BMI mean	26.786 ± 1.290	28.604 ± 1.408	26.700 ± 1.635	F 0.558	0.577
BMI categorised	21.4%	38.5%	16.7%	χ^2 3.520	0.741

Table 3.3.2.2 Summary of percentages and means relating to exercise, health variables, alcohol drinking behaviour and childhood adversities as a function of pilot study group status (*N* = 42). Percentages, numbers or mean and SEM shown.

Factor	Decline *N*, Mean or %	Control 1 *N*, Mean or %	Control 2 *N*, Mean or %	Group difference statistic	Group (*p*)
Exercise	(*N*= 16)	(*N*= 14)	(*N*= 12)		
NCDS8 regular exercise	68.8%	85.7%	83.3%	χ^2 1.503	0.472
Medical health	(*N*= 15)	(*N*= 12)	(*N*= 10)		
Registered disabled or long-term illness	53.3%	75.0%	10.0%	χ^2 9.447	**0.009**
Mental health					
Malaise inventory (high score)	(4) 25%	(1) 7.1%	0.0%	χ^2 4.541	0.103
CAPI- quality of life	27.133 ± 1.338	28.000 ± 1.062	29.333 ± 0.881	F 0.915	0.409
WEMWB survey	50.429 ± 2.429	51.571 ± 1.670	51.8333 ± 2.705	F 0.109	0.897
Alcohol drinking behaviour	(*N*= 16)	(*N*= 13)	(*N*= 12)	χ^2 2.863	0.581
Unproblematic	81.3%	69.2%	75.0%		
Harmful	12.5%	30.8%	25.0%		
Dependency	6.3%	2.8%	2.8%		
Childhood adversities					
Path Through Life Scale	(*N*= 16)	(*N*= 14)	(*N*= 12)	χ^2 15.721	0.331

3.3.3 Questionnaire responses

Table 3.3.3.1 presents a summary of responses to pilot study questionnaire items relating to health conditions and fears about developing health conditions. Examining the entire pilot study population together, Alzheimer's disease and heart disease were reported to be most worried about conditions, with approximately a third of all participants being concerned about developing one of these diseases as indicated in Graph 3.3.3.1.

With respect to whether the cohort members have ever suffered from certain health conditions, there was no difference found between the groups for any individual condition. However, when asked about whether they were worried about developing these conditions, a significantly higher percentage of control group 1 stated they had worries about having a stroke. Moreover, a higher percentage of control group 2 answered they were not worried about developing any of the above conditions, a point which is also indicated by the higher number of overall conditions worried about reported by individuals in the Decline group and control group 1. When asked in the ensuing questionnaire question about what might be the reason for worrying about developing these conditions, Decline group and control group 2 individuals were more likely to answer 'history of the condition in their family' whereas control group 1 individuals choose the 'other' answer option.

Table 3.3.3.2 presents a summary of responses to the pilot study questionnaire items relating to alcohol drinking behaviour and mental activity routines. With regards to alcohol consumption, the total derived AUDIT survey score obtained during the last survey sweep (2008/9), showed no difference between the pilot study groups in categories of drinking behaviour, i.e. unproblematic to dependent drinking behaviour ($\chi^2 = 21.187$, df = 10, $N = 42$, $p = 0.020$) However, the responses given to a questionnaire item asking about how often cohort members drink alcohol, suggested a significant difference between the groups and most evidently a higher percentage of individuals from control group 1 who report drinking alcohol 'most days' or '2 to 3 times a week'. This discrepancy might suggest that a higher number of control group 1 individuals in the pilot study drink alcohol on a more frequent basis, but their overall drinking behaviour (which takes into account, for example, how much per session and whether they are able to stop drinking) does not differ from the other groups.

Table 3.3.3.1 Responses by group status to questionnaire items concerning health conditions administered during the pilot study.

Questionnaire Item	Decline N, %	Control 1 N, %	Control 2 N, %	Total all N, %	Group (p)
Have you ever suffered from any of the following health conditions?	(N= 16)	(N= 14)	(N= 12)	(N= 42)	
Parkinson's disease	0.0%	0.0%	0.0%	0.0%	
Multiple sclerosis	0.0%	0.0%	0.0%	0.0%	
Alzheimer's disease	0.0%	0.0%	0.0%	0.0%	
Diabetes	(2) 12.5%	0.0%	0.0%	(2) 4.8%	0.182
Stroke	0.0%	0.0%	0.0%	0.0%	
Heart disease	(1) 6.3%	0.0%	0.0%	(1) 2.4%	0.435
Head trauma	0.0%	(1) 7.1%	0.0%	(1) 2.4%	0.359
Another condition affecting the nervous system, e.g. epilepsy	0.0%	0.0%	0.0%	0.0%	
Total individuals/conditions	3	1	0	4	
Do you ever worry about developing any of the following conditions?	(N= 16)	(N= 14)	(N= 12)	(N= 42)	
Parkinson's disease	(4) 25.0%	(1) 7.1%	0.0%	(5) 11.9%	0.103
Multiple sclerosis	(2) 12.5%	(1) 7.1%	0.0%	(3) 7.1%	0.446
Alzheimer's disease	(6) 37.5%	(5) 35.7%	(3) 25.0%	(14) 33.3%	0.765
Diabetes	(1) 6.3%	(4) 28.6%	(2) 16.7%	(7) 16.7%	0.262
Stroke	(2) 12.5%	(6) 42.9%	0.0%	(8) 19.0%	**0.015**
Heart disease	(6) 31.3%	(6) 42.9%	(1) 8.3%	(13) 31.0%	0.145
Other	(2) 12.5%	(1) 7.1%	0.0%	(3) 7.1%	0.446
No – none of the above	(5) 31.3%	(4) 28.6%	(9) 75.0%	(18) 25.0%	**0.029**
Total individuals/conditions	23	24	6	53	
Reasons for worrying about developing certain conditions	(N= 11)	(N= 10)	(N= 3)	(N= 24)	
There is a history of this(these) condition(s) in your family?	(6) 54.5%	(3) 30.0%	(3) 100.0%	(12) 50.0%	0.096
You have provided support for a family member or friend?	(1) 9.1%	(1) 10.0%	(2) 66.7%	(4) 16.7%	**0.046**
Other	(4) 36.4%	(7) 70.0%	0.0%	(11) 46.0%	0.071

41

Graph 3.3.3.1: Responses made by all pilot study participants to questionnaire items concerning worries about developing certain health conditions.
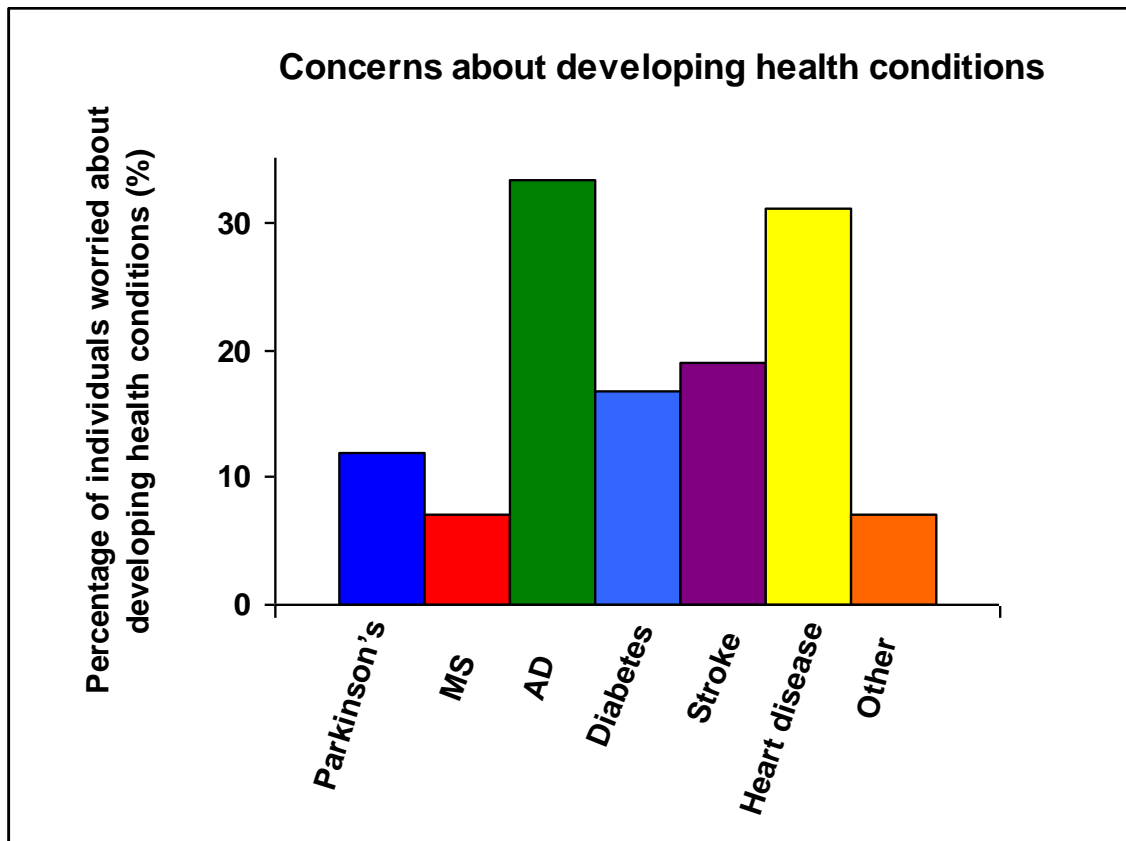
Table 3.3.3.2 Responses given by the groups to questionnaire items concerning alcohol drinking and mental activity routines administered during the pilot study.

Questionnaire Item	Decline N, %	Control 1 N, %	Control 2 N, %	Group (p)
How often do you drink alcohol?	(N= 16)	(N= 14)	(N= 12)	**0.020**
On most days	(2) 12.5%	(8) 57.1%	(4) 33.3%	
2 to 3 days a week	(4) 25.0%	(4) 28.6%	0.0%	
Once a week	(4) 25.0%	(1) 7.1%	(4) 33.3%	
2 to 3 times a month	(3) 18.8%	0.0%	0.0%	
Less often or only on special occasions	(3) 18.8%	0.0%	(4) 33.3%	
Never nowadays	0.0%	(1) 7.1%	0.0%	
Do you currently do any of the following mental activities?	(N= 16)	(N= 14)	(N= 12)	
Crossword puzzles and other puzzles such as Sudoku.	(14) 87.5%	(7) 50.0%	(5) 47.1%	**0.025**
Brain training exercises or games like 'Brain age' by Nintendo	(3) 18.8%	(3) 21.4%	(1) 8.3%	0.644
Read or write classic, scientific or educational literature	(4) 25.0%	(4) 28.6%	(4) 33.3%	0.890
Do mathematical related activities	(8) 50.0%	(5) 35.7%	(2) 16.7%	0.190
Do educational courses (e.g. IT, Open University or foreign language courses)	(3) 18.8%	(3) 21.4%	(1) 8.3%	0.644
Other mental activities, e.g. chess	(4) 25.0%	(2) 14.3%	0.0%	0.174
No – none of the above,	(2) 12.5%	(4) 28.6%	(5) 47.1%	0.214
Composite score Total individuals/activities	36	24	13	0.124
How often do you do this activity?	(N= 14)	(N= 10)	(N= 7)	0.644
Every day	(5) 35.7%	(3) 30.0%	(3) 35.5%	
2-3 days a week	(4) 28.6%	(2) 20.0%	0.0%	
4-5 days a week	(3) 21.4%	(2) 20.0%	(1) 14.3%	
Once a week	(1) 7.1%	(2) 20.0%	(3) 35.5%	
Less often	(1) 7.1%	(1) 10.0%	0.0%	

Finally, with respect to the questions regarding mental activity routines, it was predicted that the Decline group may undertake fewer mental activities in mid-life as keeping mentally active is postulated as a factor thought to be protective against cognitive decline. However, the questionnaire responses suggest that is control group 2 which appears to undertake overall fewer mental activities. In fact, the pilot study Decline group actually reports regularly doing significantly more puzzles, and as much if not more of the other mental activities listed than the other two groups [composite score (DEC vs CTL1 $t(28) = 0.954$, $p = 0.348$; DEC vs CTL2 $t(26) = 2.162$, $p = 0.40$; CTL1 vs CTL2 $t(24) = 1.137$, $p = 0.267$].

These observations could be interpreted to indicate that such a small pilot study population is perhaps not representative of the whole cohort groups. Alternatively, this finding might also be due to the decline individuals recognizing a change in their cognitive/memory ability and in response increasing they frequency of mental activities so as to remain more mentally active. Unfortunately we do not have past survey data to assess whether a change in 'mental activity' routines has occurred recently over the life course and whether Decline group individuals show differences in such behaviour evident to the two control groups.

4. Discussion

Cambridge pilot study battery findings and neuropsychological profile of the Decline group

In the current pilot study a group of NCDS cohort members was identified as having a word recall memory score more than 1 standard deviation below than that predicted from a childhood general ability test score. A subpopulation of this Decline group was subsequently shown to have poorer performance in the CANTAB visuo-spatial paired associates learning (PAL) task and Rapid Visual Information Processing (RVIP) tasks as compared to a non-decline control group matched for general ability performance in childhood. However, this Decline group showed a similar level of performance in these tasks when compared with a second non-decline control group matched for mid-life memory recall scores but with significantly lower GAT scores. In the Affective Go/No-go (AGN) task, the Decline group was also found to make the fastest responses, but made more omission errors for total correct and positive valence, but not for negative valence words, and in the GCT showed the longest deliberation time than either of the two control groups.

In addition, the results of the pilot study revealed that control group 2. who were identified as having a stable cognitive trajectory from childhood to mid-life but having a lower mean general cognitive ability, showed, as might be predicted, the worst quality of decision making in the Cambridge Gambling Task (CGT) task. Semantic memory as measured in the Graded Naming Test (GNT) task was also found to be impaired in this control group, but significantly the Decline group who showed impairment in recall function - an episodic memory measure, did not, in contrast, show a deficit in semantic memory.

The working hypothesis of this study was that the Decline group would perform significantly worse than the two matched control groups when tested on the CANTAB visuo-spatial paired associates learning (PAL) task. This hypothesis was based upon the finding of previous studies which indicated that performance in PAL is able to accurately classify individuals as belonging to Alzheimer's disease (AD), 'questionable AD' analogous to aMCI, or control/depression groups and that decline in PAL performance of patients with MCI predicts later progression to AD (Fowler et al., 1997; Swainson et al., 2001; Fowler et al., 2002; Egerhazi et al., 2007). In the present study, individuals belonging to the Decline group may hypothetically be at an even earlier stage than that typically associated with a diagnosis of aMCI, i.e. when individuals attend memory clinics. Performance in this Decline group was found however, to be poorer compared to control group (control group 1) matched for childhood general ability scores, but on par with a second control group matched for midlife memory recall. There exists two possible explanations for why the Decline group did not perform significantly worse than this second lower general cognitive ability control group.

First, it is possible that performance by low cognitive ability or IQ control individuals might generally be poorer in the PAL task. Indeed, in a recent cohort study using a large non-clinical population, level of education was found to account for a significant variance in PAL test performance with subjects who had a high level of education making significantly fewer errors than subjects who had a low level of education (Blackwell et al., unpublished findings). Although level of education is not a direct measure of cognitive ability, as found in the present pilot study, individuals who have low cognitive trajectories throughout life (e.g. control group 2 members) collectively attain fewer educational qualifications.

Second, it is also possible that individuals who belong to the control group 2 may also be experiencing some cognitive decline. Unfortunately, owing to the methodological issues of using the linear regression approach discussed in Section 2 and which results in individuals with low general ability scores being predicted to have suffered some cognitive deterioration independent of a 'real' decline in mid-life ability, it is hard to ascertain whether control group 2 individuals have in fact declined from childhood. For example and pertinent to the results, in the Cambridge pilot study subpopulation, 9 out of the 12 individuals in the control group 2 have residuals lower than -0.5 SD and half, 6 out of the 12 individuals, have residuals lower than -0.65 SD. Furthermore, as the recall scores used to identify cohort members as belonging to the three groups were attained two years ago, individuals with low cognitive ability, theorized to have low cognitive reserve, may show an accelerated trajectory in cognitive deterioration. Unfortunately, the recall scores attained in the Cambridge pilot study battery cannot be used as a comparative measure by which to clarify this issue.

In a previous study, PAL and GNT task performance in combination has been shown to accurately predict conversion from 'questionable dementia' to a later diagnosis of AD (Blackwell et al., 2004). Interestingly, in the present study although both the Decline group and control group 2 had poorer performance on PAL than control

group 1, the Decline group did not show any impairment in the GNT task. This would suggest that Decline individuals unlike the control group 2 individuals do not have a deficit in semantic memory. These findings are consistent with idea that semantic memory impairments, postulated to indicate pathology in the temporal neocortex, may provide relatively insensitive measures of decline and may only be a feature of the later stages of aMCI and AD.

Similar to the PAL task findings, both the Decline and control group 2 showed significantly poorer performance on the RVIP task. Consistent with the current results, studies of the neuropsychological profile of AD and MCI patients have also indicated that performance in RVIP is impaired and that deficits in attention and working memory, assessed in this task, are a consistent feature of Alzheimer's disease (Egerhazi et al 2007). However, as cohort members belonging to the Decline group in the present study could possibly assumed to be at an early stage of prodromal aMCI, comparison of group performance in a measure of cognition suggested to be characteristic of very early impairment, i.e. visuo-spatial memory, was of particular interest. The results of an OLS regression with RVIP 'A' as predictor, and PAL as outcome variable did indeed indicate that in contrast to the results for overall PAL and RVIP performance, the decline group performed markedly (though not significantly) worse than control group 2 in what we presume to be a purer measure of visuo-spatial memory.

One task outcome measure showed a clear Decline group specific effect. In the AGN task, the Decline group made the most no responses to a target stimulus (omissions) for total and positive valence words, but made the fewest omissions for negative words. This observation might suggest that the Decline group have a deficit in responding to positive target words (evident also when examining total omissions), but show no deficit, and hence could be interpreted to have a bias towards, negative valence words. Biases in information processing for positive and negative valence stimuli have previously been reported in patients with affective disorders and which indicated that patients in a depressed state exhibit an affective bias for negative words (Murphy et al., 1999; Erickson et al., 2005). Whether depressed mood is feature as well as a risk factor for MCI or early AD is still a matter of debate, a relationship between sub-syndromal depression and AD is certainly well documented (Starkstein et al., 2005). Further studies of aMCI patients using the AGN task might provide further insight into potential mood related information biases in early AD.

 Information processing speed or 'mental speed' is yet another cognitive measure argued to be a marker of early cognitive impairment in Alzheimer's disease (Dixon et al., 2007). In the present study, information processing speed was measured in the letter cancellation test and three CANTAB tasks; response times to target stimuli (latencies) in the AGN and RVIP tasks and deliberation time in the CGT. In the letter cancellation test and the CANTAB AGN and RVIP tasks, participants were told both speed and accuracy were important to performance, whereas in the GCT task, speed was not mentioned as an outcome measure.

The results of these tasks did not indicate a consistent deficit in processing speed for the Decline group. In the CGT task, the Decline group did take the longest time to

make a choice on which colour to bet, but in the AGN task the Decline group were the fastest to respond to correct stimuli although made the most total omission errors. An interpretation of these findings might be that when speed was specified out as a performance factor (as in the CGT) the Decline group showed the longest processing time, but when required to perform the AGN task as quickly as possible they become less accurate, i.e. there is a trade off between speed and accuracy, owing to the 'decline' brain functioning less well as compared to the control groups. This effect however, must be task and stimuli dependent because information processing speed did not differ between the Decline and control groups in the letter cancellation test or RVIP task.

Another interesting finding of this study was a gender effect found for the childhood general ability test scores evident in both the Cambridge pilot study group and whole NCDS cohort population. Females were found to have a higher mean GAT score than males, indicating a better general cognitive ability at the age of eleven. Previous studies have provided some evidence that boys' success in several academic domains lag behind that of girls (Nagy Jacklin and Martin, 1999). However, many findings specifically examining gender IQ differences in large populations have been contradictory. For example, Deary et al., investigated approximately 80,000 Scottish children tested at age 11 in 1932 and found no significant mean differences in cognitive test scores between boys and girls (Deary et al., 2003). These discrepancies between studies might be due to generation specific educational, socio-economic or cultural factors, i.e. the 1932 Scottish population were 11 years old in the 1930s where the NCDS cohort were 11 in the late 1960's.   Additional analysis of NCDS data to examine variables, which show co-variation with childhood GAT scores may provide some insight into explanations for this gender effect.

A second task outcome measure showed a clear difference between the sexes. In the CGT task, women were found to show less risk adjustment than men, but did not differ on overall risk taking. This result, that male participants exhibited a greater modulation of risk-taking in response to probability of winning, validates previous research findings by Deakin et al., who used the same CANTAB task to investigate CGT performance in a large population which ranged in age between 17- 79 years old (Deakin et al., 2004).

Factor Analysis

The results of the factor analysis performed using data from eleven key task outcome variables identified two distinct factors. One factor showed a relationship between decreases in the number of animals named in the animal naming test and total errors made during the graded naming test (both tests of semantic memory), poor performance in RVIP (RVIP A') as well as slower mean RVP and AGN correct latencies, and was found to have a significant negative correlation with childhood GAT. Therefore, a possible interpretation of the factor structure is that it represents semantic memory ability and slower reaction time processing speed.

The second factor, in contrast, accounts for co-variation in increased deliberation time, letter cancellation and AGN total omissions, number of PAL total trials to

success, and RVP mean latency; and, decreased quality of decision making, summed recall scores, RVP A' and AGN mean correct latency. A possible interpretation of the factor structure is that it represents attentional processing and learning, episodic memory as well as an information processing speed component (longer deliberation time in the CGT and faster AGN correct latencies although more AGN omissions). The fact that the Decline group had the highest overall mean score for this factor and which is significantly higher than that for control group 2 of low achievers, suggests that this factor may be pertinent as a marker of decline relevant to early mild cognitive impairment in mid-life.

Analysis of NCDS survey data

An ancillary objective of this study was to examine the NCDS survey data for potential differences between Decline and control groups in terms of demographic and epidemiological variables and which might provide support for the cognitive reserve hypothesis. Using data pertaining to individuals belonging to the three groups for the whole cohort, significant differences in socio-economic class were found. As might be predicted, control group 1 and the Decline group matched for, and showing a higher mean score in, childhood general cognitive ability, had a higher percentage of individuals in the highest socio-economic class 'professional/managerial' than compared to control group 2. However, for the lowest socio-economic class category, the Decline group and control group 2 matched for, and showing a lower mean in, mid-life memory scores, had a higher percentage of individuals not working than compared to control group 1. Most notably, the main reason given for not working by the Decline group and control group 2 related to ill health or disability, a point further discussed below.

Physical and mental health variables all showed a similar pattern, with the Decline group and control group 2 showing a significantly higher percentage of individuals with poorer health as compared to control group 1. An association between health including mental health states, and cognitive function in later life is well documented in the current literature and is consistent with the hypothesis that good physical and mental health may be a protective factor against cognitive decline (Sabia et al.,; Blaum et al., 2002; Deary and Batty, 2007; Gonzalez et al., 2008).

Significant differences between the groups were also found for highest educational qualifications. Control group 1 showed the highest percentage of individuals most qualified, whereas control group 2 had the lowest, and this pattern was reversed for percentages of individuals holding 'no qualifications'. The Decline group percentages fell between those of the two control groups for both highest and lowest qualification categories. As the Decline group was matched with control group 1 for childhood cognitive ability it might have expected that if cognitive trajectories in adult life had remained stable that the Decline group should have shown the same pattern of percentages in educational attainment as this control group. However, as the Decline group showed significant differences from control group 1, this would suggest that the decline individuals have not attained the level of educational qualification that would have been predicted.

Furthermore, even more notable was that the there was no difference in percentages between the Decline group and control group 2 when comparing educational qualifications attained only in the last ten years. This would suggest that the behaviour of the Decline group, with regards to learning and continued education, is already showing a decline as compared to predictions based on childhood ability. Indeed, education itself has been indicated to be protective of psychological performance in late life -a finding which has been related to occupational complexity and acquisition of lifelong abilities to sustain attention and conceptionalize problems (Le Carret et al., 2003; Whalley et al., 2004; Studenski et al., 2006).

As keeping mentally active is postulated to be a protective factor against cognitive decline (Valenzuela and Sachdev, 2006; Hall et al., 2009), questions regarding current mental activity routines were asked in the Cambridge pilot study questionnaire. It was predicted that the decline group may self-report undertaking fewer mental activities on a regular basis. Although numbers were small, the results suggested the Decline group actually regularly do significantly more puzzles, and as much, if not more, of the other mental activities listed than the two control groups. This surprising result could be explained by decline individuals recognizing a change in their cognitive/memory ability and in response increasing their frequency of mental activities so as to remain more mentally active. Unfortunately, there is a lack of NCDS data by which to assess whether a change in 'mental activity' routines has occurred recently over the life course. Questions specifically relating to regular mental activities would be of relevance to future NCDS surveys.

Life adversities experienced in childhood have also been implicated to impact upon early adult behaviour and cognitive ability (Bremner et al., 1995; Masten et al., 1999; Obradovic et al., 2006). Although a number of socio-economic or educational factors may interact or confound the interpretation of data, some studies have shown an effect of adverse circumstances on cognitive performance after controlling for potential mediating factors (Richards and Wadsworth 2004). In the present study, and consistent with previous findings, control group 1, who have the best mean recall performance, were found to have a significantly higher percentage of individuals who reported no childhood adversities as compared to the Decline group and control group 2. Again, these results suggest that there may be a link between adverse circumstances in early life and cognitive function in mid-life.

In addition to social and intellectual activities, physical exercise is suggested to facilitate cognitive performance and slow the rate of aging associated cognitive decline (Elwood et al., 1999; Dik et al., 2003). For example, using data from the 1946 British birth cohort, Richards et al., found that physical exercise at 36 years was associated with a significantly slower rate of decline in memory from 43 to 53 years (Richards et al., 2003). Similar to these findings, the Decline and control 2 groups who show lower memory ability in mid-life also had a significantly lower percentage of individuals who self-reported undertaking regular physical exercise at least once a month at the age of 50. However, variables relating to physical exercise undertaken during childhood do not show differences in amount of exercise undertaken between the three groups. This would suggest a lack of relationship between amount of physical exercise in early life and cognitive decline later in mid-life.

Finally, birth weight has been consistently reported to be associated with childhood intelligence and also independently associated with maternal smoking (Milberger et al., 1996; Richards et al., 2001; Shenkin et al., 2004). In the present study, control group 2, who show a lower cognitive ability in childhood and throughout life, had a significantly lower mean birth weight as compared to the other two groups. However, as no difference between the groups was found for prematurity, this would suggest that the birth weight effect is independent of whether an individual is born premature. Nevertheless, control group 2 did have a significantly higher percentage of individuals whose mothers smoked after the fourth month of pregnancy and this smoking variable shows a significant negative correlation with birth weight. Taken together, these findings replicate a positive association between birth weight and cognitive ability in childhood and in addition suggest maternal smoking later in pregnancy could be a potential contributing factor.

In summary, exploring NCDS survey data for differences between identified decline and control groups has provided some evidence in support of the cognitive reserve hypothesis. Factors such as educational attainment, routine physical and mental activities, and health status, which are postulated to influence biological processes underlying cognitive reserve, were found to show variation between the groups. Nevertheless, it is recognised that this is preliminary analysis and as yet potential gender effects, differences in factors across time, or interactions between factors (regression modelling) have not fully been taken into account. Further in depth analysis is needed for a fuller understanding of the relationship between variables and cognitive decline in mid-life.

Advantages and limitations of the study

The design of the present study has a number of advantages over previous studies which have used birth cohorts to investigate cognitive decline in adulthood. First, the Cambridge pilot study used a battery of CANTAB multi-component cognitive tasks to assess cohort members on a variety of cognitive processes. The CANTAB was originally developed for the assessment of cognitive performance in the elderly and individuals who suffer from dementia of the Alzheimer type (Robbins et al., 1994; Robbins et al., 1998). The CANTAB tasks "were designed on the premise that cognitive functions are both diverse and modular in the sense that they are supported by overlapping yet distinct sets of neural structures which may be differentially affected by different forms of CNS pathology" (Robbins et al., 1994).

Using standardised computerised tasks on a touch-sensitive tablet screen enables data to be derived, recorded and automatically analysed on several different performance outcome measures, whilst also reducing the potential for subjective bias to be inadvertently introduced by the experimenter. Furthermore, direct feedback is given automatically given to the participants during the tasks which has been shown to increase both interest and hence motivation to perform (Robbins et al., 1994). This is an important factor for both how much the cohort members enjoy and are engaged in the testing and hence for the reliability of the data if testing were to be conducted on thousands of NCDS members in the future.

Second, the same cohort members can be re-tested on PAL and other CANTAB tasks at a later stage when a clinical diagnosis, e.g. later stage MCI or Alzheimer's disease, may be relevant. Such follow-up could be administered in their home and the level of task difficulty adjusted when required, thus providing reliable and consistent data over time. Third, the preliminary analysis of survey data examined a Decline group and two control groups with numbers exceeding a thousand individuals per group. Such numbers provide the study with sufficient power to perform follow-up modelling analysis, for example, regression modelling and structural equation modelling, and which may provide important insight into the interaction between lifestyle and health related variables contributing to mid-life cognitive decline.

There are, however, a number of limitations to this Cambridge pilot study. Performance in immediate and delayed recall was shown to be notably better particularly for the Decline group and control group 2, when tested the second time as part of the CANTAB battery conducted at the Behavioural and Clinical Neuroscience Institute (BCNI). Similar findings of improved memory scores in a second sweep conducted two years after the first were also found in the English Longitudinal Study of Ageing (ELSA) study (Llewellyn et al., 2008). This effect may reflect practice effects as participants are now familiar with the recall test procedure or may be due to a quieter and more standardised testing environment at the Institute as opposed to the home setting. Unfortunately, this means that the first and second recall scores cannot be directly compared and therefore individuals, who may have declined since the first testing, especially those belonging to control group 2, cannot be identified.

The present study was designed as a proof of concept pilot study with the aims of providing an assessment of the feasibility and practicalities of conducting such cognitive tests in a larger-scale study as well as conducting analysis of the actual performance data generated. It is recognised that the pilot study is underpowered both in terms of looking at task performance between the groups, as well attempting to examine differences in variables either previously surveyed or asked in the pilot study questionnaire with group status or cognitive performance. Nevertheless, the results of this Cambridge pilot study suggested some unexpected and intriguing preliminary findings, e.g. the Decline group showing a potential affective bias for negative words, which might be worth pursuing in a larger MCI or AD clinical population.

It is to be hoped that this research might be the precursor of larger, more in depth follow-on NCDS studies which, in addition to the present design, may include genetic analysis and functional imaging to investigate early mild cognitive impairment in mid-life.

References

Abas MA, Sahakian BJ, Levy R (1990) Neuropsychological deficits and CT scan changes in elderly depressives. Psychol Med 20:507-520.

Barnett J, Sahakian B (2008) State of Science Review:SR-E4, Cognitive Reserve and Mental Capital. In: UK Government's Foresight Project, Mental Captial and Wellbeing.

Beddington J, Cooper CL, Field J, Goswami U, Huppert FA, Jenkins R, Jones HS, Kirkwood TB, Sahakian BJ, Thomas SM (2008) The mental wealth of nations. Nature 455:1057-1060.

Blackwell AD, Sahakian BJ, Vesey R, Semple JM, Robbins TW, Hodges JR (2004) Detecting dementia: novel neuropsychological markers of preclinical Alzheimer's disease. Dement Geriatr Cogn Disord 17:42-48.

Blaum CS, Ofstedal MB, Liang J (2002) Low cognitive performance, comorbid disease, and task-specific disability: findings from a nationally representative survey. J Gerontol A Biol Sci Med Sci 57:M523-531.

Bray H (2008) 2006-based national population projections for the U.K. and constituent countries. Popul Trends:8-18.

Bremner JD, Randall P, Scott TM, Capelli S, Delaney R, McCarthy G, Charney DS (1995) Deficits in short-term memory in adult survivors of childhood abuse. Psychiatry Res 59:97-107.

Brown M, Dodgeon B (2010) NCDS Cognitive Assessments at age 50: Initial Results.

Brown M, Knight HM, Dodgeon B (2010) UIBEN pilot study:Technical report concerning study design and set up. Unpublished. In.

Deakin J, Aitken M, Robbins T, Sahakian BJ (2004) Risk taking during decision-making in normal volunteers changes with age. J Int Neuropsychol Soc 10:590-598.

Deary IJ, Batty GD (2007) Cognitive epidemiology. J Epidemiol Community Health 61:378-384.

Deary IJ, Thorpe G, Wilson V, Starr JM, Whalley LJ (2003) Population sex differences in IQ at age 11: the Scottish mental survey 1932. Intelligence 31:533-542.

Deary IJ, Gow AJ, Taylor MD, Corley J, Brett C, Wilson V, Campbell H, Whalley LJ, Visscher PM, Porteous DJ, Starr JM (2007) The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. BMC Geriatr 7:28.

Dik M, Deeg DJ, Visser M, Jonker C (2003) Early life physical activity and cognition at old age. J Clin Exp Neuropsychol 25:643-653.

Dixon RA, Garrett DD, Lentz TL, MacDonald SW, Strauss E, Hultsch DF (2007) Neurocognitive markers of cognitive impairment: exploring the roles of speed and inconsistency. Neuropsychology 21:381-399.

Egerhazi A, Berecz R, Bartok E, Degrell I (2007) Automated Neuropsychological Test Battery (CANTAB) in mild cognitive impairment and in Alzheimer's disease. Prog Neuropsychopharmacol Biol Psychiatry 31:746-751.

Elwood PC, Gallacher JE, Hopkinson CA, Pickering J, Rabbitt P, Stollery B, Brayne C, Huppert FA, Bayer A (1999) Smoking, drinking, and other life style factors and cognitive function in men in the Caerphilly cohort. J Epidemiol Community Health 53:9-14.

Erickson K, Drevets WC, Clark L, Cannon DM, Bain EE, Zarate CA, Jr., Charney DS, Sahakian BJ (2005) Mood-congruent bias in affective go/no-go performance of unmedicated patients with major depressive disorder. Am J Psychiatry 162:2171-2173.

Fowler KS, Saling MM, Conway EL, Semple JM, Louis WJ (1997) Computerized neuropsychological tests in the early detection of dementia: prospective findings. J Int Neuropsychol Soc 3:139-146.

Fowler KS, Saling MM, Conway EL, Semple JM, Louis WJ (2002) Paired associate performance in the early detection of DAT. J Int Neuropsychol Soc 8:58-71.

Fowler L, Blackwell A, Jaffa A, Palmer R, Robbins TW, Sahakian BJ, Dowson JH (2006) Profile of neurocognitive impairments associated with female in-patients with anorexia nervosa. Psychol Med 36:517-527.

Gonzalez HM, Bowen ME, Fisher GG (2008) Memory decline and depressive symptoms in a nationally representative sample of older adults: the Health and Retirement Study (1998-2004). Dement Geriatr Cogn Disord 25:266-271.

Hall CB, Lipton RB, Sliwinski M, Katz MJ, Derby CA, Verghese J (2009) Cognitive activities delay onset of memory decline in persons who develop dementia. Neurology 73:356-361.

Le Carret N, Lafont S, Mayo W, Fabrigoule C (2003) The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. Dev Neuropsychol 23:317-337.

Llewellyn DJ, Lang IA, Langa KM, Huppert FA (2008) Cognitive function and psychological well-being: findings from a population-based cohort. Age Ageing 37:685-689.

Masten AS, Hubbard JJ, Gest SD, Tellegen A, Garmezy N, Ramirez M (1999) Competence in the context of adversity: pathways to resilience and

maladaptation from childhood to late adolescence. Dev Psychopathol 11:143-169.

McKenna P, Warrington EK (1980) Testing for nominal dysphasia. J Neurol Neurosurg Psychiatry 43:781-788.

Milberger S, Biederman J, Faraone SV, Chen L, Jones J (1996) Is maternal smoking during pregnancy a risk factor for attention deficit hyperactivity disorder in children? Am J Psychiatry 153:1138-1142.

Murphy FC, Sahakian BJ, Rubinsztein JS, Michael A, Rogers RD, Robbins TW, Paykel ES (1999) Emotional bias and inhibitory control processes in mania and depression. Psychol Med 29:1307-1321.

Nagy Jacklin C, Martin LJ (1999) Effects of gender on behaviour and development. In: Development - Behavioural Pediatrics, 3rd Edition (M.D L, W.B. C, A.C C, eds), pp 100-106. Philadelphia: W.B. Saunders.

Norusis MJ (1990) SPSS/PC + Statistics 4.0. Chicago:SPSS. In.

Obradovic J, Burt KB, Masten AS (2006) Pathways of adaptation from adolescence to young adulthood: antecedents and correlates. Ann N Y Acad Sci 1094:340-344.

ONS (2008) Office for National Statistics, National Population Projections 2006.

Power C, Elliott J (2006) Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol 35:34-41.

Richards M, Hardy R, Wadsworth ME (2003) Does active leisure protect cognition? Evidence from a national birth cohort. Soc Sci Med 56:785-792.

Richards M, Hardy R, Kuh D, Wadsworth ME (2001) Birth weight and cognitive function in the British 1946 birth cohort: longitudinal population based study. Bmj 322:199-203.

Richards M, Hardy R, Kuh D, Wadsworth ME (2002) Birthweight, postnatal growth and cognitive function in a national UK birth cohort. Int J Epidemiol 31:342-348.

Robbins TW, James M, Owen AM, Sahakian BJ, McInnes L, Rabbitt P (1994) Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. Dementia 5:266-281.

Robbins TW, James M, Owen AM, Sahakian BJ, Lawrence AD, McInnes L, Rabbitt PM (1998) A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging.

Cambridge Neuropsychological Test Automated Battery. *J Int Neuropsychol Soc* 4:474-490.

Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, Robbins TW (1999) Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J Neurosci* 19:9029-9038.

Rosenman S, Rodgers B (2004) Childhood adversity in an Australian population. *Soc Psychiatry Psychiatr Epidemiol* 39:695-702.

Rosenman S, Rodgers B (2006) Childhood adversity and adult personality. *Aust N Z J Psychiatry* 40:482-490.

Sabia S, Gueguen A, Marmot MG, Shipley MJ, Ankri J, Singh-Manoux A Does cognition predict mortality in midlife? Results from the Whitehall II cohort study. *Neurobiol Aging* 31:688-695.

Sahakian BJ, Owen AM (1992) Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *J R Soc Med* 85:399-402.

Sahakian BJ, Morris RG, Evenden JL, Heald A, Levy R, Philpot M, Robbins TW (1988) A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain* 111 (Pt 3):695-718.

Shenkin SD, Starr JM, Deary IJ (2004) Birth weight and cognitive ability in childhood: a systematic review. *Psychol Bull* 130:989-1013.

Starkstein SE, Jorge R, Mizrahi R, Robinson RG (2005) The construct of minor and major depression in Alzheimer's disease. *Am J Psychiatry* 162:2086-2093.

Stern Y (2002) What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc* 8:448-460.

Studenski S, Carlson MC, Fillit H, Greenough WT, Kramer A, Rebok GW (2006) From bedside to bench: does mental and physical activity promote cognitive vitality in late life? *Sci Aging Knowledge Environ* 2006:pe21.

Swainson R, Hodges JR, Galton CJ, Semple J, Michael A, Dunn BD, Iddon JL, Robbins TW, Sahakian BJ (2001) Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. *Dement Geriatr Cogn Disord* 12:265-280.

Valenzuela MJ, Sachdev P (2006) Brain reserve and dementia: a systematic review. *Psychol Med* 36:441-454.

Wadsworth M, Kuh D, Richards M, Hardy R (2006) Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *Int J Epidemiol* 35:49-54.

Waldemar G, Phung KT, Burns A, Georges J, Hansen FR, Iliffe S, Marking C, Rikkert MO, Selmes J, Stoppe G, Sartorius N (2007) Access to diagnostic evaluation and treatment for dementia in Europe. Int J Geriatr Psychiatry 22:47-54.

Wesnes K, Warburton DM (1984) Effects of scopolamine and nicotine on human rapid information processing performance. Psychopharmacology (Berl) 82:147-150.

Whalley LJ, Deary IJ, Appleton CL, Starr JM (2004) Cognitive reserve and the neurobiology of cognitive aging. Ageing Res Rev 3:369-382.

Appendices

Appendix 1. Script of UIBEN pilot study questionnaire.

National Child Development Study – Understanding Individual Differences in Learning and Memory - Questionnaire

Instructions given verbally by researcher to participants

"The final part of the session involves you completing a short self-completion computerized questionnaire similar to surveys you have completed in the past. You may read and answer the questions by yourself without me in the room or alternatively, if you wish, I can assist you. Most of the questions are about your health but we also ask you to provide some feedback about your experience of coming here today to participate in this research. The final questions are about imaging studies and I will provide you with some further details on this.

Each question will appear on the screen one at a time. There will be clear instructions on how to answer each question. When you have read the question, please indicate your response either by touching the screen or by clicking the mouse over the answer options, then press the 'next' button. (INTERVIEWER DEMONSTRATE).

On some questions you will be only be able to give one answer, other questions will allow you to give several answers. Once you have answered a question you will not be able to go back and change your answers. Occasionally you might be asked to answer a question in your own words. If you do not wish to answer a question please press 'no response' option, and if wish to stop the questionnaire at any time please press the 'exit' option.

If you have any questions please ask me. Would you be willing to have a go?"

(The first screen that participants will see will be an example page so that the researcher can demonstrate how the questions should be answered, i.e. touch screen or using the mouse to click buttons. The second screen presented to the participants will have fields in which the research provides details such as the participant's name, study ID, date of birth and whether the participants will: self-complete the questionnaire independently; do the self-completion questionnaire with the researcher assistance; or has refused to do self-complete questionnaire. The third screen will present the first question in the questionnaire.)

We would first like to ask you a number of questions about your general health.

Q1 In general, would you say your health over the last 12 months has been….

    SELECT ONE ANSWER ONLY

1. Excellent
2. Very good
3. Good
4. Fair
5. Poor

IF Q1 = 5 ASK Q2

Q2 Do you rate your health as poor owing to:

SELECT ALL THAT APPLY

1. A long standing illness/condition
2. A recent acute illness
3. A recent accident
4. Recovering from an operation
5. Recent stress, e.g. divorce, bereavement or unemployment
6. Other

Q3 Compared to one year ago, how would you rate your health in general now?

SELECT ONE ANSWER ONLY

1. Much better than one year ago
2. Somewhat better than one year ago
3. About the same as one year ago
4. Somewhat worse than one year ago
5. Much worse than one year ago

We would now like to ask you a few questions about specific health problems.

Q4. Do you currently, or have you ever suffered from any of the following health conditions?

SELECT ALL THAT APPLY

1. Parkinson's disease
2. Multiple sclerosis
3. Alzheimer's disease
4. Diabetes
5. Stroke
6. Heart disease
7. Head trauma
8. Another condition affecting the nervous system, e.g. epilepsy
9. No – none of the above.

IF Q4 = 8 ASK Q5

Q5. Would you mind telling us what condition you have?
(open question)

Q6. Do you ever worry about developing any of the following conditions?

SELECT ALL THAT APPLY

1. Parkinson's disease
2. Multiple sclerosis
3. Alzheimer's disease
4. Diabetes
5. Stroke
6. Heart disease
7. Other
8. No – none of the above

IF Q6 = 1-7 ASK Q7

Q7. Is this because.....

1. There is a history of this (these) condition(s) in your family?
2. You have provided support for a family member or friend who has/had this/these condition(s)?
3. Other

Q8. Have you consulted your GP or a psychiatrist about feeling depressed or anxious in the past 12 months?

SELECT ONE ANSWER ONLY

1. Yes
2. No

Q9. Have you been prescribed any antidepressant medication in the last 12 months

SELECT ONE ANSWER ONLY

1. Yes
2. No

IF Q9 = 1 ASK Q10

Q10. Are you still taking the medication?

SELECT ONE ANSWER ONLY

1. Yes
2. No

Q11. How often do you have an alcoholic drink of any kind? Would you say you have a drink ....

SELECT ONE ANSWER ONLY

1. On most days
2. 2 to 3 days a week
3. Once a week
4. 2 to 3 times a month
5. Once a month
6. Less often or only on special occasions
7. Never nowadays
8. Never had an alcoholic drink

IF Q11 = 1-6 ASK Q12

Q12. How often do you have six or more drinks on one occasion?

SELECT ONE ANSWER ONLY

1. Never
2. Less than monthly
3. Monthly
4. Two to three times per week
5. Four or more times a week

We would also like to ask you about any health issues which could have potentially interfered with you performing our computerised cognitive tasks.

Q13. Are you colour blind?

SELECT ONE ANSWER ONLY

1. Yes
2. No

Q14. Do you currently have any problems with your hearing? E.g. do you have a hearing aid?

SELECT ONE ANSWER ONLY

1. Yes
2. No

Q15. Do you currently have any problems with moving your fingers or hands which may prevent you from pressing buttons on computer quickly, e.g. rheumatism?

SELECT ONE ANSWER ONLY

1. Yes
2. No

Now we'd like to get a few details about your current mental activity routines

Q16. Do you currently do any of the following mental activities?

SELECT ALL THAT APPLY

1. Crossword puzzles and other puzzles such as Sudoku.
2. Brain training exercises or games like 'Brain age' by Nintendo
3. Read or write classic, scientific or educational literature
4. Do mathematical related activities
5. Do educational courses (e.g. IT, Open University or foreign language courses)
6. Other mental activities, e.g. chess
7. No – none of the above,

IF Q16 = 1-6 ASK Q17 AND Q18

Q17. You said that you do one of the following mental activities:

Which one do you do most often:

1. Crossword puzzles and other puzzles such as Sudoku.
2. Brain training exercises or games like 'Brain age' by Nintendo
3. Read or write classic, scientific or educational literature
4. Do mathematical related activities
5. Do educational courses (e.g. IT, Open University or foreign language courses)
6. Other mental activities, e.g. chess

Q18. How often do you do this activity?

SELECT ONE ANSWER ONLY

1. Every day
2. 4-5 days a week
3. 2-3 days a week
4. Once a week
5. Once a month
6. Two or three times a month
7. Less often

We realise that this taking part in this research has been a little different from anything we have asked you to do previously. This research has involved a very small number of study members but we may consider running this kind of research again with a larger number of study members so we would be very interested to hear how you have felt about being involved.

Q19. Do you feel that the letter and information sheet we sent you explained adequately what taking part in this research would involve?

PLEASE SLIDE THE POINTER TO INDICATE YOUR OPINION

Not at all   **Very much**

Q20. Do you feel that travelling here today was .....

SELECT ONE ANSWER ONLY

1. Very easy
2. Easy
3. Fairly easy
4. Fairly difficult
5. Difficult
6. Very difficult

Q21. How much have you enjoyed being involved in this research project?

PLEASE SLIDE THE POINTER TO INDICATE YOUR OPINION

Not at all   **Very much**

Q22. Which of the cognitive tasks did you enjoy undertaking the most?

SELECT ALL THAT APPLY

1. Graded naming test (naming objects test)

2. Paired Associates Learning (Patterns within boxes)

3. Rapid Visual Information Processing task (Number sequence task)

4. Affective Go/No go task (Word task)

5. Cambridge Gambling Task (gambling points task)

6. All equally

7. None
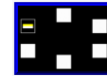
Q23. Which of the cognitive tasks did you least enjoy?

SELECT ALL THAT APPLY

1. Graded naming test (naming objects test)

2. Paired Associates Learning (Patterns within boxes)

3. Rapid Visual Information Processing task (Number sequence task)

4. Affective Go/No go task (Word task)

5. Cambridge Gambling Task (gambling points task)

Q24. If you were asked to participate in a similar research project again in the future how likely would you be to do so?

SELECT ONE ANSWER ONLY

1. Very likely
2. Fairly likely
3. Not very likely
4. Not at all likely

62

Researcher returns to testing room.

At some point in the future CLS might consider asking members of the National Child Development Study to participate in an fMRI research study. We would like to ask you a few questions about whether you might be prepared to participate in this kind of study.

Here is a picture of an fMRI scanner.
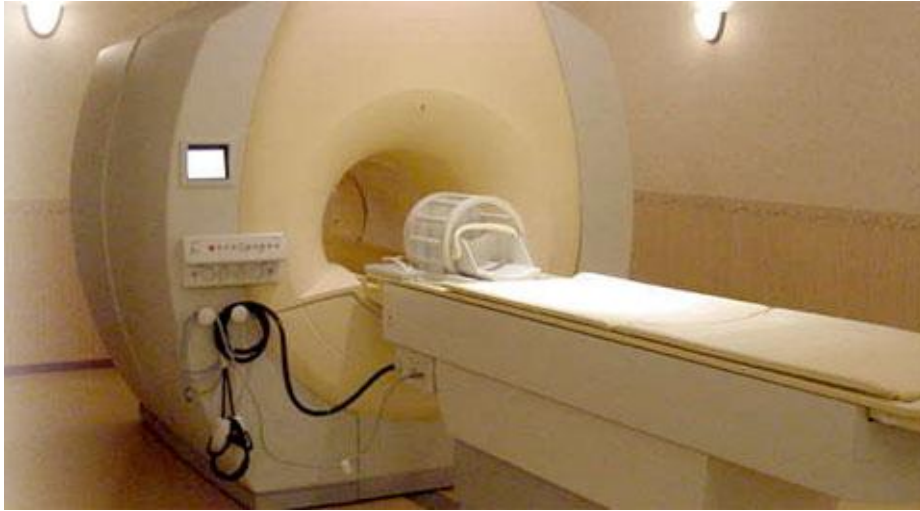
Description given verbally by researcher to participants

fMRI is a commonly used imaging technique which enables researchers to study the brain and how it is working while people are performing certain tasks. The person being scanned lies on a couch in a powerful tubular magnet for at least half an hour while the brain is scanned. Most people find the scan easy to tolerate although some people can find it noisy and slightly claustrophobic at the beginning.

Q25. Have you ever been scanned in an MRI scanner and if so was it for clinical diagnosis and/or as part of a research?

SELECT ONE ANSWER ONLY

1. No
2. Yes, for clinical diagnosis
3. Yes, as part of a research study
4. Yes, both for clinical diagnosis and as part of a research study
5. Don't know

Like faces, brains come in all shapes and sizes, so that there are many normal variations of what the scan shows. It is possible, though, that a scan could reveal something that suggests that there could be a more serious problem. This is estimated to happen in about one in forty scans. However, for the great majority of people who are scanned in research studies no significant problems will be observed.

People may differ in their views about feedback from research fMRI studies. Some people would like to know if their scan result reveals anything that might suggest they might have a serious problem, regardless of whether this may turn out to be treatable or not. Others would prefer only to know if the scan revealed clear evidence of a

serious problem that is likely to be treatable. Some do not want any feedback whatever the scan may reveal. They prefer, should they develop a serious condition, to wait until they have symptoms and then to seek a diagnosis and treatment from a doctor at that time.

Q26. If you were asked by CLS to participate in an fMRI research study do you think you would be prepared to do so?

      SELECT ONE ANSWER ONLY

1. I would not be prepared to take part in an fMRI research study
2. I would be prepared to take part in an fMRI study regardless of whether feedback was to be provided.
3. I would **only** be prepared to take part in an fMRI study which provided feedback on **all potential problems** that were observed.
4. I would **only** be prepared to take part in an fMRI study which **only** provided feedback on potential problems that were considered to be **serious and treatable**.
5. I would **only** be prepared to take part in an fMRI study which provided **no** feedback.

Q27. Please give reasons for your answer (open ended question).

Appendix 2

Table 1: Significance levels of PAL test indicators on double-sized dataset

Outcome measure	Sig (N= 84)	Sig (N = 42)
PAL t errors (adj)	0.025	0.152
PAL t errors adj Log10	0.005	0.081
PAL t errors 8 shapes (adj)	0.018	0.177
PAL mean errors to success	0.045	0.184
PAL mean trials to success	0.018	0.065
PAL mean trials to success	0.004	0.065
PAL total trials (adj)	0.005	0.053
PAL total trials adj Log10	0.002	0.051
PAL total trials 6 shapes (adj)	0.026	0.18
PAL total trials 8 shapes (adj)	0.06	0.302

Appendix 3: Correlation matrix showing of the 11 variables examined by factor analysis.

		Immediate recall and delayed 1	NUMBER OF ANIMALS MENTIONED	Letter cancellation (accuracy) omissions	PAL total errors adjusted	AGN Mean correct latency	AGN Total omissions	RVP A'	RVP mean latency	GNT total errors	CGT Quality of decision making	CGT Deliberation time	CGT Risk adjustment
Immediate recall and delayed 1	Pearson Correlation	1	.166	-.448	-.422	.210	-.359	.341	-.265	-.220	.167	-.364	.076
	Sig. (2-tailed)		.293	.003	.005	.181	.020	.027	.090	.161	.291	.018	.631
NUMBER OF ANIMALS MENTIONED	Pearson Correlation	.166	1	-.360	-.235	-.119	-.394	.507	-.345	-.407	.145	-.352	.205
	Sig. (2-tailed)	.293		.019	.133	.452	.010	.001	.025	.008	.361	.022	.193
Letter cancellation (accuracy) omissions	Pearson Correlation	-.448	-.360	1	.606	-.371	.616	-.537	.564	.171	-.532	.752	-.138
	Sig. (2-tailed)	.003	.019		.000	.016	.000	.000	.000	.279	.000	.000	.382
PAL total errors adjusted	Pearson Correlation	-.422	-.235	.606	1	-.378	.531	-.624	.478	.208	-.592	.724	-.137
	Sig. (2-tailed)	.005	.133	.000		.014	.000	.000	.001	.187	.000	.000	.387
AGN Mean correct latency	Pearson Correlation	.210	-.119	-.371	-.378	1	-.476	.049	-.133	.158	.467	-.363	.225
	Sig. (2-tailed)	.181	.452	.016	.014		.001	.756	.400	.317	.002	.018	.151
AGN Total omissions	Pearson Correlation	-.359	-.394	.616	.531	-.476	1	-.436	.463	.348	-.607	.676	-.252
	Sig. (2-tailed)	.020	.010	.000	.000	.001		.004	.002	.024	.000	.000	.108
RVP A'	Pearson Correlation	.341	.507	-.537	-.624	.049	-.436	1	-.655	-.330	.469	-.666	.355
	Sig. (2-tailed)	.027	.001	.000	.000	.756	.004		.000	.033	.002	.000	.021
RVP mean latency	Pearson Correlation	-.265	-.345	.564	.478	-.133	.463	-.655	1	.354	-.293	.642	-.338
	Sig. (2-tailed)	.090	.025	.000	.001	.400	.002	.000		.021	.060	.000	.028
GNT total errors	Pearson Correlation	-.220	-.407	.171	.208	.158	.348	-.330	.354	1	-.197	.241	-.157
	Sig. (2-tailed)	.161	.008	.279	.187	.317	.024	.033	.021		.210	.124	.321
CGT Quality of decision making	Pearson Correlation	.167	.145	-.532	-.592	.467	-.607	.469	-.293	-.197	1	-.643	.291
	Sig. (2-tailed)	.291	.361	.000	.000	.002	.000	.002	.060	.210		.000	.061
CGT Deliberation time	Pearson Correlation	-.364	-.352	.752	.724	-.363	.676	-.666	.642	.241	-.643	1	-.247
	Sig. (2-tailed)	.018	.022	.000	.000	.018	.000	.000	.000	.124	.000		.116
CGT Risk adjustment	Pearson Correlation	.076	.205	-.138	-.137	.225	-.252	.355	-.338	-.157	.291	-.247	1
	Sig. (2-tailed)	.631	.193	.382	.387	.151	.108	.021	.028	.321	.061	.116	