Available online at www.sciencedirect.com
SCIENCE DIRECT®
Intelligence 32 (2004) 113–119
INTELLIGENCE

An 'instantaneous' estimate of a lifetime's cognitive change

Ian J. Deary[a,*], Lawrence J. Whalley[b], John R. Crawford[c]

[a] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland, UK
[b] Department of Mental Health, Clinical Research Centre, Cornhill Hospital, University of Aberdeen, Cornhill Road, Aberdeen AB24 2ZD, Scotland, UK
[c] Department of Psychology, King's College, University of Aberdeen, Aberdeen AB24 2UB, Scotland, UK

Received 21 June 2002; received in revised form 19 February 2003; accepted 18 June 2003

Abstract

Change in cognitive functioning is an important aspect of human aging and a key outcome in many medical conditions. However, cognitive change can rarely be measured directly, since prior cognitive data do not exist for most people. We examined the criterion validity and one-year stability of the difference between National Adult Reading Test (NART) and Raven's Standard Progressive Matrices Test (Raven) as an estimate of cognitive change. We followed up over 80 people whose cognitive ability (using the Moray House Test [MHT]) was measured at age 11 in the Scottish Mental Survey of 1932 (SMS 1932). At age 77 and again at 78 years, they took the NART, Raven, and two Wechsler Adult Intelligence Scale-Revised (WAIS-R) subtests. The difference between NART and Raven standardised scores (*estimated* cognitive change) correlated .638 ($P < .001$) with the difference between MHT and Raven scores, and .658 ($P < .001$) with the difference between MHT and WAIS scores (two measures of *actual* cognitive change). The stability of the NART–Raven difference across a one-year period was .643 ($P < .001$). We have demonstrated the stability and criterion validity of an estimate of lifetime cognitive change that takes about half an hour to administer.

Keywords: Intelligence; Aging; Raven; NART; Cognitive change

Abbreviations: MHT, Moray House Test; Raven, Raven's Standard Progressive Matrices; NART, National Adult Reading Test; WAIS-R, brief estimate of performance ability based on Digit Symbol and Object Assembly subscales of the Wechsler Adult Intelligence Scale-Revised.

* Corresponding author. Tel.: +44-131-650-3452; fax: +44-131-650-3461.
E-mail address: I.Deary@ed.ac.uk (I.J. Deary).

0160-2896/$ - see front matter
doi:10.1016/j.intell.2003.06.001

1. Introduction

When examining a person for signs of cognitive decline, it is important either to know or at least to estimate their prior level of cognitive ability. In the ideal situation, a person's current cognitive function score would be compared with their own baseline. However, such prior cognitive ability data are rarely available, especially from a previous period of good health. Therefore, prior or premorbid mental ability is commonly *estimated* using the National Adult Reading Test (NART; Crawford, 1992; Nelson & Willison, 1991). The NART tests the pronunciation of 50 phonologically irregular English words. NART performance is resistant to the effects of normal aging (Crawford, Stewart, Garthwaite, Parker, & Besson, 1988) and is relatively resistant to the effects of neurological or psychiatric morbidity (see Crawford, 1992, 2003; O'Carroll, 1995 for reviews). We have recently shown that the NART has good retrospective validity in healthy old people. Taken at age 77 years, it correlated .73 ($P < .001$) with Moray House Test (MHT) scores from the same subjects taken at age 11 (Crawford, Deary, Starr, & Whalley, 2001).

Use of the NART as an indicator of prior or premorbid intelligence rests upon the empirical evidence that the ability to read words does not decline, as other cognitive functions do, with age and some medical conditions (Crawford, 1992; Crawford et al., 2001). One influential theory of human intelligence differences, which is highly relevant to the use of NART and to aging and cognitive functions generally, is Cattell's theory of fluid and crystallised intelligence (Cattell, 1998; Horn & Cattell, 1966). Crystallised intelligence is related to a person's stored information and cultural influences (Baltes, Staudinger, & Lindenberger, 1999). It is often tested using vocabulary- and general knowledge-type tests. It shows little change with healthy aging. Fluid intelligence is related to basic information processing and is assessed using tests comprising novel materials, often under time pressure. Scores on tests of fluid intelligence decline with age and as a result of some medical conditions. Raven's Progressive Matrices (RPM or Raven) is often used as a measure of general fluid reasoning (Carroll, 1993). In summary, NART score is often used to represent a relatively stable record of the *prior* best level of cognitive functioning, whereas Raven score is used to represent a relatively labile index of the *current* level of cognitive functioning.

At the level of the population, the difference between fluid-type and crystallised-type intelligence test scores increases with age. It is increasingly seen as important to employ this difference in the individual person to estimate potentially clinically significant cognitive change (Crawford, 2003; Sawrie, Marson, Boothe, & Harrell, 1999). The difference between an individual's NART and Raven performances has been investigated as an *estimate* of their cognitive change (Davis, Ho, Bradshaw, & Szaba, 2000; Freeman & Godfrey, 2000; van den Broek & Bradshaw, 1994), but no study until the present report has criterion validated this estimate by correlating it with their *actual* cognitive change. Here we report such a validation study after conducting two waves of follow-up on survivors of the Scottish Mental Survey of 1932 (SMS 1932).

2. Method

2.1. Participants

For the main analyses, subjects were 87 non-demented people from the Aberdeen area who took part in the SMS 1932 at age 11 years (Deary, Whalley, Lemmon, Crawford, & Starr, 2000; Scottish Council for Research in Education [SCRE], 1933). All were born in 1921. The selection of subjects for the first wave of follow-up, at age 77, was described previously (Crawford et al., 2001). The present sample comprised those subjects who undertook a further wave of follow-up cognitive testing and medical examination. Therefore, these subjects have provided cognitive data at age 11 (in 1932), age 77 (mostly in 1998), and age 78 years (mostly in 1999). None was suffering from any current major physical or mental illness, and none was taking medication known to affect cognitive functioning. The subjects for the analyses presented in Table 2 were 80 people from the same study and had the same characteristics. They had also taken Wechsler subtests in the first wave of follow-up. Over 50 of these subjects were included in the 87 subjects described above. The Joint Ethics Committee of Grampian Health Board and the University of Aberdeen gave ethical approval for the study.

2.2. Materials and procedure

The SMS 1932 used the MHT, a 76-item IQ-type test. It correlated about .80 ($N=1000$) with the Stanford Binet IQ test, and its stability across 66 years of follow-up was $r=.63$ (.73 after correction for attenuation; Deary et al., 2000). In the ages 77 and 78 follow-up examinations, subjects took the NART (Nelson & Willison, 1991), Raven's Standard Progressive Matrices (Raven, Raven, & Court, 1993), and the Digit Symbol and Object Assembly Tests from the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981). The NART, Raven, and WAIS-R subtests were administered to subjects individually using standard instructions. The time limit for the Raven test was 20 min. The NART and Raven scores were expressed as the number of correct items obtained by each subject. The Digit Symbol and Object Assembly Tests were administered according to instructions in the WAIS-R Manual and scores were converted to z scores and summed. The sum was converted to a z score to give a brief measure of WAIS-R Performance-type ability. Researchers administering the tests were trained and supervised by a chartered clinical psychologist (JRC).

(a) *Actual* lifetime (ages 11 to 77 or 78 years) cognitive change differences for the subjects were calculated using linear regression. This adjusted the current, old-age ability scores (Raven at ages 77 or 78) for the individual differences in the actual early life ability score—the MHT score in 1932—and saved the standardised residuals. A second measure of lifetime cognitive change differences was obtained by linear regression using the WAIS-R Performance Test score at age 77 and adjusting for the MHT score from 1932.

(b) *Estimated* lifetime cognitive change scores for the subjects were calculated using linear regression. This adjusted the current, old-age ability scores (Raven at ages 77 or 78) for individual differences in the estimated prior ability score—NART at ages 77 or 78—and saved the standardised residuals.

Thus, the first set of residual scores (a) represents (i) the regression-based differences between the MHT and Raven and (ii) the regression-based differences between MHT and WAIS-R Performance scores. These are, therefore, measures of *actual* changes in cognitive function across more than 66 years. The second set of residual scores (b) represents the differences between the NART and Raven and are therefore *estimates* of cognitive change tested in about 30 min. Scores (a) are the criterion against which the estimates (b) are to be validated. Sawrie et al. (1999) discuss in detail the advantages of using standardised regression-based scores of cognitive change rather than the raw differences between test scores.

There is a problem in attempting to validate the estimated cognitive change scores against the actual cognitive change scores. Two strategies were adopted to overcome this problem.

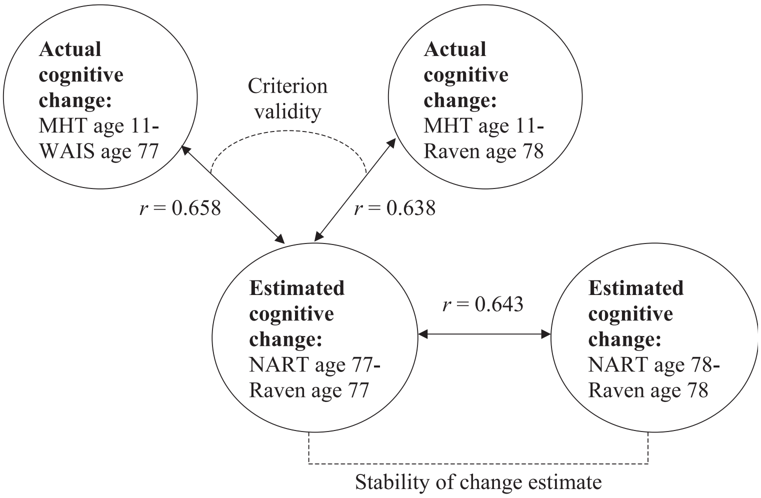

Fig. 1. Stability and criterion validity coefficients for the cognitive change estimate developed in the present study.

First, two waves of follow-up were conducted in old age, rather than one. Using the same test result (Raven score at age 77, say) as an indicator of current ability in both the *actual* (MHT vs. Raven) and *estimated* (NART vs. Raven) cognitive change scores introduces a spurious element of positive association between them. That is because any error or situation-specific variance from the same Raven test will be common to both the actual and estimated cognitive change measures. Therefore, we used (i) the Raven test score at age 78 as the current cognitive ability measure for the calculation of *actual* cognitive change, and (ii) the Raven test score at age 77 as the current ability measure for the calculation of *estimated* cognitive change (Fig. 1). As a consequence, our reported correlations between estimated and actual cognitive change are highly conservative (underestimates of the true values), because any instability in Raven scores between ages 77 and 78 will lower the achievable criterion validity coefficient. The other important advantage of having two waves is that the stability (reliability) of the measure of estimated cognitive change can be assessed by comparing estimates obtained one year apart.

However, the first approach might still be problematic. It might merely have created a residual score for Raven's matrices 'controlling' for NART at one age and MHT at another. The resulting correlation between these residual scores might be no more than the correlation between that stable variance in Raven's that is not predicted by NART or MHT. The second approach overcomes this by obtaining a second measure of cognitive change differences with MHT at age 11 and the brief Wechsler Performance score at age 77. The standardised residuals of Wechsler Performance scores at age 77 after controlling (using linear regression) for MHT scores at age 11 were correlated with Raven scores at age 77 after controlling (using linear regression) for NART scores at age 77. In this approach, the estimated and actual cognitive change differences between ages 11 and 77 do not share any tests.

3. Results

The first results concern the approach that uses Raven scores in both the actual and estimated cognitive change measures. The mean scores for Raven and NART were similar at ages 77 and 78

Table 1
Descriptive statistics and correlations for the test scores used to compute cognitive change in the present study, where actual change is based on MHT (age 11) and Raven scores (old age) ($N = 87$)

	MHT age 11	NART age 77	NART age 78	Raven age 77	Raven age 78
MHT age 11	–				
NART age 77	.741	–			
NART age 78	.687	.888	–		
Raven age 77	.423	.538	.555	–	
Raven age 78	.458	.549	.571	.784	–
Mean (S.D.)	42.5 (12.5)	35.8 (7.5)	34.2 (7.4)	29.0 (9.0)	29.1 (9.0)

All correlations are significant at $P < .001$.

years (Table 1). The one-year stability of the two tests was very high, with NART at .89 and Raven at .78. Correlations between NART and Raven were between .42 and .55. Estimated cognitive change (NART age 77 − Raven age 77) correlated .638 ($P < .001$) with actual cognitive change (MHT age 11 − Raven age 78) (Fig. 1). This large effect size is an underestimate of the true value, necessarily limited by the unreliability within the four scores that contribute to it. The stability coefficient of the estimated cognitive change ([NART age 77 − Raven age 77] correlated with [NART age 78 − Raven age 78]) across a gap of one year was .643 ($P < .001$; Fig. 1), which is again large and an underestimate of the true value.

The second results concern the approach that uses Raven scores in only the estimated cognitive change, and WAIS Performance score as the current cognition score in the actual change measure. The descriptive statistics for the variables involved are shown in Table 2. There is a high (> .8) correlation between Raven and WAIS-R Performance scores, the two tests used to measure current cognitive function in this analysis. Estimated cognitive change (NART age 77 − Raven age 77) correlated .658 ($P < .001$) with actual cognitive change (MHT age 11 − WAIS-R Performance age 77) (Fig. 1). In this analysis, there was no overlap in the tests used to measure estimated or actual cognitive change, and the coefficient is very similar to the approach that used Raven in both the estimated and actual cognitive change measures.

Table 2
Descriptive statistics and correlations for the test scores used to compute cognitive change in the present study, where actual change is based on MHT sores and Wechsler Performance scores ($N = 80$)

	MHT age 11	NART age 77	Raven age 77	WAIS age 77
MHT age 11	–			
NART age 77	.727	–		
Raven age 77	.472	.507	–	
WAIS age 77	.416	.487	.802	–
Mean (S.D.)	43.3 (12.2)	35.5 (7.5)	29.2 (8.7)	− 0.058 (1.76)

All correlations are significant at $P < .001$. WAIS = brief estimate of Performance ability based on Digit Symbol and Object Assembly subscales of the Wechsler Adult Intelligence Scale-Revised.

4. Discussion

The 'instantaneous' difference between NART and Raven scores has good criterion validity and stability as a measure of relative lifetime cognitive change. Thus, we validated a cognitive change *estimate* taking about 30 min at a single sitting against *actual* cognitive change that took about 67 years to measure. Our study involved relatively healthy old people within a narrow age range. Future studies should provide normative data on NART−Raven differences on other age and medical groups for use in clinical settings. In the present study, we did not concentrate on absolute values for the NART−Raven difference. Instead, the purpose was to ask the question whether individual differences in a brief estimate of cognitive change would correlate highly with individual differences in an actual measure of cognitive change across the life span: they did. This gives impetus to researchers to use the NART−Raven difference and apply it to the standardised regression-based method of Sawrie et al. (1999) to generate individuals' cognitive change scores. These scores will indicate whether a person is aging successfully or unsuccessfully with respect to others in their age group. With regard to the measure of cognitive ability in youth, it would have been desirable to have a valid measure recorded in young adulthood rather than at age 11. Nevertheless, two considerations mean that using the MHT scores at age 11 were not problematic here. First, the MHT scores at age 11 have an estimated true correlation of over .7 (Deary et al., 2000) and perhaps as high as .8 (Deary, Whiteman, Starr, Whalley, & Fox, in press) with IQ at age 76−80. Therefore, given the developmental and aging processes that intervene between ages 11 and about 80, they must be extremely good estimates of young adult ability differences. Second, the fact that the cognitive scores were taken at age 11, prior to full mental development, can only mean that the current high value of the validity coefficient of cognitive change is an underestimate.

Acknowledgements

This research was supported by grants from Henry Smith's Charities, the Scottish Executive Department of Health, Chief Scientist's Office, and the Biotechnology and Biological Sciences Research Council. We thank Mariesha Struth, Helen Lemmon, and Steve Leaper for gathering and compiling data. Ian J. Deary is the recipient of a Royal Society-Wolfson Research Merit Award. Lawrence J. Whalley holds a Wellcome Trust Career Development Award.

References

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology*, *50*, 471–507.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Cattell, R. B. (1998). Where is intelligence? Some answers from the triadic theory. In J. J. McArdle, & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 29–38). London: Lawrence Erlbaum Associates.

Crawford, J. R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 21–49). London: Erlbaum.

Crawford, J. R. (2003). Psychometric foundations of neuropsychological assessment. In L. H. Goldstein, & J. McNeil (Eds.), *Clinical Neuropsychology: A Practical Guide to Assessment and Management for Clinicians*. Chichester: Wiley.

Crawford, J. R., Deary, I. J., Starr, J. M., & Whalley, L. J. (2001). The NART as an index of prior intellectual functioning: A retrospective validity study covering a 66 year interval. *Psychological Medicine*, *31*, 451–458.

Crawford, J. R., Stewart, L. E., Garthwaite, P. E., Parker, D. M., & Besson, J. A. O. (1988). The relationship between demographic variables and NART performance in normal subjects. *British Journal of Clinical Psychology*, *27*, 181–182.

Davis, C., Ho, M. -Y., Bradshaw, C. M., & Szaba, E. (2000). Estimation of premorbid performance on Raven's Standard Progressive Matrices using reading-test performance: New normative data. *Clinical Neuropsychological Assessment*, *2*, 113–123.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish Mental Survey. *Intelligence*, *28*, 49–55.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (in press). The impact of childhood intelligence on later life: Following up the Scottish mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*.

Freeman, J., & Godfrey, H. (2000). The validity of the NART–RSPM index in detecting intellectual decline following traumatic brain injury: A controlled study. *British Journal of Clinical Psychology*, *39*, 95–103.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallised general intelligences. *Journal of Educational Psychology*, *57*, 253–270.

Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART) manual* (2nd ed.). Windsor, UK: NFER-Nelson.

O'Carroll, R. (1995). The assessment of premorbid ability: A critical review. *Neurocase*, *1*, 83–89.

Raven, J., Raven, J. C., & Court, J. H. (1993). *Manual for Raven's Progressive Matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.

Sawrie, S. M., Marson, D. C., Boothe, A. L., & Harrell, L. E. (1999). A method for assessing clinically relevant individual cognitive change in older adult populations. *Journal of Gerontology: Psychological Sciences*, *54B*, P116–P124.

Scottish Council for Research in Education [SCRE] (1933). *The intelligence of Scottish children: A national survey of an age-group*. London: University of London Press.

van den Broek, M. D., & Bradshaw, C. M. (1994). Detection of acquired deficits in general intelligence using the National Adult Reading Test and Raven Standard Progressive Matrices. *British Journal of Clinical Psychology*, *33*, 509–515.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.