Leading education and social research
Institute of Education
University of London

Millennium Cohort Study

Interpreting the CANTAB cognitive measures

Mark Atkinson

CLS Data Note

(First Edition)

January 2015

Centre for Longitudinal Studies
Following lives from birth through the adult years
www.cls.ioe.ac.uk

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

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

First published in 2015 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk

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

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

Contents

1. Introduction

The 5th sweep of the Millennium Cohort Study (MCS5) introduced the cohort to two cognitive tests drawn from the Cambridge Neuropsychological Test Automated Battery (CANTAB; Robbins et al., 2004; 2008). Although a subsample of the National Child Development Study (NCDS) cohort members recently undertook these tests, MC5 is the first of the British Cohort Studies to administer such tests to the full cohort sample. Two tests from the CANTAB battery were administered – the Spatial Working Memory task (SWM) and Cambridge Gambling Task (CGT).

A great advantage of these tests is that they introduce the recording of response time (RT) measures to the cognitive testing of the cohort. RT outcomes can be advantageous to studying cognition for a number of reasons. As measures of processing speed they are thought to provide an index of mental complexity. As such, they may permit a greater sensitivity to underlying processes than error or accuracy measurements. This is because they can differentiate processing capability under conditions when participants show little variation in quality of task responses, for example when a large portion of the sample makes few or no errors. This same scenario highlights another advantage that they permit the study of mental processing when tasks do not overload participants and the processes are in effect, functioning normally. RT also has the advantage of being a ratio scale with a true zero and thus has the beneficial property of quantifiable precision.

The interpretation of these outcomes is however, difficult without access to original test materials and procedures or algorithms for deriving the summary scores from each cohort member's raw data. In addition, the interviewers conducting the MCS survey recorded a number of further outcome measures pertaining to interview conditions during the test procedure. These can potentially impact upon data quality and test performance. The aim of this data note is to present an overview of the outcome measures from the two CANTAB tests in order for data analysts to successfully utilise their findings. In addition, interview factors that may impact data quality and test findings are outlined and their relevance to using and interpreting the data is described.

2. Sample and engagement

The tests took place in the homes of cohort members as part of the main interview. The CANTAB eclipse software which administered the procedure was integrated into the computer assisted personal interview (CAPI) scripts. Responses were recorded using in-built touchscreens on interviewer's CAPI machines. Those interviewers using machines that did not have this facility used an attachable touchscreen add-on alongside their CAPI machine. Interviewers were also administered with scripts to read out while demonstrating the test to cohort members. These insured all participants received a standard briefing prior to undertaking the tests themselves.

Following the test demonstration and data collection from the cohort members, interviewers were asked to complete a series of ten questions regarding the conditions within the home that the test took place in. The questions referred to whether any common technical problems had occurred, such as the touchscreen freezing. They also covered whether they felt the child was tired or hungry, any interruptions to the child during the test, possible causes of background noise such as conversation or electronics or disturbances from people entering or leaving the room/house. Finally interviewers were asked to include any other factors they thought might have influenced data collection. This section was completed outside of the family house, as close as possible to the administration of the cognitive tests.

In total, 13,287 families took part in the 5th sweep of MCS and the survey collected data from 13,469 children. 12,705 cohort members completed both tasks. There was no completed data for either CANTAB tests from 708 cohort members. Four completed only the CGT and 52 completed only the SWM.

From the full sample, 13,261 parents consented to both assessments and 13,252 children consented to both. A further 16 parents consented only to the CGT and 15 to the SWM only. In the case of the children that consented to a single task only, 21 consented to just the CGT and 20 to the solely the SWM. Non-consent by parents in the CGT was a concern, as it required cohort members to place bets and might therefore be unsuitable for religious or moral reasons. It was therefore introduced to families not as a gambling procedure, but as a decision making task. The engagement data clearly show that selective consent against the CGT was not present and indeed marginally more participants engaged in this procedure.

2.1 Normative scores

Neither of the current CANTAB tests can be compared to an extensive repository of population scores, especially at the age tested at the 5th sweep of MCS. Indeed it is hoped that the data from cohort members will contribute extensively to normative population scores in these tests. Where available, this data note will reference the available collection of normative scores for either the SWM or CGT, for each measure in the tests. Data users may also wish to refer to publications that compare samples of atypical populations with controls, in order to ascertain what may represent performance by typically developing populations.

3. The Cambridge Gambling Task

The Cambridge Gambling Task (CGT) measures risk-taking behaviour and decision-making under uncertainty. It can be contrasted with widely used tests including the Balloon Analog Risk Taking Task (BART) and Iowa Gambling Task (IGT) in that the CGT asks participants to make bets under conditions of known risk, rather than ambiguity (e.g., Bechara, Damasio, Tranel & Damasio 2005; Lejuez et al., 2002). The test therefore minimises learning, executive and working memory demands on participants, which can confound the interpretation of test scores. It also separates the decision-making - where participants choose what to bet on - from risk-taking, where participants decide how much then to bet on that choice.

On each test trial the participant is presented with ten boxes, coloured either red or blue. They are informed that a yellow token is hidden in one of the boxes. At the bottom of the screen are two response boxes, one for each colour. The participant must use these to guess whether the token is hidden under a red or blue box. The task consists of five stages, each of which is a block of trials. In the first, decision-only stage, participants simply have to guess whether the token is hidden under a red or blue box. The latter four stages are gambling stages. Following the colour decision the participant can bet a proportion of their points (from an initial 100 on each stage) on their confidence in the location of the yellow token. Two of the gambling stages are practice sessions undertaken prior to a test session, so that the participants' performance is ultimately assessed by the two test gambling stages. Bets are made by touching the desired amount in the stake box. The possible values of bets are displayed sequentially at 5 second intervals and the participant then touches the stake box to select their desired bet value. Importantly, one of the test stages will display these bets in ascending order, the other in descending order. Participants are informed that correct bets will be added onto their points score and incorrect ones taken away and that they should try to win as many points as possible. They are also briefed regarding how the bets will be displayed (ascending or descending), which stages are practice blocks and which involve gambles.

The administration script of the test used the term 'decision-making task' and avoided the language of bets and gambling, which might no have been considered an appropriate by families, for religious or ethical reasons. Nonetheless, five participants who did not complete the task were recorded by interviewers as having problems with the term 'decision-making task'. The test can be administered with either the ascending or descending gambling stage undertaken first. All the cohort members in the survey took the test with the ascending stage presented first.

Table 2: Key outcome measures and relevant variable names in the Cambridge Gambling Task

Outcome	**Variable**	**Description**
Deliberation time (milliseconds) [a]	CGTTTIME	*Mean time taken to make a box colour response (milliseconds).*
Quality of decision making [a]	CGTQOFDM	*The mean proportion of trials where the participants selects the correct colour outcome.*
Delay aversion*	CGTDELAY	*Difference in percentage bet in ascending verses descending condition.*
Overall proportion bet[a]	CGTOPBET	*The mean proportion of points bet across all trials.*
Risk adjustment	CGTRISKA	*The extent to which betting behaviour is moderated by the ratio of boxes.*
Risk taking[a]	CGTRISKT	*The mean proportion of points bet on trials where the most likely outcome was chosen.*
Test duration (seconds)	CGTDTIME	*Length of test (seconds)*

[a] Key outcome

3.1 Outcome measures

The test outcome measures, relevant variable names and outcome descriptions shown in Table 2. The Cambridge Gambling Task provides six distinct outcome measures as well as indicators of the test outcome and the duration of the test. The table identifies six key outcomes which are of substantive interest as measures of risk taking and/or decision making (deliberation time, quality of decision making, delay aversion, overall proportion bet, risk adjustment and risk taking). The remaining variable (test duration) pertains to the duration of the test procedure. Distributions are given for each of the key variables, indicating the frequency of all scores across participants. The large cohort sample tested presently means that analyses should be fairly robust to variables departing from normal distributions. Nonetheless, measures covering a range of responses will be most likely to enable researchers to analyse the predictors of a particular variable. The data in such variables may therefore be the most useful to users interested in cognitive outcomes.

3.1.1 Deliberation time

Deliberation time is an RT measure indicating the participant's latencies in making a choice response on which colour to bet upon. Longer deliberation times have previously been associated with neurocognitive damage form long-term alcohol consumption and acquired damage through accident and injury (e.g. Lawrence, Luty, Bogdan, Sahakian & Clark, 2009; Newcome et al., 2011). Shorter RTs may indicate impulsive decision making, however in the Cambridge Gambling Task delay does not increase the available information for decision making with elapsed time. Therefore longer RTs do not mean that decisions were made in lieu of all available information. They provide a measure of pre-motor processing and movement time after the decision making information is presented, from the initial presentation of the boxes, until the screen is touched by the participant.

Figure 1: Distribution of Deliberation time in the Cambridge Gambling task

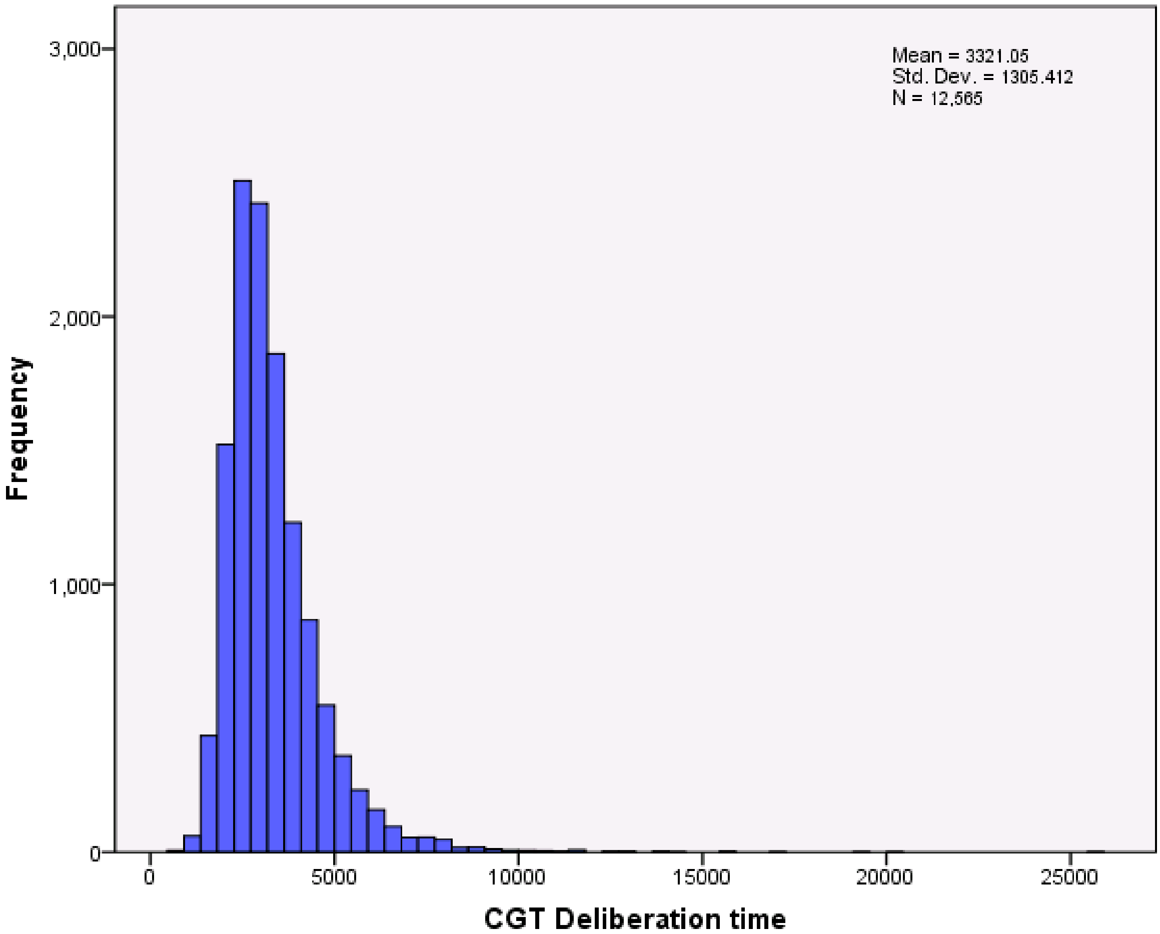


Figure 1 shows the distribution of deliberation time scores. Typically, RTs are characterised by positively skewed distributions and this is evident from the present data (Whelan, 1998). The extended, flat tail at the positive end of the distribution indicates that the mean value of this variable may also be influenced by extreme outliers. There was no time-out for trials and some participants could potentially have recorded deliberation times far in excess of the general mean.

3.1.2 Risk adjustment

Participants are expected to gamble an increased amount of their points when the odds are in their favour. Specifically, higher bets should be evident when the majority of the boxes are that of the colour chosen. Risk adjustment measures this tendency to gamble higher proportions of points on trials where a larger proportion of the boxes the participant's chosen colours.

It is calculated using the following formula: $(2a + b - c - 2d)/e$

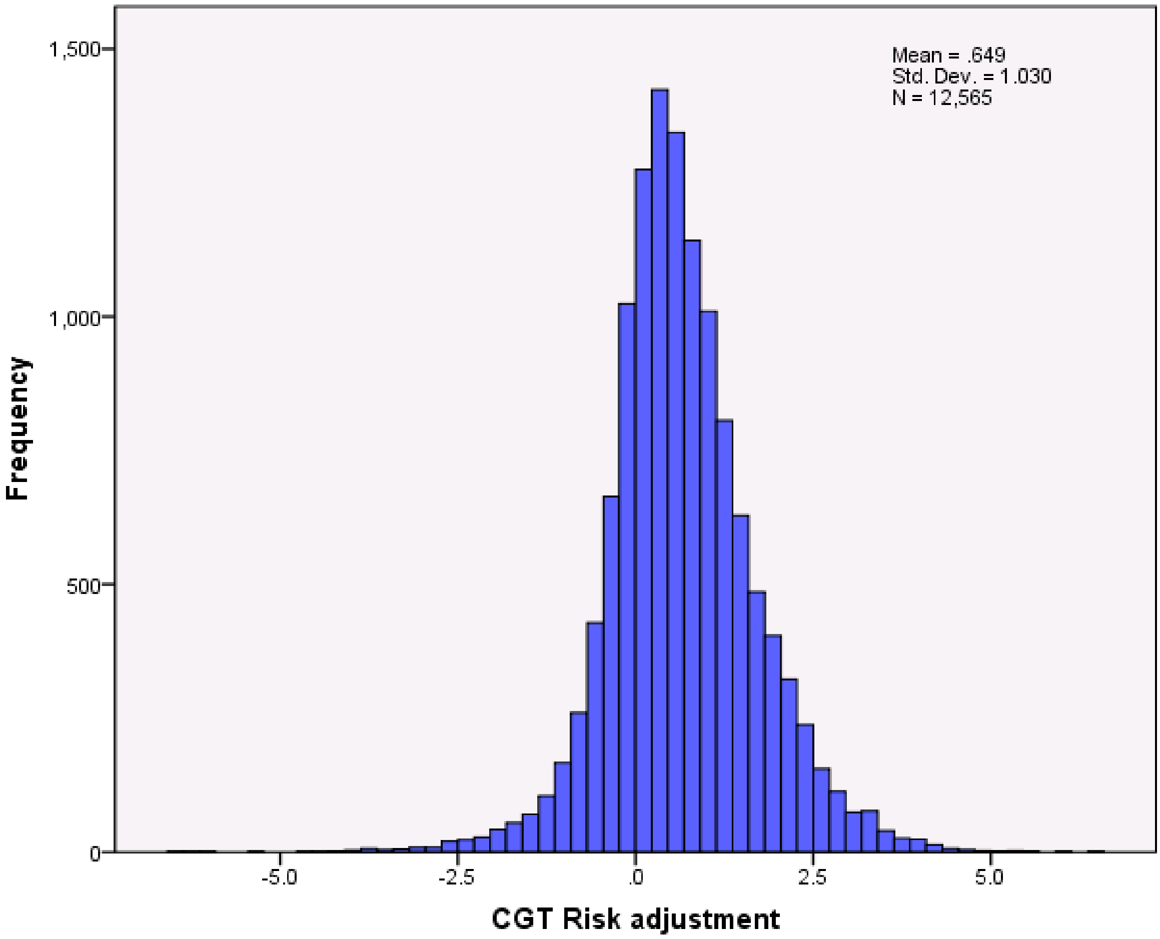
Where:
- a = mean proportion where chosen colour ratio is 9:1
- b = mean proportion where chosen colour ratio is 8:2
- c = mean proportion where chosen colour ratio is 7:3
- d = mean proportion where chosen colour ratio is 6:4

e = mean proportion risked over all trials

Higher scores therefore represent a higher proportion of the overall mean proportions that are bet when the majority of boxes are congruent with the colour chosen.

Figure 2 shows the distribution of risk adjustment scores. This distribution is slightly peaked, with the majority of scores clustering just above 0, indicating that most of the cohort did bet higher amounts of points when the ratio of boxes was in their favour.

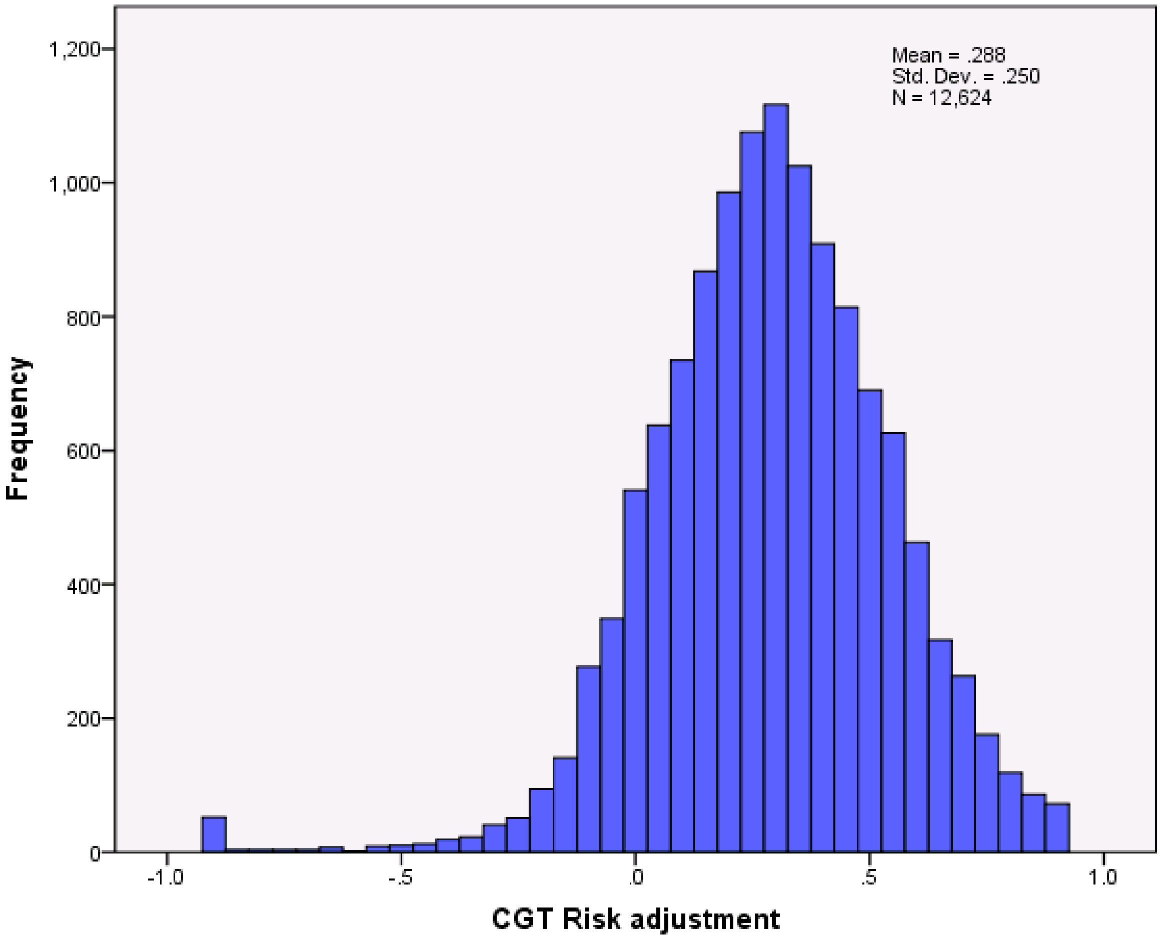
Figure 2: Distribution of Risk adjustment in the Cambridge Gambling task



3.1.3 Delay aversion

If participants are unwilling or unable to wait to make a decision, then they will be more likely to bet larger amounts when the possible bets are displayed in descending rather than ascending order. The delay aversion outcome measures this behaviour. The measure is calculated by subtracting the risk-taking score from ascending trials from that of the descending trials. Higher scores indicate the size of bets is determined more by their presentation sequence than by deliberation. Figure 3 shows the distribution of scores for the Delay Aversion variable. The distribution is broadly normal in shape; however notable is the spike at the negative tail, which in turn gives the distribution a slightly negative skew. The peak of the distribution is around .3, suggesting that there is some delay aversion in general. The noticeable spike in left tail of the distribution for scores of -.9 may represent around 50 cases where the child stopped the test before taking any descending trials. As the descending trials would then have a value of 0, subtracting the ascending score would result in a negative number.

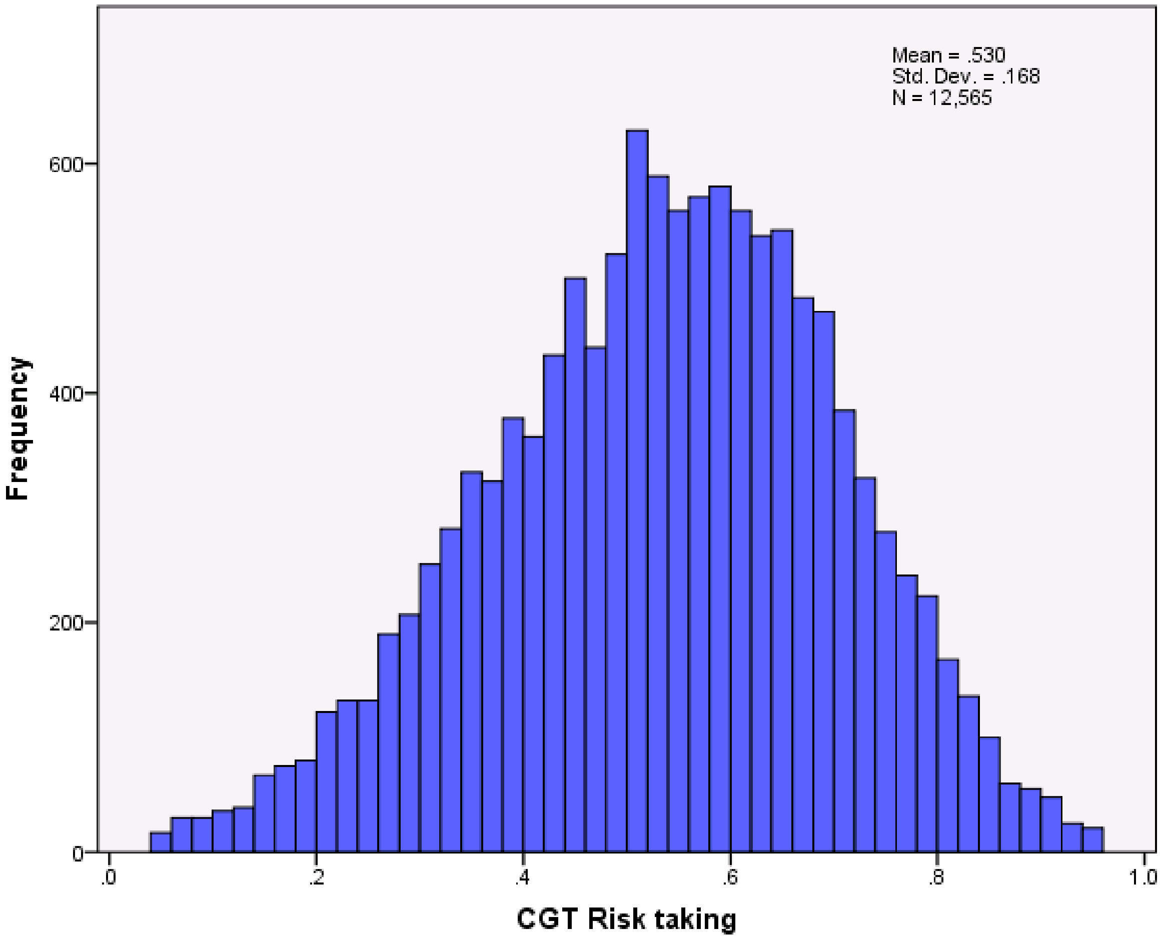
Figure 3: Distribution of Delay aversion in the Cambridge Gambling task



3.1.4 Risk taking

The risk taking outcome is the mean proportion of the current points total that the participant chooses to gamble on trials when they have selected the most likely outcome. Figure 4 shows the distribution of the risk taking measure. As the nominal percentage displayed in the stake box lies between 5 per cent and 95 per cent, the limits of the risk taking distribution are .05 and .95. The scores are therefore broadly normally distributed with only slightly short tails. The majority of scores for this measure are just over .5.

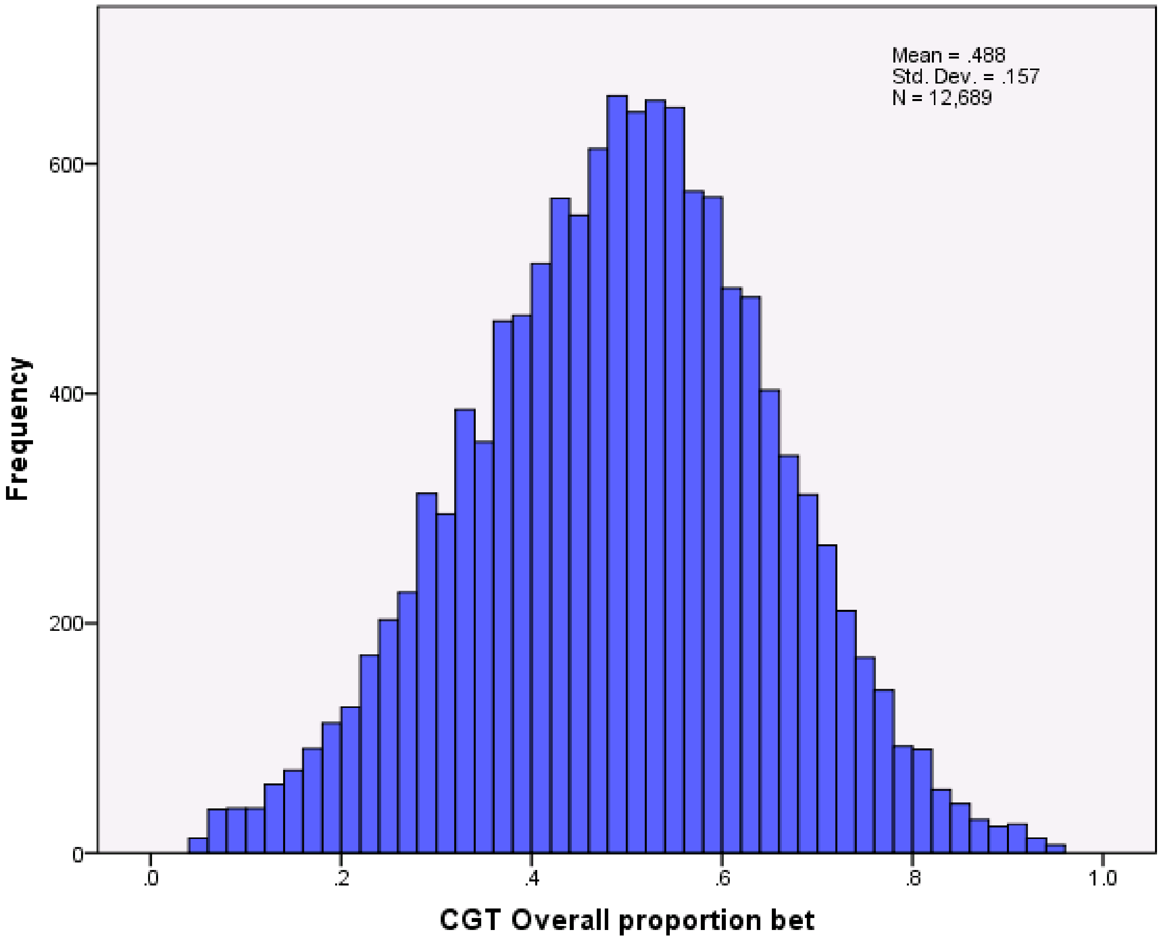
Figure 4: Distribution of risk taking in the Cambridge Gambling task.



3.1.5 Overall proportion bet

This measure is the mean proportion of the participant's current points total across all trials. This includes those trials where they bet on the less likely outcome or that where both outcomes were equally likely. This outcome is broadly normally distributed with the majority of scores falling in the centre of the distribution. Like risk taking, the limits of the scores of .05 and .95 give the distribution marginally shortened tails. However, the distribution of the overall proportion bet is more broadly normal, i.e. more centred around zero, than risk taking, reflecting that a slightly higher proportion bet on more likely outcomes.

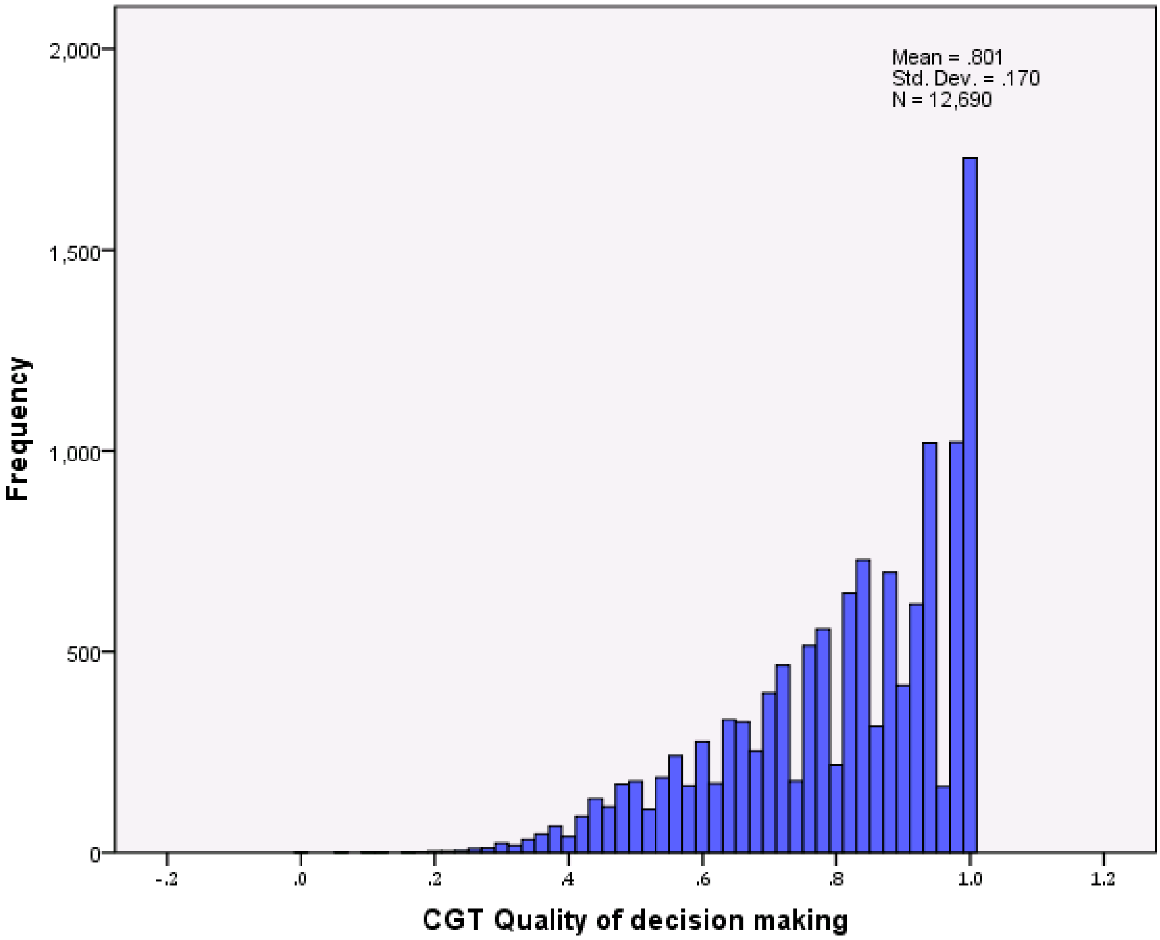
Figure 5: Distribution of overall proportion bet in the Cambridge Gambling Task



3.1.6 Quality of decision making

The quality of decision making measure is the proportion of trials where participants bet on the most likely outcome. Figure 6 shows the distribution for quality of decision making. The distribution for the quality of decision making variable is extremely negatively skewed in shape, with the majority of responses being bet on the most likely outcome. This distribution is indicative of a ceiling effect in the measure.

Figure 6: Distribution of quality of decision making scores in the Spatial Working Memory task



4. Spatial Working Memory

The assessment of spatial working memory (SWM) measures participants' ability to preserve spatial information, including the use of heuristic memory strategies. Although it is sensitive to spatial working memory capacity it is also thought to be measure executive function.

The aim of the task for each participant is to find the blue tokens which are hidden under the coloured boxes in the display. They must search the boxes by touching them in sequence to determine if there is a blue token hidden underneath. Once they have found a token, they can move the token to the right-hand 'home' column, on the right-hand side of the screen by touching this area. They then proceed to search the next token. All of the touch responses required to find a single token are known as a search.

A single trial consists of the searches needed to find a blue token under each box and place them in the home column. Once the first token has been found, subsequent tokens may be in any of the boxes that have not previously contained a token. Participants are informed that touching a box that has already contained a token in that trial is an error and that they must avoid doing this. They must also avoid returning to touch a box that has found to be empty within that search. The number of boxes to be searched (not including errors) is determined by the software.

A block contains a number of trials, each presenting the same number of boxes to the participant. The colour and location of the boxes changes between each trial in order to discourage the employment of repetitive or stereotyped search strategies.

Table 3: Outcome variables for the Spatial Working Memory task

Outcome	Variable	Description
Between errors	SWMBERRS	*Total frequency participants touched a box found to contain a token*
Between errors 4 boxes	SWMBE4BX	*Frequency participants touched a box found to contain a token on 4 box trials*
Between errors 4 to 8 boxes	SWMBE8BX	*Frequency participants touched a box found to contain a token on 4, 6 and 8 box trials*
Within errors	SWMWERRS	*Total frequency participants touched a box already touched in same search*
Within errors 4 boxes	SWMWE4BX	*Frequency participants touched a box already touched in same search for 4 box trials*
Within errors 4 to 8 boxes	SWMWE8BX	*Frequency participants touched a box already touched in same search for 4, 6 and 8 box trials*
Double errors	SWMDERRS	*Total frequency participants touched a box already touched in same search that already found to contain a token*
Double errors 4 boxes	SWMDERR4	*Frequency participants touched a box already touched in same search that already found to contain a token for 4 box trials*
Total errors 4 boxes	SWMTE4BX	*Frequency all errors on 4 box trials*
Total errors 4 to 8 boxes [a]	SWMTE8BX	*Frequency all errors on 4, 6 and 8 box trials*
Strategy 4 to 8 boxes	SWMS48	*Degree to which searches employ sequential heuristic for 6 and 8 box trials*
Strategy [a]	SWMSTRAT	*Overall degree to which searches employ sequential heuristic*
Mean time to first response (milliseconds)	SWMMTOFR	*Latency from trial onset to first touch response*
Mean time to last response (milliseconds) [a]	SWMMTTLR	*Latency from trial onset to touch response on final blue token*
Test outcome	SWMOUTCM	*Whether test completed*
Test duration (seconds)	SWMTTIME	*Length of test (seconds)*
Problem reached	SWMPROBR	*Which trial participants reached*

[a] Key Outcome

4.1 Outcome measures

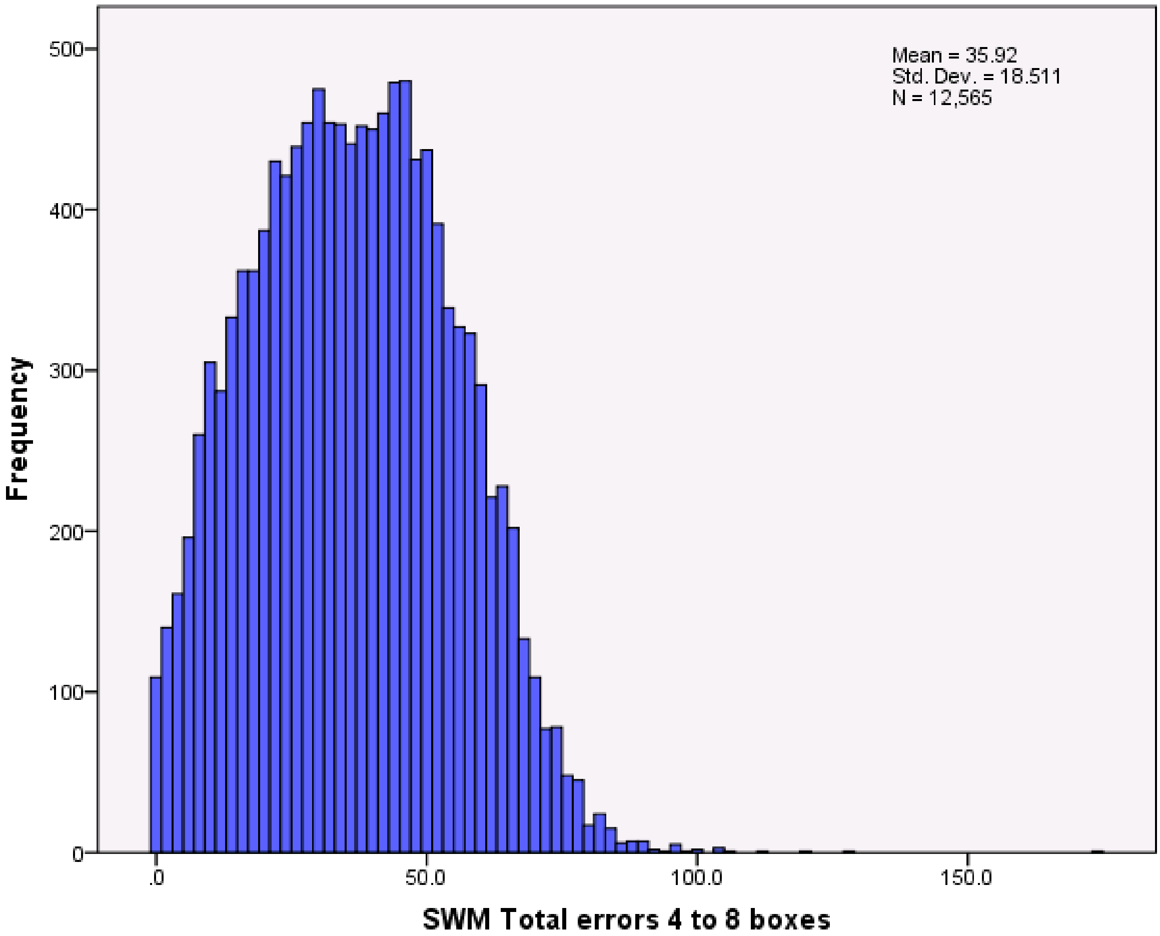
The task consists of a practice block of 3 trials, containing 3 boxes and tokens each. There are then 3 further blocks of test trials. The first block contains 4 trials with 4 boxes. The second block includes 4 trials with 6 boxes and finally the last block contains 4 trials including 8 boxes. The latter blocks are therefore more difficult and performance can be enhanced by the use of a search strategy. Table 2 describes the outcome variables and variable names from the SWM task. The key outcomes of this test are the total errors, time taken until last response and strategy. In the case of the all three variables, higher scores are indicative of worse performance.

4.1.1 Total errors 4 to 8 boxes

Total errors are the number of times a subject touches a box that is certain not to contain a token. It is therefore the sum of errors made within searches (within errors, i.e. whether they revisited a box known to be empty) and the number of errors made between searches (between errors, i.e. whether they revisited a box where a blue token had already been found). As some touch responses can be both errors made within and between searches (double errors) these are removed from the total errors score so a single response does not generate more than one type of error. The subcomponent variables for total errors are available for each error type, for the trials in each box number block.

Figure 7 shows the distribution of total errors by the cohort members. The distribution is somewhat positively skewed and flattened in shape, with a minority of scores reaching a high number of errors. These participants may have failed to engage with the task or not adhered to interviewer’s instructions.

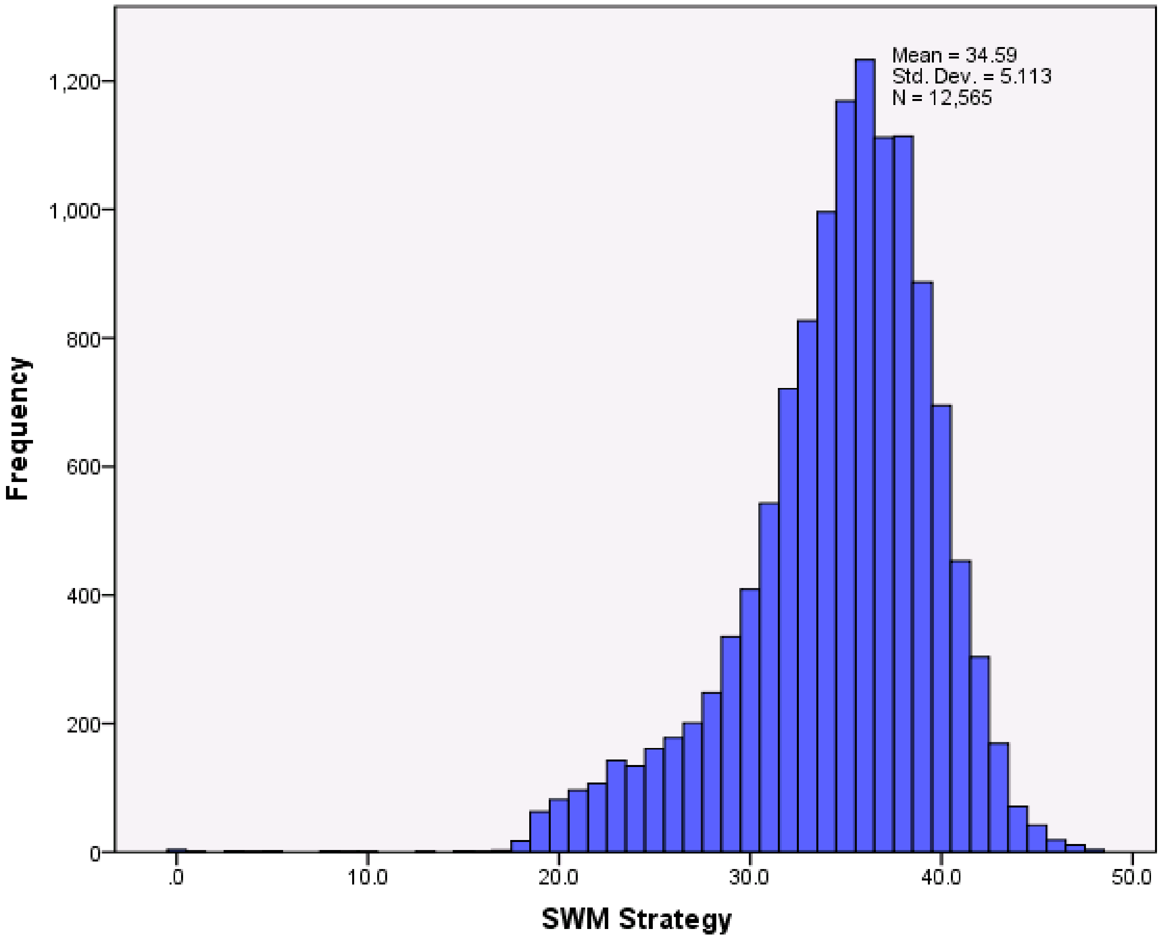
Figure 7: Distribution of Total errors 4 to 8 boxes in the Spatial Working Memory task



4.1.2 Strategy

The strategy measure is an estimate of the use of a heuristic search sequence in order to complete the task in the most efficient way. It has been suggested that this strategy would adopt a fixed sequence where a box is chosen to begin the search and once a blue token has been found, that original box is used as a start point for the next search sequence (Owen et al., 1990).

Figure 8: Distribution of Strategy scores in the Spatial Working Memory task



This strategy is estimated by counting the frequency of first responses that begin on a different box to that of the first response of the previous search. Thus a high strategy score indicates a less optimal use of this strategy and a lower scorer a more effective use. The strategy outcome is calculated using 6 and 8 box trials only. The minimum possible strategy score is therefore 1 for each test stage (a total of 8) and the maximum possible score is 1 for each search (a total of 56).

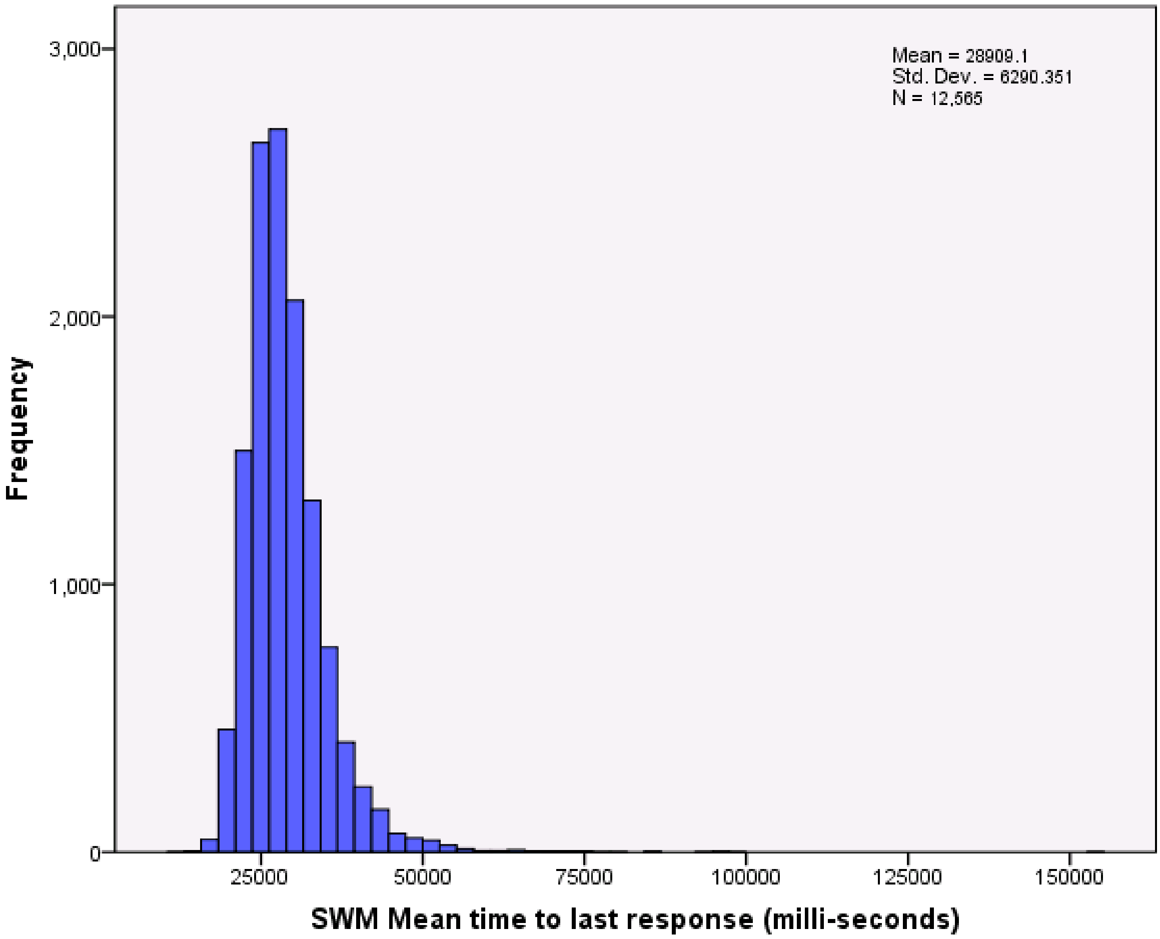
Figure 8 depicts the distribution of strategy scores among participants. The distribution has a moderate negative skew and is peaked in shape with short tails. The majority of scores are found just above the centre of the distribution, indicating more cases employed a less optimal strategy. Very few cases scored less than 15, although there was a very small cluster that scored 0, indicating a consistent return to the same box as previous searches.

4.1.3 Mean time to last response

Mean time to last response measures the mean RT from each trial's boxes being presented to the last screen touch to locate the final token of the trial. This mean latency is calculated across 4, 6 and 8 box trials.

Figure 9 shows the distribution of mean time to last response measurements. RTs again show a typically positively skewed distribution, similar in shape to the deliberation time measures in the CGT. Again the lack of a time-out mechanism in the test may further exaggerate the mean scores, due to extreme outliers in the positive tail.

Figure 9: Distribution of mean time to Last Response in the Spatial Working Memory task



5. The influence of interview factors

Following the administration of the tests, the interviewers were prompted to record observation of technical, environmental or personal factors which may have affected the administration of the test. These factors did affect some of the initial outcomes of the tests among cohort members and users of the data should be aware of factors which modulated the key variables. The SWM task was particularly affected by most of the test factors described below. The SWM test may be particularly sensitive to disruptions, due to the reduction of cognitive resources. Some of the variables of the CGT were less predictable in terms of its sensitivity to disruptions as they do not measure cognition per se but are also affected by trait factors such as personality and cognitive style. These measures may additionally be influenced by state variables such as affective arousal and state anxiety. This is the case especially for the risk-taking variables which are concerned with the quantities bet, where test disruptions may affect performance but their direction is less clear. In the case of decision-making variables such as quality of decision making and deliberation time, disruptions during the test should affect performance in a similar direction to the SWM outcomes.

5.1 Apparatus

5.1.1 Use of attached touchscreen

For the administration of 48.1 per cent of cases tested, interviewers were able to employ the built-in touch screen facility of their laptops. For 50.8 per cent of cases, interviewers did not have this facility and had to use an external attachment. Of particular concern was the effect of any systematic difference in the registration of response time for latency variables. Across both tasks, only CGT deliberation time was significantly affected by the use of the touch screen add-on. The mean difference was 70.91 ms, around 1/19 of a standard deviation. When considered alongside group differences typically detected in this task, this can be considered a negligible effect. Users can therefore be confident that the use of the external apparatus does not need to be controlled for when examining the test outcomes.

5.1.2 Freezing and recalibrating the touchscreen

Data collected from cases where the touchscreen froze and had to be recalibrated differed significantly from those without this technical problem for a number of variables in both tasks. Overall 332 productive cases experienced this technical issue. The majority (399 cases) used the external touch screen, whilst a remaining 35 cases were tested using built-in apparatus. Table 4 and 5 describe the effect of the touchscreen freezing on the key test variables from both tasks.

There was a significant effect of touch screen freezes on all of the CGT variables, with the exception of deliberation time and risk adjustment ($p < 0.05$). There was also a significant effect on both total errors and strategy in the SWM. This technical problem had the most pronounced effect on quality of decision and delay aversion measures of the CGT. For the SWM, it appeared to increase the strategy measure and reduce the total errors. This was the only disruption that did not affect response times in either task.

Table 4: Mean group differences and parameter estimates for the touchscreen freezing across key Cambridge Gambling Task variables

Variable	Touchscreen froze?	Mean	Standard Error	Upper CI	Lower CI	Observations
Deliberation time (milliseconds)	No	3328.92	12.02	3305.36	3352.48	12274
	Yes	3467.71	112.02	3247.32	3688.09	324
	Total	3332.49	12.06	3308.85	3356.13	12598
Quality of decision making* **	No	0.80	0.00	0.80	0.80	12274
	Yes	0.83	0.01	0.81	0.85	324
	Total	0.80	0.00	0.80	0.80	12598
Risk taking**	No	0.53	0.00	0.53	0.53	12273
	Yes	0.55	0.01	0.54	0.57	324
	Total	0.53	0.00	0.53	0.53	12597
Overall proportion bet***	No	0.49	0.00	0.48	0.49	12273
	Yes	0.52	0.01	0.50	0.53	324
	Total	0.49	0.00	0.49	0.49	12597
Delay aversion	No	0.29	0.00	0.28	0.29	12213
	Yes	0.29	0.01	0.26	0.32	319
	Total	0.29	0.00	0.28	0.29	12532
Risk adjustment	No	0.65	0.01	0.63	0.67	12273
	Yes	0.59	0.05	0.50	0.69	324
	Total	0.65	0.01	0.63	0.66	12597

*Note:**** p < .001 ** p < .01 * p < 05

Table 5: Mean group differences and parameter estimates for the touchscreen freezing across key Spatial Working Memory Task variables

Variable	Touchscreen froze?	Mean	Standard Error	Upper CI	Lower CI	Observations
Total errors* **	No	35.79	0.168	35.46	36.12	12311
	Yes	31.83	1.18	29.51	34.15	344
	Total	35.68	0.167	35.35	36	12655
Mean time to last response	No	28904.25	56	28794	29015	12256
	Yes	29480.43	477	28541	30420	290
	Total	28917.57	56	28807	29028	12546
Strategy* **	No	34.47	0.049	34.37	34.57	12311
	Yes	29.13	0.728	27.7	30.57	344
	Total	34.33	0.053	34.22	34.43	12655

*Note:**** $p < .001$ ** $p < .01$ * $p <$ 05

5.2 Test conditions

5.2.1 Interruptions

Unsurprisingly, interruptions during the tests by another adult or child significantly affected all of the test variables ($p < 0.01$). Tables 6 and 7 show the group mean differences for increasing numbers of interruptions on the test variables. Differences in outcomes due to interruptions were most evident in the delay aversion and risk adjustment measures of the CGT. In particular, a strong linear effect was present for risk adjustment, reducing this measure with increasing interruptions. For the SWM, interruptions considerably increased the amount of total errors recorded and the time to last response. Interestingly there was a linear effect of increasing the strategy index with increasing numbers of interruptions, however this may have been due to the greater cognitive effort required to undertake the test with interruptions and therefore the increased use of heuristic strategies.

Table 6: Mean group differences and parameter estimates for the number of interruptions across key Cambridge Gambling Task Variables

Variable	Interuptions	Mean	Standard Error	Upper CI	Lower CI	Observations
Deliberation time (milliseconds) * **	None	3302.14	13.23	3328.07	3276.20	10244
	1 or 2	3428.86	31.15	3489.95	3367.78	1806
	3 or more	3482.11	79.60	3638.56	3325.66	429
	Total	3326.67	12.08	3350.35	3302.98	12479
Quality of decision making* **	None	0.80	0.00	0.81	0.80	10244
	1 or 2	0.79	0.00	0.80	0.78	1806
	3 or more	0.79	0.01	0.80	0.77	429
	Total	0.80	0.00	0.80	0.80	12479
Risk taking* **	None	0.53	0.00	0.53	0.52	10243
	1 or 2	0.54	0.00	0.54	0.53	1806
	3 or more	0.56	0.01	0.57	0.54	429
	Total	0.53	0.00	0.53	0.53	12478
Overall proportion bet* **	None	0.48	0.00	0.49	0.48	10243
	1 or 2	0.50	0.00	0.50	0.49	1806
	3 or more	0.52	0.01	0.53	0.50	429
	Total	0.49	0.00	0.49	0.48	12478
Delay aversion* **	None	0.29	0.00	0.29	0.28	10189
	1 or 2	0.30	0.01	0.31	0.29	1801
	3 or more	0.33	0.01	0.35	0.30	425
	Total	0.29	0.00	0.29	0.28	12415
Risk adjustment* **	None	0.67	0.01	0.69	0.65	10243
	1 or 2	0.58	0.02	0.63	0.53	1806
	3 or more	0.42	0.05	0.52	0.32	429
	Total	0.65	0.01	0.67	0.63	12478

*Note:**** p < .001 ** p < .01 * p < 05

Table 7: Mean group differences and parameter estimates for the number of interruptions across key Spatial Working Memory variables

Variable	Interruptions	Mean	Standard Error	Upper CI	Lower CI	Observations
Total errors* **	None	34.92	0.184	35.28	34.56	10290
	1 or 2	38.6	0.442	39.47	37.74	1811
	3 or more	42.29	0.868	44	40.59	433
	Total	35.71	0.168	36.04	35.38	12534
Mean time to last response* **	None	28822.35	61.462	28942.82	28701.87	10203
	1 or 2	29378.63	165.894	29704	29053.26	1798
	3 or more	29510.43	272.081	30045.22	28975.64	427
	Total	28926.47	56.683	29037.57	28815.36	12428
Stategy* **	None	34.17	0.058	34.28	34.05	10290
	1 or 2	35.1	0.13	35.35	34.84	1811
	3 or more	35.09	0.302	35.69	34.5	433
	Total	34.33	0.053	34.44	34.23	12534

*Note:**** $p < .001$ ** $p < .01$ * $p < 05$

5.2.2 Background noise

Ambient noise also significantly affected test performance on all key outcomes, although somewhat to a lesser extent than direct interruptions ($p < 0.05$). Group differences in the outcome measures based upon noise levels from conversation, electronics and other background disturbances are shown in tables 8 and 9 for each CANTAB test. The presence of noise slowed deliberation time relative to cases where none was recorded. It also notably reduced the risk taking and risk adjustment measures. Like interruptions, the background noise increased total errors and mean time to last response during the SWM, while also slightly increasing strategy scores.

Table 8: Mean group differences and parameter estimates for the level of background noise across key Cambridge Gambling Task variables

Variable	Noise	Mean	Standard Error	Upper CI	Lower CI	Observations
Deliberation time ***	None	3283.32	15.00	3312.72	3253.92	7419
	Audible	3380.10	20.99	3421.26	3338.94	4408
	Loud	3456.77	63.94	3582.32	3331.21	658
	Total	3326.63	12.08	3350.31	3302.95	12485
Quality of decision making***	None	0.81	0.00	0.81	0.80	7419
	Audible	0.80	0.00	0.80	0.79	4408
	Loud	0.77	0.01	0.79	0.76	658
	Total	0.80	0.00	0.80	0.80	12485
Risk taking***	None	0.52	0.00	0.53	0.52	7419
	Audible	0.54	0.00	0.54	0.53	4407
	Loud	0.54	0.01	0.55	0.53	658
	Total	0.53	0.00	0.53	0.53	12484
Overall proportion bet***	None	0.48	0.00	0.49	0.48	7419
	Audible	0.49	0.00	0.50	0.49	4407
	Loud	0.50	0.01	0.51	0.49	658
	Total	0.49	0.00	0.49	0.48	12484
Delay aversion**	None	0.28	0.00	0.29	0.28	7377
	Audible	0.29	0.00	0.30	0.28	4390
	Loud	0.32	0.01	0.33	0.30	653
	Total	0.29	0.00	0.29	0.28	12420
Risk adjustment***	None	0.70	0.01	0.72	0.68	7419
	Audible	0.59	0.02	0.62	0.56	4407
	Loud	0.47	0.04	0.55	0.39	658
	Total	0.65	0.01	0.67	0.63	12484

*Note:**** $p < .001$ ** $p < .01$ * $p <$ 05

Table 9: Mean group differences and parameter estimates for the level of background noise across key Spatial Working Memory Task variables

Variable	Interruptions	Mean	Standard Error	Upper CI	Lower CI	Observations
Total errors* **	None	34.24	0.217	34.67	33.82	7453
	1 or 2	37.33	0.28	37.88	36.78	4426
	3 or more	41.31	0.707	42.7	39.92	662
	Total	35.7	0.168	36.03	35.38	12541
Mean time to last response* **	None	28758.63	73.089	28901.9	28615.35	7383
	1 or 2	29121.6	96.6	29310.98	28932.21	4395
	3 or more	29429.09	232.758	29886.13	28972.05	656
	Total	28922.3	56.596	29033.23	28811.36	12434
Strategy* **	None	34	0.07	34.14	33.86	7453
	1 or 2	34.75	0.084	34.91	34.58	4426
	3 or more	35.29	0.225	35.73	34.84	662
	Total	34.33	0.053	34.43	34.23	12541

*Note:**** $p < .001$ ** $p < .01$ * $p < 05$

5.3 Sample and individual differences

5.3.1 Fatigue

A number of interviewers coded whether they felt the child showed signs of tiredness during the test. The effect of this fatigue coding on the group means of the test variables are shown in Tables 10 and 11. There was a significant effect of child fatigue in all of the test variables, with the exception of delay aversion in the Cambridge Gambling task (p < 0.05). Children coded as tired tended to have longer deliberation times in the CGT and accordingly, reduced delay aversion. In the SWM, total errors and mean time to last response were increased, while strategy scores also slightly increased with fatigue.

Table 10: Mean group differences and parameter estimates for the level of fatigue across key Cambridge Gambling Task variables

Variable	Fatigue	Mean	Standard Error	Upper CI	Lower CI	Observations
Deliberation time* **	Tired	3437.59	37.61	3511.36	3363.83	1645
	Awake	3310.15	12.69	3335.02	3285.28	10813
	Total	3326.98	12.09	3350.66	3303.29	12458
Quality of decision making* **	Tired	0.79	0.00	0.80	0.78	1645
	Awake	0.80	0.00	0.81	0.80	10813
	Total	0.80	0.00	0.80	0.80	12458
Risk taking*	Tired	0.54	0.00	0.55	0.53	1645
	Awake	0.53	0.00	0.53	0.52	10812
	Total	0.53	0.00	0.53	0.53	12457
Overall Proportion bet*	Tired	0.50	0.00	0.50	0.49	1645
	Awake	0.49	0.00	0.49	0.48	10812
	Total	0.49	0.00	0.49	0.48	12457
Delay aversion	Tired	0.29	0.01	0.30	0.28	1628
	Awake	0.29	0.00	0.29	0.28	10766
	Total	0.29	0.00	0.29	0.28	12394
Risk adjustment* **	Tired	0.57	0.03	0.62	0.52	1645
	Awake	0.66	0.01	0.68	0.64	10812
	Total	0.65	0.01	0.67	0.63	12457

*Note:**** $p < .001$ ** $p < .01$ * $p < 05$

Table 11: Mean group differences and parameter estimates for the level of fatigue across key Spatial Working Memory task variables

Variable	Fatigue Level	Mean	Standard Error	Upper CI	Lower CI	Observations
Total errors***	Tired	39.04	0.47	39.96	38.12	1660
	Awake	35.19	0.80	35.54	34.84	10853
	Total	35.7	0.17	36.03	35.37	12513
Mean time to last response***	Tired	29801.03	163.75	30122.22	29479.84	1639
	Awake	28791.37	60.34	28909.65	28673.08	10768
	Total	28924.75	56.74	29035.97	28813.52	12407
Strategy***	Tired	34.81	0.15	35.11	34.52	1660
	Awake	34.26	0.06	34.37	34.15	10853
	Total	34.33	0.05	34.44	34.23	12513

*Note:**** $p < .001$ ** $p < .01$ * $p < 05$

6. Conclusions

This data note presented an overview of the measures obtained by two CANTAB neuropsychological tests, which were administered to participants of the 5th sweep of the Millennium Cohort Study. Both the Cambridge Gambling Task and Spatial Working Memory task test cognitive style and task strategy. The interpretations of the key measures were set out, alongside the distributions of corresponding variables and descriptive statistics. In addition, the effects of a number of key circumstances in the cognitive tests, such as technical problems, environmental disruption and participant fatigue, were presented.

Researchers may wish to consider controlling for some of the cognitive circumstances which have notable effects upon the mean values. For example and in particular, high levels of background noise and interruptions have large effects on risk adjustment in the CGT and total errors in the SWM. Removing participants who had highly disruptive tests should increase the available power of analyses and better approximate lab, clinical and educational settings for which these tests were originally intended.

7. References

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in cognitive sciences*, *9*(4), 159-162.

Lawrence, A. J., Luty, J., Bogdan, N. A., Sahakian, B. J., & Clark, L. (2009). Problem gamblers share deficits in impulsive decision-making with alcohol-dependent individuals. *Addiction*, *104*(6), 1006-1015.

Luciana, M. (2003). Practitioner review: computerized assessment of neuropsychological function in children: clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). *Journal of Child Psychology and Psychiatry*, *44*(5), 649-663.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L. & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75.

Newcombe, V. F., Outtrim, J. G., Chatfield, D. A., Manktelow, A., Hutchinson, P. J., Coles, J. P. & Menon, D. K. (2011). Parcellating the neuroanatomical basis of impaired decision-making in traumatic brain injury. *Brain*, *134*(3), 759-768.

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders*, *5*(5), 266-281.

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. (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. *Journal of the International Neuropsychological Society*, *4*(05), 474-490.

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *Psychological Record*, *58*(3).

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