

Airwave cognitive data

Clive P. Mitchell

Contents

Introduction.....	1
Raw data - header and summary rows.....	2
1. Reaction (Two-choice reaction time).....	3
2. Pairing (Paired associate learning – episodic memory).....	5
3. Quiz (Fluid intelligence).....	7
4. Stroop (Mild interference/Stroop interference).....	8
5. Number (Working Memory Test - Forward digit span).....	9
Data conversion.....	11
Database build.....	15
Data extraction.....	15
Testing.....	16
Appendix A. Screen shots.....	17
Appendix B. Timing precision.....	24
Appendix C. Quiz questions.....	27

Introduction

A version of the Cardiff Cognitive Battery was developed for the Airwave study. This document describes the tests, the collected raw data and the converted data.

The cognitive tests were displayed in portrait mode on the touch screen of a tablet with the interviewer and user ids in the top-right, the screen id in the bottom-left and a copyright notice in the bottom-right. Each time the screen arrangement changed, the screen id was updated, permitting the point in the program where an issue of interest occurred to be identified easily.

The battery consisted of five tests which measured reaction time, episodic memory, fluid intelligence, Stroop interference and working memory. Apart from one brief period in January 2008 when the order of tests was randomized, the order of presentation was as stated above.

Raw data - header and summary rows

The raw data was stored in a .csv file as a set of rows per test grouped by participant sessions, with one set for each test per session.

The first row of data for each test was a header row containing column names. Padding fields were used to ensure all rows had the same number of columns. The header row was for convenience only. All header row for a particular type of test were identical. Columns 0-7 and column 14 were common across all rows in the file, as shown in Table 0.1.

Column	Title	Description
0	(Row type for header row)	Each test produced three two-digit record types: header record (ending in zero), data record (ending in 1), summary record (ending in 2). The first digit denoted the test type. The header row types were: 10 – Reaction time 20 – Pairing (episodic memory) 30 – Reasoning Quiz (fluid intelligence) 40 – Stroop (interference) 50 – Working Memory (forward-digit span)
1	Test	The short name for the test.
2	Version	The version of the individual test.
3	Clinic	The identifier of the clinic running the test.
4	InterviewerId	The name or identifier of the member of staff at the clinic overseeing the participant.
5	UserId	The participant user id.
6	UserStartDateTime	The date and time when the participant started the cognitive test battery.
7	ScreenId	The identifier of the screen producing the line of data. If the identifier is “Level” then the row is a summary row.
14	Dummy	As some applications truncate .csv file rows where they end in empty fields, a dummy column was populated to prevent this.

Table 0.1. Header columns

Each trial produced one row of data, stored immediately after the trial. The number of fields used in each type of test varied, although the overall number of columns in the file stayed the same. These will be described with the individual tests.

Each level of a test produced a summary row, described in Table 0.2.

Column	Title	Description
0	Summary row type	12 – Reaction time 22 – Pairing 32 – Reasoning Quiz 42 – Stroop 52 – Working Memory
7	Level label	If present, contains “Label” except for early versions of Working Memory which contain “EndTime”
8	Level	“box” or “ink” for the Stroop interference test, “7” for pairs and “1” for other tests. The level only has relevance for the Stroop interference test. Early versions of Working Memory contained the date and time at which the test ended.
9	EndTime label	Contains “EndTime” as the title to the following field on the row.
10	EndTime	The date and time at which the test ended.
11	StartDates label	Contains “StartDates” as the title to the following two fields in the row. Earlier versions of the tests did not contain columns 11-13 in the summary rows.
12	InstructionsStart	The date and time at which the instructions and practice (if any) started.
13	TestStart	The date and time at which the recorded test begins. Subtracting InstructionsStart gives the time taken to read the instructions and complete the practice levels (if any).
14	Dummy	As some applications truncate .csv file rows where they end in empty fields, a dummy column was populated to prevent this.

Table 0.2. Summary row fields.

1. Reaction (Two-choice reaction time)

This test measured the time taken to select between a left or right response to match a left or right stimulus. The data collected could be used to calculate a mean response time and to reveal the effect of varying the response stimulus interval (RSI), fatigue effects and loss of concentration instances.

The primary measure is the mean reaction time but this can be affected by a few trials where the participant loses concentration and returns a disproportionately large reaction

time but this effect is diluted by the large number of 60 trials.

Displayed instructional text is shown in the screen shots in Appendix A, Figures A1.1 to A1.5.

Two black panels were displayed, with a large arrow button beneath each panel.

Each trial stimulus was displayed as a single white square in one of a sequence of three positions in one of the panels. The left panel stimuli appear in the sequence: top left, bottom left, right. The right panel stimuli appear in the sequence: top, bottom right, left. In this way the same stimulus did not appear twice in succession which could have prevented the user from noticing a change of stimulus.

All of the stimuli positions are displayed in Figure 1.

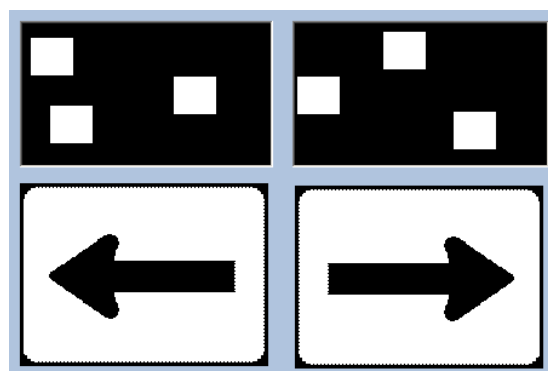


Figure 1. Reaction stimuli positions

The left/right stimuli sequence was identical for every test.

Response stimulus intervals in milliseconds followed the sequence 350, 650, 150, 450, 750, 250, 550, 50. For the 60 recorded trials, the first four intervals were encountered eight times and the last four, seven times.

The first eight screens consist of instructions followed by five practice trials with feedback, and an introduction to the recorded test. The practice trials were not recorded.

The first recorded trial took place at screen RT08 in the program and the last at screen RT67, i.e., 60 recorded trials in all.

Column	Title	Description
0	Trial data row type	"11"
7	ScreenId	"RT08"- "RT67"
8	IfCorrect1Else0	This gives the mark for a trial (1 for correct, 0 for incorrect). Summing the marks gives a score for accuracy.
9	ElapsedTime_ms	The elapsed time recorded in milliseconds but not to millisecond accuracy [Appendix B].
10	CorrectAnswer	The expected correct answer, "Left" or "Right".

		In a normalised database, this would be held in a separate table as the sequence of stimuli does not change.
11	ResponseStimulusInterval_ms	The target response stimulus interval recorded in milliseconds but not to millisecond accuracy [Appendix B]. In a normalised database, this would be held in a separate table as the sequence of response stimulus intervals does not change.
12	Empty	Not used
13	Empty	Not used
14	Dummy	As some applications truncate .csv file rows when they end in empty fields, a dummy column was populated to prevent this.

Table 1. Reaction time trial raw data

Program source: Reaction.h

Versions:

- 1.2 16/07/2007 Earliest released version
- 1.2.1 11/04/2008 Start date/times for instructions and timed test added to summary data line
- 1.2.3 17/02/2009 Enable buttons as well as Show() to rectify driver problem

2. Pairing (Paired associate learning – episodic memory)

This test requires the user to recall the positions of a number of hidden pictures.

The primary measure is the number of attempts required to complete the task.

The test consisted of two similar programs – a practice program and a test program. As only data from the test program was recorded, the practice program will not be described fully here.

A selection of screens are presented in Figure A2.1 to A2.4.

The participant was presented with instructions, a circle of seven blank disabled buttons beneath and a Start button near the bottom of the screen. From the top going clockwise, the buttons are referred to as TopMiddle, RightTop, RightBottom, BottomRight, BottomLeft, LeftBottom and LeftTop. On tapping the Start button, each of the seven buttons had its associated picture displayed for 2000ms and then hidden, in the sequence TopMiddle, RightBottom, LeftTop, BottomLeft, BottomRight, RightTop, LeftBottom.

After the button picture display sequence was completed, a stimulus picture was revealed

in the centre of the circle of buttons and the seven buttons enabled. If the participant clicked an incorrect button, a red X was displayed on that button for 600ms and the participant would keep choosing until the correct button was selected. When the correct button was tapped, the associated picture was displayed on the button for 600ms then the button and stimulus picture were blanked. After 500ms the next stimulus picture was revealed.

The test ended when the participant had found all seven pictures.

Regrettably, the pictures associated with the buttons were embedded into a resx file by Microsoft Visual Studio and so the associated .gif files cannot simply be determined from the source code.

The practice program presented only four pictures and took approximately one minute to complete, as calculated by subtracting the WorkingMemory EndDateTime from the Pairing InstructionsStart.,

Column	Title	Description
7	ScreenId	P7A-P7G
8	NumberOfGuesses	The number of attempts taken to find the button with the same associated picture as the current stimulus.
9	ElapsedTime_ms	The elapsed time taken from the stimulus reveal to selecting the appropriate button recorded in milliseconds but not to millisecond accuracy [Appendix B].
10	Level	As the four-picture practice is not recorded, this is always seven and thus redundant.
11	CorrectAnswer	The position of the correct button. As this is constant for a given screen, this should be contained in a separate table in a normalised database.
12	SelectedAnswer	The button selected by the participant. Originally the participant was given only one chance to find each button, so there was merit in recording an incorrect response. When the test was changed to permit multiple attempts and only a count of incorrect responses was held instead of a binary score, this item was always the same as CorrectAnswer and thus became redundant.
13	Empty	Not used
14	Dummy	Prevents row truncation.

Table 2. Pairs raw data.

Program source: pairing7.h (preceded by the practice program Pairing.h).

Versions

- 1.3.2 27/07/2007 Change the summary level from 1 to 7.
- 1.3.3 11/04/2008 Change summary line to include start instructions and test dates/times.
- 1.3.4 23/04/2008 Prevent clicking too soon.
- 1.3.5 17/02/2009 Rectify driver bug.

3. Quiz (Fluid intelligence)

This reasoning quiz presents multiple-choice questions, alternating between numerical and verbal problems.

The primary measure is the number of correct answers entered within the time limit.

The test is preceded by instructions and an example question. These are shown in Appendix A, Figures A3.1 to A3.2. The example question results are not recorded.

The participant was given two minutes to answer up to 13 questions. The test ended when either the last question was answered or the two minutes expired.

The questions and correct answers are given in Appendix C.

For the latter versions of the test, the first recorded trial took place at screen RQ00 in the program and the last possible at screen RQ12K, i.e., a maximum of 13 recorded trials. The screen id consisted of “RQ” followed by a two-digit question number and usually a letter identifying the question content. During piloting, the order of questions was changed to reflect ascending empirical difficulty of the numerical and verbal streams and so renumbered questions gained an identifying letter.

Prior to version 1.4.2, the first recorded trial took place at screen RQ01F. By inserting question RQ00 at version 1.4.2, the time available for subsequent questions was reduced. Therefore it is not valid to compare the scores of tests which included RQ00 with those that did not, even if the marks for RQ00 are discarded.

Prior to version 1.4.3, question RQ12K was incorrect. As this flaw may have prevented participants from answering the question correctly, scores prior to version 1.4.3 which include RQ12K marks cannot be trusted. At the time of writing, no test at version 1.4.2 has shown an attempt at question RQ12K. If this remains the case, marks for question RQ12K should be discarded prior to version 1.4.2.

For the above reasons, results for tests prior to version 1.4.2 should either be discarded or analysed separately from results for version 1.4.2 or later.

Column	Title	Description
0	Trial data row type	
7	ScreenId	Screen id – see Appendix C.

8	IfCorrect1Else0	This gives the mark for a question (1 for correct, 0 for incorrect). Summing the marks gives the score.
9	ElapsedTime_ms	The elapsed time to answer the question recorded in milliseconds but not to millisecond accuracy [Appendix B].
10	CorrectAnswer	The expected correct answer. In a normalised database, this would be held in a separate table as the answers do not change.
11	SelectedAnswer	The answer selected by the participant.
12	Empty	Not used
13	Empty	Not used
14	Dummy	Prevents row truncation.

Table 3. Quiz raw data.

Program source: Quiz.h

Versions:

- 1.4 15/02/2008 Add timer to cut off questions
- 1.4.1 11/04/2008 Start dates/times for instructions and timed test added to summary data line
- 1.4.2 11/04/2008 Insert very easy question after practice question (RQ00)
- 1.4.3 20/05/2008 RQ12K: John is forty-seven, not forty-six. All RQ12K data (if any) preceding this version is based on an incorrect question, although the answer is the same.
- 1.4.4 17/02/2009 Rectify driver problem

4. Stroop (Mild interference/Stroop interference)

This test measured the effect on two-choice reaction time of increasing choice complexity.

The primary measure for each of the two levels was the average response time. The primary measure of the test was the difference between the average response times.

A selection of screens are presented in Figure A4.1 to A4.8.

The single program treated the two levels separately, producing two sets of data with associated summary lines. Unlike the Reaction test, there was no varied response stimulus interval.

Following an introduction, the first (box) level showed instructions and a coloured panel and two names of colours to be chosen between, one of which matched the panel colour. After three unrecorded practice rounds with feedback, a timed test of 30 trials was performed.

The second (ink) level showed the name of a colour displayed in text of a different colour. The participant was presented with two colour names and was requested to reject the word matching the name and select the word matching the colour of the text. After three practice trials with feedback, a timed test of 30 trials was performed.

To accommodate colour blindness, the colours in the test were restricted to black, white, red and blue. The background was pale blue which could appear as pale grey achromatically. Red was brightened to red/orange to aid those who see red as dark. Red and blue were not paired together, as they could both appear dark grey achromatically. Neither the stimulus nor the left or right responses remained the same for adjacent trials.

Column	Title	Description
7	ScreenId	Box – ST[E-X]*; Ink – ST2[E-X]* (in alphanumeric order but not sequential).
8	IfCorrect1Else0	This indicates a correct response for a trial (1 for correct, 0 for incorrect). The mean gives an accuracy rate.
9	ElapsedTime_ms	The elapsed time taken from the stimulus reveal to selecting the appropriate arrow, recorded in milliseconds but not to millisecond accuracy [Appendix B].
10	Level	This is “box” for the coloured box test and “ink” for the Stroop test.
11	CorrectAnswer	The correct colour response, from “black”, “white”, “orangered”, “blue”. In a normalised database this would be in a separate table associated with the screen id.
12	LeftColour	The response word on the left, from “black”, “white”, “red”, “blue”. In a normalised database this would be in a separate table associated with the screen id.
13	RightColour	The response word on the right, from “black”, “white”, “red”, “blue”. In a normalised database this would be in a separate table associated with the screen id.
14	Dummy	Prevents row truncation.

Table 4. Stroop raw data.

Program source: StroopPage.h

Versions

- 1.2.3 18/02/2008 Change shape of Stroop buttons
- 1.2.4 11/04/2008 Change summary line to include start instructions and test dates/times.
- 1.2.5 22/04/2008 Add 10 more trials to each of box and ink. Lighten Red to

5. Number (Working Memory Test - Forward digit span)

This test measures the number of digits which can be recalled by the participant, beginning with a single digit and increasing the length of digit strings presented until the participant cannot recall a string of a particular length in two attempts. The test assumes that eventually every participant will be able to remember one- and two-digit strings.

The primary measure is the maximum number of digits recalled correctly.

A selection of screens are presented in Figure A5.1 to A5.4.

The participant was presented with a string of digits for 2 seconds plus 0.5 seconds for each digit. The string was replaced with “Wait...” and the on-screen keyboard disabled for 3 seconds. The keyboard was enabled and the participant was given 5 seconds plus 1 second per digit to enter the string before the trial was timed out and treated as a failure. If the participant enters a string of the required length, the program waits for 0.6 second before hiding the entered string and disabling the keyboard, then another 0.6 second before showing the next string of digits.

Successive strings did not start with the same digit and a string did not contain repeated digits set adjacently or separated by another single digit.

Column	Title	Description
7	ScreenId	“WMA” - the test is performed on a single screen
8	Digits	The number of digits in the presented string. Recalling the maximum of 25 digits correctly ends the test.
9	ElapsedTime_ms	The elapsed time taken to enter the numeric string recorded in milliseconds but not to millisecond accuracy [Appendix B].
10	Number	The string of digits presented to the participant.
11	Entered	The string of digits entered by the participant.
12	Strikes	The count of successive failures at this string length up to this point. A success resets the count. Single-digit or two-digit strings may have any number of failures but only the last two trials are recorded. Strings longer than two digits end the test on the second failure.
13	Empty	Not used
14	Dummy	Prevents row truncation.

Table 5. Number raw data.

Program source: Numbers.h

Versions

- 1.2 15/02/2008 Blank keypad until required.
- 1.2.1 14/04/2008 Prevent same initial digit on successive numbers.
- 1.2.2 21/04/2008 Disable failure on one digit.

Data conversion

Data conversion is performed by inserting cleaned data into a relational database.

The database has the table arrangement shown in Figure 2:

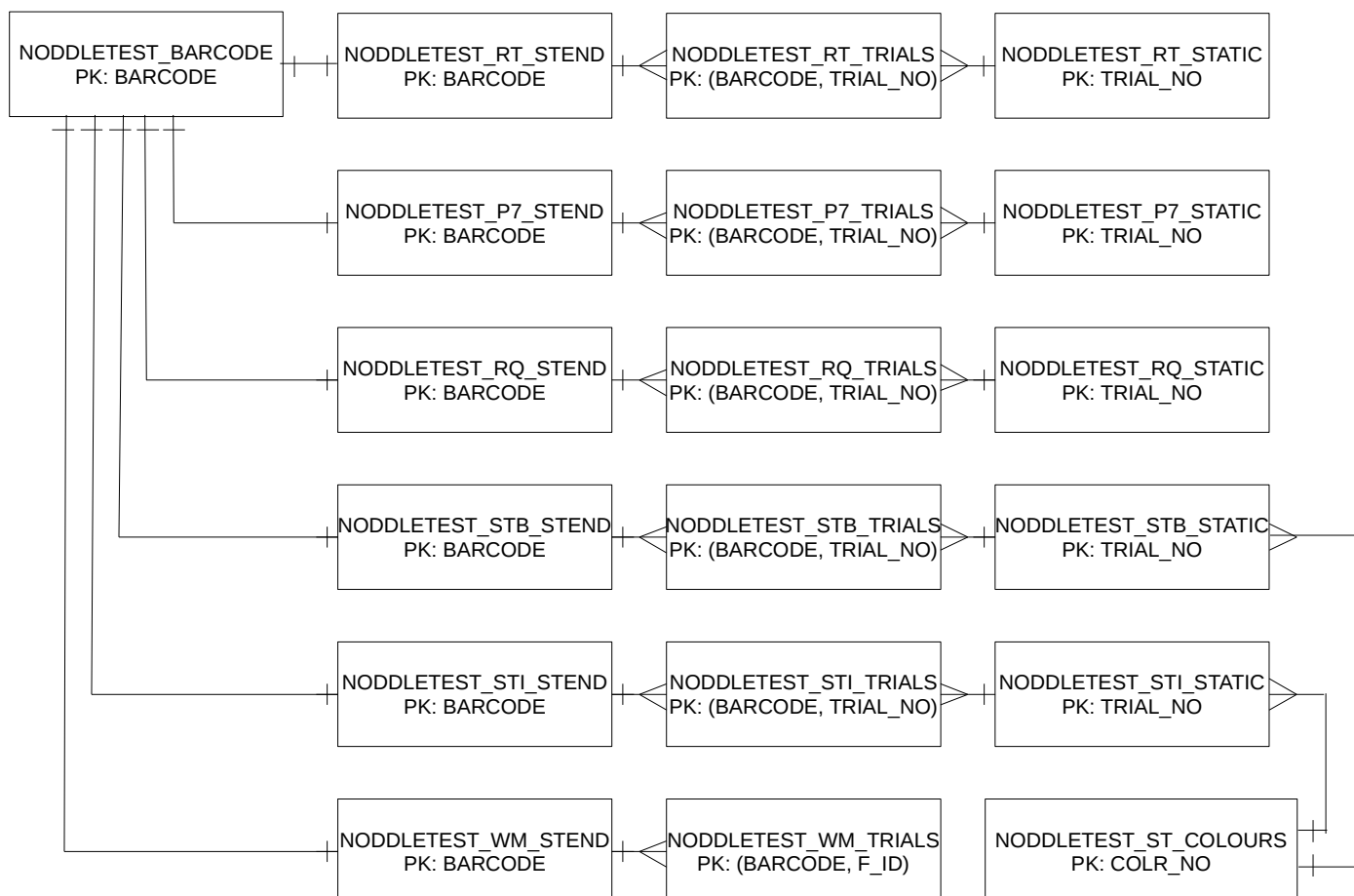


Figure 2. Noddletest cognitive battery schema outline.

The table “NODDLETEST_BARCODE” contains test session data, keyed on the barcode. This data includes the test session start time (which may be used for calculating the length of testing prior to a particular test as an estimate of fatigue), the clinic and the administrator (for detecting biases caused by clinic- or administrator-specific practices). This table is linked to all current tests.

Table	Attribute	Description
NODDLETEST_BARCODE	BARCODE	Session identifier.
	BARCODE_DT	test session start time
	CLINIC	Clinic name
	ADMINISTRATOR	Test administrator name

Table 6. NODDLETEST_BARCODE items

The tables with names ending “_STEND” contain test-specific data which include start and end times and statistical summary data (for convenience).

Table	Attribute	Description
NODDLETEST_XX_STEND Common to xx= RT, P7, RQ, STB, STI or WM	BARCODE	Session identifier.
	F_ID	Taken from the input table to aid error resolution.
	xx_VERSION	Textual test version
	INSTR_ST	Start date/time of test instructions/practice.
	TEST_ST	Start date/time of timed test.
	TEST_EN	End date/time of test
NODDLETEST_RT_STEND	RT_SCORE	Count of correct responses
	RT_MEAN	Mean of elapsed times (including outliers)
	RT_MEDIAN	Median of elapsed times
	RT_STDDEV	Standard deviation of elapsed times.
NODDLETEST_P7_STEND	P7_TOT_ATTEMPTS	Sum of attempts needed to find all seven pictures.
	P7_TOT_ELAP	Sum of all elapsed times
NODDLETEST_RQ_STEND	RQ_SCORE	Number of correct answers for version 1.4.2 or later, otherwise null.
	RQ_MAXTRIAL	Last question answered if version 1.4.2 or later, otherwise null.
NODDLETEST_STB_STEND	STB_SCORE	Count of correct responses
	STB_MEAN	Mean of elapsed times (including outliers)
	STB_MEDIAN	Median of elapsed times
	STB_STDDEV	Standard deviation of elapsed times.
NODDLETEST_STI_STEND	STI_SCORE	Count of correct responses
	STI_MEAN	Mean of elapsed times (including outliers)
	STI_MEDIAN	Median of elapsed times
	STI_STDDEV	Standard deviation of elapsed times.
NODDLETEST_WM_STEND	WM_SCORE	Maximum digits recalled

Table 7. *_STEND table items.

The tables with names ending in “_TRIALS” contain one row per trial. The data collected are of the participant responses.

Table	Attribute	Description
NODDLETEST_xx_TRIALS Common to xx= RT, P7, RQ, STB, STI or WM	BARCODE	Session barcode
NODDLETEST_RT_TRIALS	TRIAL_NO	Trial no. 0-59
	RT_ELAP	Reaction time
	RT_MARK	1: Correct, 0: incorrect
NODDLETEST_P7_TRIALS	TRIAL_NO	Trial no 0-6
	P7_ELAP	Time to locate correct picture
	P7_ATTEMPTS	No. of attempts taken to locate correct picture
NODDLETEST_RQ_TRIALS	TRIAL_NO	Question no. 0-12 (max)
	RQ_ELAP	Time to complete question
	RQ_SELECTED	Question selected
	RQ_MARK	1: Correct, 0: incorrect
NODDLETEST_STB_TRIALS	TRIAL_NO	Trial no. 0-29
	STB_ELAP	Reaction time
	STB_MARK	1: Correct, 0: incorrect
NODDLETEST_STI_TRIALS	TRIAL_NO	Trial no. 0-29
	STI_ELAP	Reaction time
	STI_MARK	1: Correct, 0: incorrect
NODDLETEST_WM_TRIALS	F_ID	Used with barcode to create unique primary key as no trial no.
	WM_ELAP	Time to enter number.
	DIGITS	Length of presented number.
	SHOWN	String of digits presented.
	ENTERED	String of digits entered.
	STRIKES	Successive incorrect attempts for no. of digits. Correct answer always 0.

Table 8. *_TRIALS table items.

The tables with names ending “_STATIC” contain data pertinent to an individual test screen, such as the format of the options and the expected result. As all of the Working Memory trials take place on the same screen and are randomized, there is no static data for this test.

Table	Attribute	Description
NODDLETEST_RT_STATIC	TRIAL_NO	Trial no. 0-59
	RT_SCR	Textual screen id.
	CORRECT	Expected response. 0: Left, 1: Right
	RSI	Response stimulus interval. One of 50, 150, 250, 350, 450, 550, 650, 750ms.
NODDLETEST_P7_STATIC	TRIAL_NO	Trial no 0-6
	P7_SCR	Textual screen id.
	POSITION	Textual description of target picture's position.
	POSN_NUM	Enumerated picture position, starting at the top with zero, counting clockwise.
NODDLETEST_RQ_STATIC	TRIAL_NO	Question no. 0-12 (max)
	RQ_SCR	Textual screen id.
	CORRECT_ANS	Expected answer.
	NUM_VERB	Question type: 0: numeric, 1: verbal.
NODDLETEST_STB_STATIC *(trials 6, 7, 11, 16, 23, 24, 25, 28 & 29 missing prior to version 1.2.5)	TRIAL_NO	Trial no. 0-29 *
	STB_SCR	Textual screen id.
	CORRECT_COLR	Id of expected colour. 0-3.
	L_COLR	Id of left colour. 0-3
	R_COLR	Id of right colour. 0-3
	CORRECT_SIDE	Expected response. 0: left, 1: right.
NODDLETEST_STI_STATIC **(trials 20-29 missing prior to version 1.2.5)	TRIAL_NO	Trial no. 0-29 **
	STI_SCR	Textual screen id.
	CORRECT_COLR	Id of expected colour. 0-3.
	L_COLR	Id of left colour. 0-3
	R_COLR	Id of right colour. 0-3
	CORRECT_SIDE	Expected response. 0: left, 1: right.

Table 9. *_STATIC table items.

The Stroop test colours are recoded for numerical analysis. The table “NODDLETEST_ST_COLOURS” maps the numerical values onto appropriate verbal names.

Table	Attribute	Description
NODDLETEST_ST_COLOURS	COLR_NO	Colour id.
	COLR_NAME	Textual colour name.

Table 10. NODDLETEST_ST_COLOURS table items.

Database build

The following sql scripts are located in the “Database build scripts” folder.

The tables are created in an Oracle database under the JMS user by running the PL/SQL script “create_noddletest_tables_v1.3.sql”. This script creates tables and inserts static data. It is run once.

Under the assumption that the raw data has been cleaned and imported into the table “NODDLETEST_2012”, the script “insert_from_noddle_2012_v1.7.sql” converts the raw data and inserts it into the database, along with summary data. The script is re-runnable, ignoring data which already have entries. If a new data set needs to be inserted, copy the insert script and perform a global replacement on “NODDLETEST_2012” with the name of the table which holds the new data.

Reaction time and Stroop trials which contain elapsed times equal to zero are ignored as they are created in error - the tests should not permit short reaction times and such data rows are usually followed immediately by another row for the same trial with a valid elapsed time. Erroneous zero measures would have a greater effect on mean values than would nulls as zero values are included in aggregate functions.

If needed, all data associated with a barcode may be removed by running the script “delete_barcode_v1.1.sql”.

Data extraction

The following scripts are located in the “Data extract scripts” folder.

Summary data is extracted with the script summary_output_v1.1.sql to a .tsv file using Oracle SQL Developer. This provides a single row per barcode containing summary data for all tests.

Trial data may be extracted as a single row per test using the scripts RT_trials.sql, P7_trials.sql, RQ_trials.sql, STB_trials.sql and STI_trials.sql, as appropriate. Due to the

open-ended nature of the working memory test, the format of a meaningful WM_trials data extraction would require further discussion.

Testing

The data extract supplied for database development was used to run the database build and extraction processes on an Oracle Express database. A set of scripts were employed to recreate the various rows of the original .csv extract from the database for comparison with the raw data. Scripts used for testing are located in the “Test scripts” folder.

To combat the introduction of typos in repeated lines of code with incrementing variable names, open document spreadsheets (e.g. “RT trials sql calcs.ods”) were used in writing database build and data extraction sql scripts.

Extracts from the test database are located in the “Sample extracts” folder as .tsv files.

Appendix A. Screen shots

The following example screen shots were gleaned from available archives and are not comprehensive.

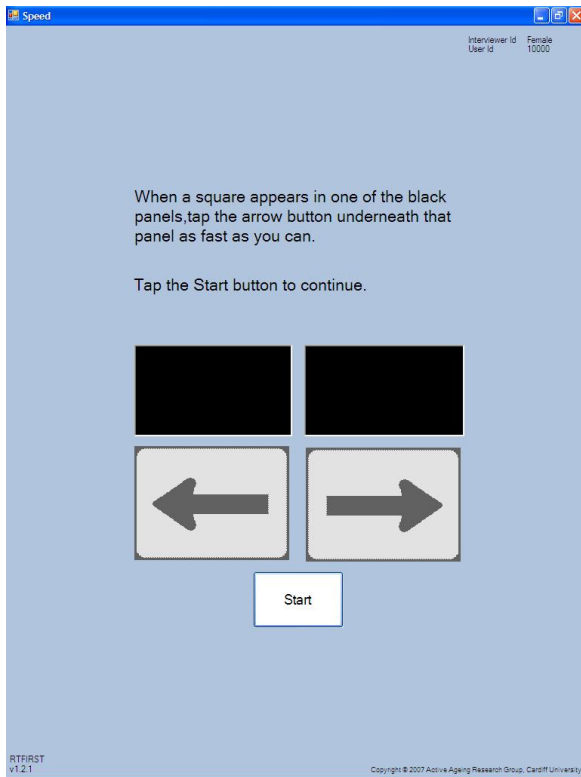


Figure A1.1 Reaction time instructions

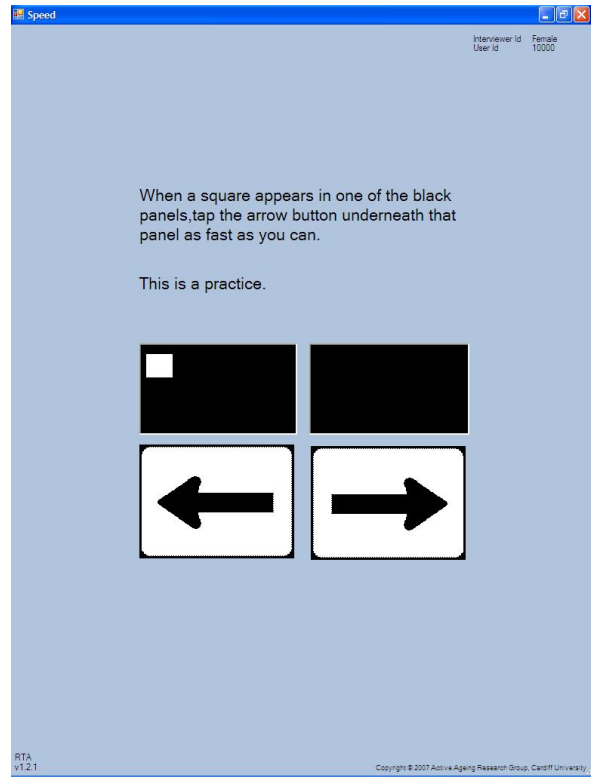


Figure A1.2 Reaction time practice

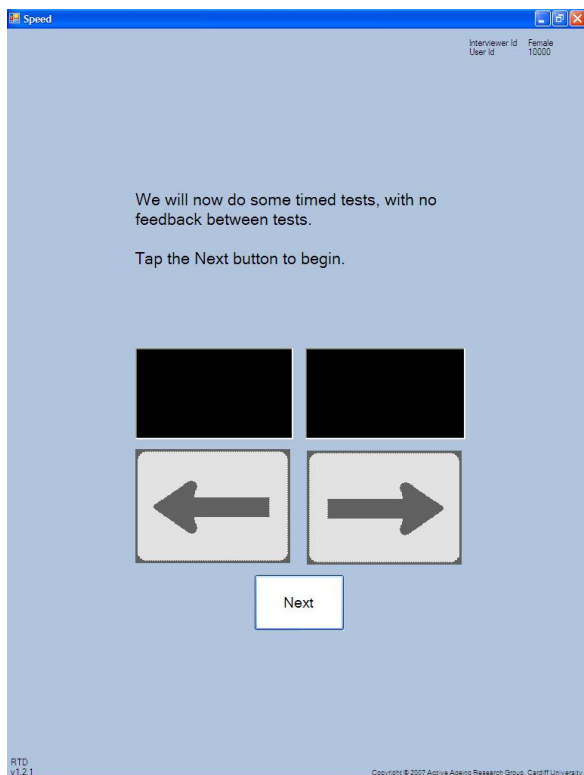


Figure A1.3 Timed reactions introduction

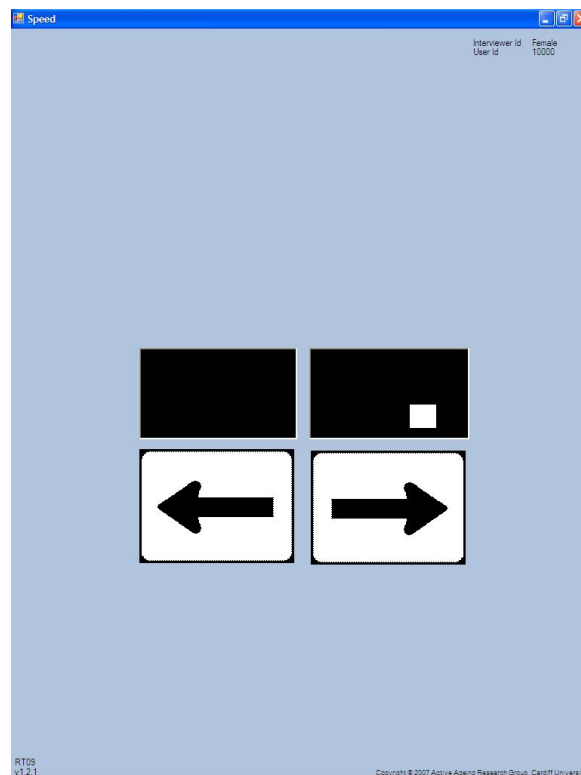


Figure A1.4 Reaction time test

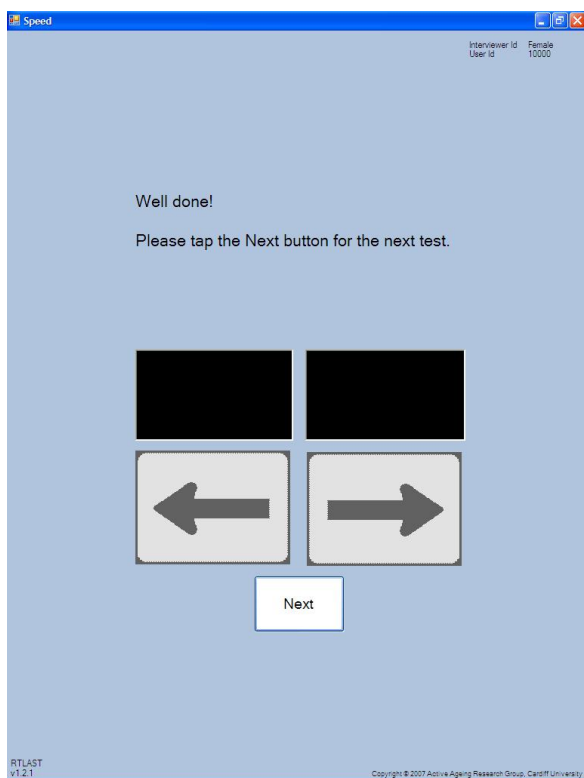


Figure A1.5 Reaction time end

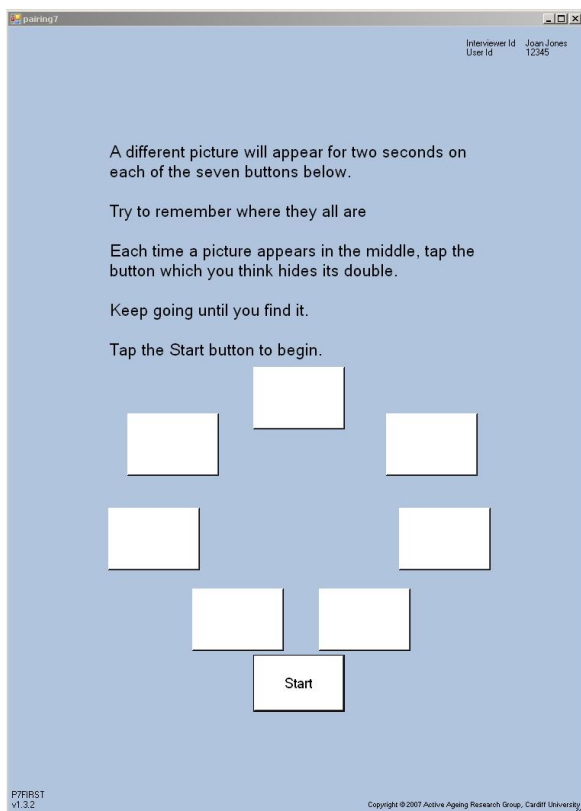


Figure A2.1 Pairs instructions

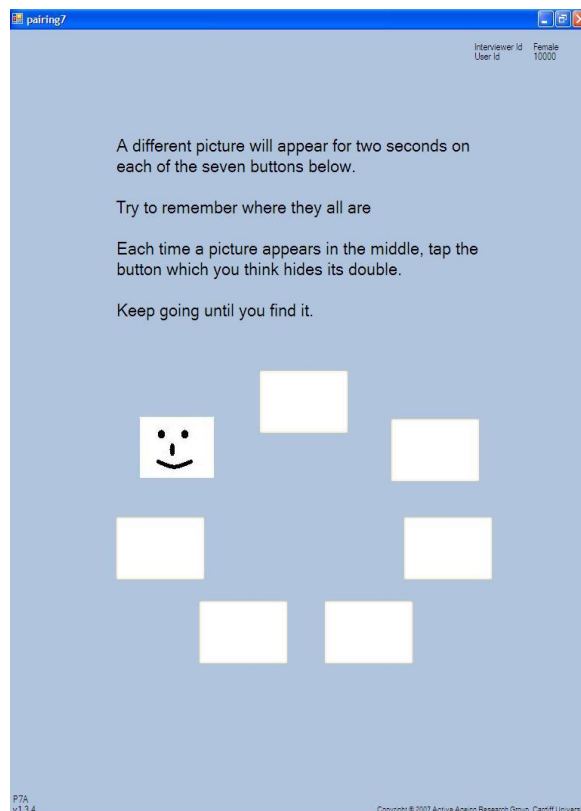


Figure A2.2 Pairs initial show

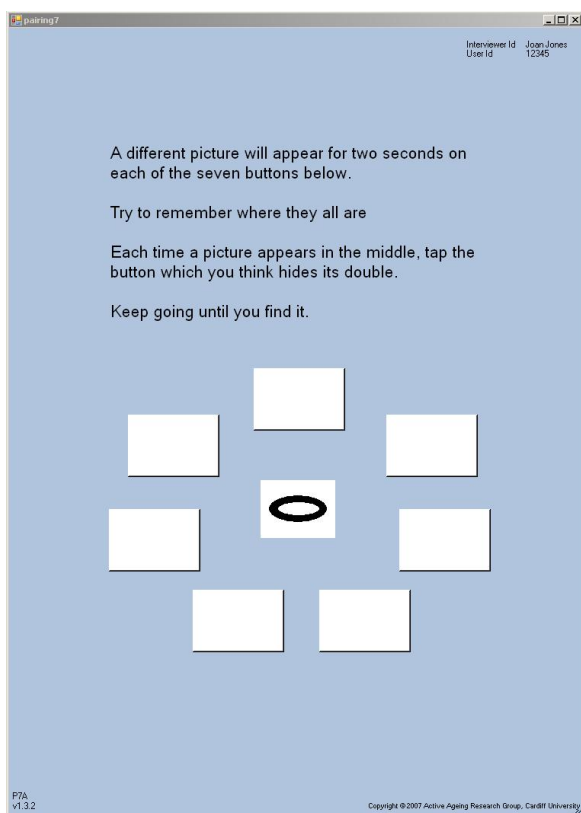


Figure A2.3 Pairs stimulus

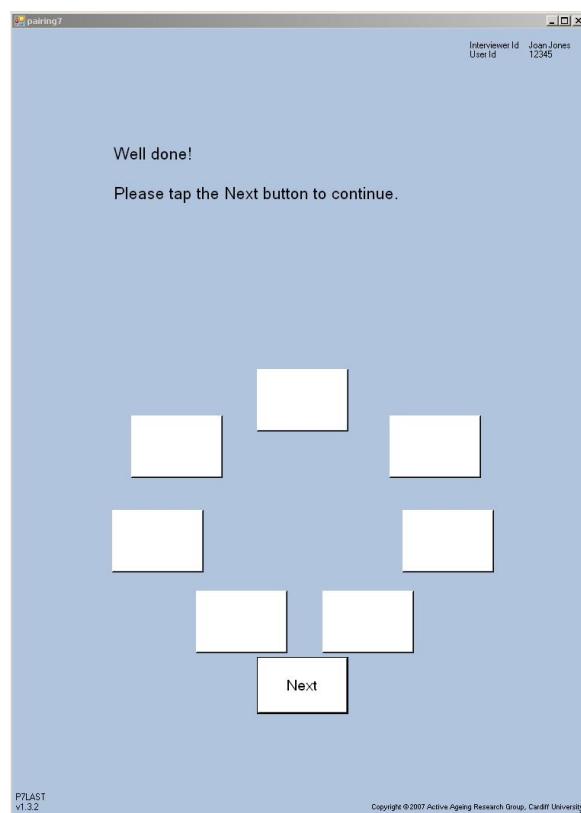


Figure A2.4 Pairs finish

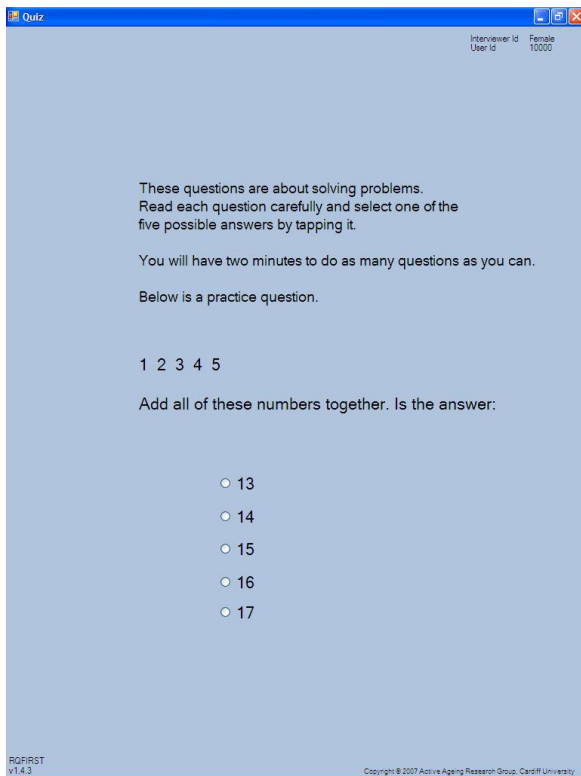


Figure A3.1 Quiz example

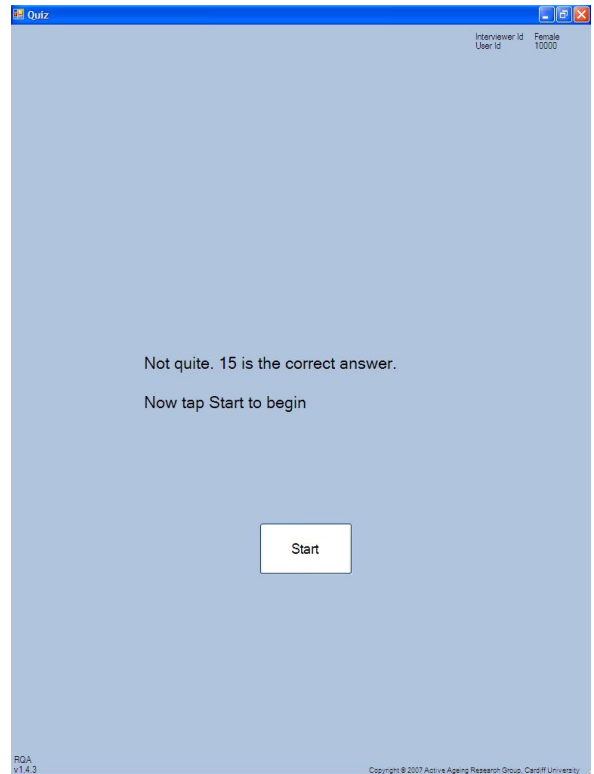


Figure A3.2 Example feedback

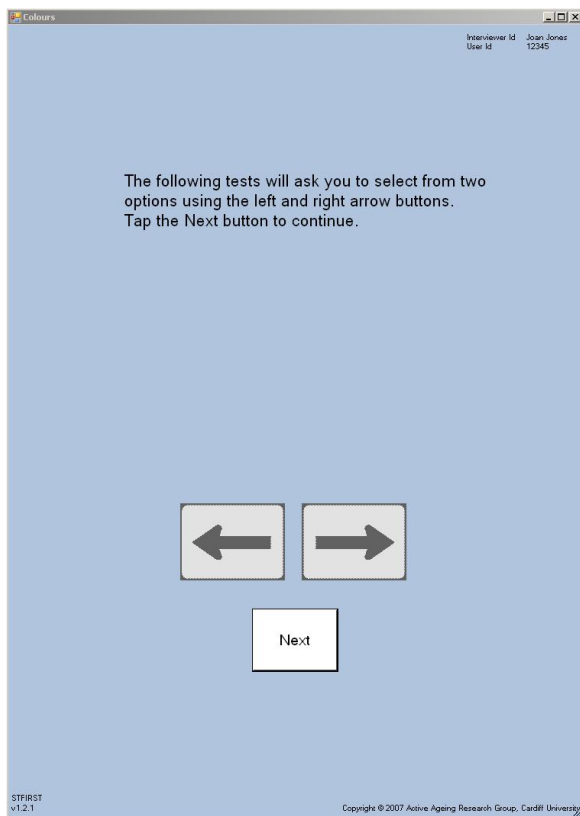


Figure A4.1 Box level introduction

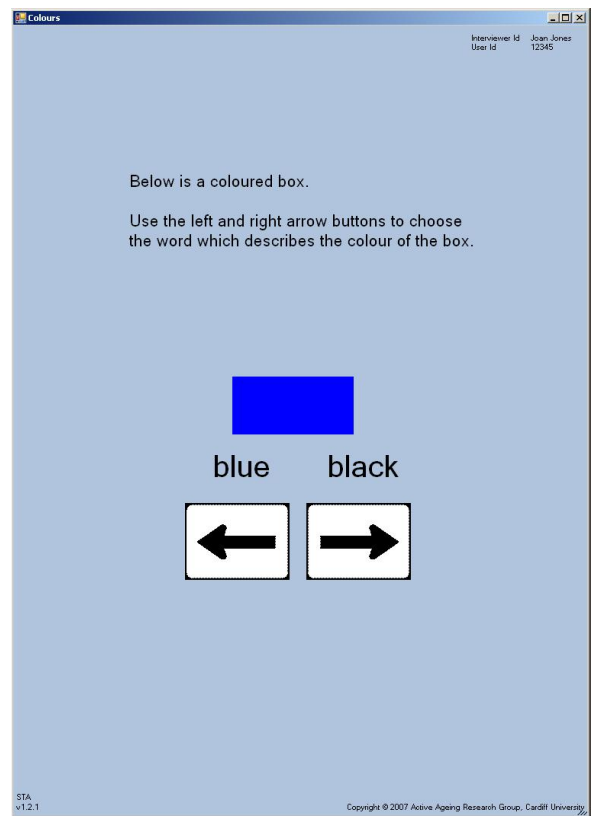


Figure A4.2 Box instructions

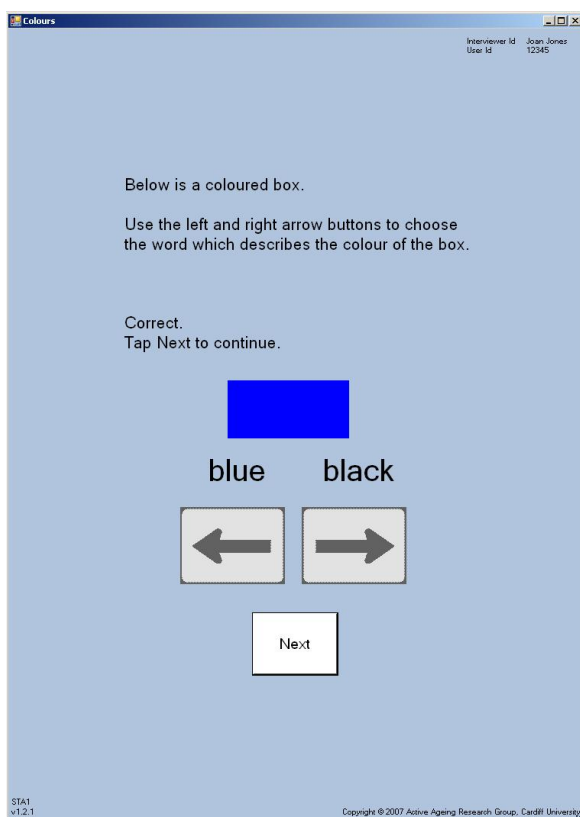


Figure A4.3 Box practice feedback

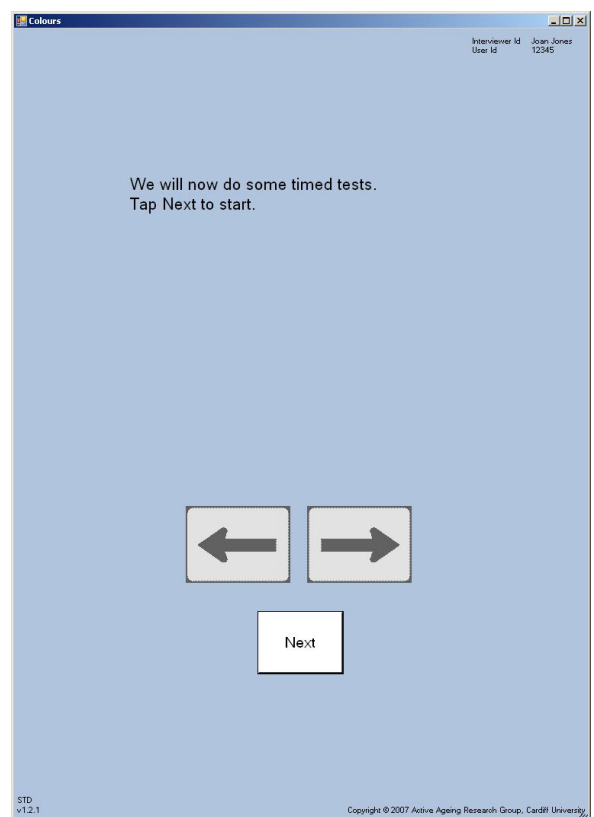


Figure A4.4 Box practice end

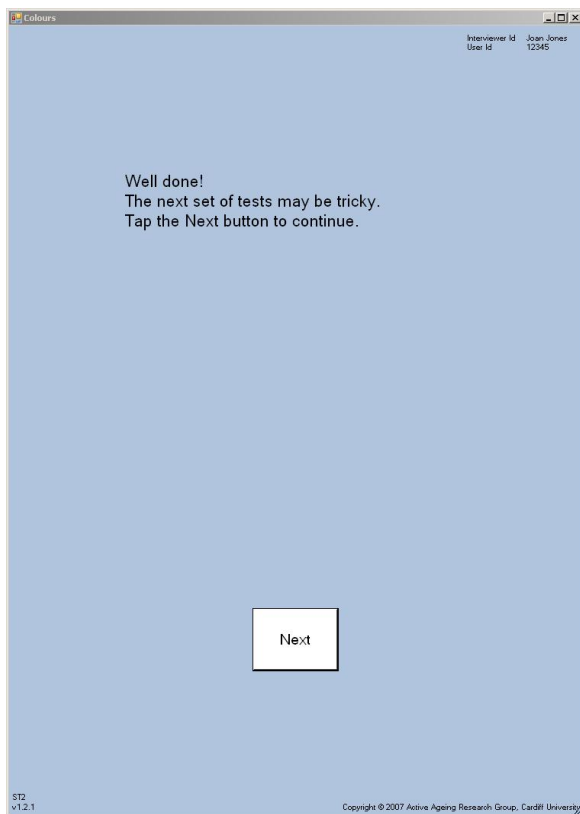


Figure A4.5 Stroop continuity page

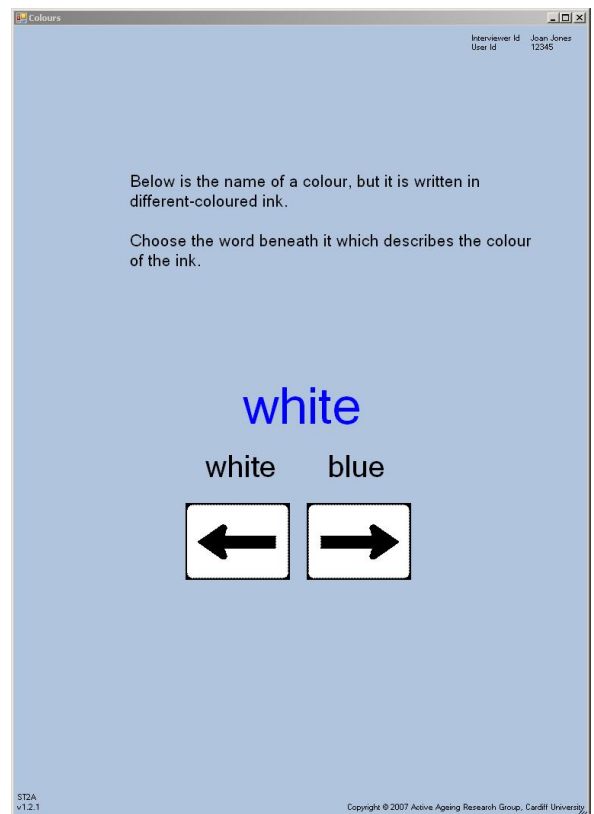


Figure A4.6 Stroop instructions

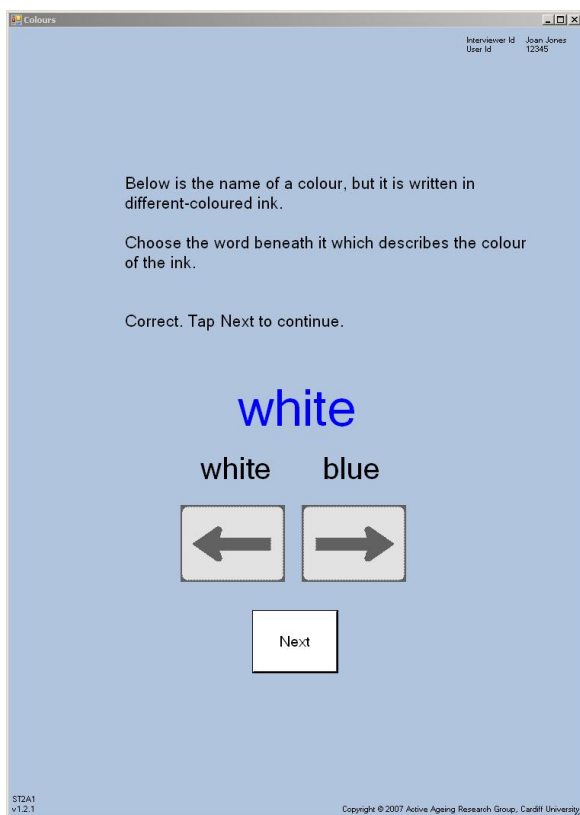


Figure A4.7 Stroop practice feedback

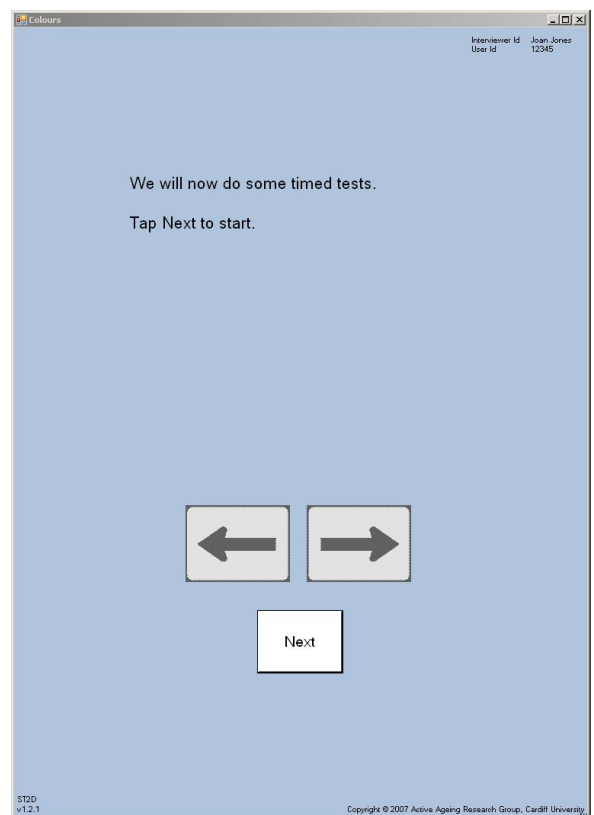


Figure A4.8 Stroop practice end

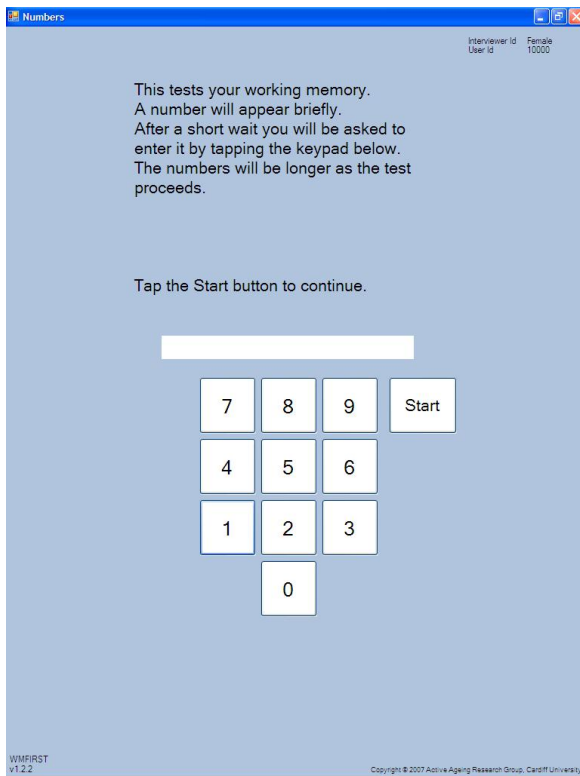


Figure A5.1 Working memory intro

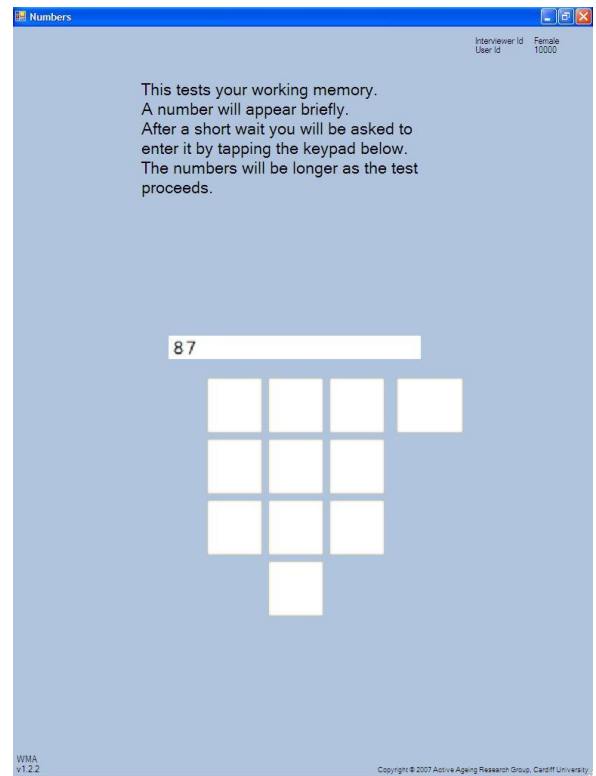


Figure A5.2 2-digit number reveal

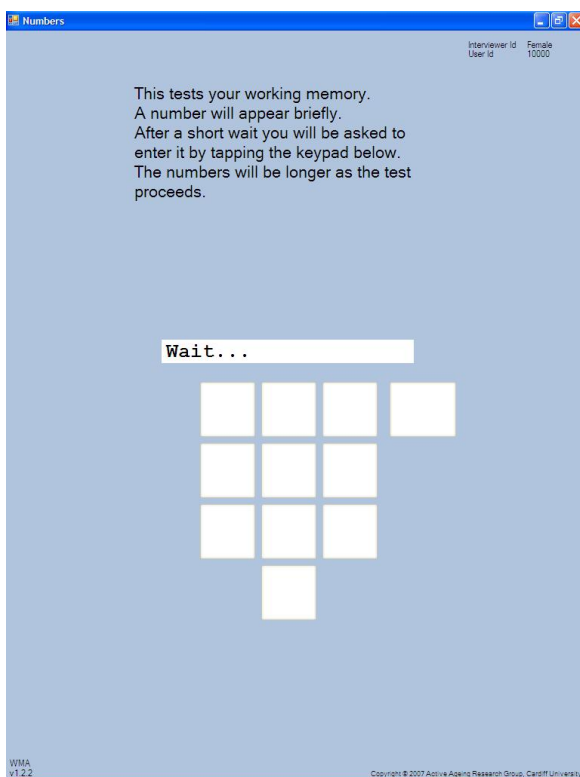


Figure A5.3 Number hidden

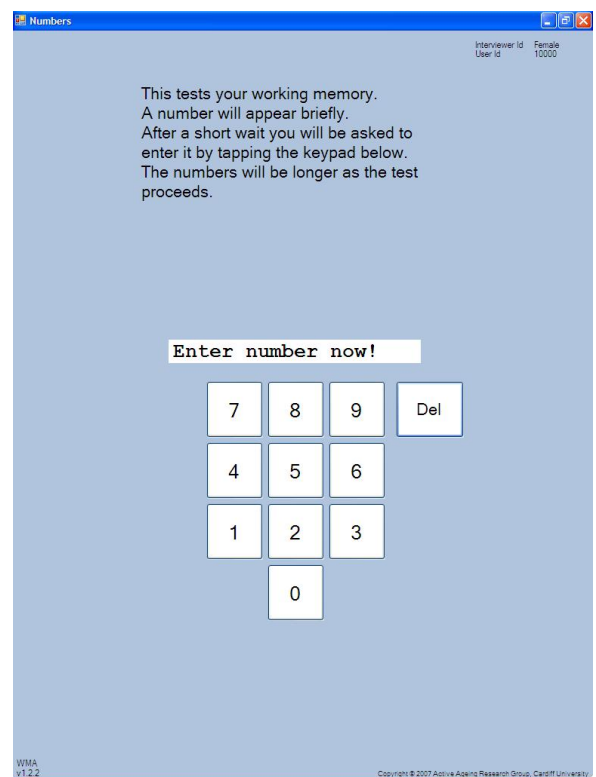


Figure A5.4 Number entry

Appendix B. Timing precision

The ideal cognitive measurement system would measure, to the nearest millisecond, the elapsed time between when a stimulus was presented to the participant and when the participant responded. The Windows XP and later operating systems work on the basis of regular clock ticks to regulate system functions, such as updating the program-accessible clock and allocating timeslices to programs. The Windows XP clock interval was 1/64 second, or 15.625ms. Changing this interval can risk having detrimental effects on processes which take time to complete, especially input and output. Other system timers could record times to the nearest millisecond and the compiled program could perform its processing fast enough to have little impact on the results. The idea of millisecond precision is, however, flawed.

The concept of millisecond accuracy assumes that the program can accurately record the times when a participant receives a stimulus and gives a response. In reality, the program can record the times when the stimulus was dispatched by the program and the response received by the program.

The program is dependent upon the operating system permitting it to run when a timing needs to be taken. Higher priority processes may cause the program to be switched out for several milliseconds creating a delay in receiving input.

The frame rate is the rate at which the graphics processor provides new visual frames to the display. This is commonly around 24-30 frames per second, although high-end games processors can generate 144 frames per second. This introduces a maximum potential delay of between 7 and 41 milliseconds between the stimulus being generated by the program and being submitted by the graphics processor to the screen buffer.

The screen refresh rate is the frequency at which the display copies the contents of the frame buffer to the screen and is never less than the frame rate. Screen refresh usually happens several times for each frame. The refresh rate will not be below 60Hz (as flicker becomes noticeable) and around 100Hz is common. This means that the display can take 10 or more milliseconds to update. If the stimulus is in the centre of the screen the delay between the frame reaching the buffer and appearing on the screen will be around 5ms.

The touch screen scan rate is the rate at which the tablet is scanned for input. 60Hz is common. If the input device waits for a scan to complete before submitting its input to the program or if the input area is at a consistent point in the scan, then the input times will have a granularity of around 16ms, although recent smart phone models may have scan rates of 200Hz giving a granularity of 5ms.

As can be seen from the above, random multi-millisecond delays can be introduced by the system into the interaction between the cognitive test program and the participant. Only bespoke electronic equipment using non-scanned stimuli (e.g. an individual LED or a display which is instantaneously illuminated after updating completes), an input method which requires minimal physical movement (e.g. short-travel buttons) and the ability to receive instructions from the host computer, perform the trial in isolation and return the results to the host computer can eliminate these delays.

Table 11 shows the frequency of recorded elapsed times for an individual reaction time test (barcode 1081). Rather than a smooth distribution of measurements, the system recorded timings which formed clusters 1-3ms wide and separated by multiples of about 15ms.

The participants themselves can introduce delays in addition to their actual reaction time. The technique for positioning the stylus prior to each trial can have a marked effect. During piloting, a subject was observed to hold the stylus just above the tablet surface between the two target areas, giving a minimal travel distance. Another was seen to bring the stylus up off the tablet almost to their ear after each trial producing a much greater travel distance and consequently a longer response time which did not reflect the underlying reaction time. An individual's reaction time distribution can vary from session to session due to mood, etc.

Ideally, a reaction time should be a consistent measure. Unfortunately, the data does not support this as few individual participants produce a narrow range of times (with allowance for the above delays). Of particular note is apparent loss of concentration when a trial produces an elapsed time two or more times greater than adjacent trials, having a substantial impact on the mean elapsed time, especially when there are multiple lapses of concentration. For example, barcode 37504 revealed several apparent lapses in concentration, the longest resulting in a response time of 9907ms. Defining lapses in concentration by elapsed times exceeding an agreed number of standard deviations (which will be affected by the magnitude and frequency of lapse times) or exceeding an agreed multiple of the median value (which probably would be less affected) would need discussion. Whether concentration lapses have any bearing on cognitive decline may deserve consideration.

In conclusion, the Noddletest Windows XP-based reaction time tests do not produce precise, consistent measures. A large number of trials for each test alleviates this giving more precise aggregate results but the tests were designed for large epidemiological studies rather than than to clinical studies.

Reaction time/ms	Frequency	Group mean	Gap from previous mean
298	1	298	0
328	3	328.4	30.4
329	2		
343	1	344.17	15.77
344	3		
345	2		
360	2	360.5	16.33
361	2		
375	3	375.25	14.75
376	1		
390	1	391.25	16
391	1		
392	2		
406	1	406.75	15.5
407	3		
421	1	422.4	15.65
422	1		
423	3		
439	1	439	16.6
454	4	454	15
485	1	485	31
501	4	501	16
516	1	516.67	15.67
517	2		
532	1	532	15.33
547	1	547.67	15.67
548	2		
563	1	563	15.33
578	1	578.67	15.67
579	2		
594	1	594	15.33
657	1	657	63
734	1	734.5	77.5
735	1		
969	1	969	234.5
1235	1	1235	266

Table 11 – Reaction time frequencies of participant session 10801

Appendix C. Quiz questions

The questions and correct answers are as follows:

Screen id: RQ00

Which number is the largest?

1: 642

2: 308

3: 987

4: 714

5: 253

correct = 3

Screen id: RQ01F

Bud is to flower as child is to

1: Grow

2: Develop

3: Improve

4: Adult

5: Old

correct = 4

Screen id: RQ02C

11 12 13 14 15 16 17 18

Divide the sixth number to the right of twelve by three. Is the answer

1: 5

2: 6

3: 7

4: 8

5: 9

correct = 2

Screen id: RQ03H

If Truda's mother's brother is Tim's sister's father,
what relation is Truda to Tim?

1: Aunt

2: Sister

3: Niece

4: Cousin

5: No relation

correct = 4

Screen id: RQ04E

If sixty is more than half of seventy-five,
multiply twenty-three by three.

If not subtract 15 from eighty-five.

Is the answer

1: 68

2: 69

3: 70
4: 71
5: 72
correct = 2

Screen id: RQ05D

Stop means the same as

1: Pause
2: Close
3: Cease
4: Break
5: Rest
correct = 3

Screen id: RQ06I

If David is twenty-one and Owen is nineteen and Daniel is nine years younger than David, what is half their combined age? Is it

1: 25
2: 26
3: 27
4: 28
5: 29
correct = 2

Screen id: RQ07J

Age is to years as height is to...

1: Long
2: Deep
3: Top
4: Metres
5: Tall
correct = 4

Screen id: RQ08G

150 137 125 114 104

What comes next?

1: 96
2: 95
3: 94
4: 93
5: 92
correct = 2

Screen id: RQ09B

Relaxed means the opposite of

1: Calm
2: Anxious
3: Cool
4: Worried

5: Tense
correct = 5

Screen id: RQ10

100 99 95 86 70

What comes next?

1: 50

2: 49

3: 47

4: 46

5: 45

correct = 5

Screen id: RQ11

If some flinks are plinks and some plinks are stinks, then some flinks are definitely stinks

1: False

2: True

3: Neither true nor false

4: Not sure

correct = 1

Screen id: RQ12K

If 'Anne' is thirty-four and 'John' is forty-seven, what number is 'that'? Is it

1: 49

2: 50

3: 51

4: 52

5: 53

correct = 1