

A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task

Ian J. Deary · David Liewald · Jack Nissan

Published online: 24 November 2010
© Psychonomic Society, Inc. 2010

Abstract Reaction time tasks are used widely in basic and applied psychology. There is a need for an easy-to-use, freely available programme that can run simple and choice reaction time tasks with no special software. We report the development of, and make available, the Deary-Liewald reaction time task. It is initially tested here on 150 participants, aged from 18 to 80, alongside another widely used reaction time device and tests of fluid and crystallised intelligence and processing speed. The new task's parameters perform as expected with respect to age and intelligence differences. The new task's parameters are reliable, and have very high correlations with the existing task. We also provide instructions for downloading and using the new reaction time programme, and we encourage other researchers to use it.

Keywords Reaction time · Ageing · Intelligence · Processing speed · Psychometric testing

Introduction

Reaction time has been used as a psychological task since the mid-19th century. Originally a result of astronomers' noticing that observers made different responses to star transit times, Donders (1868, 1969) was early in introducing the technique to scientific psychology. Thereafter, it became

a staple of scientific study in psychology. Famously, James McKeen Cattell (1890) suggested reaction time as one of the 'mental tests' that he introduced in 1890. This received endorsement from Francis Galton (1890), who used reaction time to test thousands of subjects (see Johnson et al., 1985). The use of reaction time grew and has persisted during the whole of the 20th century and into the 21st century (for example, as described in Aufdembrinke, Hindmarch and Ott 1988; Deary, 2000; Jensen, 2006). There are many different reaction time devices, and reaction times are taken in response to many psychological and other manipulations. However, two common and useful procedures are to measure simple reaction time and choice reaction time (here, we shall concentrate on four-choice reaction time). Simple reaction time involves making a response as quickly as possible in response to a single stimulus. Choice reaction time is complicated by requiring the subject to make the appropriate response to one of a number of stimuli. The experimental variables that are most commonly derived from both of these are some measure of the central tendency (mean or median usually), and a measure of intraindividual variability, typically the raw standard deviation of a number of trials or the coefficient of variation (Hultsch, MacDonald, & Dixon, 2002).

Simple and choice reaction times are relatively straightforward in conception and to perform, compared to many other mental tasks that are used within experimental and differential psychology. Of course, this should not be taken to mean that even such simple psychological tasks are not founded on a number of more basic psychological operations and parameters, which can be bound in complex models (e.g., Luce, 1991; Ratcliff, 2008). The stimulus-response contingencies of reaction time procedures are such that, when no time pressure is applied, errors are rare, and the time to complete an item is much less than a typical IQ-type test item. Despite the apparent lack of cognitive

Electronic supplementary material The online version of this article (doi:10.3758/s13428-010-0024-1) contains supplementary material, which is available to authorized users.

I. J. Deary (✉) · D. Liewald · J. Nissan
Centre for Cognitive Ageing and Cognitive Epidemiology,
Department of Psychology, University of Edinburgh,
7 George Square,
Edinburgh EH8 9JZ Scotland, UK
e-mail: I.Deary@ed.ac.uk

demand required to perform reaction time tasks, they have produced an interesting set of findings. Reaction times—especially choice reaction times—show marked slowing with age, which begins from young adulthood and accelerates after middle adulthood (Deary & Der, 2005a; Der & Deary, 2006). Indeed, reaction times have been viewed as capturing the capacity of processing speed that is a major foundation of the age-related declines in higher-level cognitive functions (Madden, 2001; Salthouse, 1996). Reaction times—especially choice reaction times—are moderately to strongly correlated with measures of general fluid intelligence (Jensen, 2006). For example, in one large ($n = 900$), representative sample of 55-year-olds in Scotland, four-choice reaction time correlated 0.49 with a measure of general intelligence (the Alice Heim 4 test; Deary, Der, & Ford, 2001). Reaction times—simple and choice, and their means and individual variability, are associated with survival. For example, in the same large group of 55-year-olds from Scotland, four-choice reaction time mean was strongly associated with survival over the next 15 years (Deary & Der, 2005b); and this was replicated in a sample of about 7,000 individuals aged from 18 to 80 (Shibley, Der, Taylor, & Deary, 2006). These are just a few empirical associations that make reaction time valuable in studying aspects of human psychology and health. In addition to these, reaction times are widely used in experimental psychology, psychopharmacology, medical studies, and areas beyond these (e.g., Strachan et al., 2001). Therefore, reaction time is a much-valued predictor and outcome variable in psychology. The examples cited above are just a few—using some from our own work—to provide examples of the range of psychological research—basic and applied—situations in which reaction times are used.

In view of the long period over which reaction times have been used, and their importance with regard to key aspects of human life, it is surprising that there is no standard reaction time measure. For example, when we reviewed the literature on something as straightforward as reaction time and age, it was remarkable that each study had used a different reaction time procedure, making comparisons difficult or impossible (Deary & Der, 2005a; Der & Deary, 2006). Therefore, it would be useful for a broad range of psychological disciplines and applications if there were a freely available reaction time test with some basic stimulus-response associations, a set of parameters which could be varied, and all set on a common platform. This lack and need were argued strongly by Jensen (2006, p. 241): “it would also be advantageous to provide standardized computer programs for a number of classical paradigms, which were originally intended to measure the speed of various information processes”. The purpose of the present study is to fill this gap. It aims to provide a free-to-all, easy-to-use programme that will allow means and

standard deviations to be derived from simple and four-choice reaction times. We provide some initial reliability and validity data for the task. We also provide a location from which other researchers can download the reaction time programme and instructions.

Method

Participants

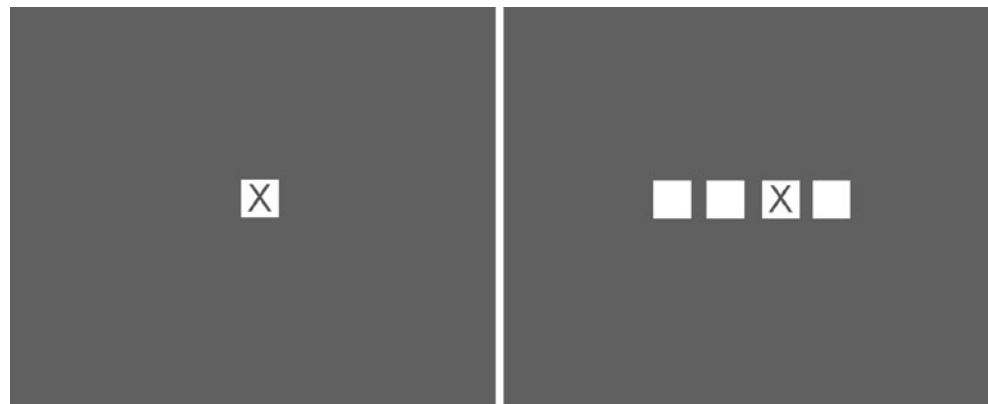
Fifty young adults aged between 18 and 25 years (mean = 20.5, SD = 2.6), 50 middle-aged adults aged between 45 and 60 (mean = 53.7, SD = 4.9), and 50 older adults aged between 61 and 80 (mean = 69.1, SD = 6.2) took part in the study. Participants were either students at the University of Edinburgh or residents from the City of Edinburgh. The students received course credit for their participation and all other adults were paid a small sum for taking part.

Reaction time tasks and other mental tests

The digit-symbol coding subtest of the Wechsler adult intelligence scale III (Wechsler, 1997), the matrix reasoning subtest of the Wechsler abbreviated scale of intelligence (Psychological Corporation (The), 1999), and the Wechsler test of adult reading (WTAR) (Psychological Corporation (The), 2001) were used as higher-level cognitive measures. Digit-symbol coding was included as a test of processing speed, matrix reasoning as a fluid-type (age-sensitive) intelligence task, and WTAR as a test of crystallised-type (age-insensitive) intelligence. The tests were applied according to instructions in the tests' manuals.

Two reaction time tasks were used. These will be referred to as the Deary-Liewald reaction time task, and the numbers reaction time box. The Deary-Liewald task is the new, computer-based task of principal interest. The numbers reaction time box was employed for comparison, because there is much previous information about it: it has been used in large, epidemiological surveys in the UK, and its parameters' associations with age, intelligence and mortality are known and replicated (Cox, Huppert, & Whittington, 1993; Deary et al., 2001; Deary & Der, 2005a, 2005b; Der & Deary, 2006; Huppert & Whittington, 1993; Shibley et al., 2006). Simple reaction time (SRT) and four-choice reaction time (CRT) means and standard deviations were measured for each participant on both tasks. In the SRT, participants had to press a button or key in response to a single stimulus. In the CRT, there were four stimuli and participants had to press the button that corresponded to the correct response. For both reaction time tasks, the SRT involved eight practice trials and twenty test trials. The CRT for both tasks involved eight practice

Fig. 1 Screen shots of the Deary-Liewald task for the simple reaction time task (*left*) and the choice reaction time task (*right*)



trials and forty test trials. Subjects undertook a third reaction time task, but it is not reported further here.

Deary-Liewald reaction time task

This was designed by IJD and programmed by DL, with several iterations between the initial design and the final programme that was used here. The programme was run on a screen with a vertical refresh rate of 60 Hz. For the SRT, one white square was positioned approximately in the centre of a computer screen, set against a blue background (see Fig. 1). The stimulus to respond is the appearance of a diagonal cross within the square. Each time a cross appeared, participants had to respond by pressing a key as quickly as possible. Each cross remained on the screen until the key was pressed, after which it disappeared and another cross appeared shortly after. The inter-stimulus interval (the time interval between each response and when the next cross appeared) ranged between 1 and 3 s and was randomised within these boundaries.¹ The computer programme recorded the response time and the inter-stimulus interval for each trial.

For the CRT, four white squares were positioned in a horizontal line across approximately the middle of the computer screen, set against a blue background (see Fig. 1). Four keys on a standard computer keyboard corresponded to the different squares. The position of the keys corresponded in alignment to the position of the squares on the screen: the ‘z’ key corresponded to the square on the far left, the ‘x’ key to the square second from the left, the ‘comma’ key to the square second from the right and the ‘full-stop’ key to the square on the far right. The stimulus to respond was the appearance of a diagonal cross within one of the squares. Participants were instructed to gently rest the index and middle fingers of their

left hand on the ‘z’ and the ‘x’ keys, and the index and middle fingers of their right hand on the ‘comma’ and ‘full stop’ keys. A cross appeared randomly in one of the squares and participants were asked to respond as quickly as possible by pressing the corresponding key on the keyboard. Each cross remained on the screen until one of the four keys was pressed, after which it disappeared and another cross appeared shortly after. The inter-stimulus interval ranged between 1 and 3 s and was randomised within these boundaries. The computer programme recorded the response times for each cross, the inter-stimulus interval for each trial, which key was pressed and, in the case of four-choice reaction time, whether the response was correct or wrong. It also calculated the mean, median, variance, standard deviation, skewness, and kurtosis of the response times.

Numbers-based reaction time box

The numbers reaction time box was a rectangular, stand-alone box, originally designed for the UK Health and Lifestyle Survey (Cox et al., 1993; Fig. 2). It provided the data on ageing, correlations with intelligence, and associations with mortality that were summarised in the Introduction. On the top surface, there was a liquid crystal display (LCD) screen and five response buttons, each with a number written above

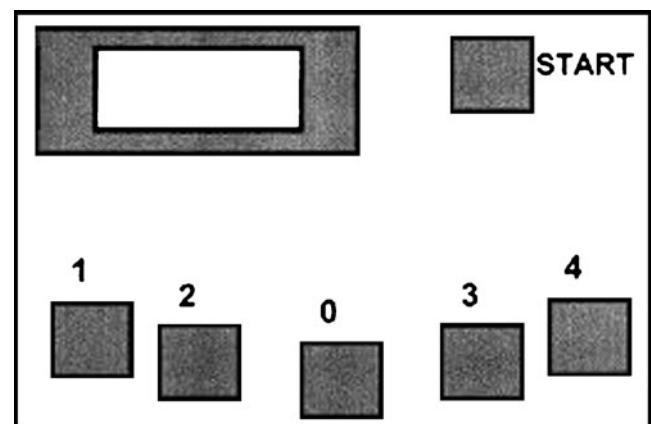


Fig. 2 Illustration of the top surface of the numbers task box

¹ Analysis of the distribution of inter-stimulus intervals (ISIs) highlighted that the same sequence of randomly generated ISIs for the SRT were given to a number of young and middle-aged participants. While this should not affect the results, the programme has been amended so that a new random sequence of ISIs is generated for each participant.

it. The buttons were arranged underneath the LCD screen in a gentle curve to fit the natural position of the participant's fingers. From left to right, the buttons were labelled with the numbers 1, 2, 0, 3, and 4 (see Fig. 2). The stimulus for response was the appearance of a number on the LCD screen. Subjects were asked to respond as quickly as possible when a number appeared. A number remained on the screen until participants made a response, after which it disappeared and another number appeared shortly after. The inter-stimulus interval ranged between 1 and 3 s and was randomised within these boundaries.

For the SRT, only the number '0' appeared on the screen. Participants were instructed gently to rest the index finger of their preferred hand on the button labelled '0', and told that they would only be using this button. For the CRT, one of the numbers 1, 2, 3, or 4 appeared on the screen. Participants were instructed gently to rest the index and middle fingers of their left hand on the buttons labelled '1' and '2', and the index and middle fingers of their right hand on the buttons labelled '3' and '4', and to press the button which corresponded to the number that appeared on the screen. For the SRT, the box recorded mean and standard deviation of response times. For the CRT, the box recorded the number of errors and the means and standard deviations of response times for correct and incorrect responses. The numbers box does not record individual trial data.

Procedure

Participants first completed a short social and demographic questionnaire which asked questions about their age, gender, education (number of years in full-time education), and occupation (graded according to the SOC2000, based on the UK's standard classification of occupations; Rose & Pevalin, 2003). The younger group was asked about their parents' occupations. They then completed the tasks in the following order: reaction time task (a), matrix reasoning, reaction time task (b), WTAR, digit-symbol coding, reaction time task (c). The order in which the different reaction time tasks were completed was varied equally among the participants.

Results

Background and cognitive measures

Table 1 describes the means (SD) and Table 2 describes the frequencies for the background measures, cognitive measures and the reaction time results for the total sample and for different age groups. Percentiles of the Deary-Liewald reaction time task scores for the different age groups are shown in Appendix 1. The mean (SD) overall age was 47.7 years (20.9). The mean (SD) number of years in full time

education was 15.1 (2.9). There was a significant difference between the age groups with regard to the Standard Occupational Classification (SOC2000; $\chi^2[12, n = 150] = 24.46, p < 0.009$; see Table 2). With regard to the cognitive measures, the mean (SD) total score for the WTAR was 44.3 (5.4), the mean (SD) total score for the Matrix Reasoning test was 24.6 (4.7), and the mean (SD) total score for the Digit-Symbol Coding test was 74.8 (15.4). One-way ANOVAs with a between subjects factor of age (three levels: young, middle-aged and old) revealed a significant effect of Age on the WTAR ($F[2, 147] = 13.05, p < 0.01, \eta^2 = .15$), the Matrix Reasoning test ($F[2, 147] = 33.73, p < 0.01, \eta^2 = .32$), and the Digit-Symbol Coding test ($F[2, 147] = 22.73, p < 0.01, \eta^2 = .24$). Younger adults scored higher on the Matrix Reasoning and Digit-Symbol Coding tests, and lower on the WTAR, than the middle-aged and older adults. There was no difference between the middle-aged and old groups in any of these tests (see Table 1). The full correlation matrix for these variables is shown in Table 3. Most notable are the strong inverse correlations between age and Matrix Reasoning and Digit-Symbol Coding tests, and a substantial positive correlation between age and WTAR.

Reaction time tasks

Comparison of the two reaction time tasks

With regard to the SRT measures, repeated measures t-tests revealed that the mean response time for the Deary-Liewald task (274.4 ms) was significantly longer than the numbers task (255.7 ms; $t[149] = -6.30, p < 0.01$). The mean SRT SD was lower for the Deary-Liewald task (45.3 ms) than for the numbers task (49.7 ms); $t[149] = 2.24, p < 0.05$. With regard to the CRT measures, mean response time was lower for the Deary-Liewald task (474.5 ms) than the numbers box (555.8 ms; $t[149] = 18.08, p < 0.01$). This may be due to the different stimuli used in the two tasks. The stimulus-response arrangement in the Deary-Liewald task was designed to rely on spatial coding, and hence may have been more straightforward than the numbers box, which required participants to recode a centrally placed number into the appropriate response. The mean SD of CRT response times was slightly lower for the Deary-Liewald task (100.1 ms), than the numbers task (108.2 ms; $t[149] = 3.25, p < 0.01$). The mean number of errors made with the Deary-Liewald task was 2.4, and with the numbers box was 2.5; there was no significant difference between them.

The correlations between the reaction time measures are shown in Table 4. With regard to the Simple reaction time (SRT) tasks, there was a large, significant positive correlation between the mean response times of the Deary-Liewald task and the numbers task ($r[148] = .68, p < 0.01$). There was also a significant positive correlation between the

Table 1 Means and standard deviations (SD) for background, cognitive and reaction time task measures

Variable	Age								
	18–25		45–60		61–80		Total		ANOVA
	<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)	<i>P</i> value
Age	50	20.5 (2.6)	50	53.7 (4.9) ^a	50	69.1 (6.2) ^{b c}	150	47.7 (20.9)	< 0.001
Education	50	14.8 (2.3)	50	15.5 (3.2)	50	14.9 (3.2)	150	15.1 (2.9)	.37
WTAR	50	41.5 (4.2)	50	45.1 (6.5) ^a	50	46.4 (4.0) ^b	150	44.3 (5.4)	< 0.001
Matrix reasoning	50	28.4 (3.6)	50	23.2 (4.3) ^a	50	22.4 (3.9) ^b	150	24.6 (4.7)	< 0.001
Digit-symbol coding	50	85.0 (13.7)	50	72.1 (13.6) ^a	50	67.3 (13.4) ^b	150	74.8 (15.4)	< 0.001
NS mean	50	230.2 (17.5)	50	269.1 (30.4) ^a	50	267.7 (45.2) ^b	150	255.7 (37.5)	< 0.001
NS SD	50	40.8 (15.2)	50	54.0 (23.1) ^a	50	54.2 (23.1) ^b	150	49.7 (21.6)	.001
NC mean	50	459.4 (42.5)	50	581.2 (66.3) ^a	50	626.8 (63.0) ^{b c}	150	555.8 (91.5)	< 0.001
NC SD	50	80.8 (20.0)	50	115.5 (28.3) ^a	50	128.2 (33.4) ^{b c*}	150	108.2 (34.2)	< 0.001
NC errors	50	3.6 (3.4)	50	1.6 (2.1) ^a	50	2.2 (2.6) ^{b*}	150	2.5 (2.8)	.001
DLS mean	50	243.1 (17.6)	50	283.9 (38.0) ^a	50	296.1 (63.9) ^b	150	274.4 (49.4)	< 0.001
DLS SD	50	32.9 (14.1)	50	50.6 (23.0) ^a	50	52.4 (22.8) ^b	150	45.3 (22.1)	< 0.001
DLC mean	50	388.0 (45.0)	50	492.4 (68.0) ^a	50	543.2 (85.3) ^{b c}	150	474.5 (93.7)	< 0.001
DLC SD	50	69.4 (20.3)	50	107.8 (34.4) ^a	50	123.1 (33.0) ^{b c*}	150	100.1 (37.4)	< 0.001
DLC errors	50	3.0 (2.6)	50	2.0 (2.4)	50	2.1 (2.3)	150	2.4 (2.4)	.10

^a Significant difference between age groups 18–25 and 45–60 at $p < 0.01$

^b Significant difference between age groups 18–25 and 61–80 at $p < 0.01$

^c Significant difference between age groups 45–60 and 61–80 at $p < 0.01$

* Significant at $p < 0.05$

WTAR Wechsler test of adult reading; NS Numbers box, Simple reaction time task; NC Numbers box, Choice reaction time task; DLS Deary-Liewald task, Simple reaction time task; DLC Deary-Liewald task, Choice reaction time task; Errors Percentage of incorrect responses

Table 2 Frequencies, percentages and non-parametric tests for gender, handedness and occupational classification

Variable	Age				Non-parametric tests <i>p</i> value
	18–25 <i>n</i> (%)	45–60 <i>n</i> (%)	61–80 <i>n</i> (%)	Total <i>n</i> (%)	
Gender					
Male	24 (48)	17 (34)	17 (34)	58 (39)	0.25 ^a
Female	26 (52)	33 (66)	33 (66)	92 (61)	
Handedness					
Right	46 (92)	46 (92)	45 (90)	137 (91)	0.92 ^b
Left	4 (8)	4 (8)	5 (10)	13 (9)	
SOC2000*					
1	20 (40)	10 (20)	6 (12)	36 (24)	0.009 ^c
2	23 (46)	17 (34)	27 (54)	67 (45)	
3	2 (4)	9 (18)	8 (16)	19 (13)	
4	4 (8)	7 (14)	7 (14)	18 (12)	
5	0 (0)	1 (2)	0 (0)	1 (1)	
6	1 (2)	4 (8)	2 (4)	7 (5)	
7	0 (0)	2 (4)	0 (0)	2 (1)	
8	0 (0)	0 (0)	0 (0)	0 (0)	
9	0 (0)	0 (0)	0 (0)	0 (0)	

* Standard Occupational Classification 2000: 1 Managers and senior officials, 2 Professional occupations, 3 Associate professional and technical occupations, 4 Administrative and secretarial occupations, 5 Skilled trades occupations, 6 Personal service occupations, 7 Sales and customer service occupations, 8 Process, plant and machine operatives, 9 Elementary occupation

^a Chi-squared test

^b Exact test

^c Monte Carlo test: based on 10,000 sampled tables with starting seed of 2,000,000

Table 3 Pearson correlations among background and cognitive measures

	2	3	4	5	6
1. Age	.08	.25**	.40**	-.57**	-.53**
2. Education	—	-.25**	.50**	.30**	.05
3. SOC2000 ^a		—	-.08	-.29**	-.21**
4. WTAR ^b			—	.10	-.18*
5. Matrix reasoning				—	.42**
6. Digit-symbol coding					—

** $p < 0.01$; * $p < 0.05$

^a Standard Occupational Classification 2000

^b Wechsler Test of Adult Reading

standard deviations (SD) of response times of the Deary-Liewald task and the numbers task ($r[148] = .40$, $p < 0.01$). The correlations between the means and SDs within both reaction time tasks were also significant: Deary-Liewald task ($r[148] = .56$, $p < 0.01$); numbers task ($r[148] = .56$, $p < 0.01$).

With regard to the choice reaction time (CRT) tasks, there was a very large, significant positive correlation between the mean response times of the Deary-Liewald task and the numbers task ($r[148] = .82$, $p < 0.01$). There was a large, significant positive correlation between the standard deviations (SD) of response times for the Deary-Liewald task and the numbers task ($r[148] = .64$, $p < 0.01$). The correlations between the means and SD within each task were also large and significant: Deary-Liewald task ($r[148] = .82$, $p < 0.01$); numbers task ($r[148] = .78$, $p < 0.01$). Faster participants were less variable. There was a small, significant positive correlation between the number of errors made in the Deary-Liewald task and the numbers task ($r[148] = .18$, $p < 0.05$). There were few errors overall. The number of errors and mean response times within each task were slightly negatively

correlated: Deary-Liewald task ($r[148] = -.24$, $p < 0.01$); numbers task ($r[148] = -.25$, $p < 0.01$). Faster participants made more mistakes.

Reliability of the Deary-Liewald task

Internal consistency for the Deary-Liewald task was measured using Cronbach's alpha and was very high for both the SRT ($\alpha = .94$) and for correct responses on the CRT ($\alpha = .97$). Reliability of the SD of response times was measured using a split-half analysis. A correlation was conducted between the SD of the first half of responses and the SD of the second half of responses, which revealed a high significant correlation for correct responses on the CRT ($r[148] = .64$, $p < 0.01$). The correlation was not significant for the SRT ($r[148] = .15$, $p = 0.07$). A further experiment on 20 participants was conducted to provide another measure of period-free reliability. Each participant completed the SRT and CRT twice immediately one after the other. Means and SDs for these tests are shown in Table 5. Correlations between the first test and second test were significant for the SRT mean ($r[18] = .64$, $p < 0.01$) and SRT SD ($r[18] = .47$, $p < 0.05$), and highly significant for the CRT mean ($r[18] = .83$, $p < 0.01$) and CRT SD ($r[18] = .62$, $p < 0.01$). The correlation was not significant for the number of errors made in the CRT ($r[18] = .34$, $p = 0.14$).

Reaction time correlations with age and intelligence

Table 6 shows the correlations between the background and cognitive variables with the measures from the two reaction time tasks. Age correlated significantly with all of the reaction time measures. Older people were slower and more variable, and made fewer errors. Education did not correlate

Table 4 Correlations among the measures of the simple and choice reaction time tasks for the Deary-Liewald task and numbers task

	2	3	4	5	6	7	8	9	10
1. NS Mean	.56**	.54**	.32**	-.19*	.68**	.47**	.48**	.31**	-.15
2. NS SD	—	.33**	.26**	-.12	.35**	.40**	.36**	.35**	-.04
3. NC Mean		—	.78**	-.25**	.61**	.51**	.82**	.73**	-.21**
4. NC SD			—	-.15	.39**	.43**	.61**	.64**	-.07
5. NC Errors				—	-.18*	-.24**	-.30**	-.24**	.18*
6. DLS Mean					—	.56**	.61**	.39**	-.20*
7. DLS SD						—	.49**	.52**	-.12
8. DLC Mean							—	.82**	-.24**
9. DLC SD								—	-.13
10. DLC Errors									—

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

NS Numbers box, Simple reaction time task; NC Numbers box, Choice reaction time task; DLS Deary-Liewald task, Simple reaction time task; DLC Deary-Liewald task, Choice reaction time task; Errors Percentage of incorrect responses

Table 5 Means and standard deviations (SD) for the test-retest reliability study for the Deary-Liewald task

	<i>n</i>	SRT mean	SRT SD	CRT mean	CRT SD	CRT errors
First test	20	282.6	56.3	420.0	82.7	1.6
Second test	20	287.0	46.6	427.3	80.5	1.8

SRT Simple reaction time task; CRT Choice reaction time task; Errors Percentage of incorrect responses

significantly with any reaction time measure. People in more professional occupations (SOC2000) had faster SRT and CRT, and less variable CRT in both tasks. For the cognitive measures (WTAR, matrix reasoning and digit-symbol coding), we report both raw and age-adjusted correlations because of these measures' different correlations with age (see Tables 3 and 5). The WTAR showed near-to-zero raw correlations. When age-adjusted, there were significant negative correlations with the CRT means and SDs for the Deary-Liewald and numbers tasks, and the SRT variables in the Deary-Liewald task. Matrix Reasoning was negatively correlated with most of the SRT and CRT variables. The effect sizes were reduced when age was controlled. Digit-symbol coding correlated negatively with the majority of reaction time measures, except errors, and these persisted, though reduced in effect size, when age was controlled. In all instances, the correlations with cognitive tasks were very similar for the Deary-Liewald task and the numbers task.

Discussion

We have devised a new reaction time programme that allows the user to conduct simple and four-choice reaction time procedures. It allows certain experimental parameters to be adjusted. It collects data in a file that is straightforward to transfer for analysis. The programme is free, easy to use, and needs no special software. This report aims to let people know about the programme and invites them to use it. It also reports some data from a wide range of ages, spanning 18 to 80 years. The Deary-Liewald reaction time task provides reliable and valid measures. We found the expected associations between reaction time and age, and similarly with fluid intelligence and a psychometric test of processing speed. As expected, there was less association with crystallised intelligence. The associations with the same parameters from a very well studied reaction time device were very high, especially for choice reaction time.

With respect to investigations in intelligence differences (Der & Deary, 2003), ageing (Der & Deary, 2006), mortality (Shiple et al., 2006), and psychopharmacology (Strachan et al., 2001), it would be the four-choice reaction time measures (mean and standard deviation) that are recommended. Simple reaction time measures have lower associations with other variables generally, the distribution of simple reaction time means is less normal and the bivariate distribution with intelligence more problematic (Der & Deary, 2003), and simple reaction time standard deviations (intraindividual variability) have lower reliability here and elsewhere (Deary & Der, 2005a).

Table 6 Correlations between background and cognitive variables and the measures of the simple and choice reaction time tasks for the Deary-Liewald task and numbers task

	NS Mean	NS SD	NC Mean	NC SD	NC Errors	DLS Mean	DLS SD	DLC Mean	DLC SD	DLC Errors
Age ^a	.43**	.30**	.76**	.58**	-.26**	.45**	.39**	.71**	.65**	-.16*
Education ^a	-.04	.02	-.11	-.09	-.06	-.16	-.11	-.13	-.11	.07
SOC2000 ^b	.27**	.16	.37**	.32**	-.08	.31**	.19*	.36**	.28**	.02
WTAR ^a										
Full Sample	.10	.08	.11	.05	-.06	.01	-.08	.09	.06	-.08
(Age-adjusted)	(-.09)	(-.05)	(-.33**)	(-.25**)	(.05)	(-.22**)	(-.28**)	(-.31**)	(-.29**)	(-.02)
Matrix reasoning ^a										
Full sample	-.35**	-.18*	-.56**	-.38**	.19*	-.43**	-.40**	-.53**	-.48**	.13
(Age-adjusted)	(-.14)	(-.01)	(-.24**)	(-.08)	(.06)	(-.24**)	(-.23**)	(-.21*)	(-.17*)	(.05)
Digit-symbol coding ^a										
Full sample	-.41**	-.37**	-.62**	-.46**	.15	-.44**	-.42**	-.60**	-.54**	.17*
(Age-adjusted)	(-.24**)	(-.26**)	(-.38**)	(-.22**)	(.01)	(-.27**)	(-.27**)	(-.37**)	(-.30**)	(.10)

** $p < 0.01$; * $p < 0.05$

^a Pearson's Correlations

^b Spearman's Correlations

SOC2000 Standard Occupational Classification 2000; WTAR Wechsler Test of Adult Reading; NS Numbers box, simple reaction time task; NC Numbers box, choice reaction time task; DLS Deary-Liewald task, Simple reaction time task; DLC Deary-Liewald task, Choice reaction time task; Errors Percentage of incorrect responses

This report is intended to meet the need for a reaction time platform that is easily accessible to all relevant researchers. It also attempts to negotiate a tricky combination: of, on the one hand, being flexible enough to allow different researchers to run the test that they wish; and, on the other hand, of being sufficiently restricted so that different researchers can compare data because they are running the same basic task. Intentionally, there is no special software needed to run the test. We understand that many psychologists will wish to use reaction times that are tailor-made, with their own stimulus-response contingencies and manipulations, in order to test specific hypotheses. The Deary-Liewald task is not intended for them. It is intended for the large group of researchers who wish to have a standard simple or four-choice reaction time test to be used as a predictor or outcome variable.

We do not provide norms, and neither should we. We envisage slight between-study differences in overall levels of reaction times, based on their hardware (but see Appendix 3). However, within studies that use the same equipment for all subjects, the results will be useful: for making between-group comparisons, and for examining correlations.

We encourage researchers to download and use this reaction time programme in their studies (Appendix 2) and we offer to provide a summary of their findings on our website to provide a cumulative record of the findings with the task. As it becomes widely used, the validity and reliability data will accrue, and, after more than a century, it will be possible to compare studies that have used basically the same reaction time task.

Characteristics of the Deary-Liewald reaction time programme

The programme is designed to run on all laptop and desktop computers, requiring no special software. We recommend using a monitor with a vertical refresh rate of 60 Hz or better and with a pixel response time of 5 ms or faster (nearly all modern monitors fit this description). A simple, single screen page for the experimenter provides the following with respect to task set-up. The subject identity can be entered and the location for the saved data file. For SRT the experimenter can: indicate the number of practice and experimental trials required, the range (in milliseconds) for acceptable responses, and the range for the inter-stimulus interval. The experimenter can select to run a practice or the experiment proper. For CRT, the experimenter has the same control. Additionally, the response keys that correspond to each stimulus box may be programmed, simply by typing them into boxes on the screen. The programme allows the experimenter to save default settings. Data from the programme are saved to a database on the computer, from where they can be exported easily to a .csv file.

The location for downloading this programme is given in Appendix 2. There, the user will find the fully operational programme and brief instructions for use. The standard operating procedure for this task is in the supplementary online information.

Acknowledgement This work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative. Funding from the BBSRC, EPSRC, ESRC and MRC is gratefully acknowledged.

Appendix 1

Table 7 Percentiles (25, 50, 75) for the Deary-Liewald reaction time task

Variable	Age 18–25			Age 45–60			Age 61–80					
	Percentiles			Percentiles			Percentiles					
	<i>n</i>	25	50	75	<i>n</i>	25	50	75	<i>n</i>	25	50	75
DLS mean	50	230.6	243.1	250.7	50	258.5	280.6	301.0	50	256.4	286.6	306.6
DLS SD	50	23.0	28.3	40.8	50	33.9	42.6	60.1	50	34.1	51.7	63.0
DLC mean	50	355.9	381.8	418.9	50	439.2	481.0	524.1	50	478.3	548.2	605.2
DLC SD	50	53.2	68.5	83.2	50	83.1	100.9	124.6	50	95.9	119.0	148.1
DLC Errors	50	0.0	2.5	5.0	50	0.0	2.5	2.5	50	0.0	2.5	2.5

DLS Deary-Liewald task, simple reaction time task; *DLC* Deary-Liewald task, choice reaction time task; *Errors* Percentage of incorrect responses

Appendix 2 Instructions for downloading and using the Deary-Liewald reaction time programme

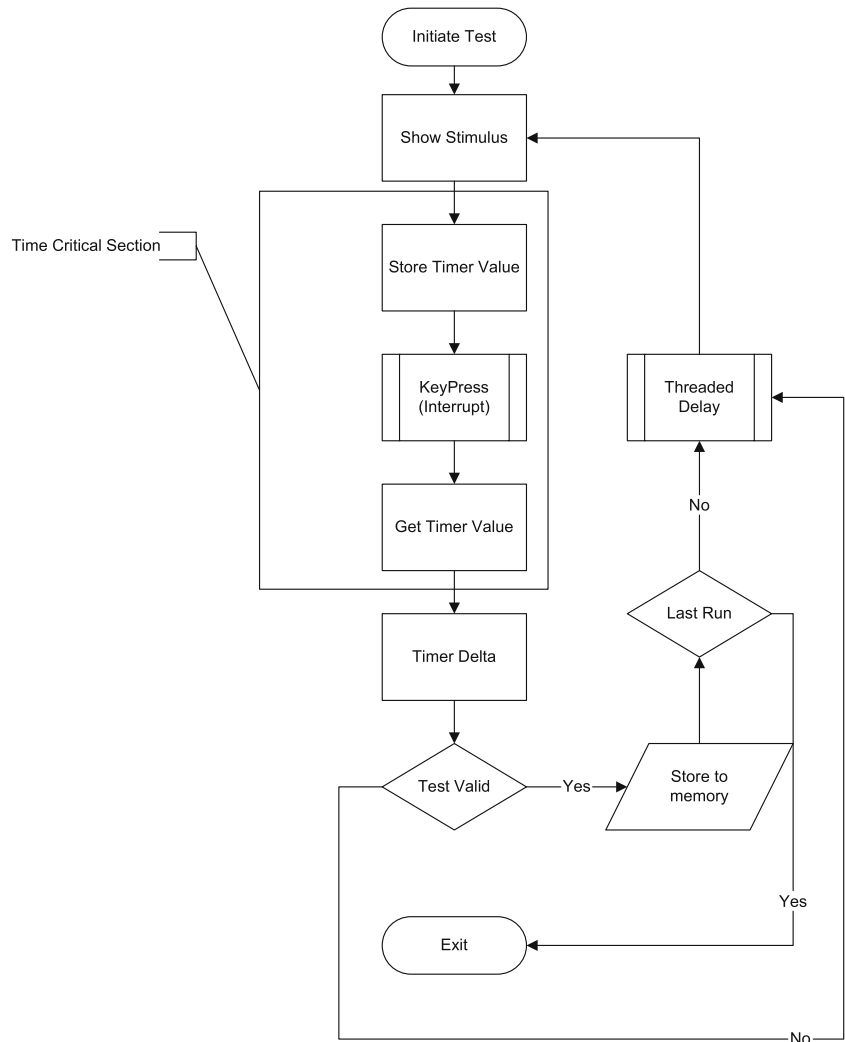
The Deary-Liewald reaction time is donation ware and can be downloaded, after registration, from the CCACE software repository site at www.ccace.ed.ac.uk/software. To register, click the “Register” option on the “Main Menu” on the left hand side off the page, fill out the form and press the “Register” button to complete registration. Once you have registered, you must log in to download the programme. After you have logged in, click on “software downloads” under the main menu on the left. Under the “Deary-Liewald Reaction Time Task” click on “Software Versions”. Click on the link to the zip file for the Deary-Liewald Reaction Time Task. Click “Download” and follow instructions.

This site also contains a help/feedback forum and a bug reporting/tracking system. Please use these utilities for support and/or functionality requests. Alternatively, the programme can be requested by emailing the first or second authors. The instructions for installing and running the programme are also available from the website and from Supplementary Materials to the present paper.

Appendix 3 The timing of operations in the Deary-Liewald reaction time programme

The nature of the Windows operating system is that it is multitasking and multithreaded. It achieves this by running a single message loop and queuing messages to this loop.

Fig. 3 Deary-Liewald Reaction time – time critical section flowchart



This model makes accurate timing in standard Windows programming a difficult task. Normal Windows timer events are dependent upon messages and therefore are dependent on message loop queuing. This makes them unreliable and unpredictable. This timing problem was identified very early on in the evolution of Windows and a solution was provided by the processor manufacturers by placing a number of high-resolution timers on the CPU. These are hardware-based timers and completely independent of the operating system being used on a particular computer. They were first implemented in the 386 CPU architecture and do not exist in previous versions of the chip. The code to access these timers has been built into the Kernel32.dll of the Windows operating system and is quite easily invoked from any language.

The easiest of these timers to use is the QueryPerformance Timer. This timer is widely used by gaming coders to control time critical animations. It provides sub-millisecond accuracy. The timer frequency is obtained by calling the QueryPerformanceFrequency function. The resolution of the timer varies, but it is sufficient to provide, in theory, sub-microsecond timing. The current value of the high-resolution timer is obtained by calling QueryPerformanceCounter. The returned value is a 64-bit integer. To use the high-resolution timer to get the starting value, we run the code that is to be timed, and then get the ending value. Subtracting the starting value from the ending value enables us to find how many timer ticks elapsed, and we divide by the performance frequency to obtain the number of seconds elapsed.

$$\text{Elapsed time}(s) = (\text{Endcount} - \text{StartCount}) / \text{Frequency}$$

This time is software-independent and gives sub-millisecond accuracy.

New processor architectures (Multi-core) can cause a problem with this timing model as there could be a timing mismatch in the timers on the two cores and, unless the Kernel32 is completely up to date with the latest patches from Microsoft, there is no guarantee which processor the startcount and endcount will be retrieved from. It is therefore critical that this software only be run on post-386 Windows systems that have all of the latest kernel patches applied. This is the only way to ensure the accuracy of this timing process.

The program itself has a timed loop with a time critical section (Fig. 3). The main process loop is controlled by a standard Windows timer placed on a time critical thread. This timer is triggering relatively slow events, and placing it on its own time critical thread gives it sufficient

precedence in the windows message queue to give it a 0.1 sec accuracy. The time critical section however is timed using the QueryPerformanceCounter to ensure the accuracy of timing the users' response to the stimulus.

References

- Aufdembrinke, B., Hindmarch, I., & Ott, H. (Eds.). (1988). *Psychopharmacology and reaction time*. Chichester, UK: John Wiley & Sons.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, *15*, 373–380.
- Cox, B. D., Huppert, F. A., & Whichelow, M. J. (1993). *The health and lifestyle survey: seven years on*. Aldershot, UK: Dartmouth.
- Deary, I. J. (2000). *Looking down on human intelligence: from psychometrics to the brain*. Cambridge, UK: Cambridge University Press.
- Deary, I. J., & Der, G. (2005a). Reaction time, age, and cognitive ability longitudinal findings from age 16 to 63 years in representative population samples. *Aging Neuropsychology and Cognition*, *12*, 187–215.
- Deary, I. J., & Der, G. (2005b). Reaction time explains IQ's association with death. *Psychological Science*, *16*, 64–69.
- Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: a population-based cohort study. *Intelligence*, *29*, 389–399.
- Der, G., & Deary, I. J. (2003). IQ, reaction time and the differentiation hypothesis. *Intelligence*, *31*, 491–503.
- Der, G., & Deary, I. J. (2006). Reaction time age changes and sex differences in adulthood. Results from a large, population-based study: the UK Health and Lifestyle Survey. *Psychology & Aging*, *21*, 62–73.
- Donders, F. C. (1868). Die schnelligkeit psychischer processe. *Archiv für Anatomie und Physiologie*, 657–681. [Reprinted as Donders, F. C. [1969]. On the speed of mental processes. *Acta Psychologica*, *30*, 412–431.]
- Galton, F. (1890). Remarks on 'Mental tests and measurements' by J McK. Cattell. *Mind*, *15*, 380–381.
- Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *Journal of Gerontology Psychological Sciences*, *57B*, P101–P105.
- Huppert, F. A., & Whittington, J. E. (1993). Changes in cognitive function in a population sample. In B. D. Cox, F. A. Huppert, & M. J. Whichelow (Eds.), *The Health and Lifestyle Survey: seven years on*. Aldershot, U.K.: Dartmouth.
- Jensen, A. R. (2006). *Clocking the mind: mental chronometry and individual differences*. Amsterdam, The Netherlands: Elsevier.
- Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M., & Cole, R. E. (1985). Galton's data a century later. *The American Psychologist*, *40*, 875–892.
- Luce, R. D. (1991). *Response times: their role in inferring elementary mental organisation*. Oxford, UK: Oxford University Press.
- Madden, D. J. (2001). Speed and timing of behavioural processes. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 288–312). San Diego: Academic Press.
- Psychological Corporation (The) (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio: Author.

- Psychological Corporation (The). (2001). *Wechsler test of adult reading*. San Antonio, TX: Harcourt Assessment.
- Ratcliff, R. (2008). Modeling aging effects on two-choice tasks: response signal and response time data. *Psychology and Aging*, 23, 900–916.
- Rose, D., & Pevalin, D. J. (Eds.). (2003). *A researcher's guide to the National Statistics Socio-economic Classification*. London, UK: Sage.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Shiple, B. A., Der, G., Taylor, M. D., & Deary, I. J. (2006). Cognition and all-cause mortality across the entire adult age range: health and lifestyle survey. *Psychosomatic Medicine*, 68, 17–24.
- Strachan, M. W. J., Deary, I. J., Ewing, F. M. E., Ferguson, S. S. C., Young, M. J., & Frier, B. M. (2001). Acute hypoglycaemia impairs functions of the central but not the peripheral nervous system. *Physiology & Behavior*, 72, 83–92.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III*. San Antonio: The Psychological Corporation.