The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval

J. R. CRAWFORD,[1] I. J. DEARY, J. STARR AND L. J. WHALLEY

From the Departments of Psychology and Mental Health, University of Aberdeen; and Department of Psychology, University of Edinburgh and Department of Geriatric Medicine, Royal Victoria Hospital, Edinburgh

ABSTRACT

Background. The National Adult Reading Test (NART) is widely used in research and clinical practice as an estimate of pre-morbid or prior ability. However, most of the evidence on the NART's validity as a measure of prior intellectual ability is based on concurrent administration of the NART and an IQ measure.

Method. We followed up 179 individuals who had taken an IQ test (the Moray House Test) at age 11 and administered the NART and the Mini-Mental State Examination (MMSE) at age 77. A subset ($N = 97$) were also re-administered the original IQ test.

Results. The correlation between NART performance at age 77 and IQ age 11 was high and statistically significant ($r = 0.73$; $P < 0.001$). This correlation was comparable to the correlation between NART and current IQ, and childhood IQ and current IQ, despite the shared influences on the latter variable pairings. The NART had a significant correlation with the MMSE but this correlation fell to near zero ($r = 0.02$) after partialling out the influence of childhood IQ.

Discussion. The pattern of results provides strong support for the claim that the NART primarily indexes prior (rather than current) intellectual ability.

INTRODUCTION

The detection and quantification of cognitive impairment in the individual case is problematical because of the wide variability in cognitive abilities in the general adult population. Scores on cognitive measures that are average, or even-above average, can still represent a significant impairment for an individual of high pre-morbid ability (Crawford, 1992; Deary, 1995). Conversely, for individuals with modest pre-morbid resources, test scores that fall well below the mean may be entirely consistent with their prior level of functioning (Crawford *et al.* 1998). Because of this, normative comparison standards are of limited utility in the detection

of impairment; and are supplemented with individual comparison standards when attempting to assess acquired deficits (Lezak, 1995; O'Carroll, 1995; Crawford, 1996). Ideally, these individual standards would be obtained from cognitive test scores obtained in the pre-morbid period. However, this is rarely a viable option as most individuals have had no prior formal testing and even where they have, it is difficult or impossible to obtain the results. Because of these difficulties, clinicians normally have to settle for some means of estimating an individual's pre-morbid ability.

One approach to this problem is to use demographic variables (i.e. years of education, occupational status, age etc.) to provide a broad estimate of pre-morbid intelligence (Barona *et al.* 1984; Crawford & Allan, 1997). A more common alternative is to use a measure of

[1] Address for correspondence: Professor John R. Crawford, Department of Psychology, King's College, University of Aberdeen, Aberdeen AB24 2UB.

present ability that is relatively unaffected by neurological or psychiatric disorder. Currently, the test most commonly used to estimate pre-morbid intelligence is the National Adult Reading Test (NART; Nelson, 1982; Nelson & Willison, 1991). The NART is 50-item single word reading test of graded difficulty. The development of the NART stemmed from: (1) the assumption that, in the normal adult population, level of reading ability is closely related to general intellectual level; and (2) from the clinical observation that, in patients with dementia, oral reading was relatively well preserved until late in the dementing process (Nelson & McKenna, 1975).

It is argued that the NART makes minimal demands on current cognitive capacity because it requires only oral reading of short, single words (Nelson & O'Connell, 1978). In addition all the words are irregular, that is, they violate grapheme–phoneme correspondence rules (e.g. *ache*, *thyme*, *topiary*). The supposition is that, as a result, the test depends on prior or pre-morbid ability because a testee must have prior knowledge of a word's pronunciation; deployment of current cognitive resources (e.g. the application of grapheme–phoneme correspondence rules) will not result in a correct pronunciation.

To be considered valid, any putative measure of pre-morbid intelligence must fulfil three criteria (Crawford, 1992). First, like any psychological test, it must possess adequate reliability. Secondly, it must have adequate criterion validity, that is, it must correlate highly with measures of psychometric intelligence. Thirdly, performance on the measure must be largely impervious to the effects of neurological or psychiatric disorder. The NART fulfils the first criterion in that it possesses high internal consistency (Crawford *et al.* 1988; Nelson & Willison, 1991), test–retest reliability (Crawford *et al.* 1989) and inter-rater reliability (O'Carroll, 1987; Crawford *et al.* 1989).

The results from studies that have addressed the criterion validity of the NART have been generally positive. For example, in the original NART standardization sample (Nelson, 1982) which consisted of 120 healthy adults, the NART predicted 55 % of the variance in IQ as measured by a short-form Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955). Crawford *et al.* (1989) used a full-length WAIS and reported

that the NART explained 66 % of IQ variance in a census matched sample of healthy adults.

A feature of all these studies is that, although the NART is intended to measure prior or pre-morbid intelligence, the validity information they provide is based solely on concurrent administration of the NART and an IQ measure. A notable exception to this is an important study by Berry *et al.* (1994) of an elderly sample in which an North American variant of the NART (Blair & Spreen, 1989) had a correlation of 0·70 with WAIS-R Full Scale IQ scores (Wechsler, 1981) obtained 3·5 years earlier.

In the present study we had access to IQ scores obtained at age 11 for a sample of 77 year-old participants and we can therefore extend this approach by examining the retrospective validity of the NART over a 66 year interval. As the NART is intended as a measure of prior ability, we therefore hypothesized that NART scores would be significantly correlated with earlier IQ. Further, in a subgroup of this sample we were also able to re-administer the original IQ test at age 77. This permitted us to test additional competing hypotheses concerning the measurement characteristics of the NART.

Does the NART index current or prior ability?

If the NART is primarily an index of current intelligence, then it follows that it should correlate more highly with IQ measured concurrently than with earlier IQ. It also follows that earlier IQ should correlate more highly with IQ scores age 77 than with the NART (given that, in the former case, the same test was used on both occasions). Conversely, if the NART is indeed primarily an index of prior intelligence, then it follows that it should correlate as highly, or even more highly, with earlier IQ scores than with current IQ scores. It follows also that the correlation between earlier IQ scores and NART would be as high, or again, possibly higher, than the correlation between the two sets of IQ scores.

Is the NART sensitive or impervious to cognitive impairment?

As previously noted, the third criterion for a valid measure of prior or pre-morbid intelligence is that performance should be largely impervious to the effects of neurological or psychiatric disorder (we would suggest that the measure

should also be insensitive to individual differences in cognitive decline associated with normal ageing). The most common way of assessing the extent to which the NART fulfils this criterion has been to compare NART performance in neurological or psychiatric samples with control participants matched for demographic characteristics. Extensive reviews of these studies can be found in Crawford (1992), O'Carroll (1995) and Franzen *et al.* (1997). Briefly, to summarize this literature, NART performance appears to be largely impervious to the effects of many neurological and psychiatric disorders, as judged by the absence of significant differences between the clinical and control samples. However, the balance of evidence suggests that the NART is commonly affected in many cases of severe, or even moderate, Alzheimer's disease, and in some other specific conditions (e.g. Huntington's disease).

Another means of assessing the robustness of the NART is to examine the correlation between an impairment sensitive measure and NART performance (Crawford, 1992; O'Carroll, 1995). The studies that have employed this approach have been concerned with the relationship between the NART and mental status instruments in dementia samples. The rationale is that a significant correlation between these variables would indicate that NART performance is affected by the severity of dementia whereas the absence of a correlation would suggest that NART performance is independent of current mental status. Results using this approach have been contradictory. For example, O'Carroll & Gilleard (1986) found that the NART did not correlate with mental status scores on the Clifton Assessment Procedures for the Elderly (Pattie, 1981) in a sample of dementia patients of mixed aetiology. In contrast, Patterson *et al.* (1994) assessed 45 patients with probable Alzheimer's disease and found that NART performance was significantly correlated ($r = 0.56$) with performance on the Mini-Mental State Examination (MMSE; Folstein *et al.* 1975).

A difficulty with the rationale of this approach is that, although any observed relationship between the NART and mental status could reflect a shared sensitivity to current level of impairment, it could also arise because performance on the mental status instruments is not independent of pre-morbid ability (Crawford, 1989). Consistent with this interpretation, Starr *et al.* (1992) reported a highly significant correlation ($r = 0.49$) between the NART and MMSE in a large sample of healthy elderly participants. However, even in this study, it is still possible that the observed relationship arose because both tests were sensitive to individual differences in age associated cognitive decline, or because both were indexing deficits arising from undiagnosed pathology.

The present study offers a means of evaluating these competing explanations. Both explanations lead to the prediction that the NART will be significantly correlated with the MMSE. However, if the relationship stems from a shared sensitivity to cognitive decline, then statistically controlling for childhood intellectual ability should leave the correlation unaffected. In contrast, if the shared variance reflects the influence of prior intellectual ability, then controlling for childhood IQ should abolish the relationship. It can be noted that the prior ability explanation predicts the null hypothesis, i.e. that the correlation will not be significant after partialling out prior IQ. However, a positive test of this explanation can be formed by testing whether the partial correlation is significantly lower than the raw correlation.

The Scottish Mental Survey 1932

The present study was made possible because of the Scottish Mental Survey conducted in 1932. This study, which was run by the Scottish Council for Research in Education (SCRE), sought to 'obtain data about the whole distribution of the intelligence of Scottish pupils from one end of the scale to the other' (SCRE, 1933). On 1 June 1932, all children at school in Scotland and born in the calendar year 1921 undertook a group-administered mental ability test, including some practice items. Children were tested in classrooms by teachers who followed detailed instructions. The number of children tested was 87498 (44210 boys and 43288 girls). A very small number of private schools and those children absent owing to sickness were the only 1921-born children not tested.

The group mental ability test used in the Survey is referred to in the original publication

as the 'Verbal Test' (SCRE, 1933). This however is something of a misnomer as the test comprises a variety of types of item as follows: following directions (14 items); same-opposites (11); word classification (10); analogies (8); practical items (6); reasoning (5); proverbs (4); arithmetic (4); spatial items (4); mixed sentences (3); cypher decoding (2); and other items (4). The test has 71 numbered items, 75 items in total, and the maximum possible score is 76. The test was very closely related to the Moray House Test No. 12 that was used in 'eleven-plus' examinations in England. We shall refer to the test as the Moray House Test (MHT). The scores on the MHT in 1932 were validated by individually testing a representative sample of 1000 of the children (500 boys, 500 girls) on the Stanford-Binet Intelligence Scale (McNemar, 1942). The correlation between the two tests for boys ($r = 0.81$) and girls ($r = 0.78$) established that the MHT had high concurrent validity as a measure of IQ.

METHOD

Participants

The Scottish Council for Research in Education made the complete data set for the 1932 Scottish Mental Survey available to the authors. From January to May 1998 we (L.J.W. and I.J.D.) traced local (North-East Scotland) survivors of the 1932 Scottish Mental Survey. With the approval of the Joint Ethics Committee of Grampian Health Board and the University of Aberdeen, we contacted 199 survivors randomly selected from the Community Health Index and 35 other locals who volunteered on hearing media reports of the study. Of the 234 potential subjects, 208 people agreed to a full physical and mental health assessment. From this sample we obtained 196 participants, none of whom was suffering from any major physical or mental illness or medication known to affect cognitive functioning. A subsample ($N = 97$) also agreed to re-take the Moray House Test 66 years after they had originally taken the test.

Materials and procedure

During 1998 the 196 participants were administered the NART and MMSE according to standard procedures. These tests were administered to subjects individually by trained researchers under the supervision of a qualified

clinical psychologist (J.R.C.). The tests were scored according to standard instructions except that NART performance was expressed as number of items correct rather than number of errors (this was done so that all psychological tests were scaled in the same direction i.e. high scores indicated good performance). The MHT scores the participants obtained in 1932 were identified from the SCRE records.

Of the 97 participants who agreed to retake the MHT, 71 retook the test at a group testing session in a large public hall in Aberdeen town centre on 1 June 1998, precisely 66 years to the day after the first sitting. The remaining 16 participants attended on later dates convenient to themselves. The same instructions as used in 1932 (SCRE, 1933) were employed. Forty-five minutes were allowed for the completion of the test. Only two of the tests' items required minor altering from the 1932 version. A question involving shillings and pence was altered to feet and inches because, whereas money altered to a decimal format in 1971, the measurement of distance is still principally duodecimal in the United Kingdom, especially among old people. Another question referring to 'vitamine' was changed to read 'vitamins'.

RESULTS

The summary statistics (means, S.D.s and ranges) for the NART and MMSE scores obtained in 1998, and scores on the Moray House Test obtained in 1932, are reported in Table 1. This table also presents the same summary statistics for the subsample ($N = 97$) re-administered the Moray House Test in 1998 and includes the statistics for the 1998 scores.

A scatterplot of the relationship between NART performance and scores on the MHT in

Table 1. *Summary statistics for the NART, MMSE, and MHT in the full sample and subsample readministered the MHT in 1998*

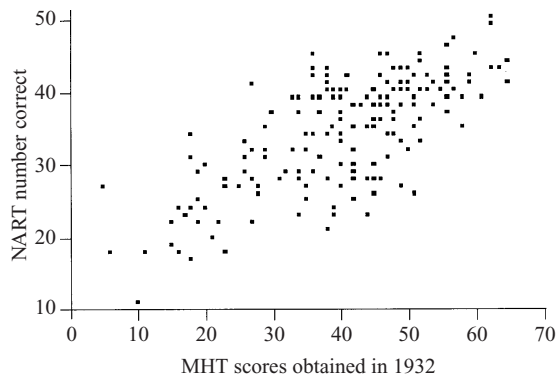|	Full sample ($N = 179$)			Subsample ($N = 97$)		
	Mean	(S.D.)	Range	Mean	(S.D.)	Range
NART (no. correct)	34·27	(7·58)	11–50	36·09	(7·14)	18–50
MMSE	28·31	(1·97)	18–30	28·85	(1·27)	24–30
MHT-1932	40·33	(13·02)	5–65	43·20	(12·02)	6–65
MHT-1998	—	(—)	—	54·01	(12·04)	17–74

FIG. 1. Scatterplot of the relationship between MHT scores obtained in 1932 and NART scores obtained in 1998.

1932 is presented as Fig. 1. The Pearson Product Moment correlation between these variables was 0·73, $P < 0.001$ (95% CLs = 0·65, 0·79). Hierarchical linear regression revealed that polynomial functions of MHT scores (MHT^2 and MHT^3) did not significantly increase the percentage of NART variance explained over that explained by MHT scores alone; this confirms the impression of visual inspection that the relationship is linear.

Comparison of correlations between NART, childhood IQ, and concurrent IQ scores

In the subsample who retook the MHT ($N = 97$) in 1998, the correlation between the NART and MHT scores from 1932 was similar to that obtained in the full sample; $r = 0.69$, $P < 0.001$. The correlation between the NART and MHT scores obtained in 1998 was 0·63 ($P < 0.001$). Thus, the NART was as highly correlated with prior IQ scores than it was with IQ scores concurrently. A significance test for non-independent correlations (Williams, 1959; Steiger, 1980) revealed that these two correlations were not significantly different; $t = 1.05$, df = 94, $P = 0.30$. As reported elsewhere (Deary *et al.* 2000), the correlation between MHT scores in 1932 and 1998 was 0·64. Thus, the correlation between NART and prior IQ scores (0·69) was comparable to the correlation between scores on the same IQ test administered on two occasions. William's test revealed that these two correlations were not significantly different; $t = 0.92$, df = 94, $P = 0.36$. The relationships between these three variables are represented in Fig. 2.

Correction for attenuation

Although the correlation between IQ in 1932 and NART in the full sample is substantial, it is an underestimate of the population correlation because of attenuation arising from restriction in the range of ability. The population standard deviation for the MHT in 1932 was 15·5 in comparison with the s.D. of 13·0 for the 1932 scores in the current sample. The variance in the sample was therefore 70% of that found in the population. When the correlation between NART and MHT scores was corrected for the effects of restriction of range, using the standard formula (Nunnally & Bernstein, 1994), it rose to 0·78.

Relationship between NART performance and mental status

Because the present sample consisted of healthy participants, the MMSE had limited variance and was subject to ceiling effects (see Table 1). Despite this, NART scores had a modest but highly significant correlation with scores on the MMSE; $r = 0.25$, $P < 0.001$ (95% CLs = 0·12, 0·38). The partial correlation between the NART and MMSE, controlling for MHT scores in 1932, was essentially zero and was not statistically significant; $r = 0.02$, df = 176, $P = 0.802$. A test developed by Steiger (1980) was used to test whether partialling out the effects of MHT scores significantly reduced the correlation between the NART and MMSE (as noted, this provides a positive test of the prior ability hypothesis). This procedure tests whether $\rho_{12} = \rho_{34}$ where, in the present case, 1 = the MMSE, 2 = the NART, 3 = the residuals obtained after
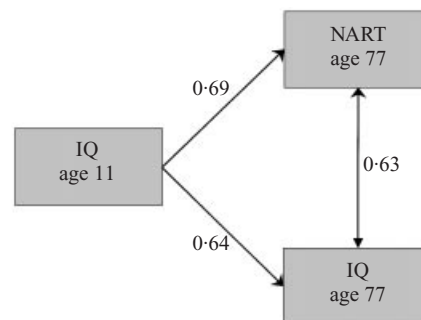

FIG. 2. Relationship between the NART and measures of concurrent and prior IQ.

predicting MMSE scores from the MHT, and 4 = the residuals obtained after predicting NART scores from the MHT (ρ_{34} therefore represents the partial correlation). This procedure revealed that the partial correlation was significantly lower than the raw correlation; $z = 3 \cdot 99$, $P < 0 \cdot 001$.

DISCUSSION

The NART, a brief 50-item oral reading test, administered at age 77, was able to retrospectively account for over 50 % of the variance ($r = 0 \cdot 73$) in psychometric intelligence measured at age 11. This is a remarkable finding as it would be expected that any relationship between these two variables would have been severely attenuated by individual differences in exposure to a myriad of environmental factors in the intervening 66 years. Further, at age 11, the participants would not have been exposed to most of the vocabulary subsequently sampled by the NART.

Restriction of range strongly attenuates correlation coefficients. Therefore, we took the advice of Detterman (1989) and corrected the correlation between the NART and childhood MHT scores to adjust for range restriction in the present sample. This corrected coefficient (0·78) gives a better estimate of the population correlation coefficient and indicates that the NART retrospectively accounts for a substantial 61 % of the variance. The caveat is that use of the correction assumes that the relationship between the variables remains linear beyond the range observed in the sample. However, this assumption is not unreasonable. The results of the hierarchical regression and visual inspection of the scatterplot give no indication that the relationship departs from linearity at the extremes of the range. In addition, Crawford *et al.* (1989) reported that the relationship between the NART and IQ (as measured concurrently by the WAIS-R) did not depart from linearity in a sample with a wide range of IQ (75–140).

Given that the NART is used in clinical practice and research as an index of prior or premorbid intelligence, it is critical to establish retrospective validity rather than rely on existing evidence of concurrent validity. Berry *et al.* (1994) demonstrated that the NART had very adequate retrospective validity over a 3·5 year interval in a healthy elderly sample. The results discussed above demonstrate that very high retrospective validity is also found even with a very long interval.

The data from the subsample readministered the MHT at age 77 (see Fig. 2) allowed us to test competing hypotheses concerning the determinants of NART performance. If the NART is primarily an index of current ability, then it would be expected that its correlation with MHT scores obtained concurrently would be significantly higher than its correlation with scores obtained 66 years earlier. In fact the correlation between NART and IQ aged 11 was higher (0·69), although not significantly so, than the correlation between the NART and concurrent IQ scores (0·64). This pattern of results was obtained despite the fact that the NART and concurrent MHT scores were both subject to the potential effects of any cognitive decline associated with normal ageing, and also to the effect of educational and general socio-economic influences encountered after 11 years of age.

Were the NART primarily an index of current ability, then it would also be expected that its correlation with MHT scores at age 11 would be significantly exceeded by the correlation between MHT scores at ages 77 and 11. If the NART does not tap prior ability there would be no reason to expect otherwise; the correlation between a test administered on two occasions should substantially exceed the correlation between the test and another instrument administered on a second occasion. Again the pattern of correlations was the opposite of that predicted by the hypothesis that the NART is primarily an index of current ability. The correlation between the NART and MHT scores was higher, although not significantly so, than the correlation between the sets of MHT scores. This is also a remarkable finding. The NART samples a single, specific, cognitive ability; the oral reading of irregular words. In contrast MHT scores on the two occasions represent performance over a broad range of cognitive tasks and were obtained using identical items (with the minor exceptions noted earlier).

Examination of the relationship between the NART and impairment sensitive measures in clinical samples has been used to assess whether NART performance is impervious to current level of impairment. The assumption has been

that the presence of a significant correlation indicates that the NART is sensitive to current level of impairment. However, as noted, the presence of such a correlation could also indicate that performance on the impairment sensitive measure (normally a screening test for mental status) is partly determined by an individual's pre-morbid or prior level of intelligence (Crawford, 1989). Both explanations predict that the NART would be significantly correlated with mental status. In the present study these two variables were significantly correlated ($r = 0.25$, $P < 0.001$), despite the restricted variance in MMSE scores. However, when childhood IQ was controlled for, the correlation was no longer significant ($P = 0.802$) and was only marginally different from zero ($r = 0.02$). Furthermore, the partial correlation was significantly lower than the raw correlation ($P < 0.001$). These findings are consistent with the view that shared variance between NART and MMSE reflects the influence of prior ability. They are not consistent with the view that the shared variance reflects individual differences in age associated cognitive decline, as partialling out childhood IQ should have left the relationship unaffected.

The present findings are based on a non-clinical sample and therefore demonstrate only that a relationship between NART and mental status can arise because performance on mental status instruments are not independent of prior intellectual level. They do not rule out the possibility that a relationship could be observed in other samples because of a shared sensitivity to impairment. It may be that observed relationships between NART and mental status reflect the operation of different mechanisms in different populations. The two may be related in the healthy population and in the mildly impaired because of the common influence of prior intelligence, they may become dissociated from each other in more moderately impaired populations (as the covariance attributable to pre-morbid intelligence becomes swamped by the influence of pathology on mental status), and, finally, the two may become associated again in severely impaired populations because NART performance also becomes impaired.

The NART has mainly been used in clinical populations to provide an estimate of prior or pre-morbid ability. However, we believe the NART can play a useful role in the study of cognitive changes associated with normal ageing. Most obviously, it can be used as a means of matching elderly and younger participants on prior ability. This should help to minimize the possibility that any differences observed on the variables of interest are simply a reflection of pre-existing differences in cognitive ability rather than the results of ageing. Furthermore, in the study of the effects of ageing on cognitive performance (and in more complex models involving these variables), prior or pre-morbid ability can be seen as a classic example of a suppressor variable (Darlington, 1990); i.e. the relationship between age and cognitive performance is partly obscured by the influences of individual differences in prior ability that are independent of age. The NART could be used to statistically control for the effects of such differences.

REFERENCES

Barona, A., Reynolds, C. R. & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Consulting and Clinical Psychology* **52**, 885–887.

Berry, D. T. R., Carpenter, G. S., Campbell, D. A., Schmitt, F. A., Helton, K. & Lipke-Molby, T. (1994). The new adult test-revised: accuracy in estimating WAIS-R IQ scores obtained 3·5 years earlier from normal older persons. *Archives of Clinical Neuropsychology* **9**, 239–250.

Blair, J. R. & Spreen, O. (1989). Predicting premorbid IQ: a revision of the National Adult Reading Test. *Clinical Neuropsychologist* **3**, 129–136.

Crawford, J. R. (1989). Estimation of premorbid intelligence: a review of recent developments. In *Developments in Clinical and Experimental Neuropsychology* (ed. J. R. Crawford and D. M. Parker), pp. 55–74. Plenum Press: New York.

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

Crawford, J. R. (1996). Assessment. In *The Blackwell Dictionary of Neuropsychology* (ed. J. G. Beaumont, P. M. Kenealy and M. J. Rogers), pp. 108–116. Blackwell: London.

Crawford, J. R. & Allan, K. M. (1997). Estimating premorbid IQ with demographic variables: regression equations derived from a U.K. sample. *Clinical Neuropsychologist* **11**, 192–197.

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.

Crawford, J. R., Stewart, L. E., Parker, D. M., Besson, J. A. O. & De Lacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: cross-validation and extension. *British Journal of Clinical Psychology* **28**, 267–273.

Crawford, J. R., Venneri, A. and O'Carroll, R. E. (1998). Neuropsychological assessment of the elderly. In *Comprehensive Clinical Psychology, Vol. 7: Clinical Geropsychology* (ed. A. S. Bellack and M. Hersen), pp. 133–169. Pergamon: Oxford, UK.

Darlington, R. B. (1990). *Regression and Linear Models.* McGraw-Hill: New York.

Deary, I. J. (1995). Age-associated memory impairment: a suitable case for treatment. *Ageing and Society* **15**, 393–406.

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.

Detterman, D. K. (1989). The future of intelligence research. *Intelligence* **13**, 199–203.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). 'Mini Mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189–198.

Franzen, M. D., Burgess, E. J. & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology* **12**, 711–738.

Lezak, M. D. (1995). *Neuropsychological Assessment (3rd edn.).* Oxford University Press: New York.

McNemar, Q. (1942). *The Revision of the Stanford-Binet Scale.* Houghton Mifflin: Boston.

Nelson, H. E. (1982). *National Adult Reading Test (NART): Test Manual.* NFER-Nelson: Windsor.

Nelson, H. E. & McKenna, P. (1975). The use of current reading ability in the assessment of dementia. *British Journal of Social and Clinical Psychology* **14**, 259–267.

Nelson, H. E. & O'Connell, A. (1978). Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex* **14**, 234–244.

Nelson, H. E. & Willison, J. (1991). *National Adult Reading Test Manual (2nd edn.).* NFER-Nelson: Windsor.

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory (3rd edn.).* McGraw-Hill: New York.

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

O'Carroll, R. E. (1987). The inter-rater reliability of the National Adult Reading Test (NART): a pilot study. *British Journal of Clinical Psychology* **26**, 229–230.

O'Carroll, R. E. & Gilleard, C. J. (1986). Estimation of premorbid intelligence in dementia. *British Journal of Clinical Psychology* **25**, 157–158.

Patterson, K., Graham, N. & Hodges, J. R. (1994). Reading in dementia: a preserved ability? *Neuropsychology* **8**, 395–407.

Pattie, A. H. (1981). A survey version of the Clifton Assessment Procedures for the Elderly (CAPE). *British Journal of Clinical Psychology* **20**, 173–178.

SCRE (1933). *The Intelligence of Scottish Children: A National Survey of an Age Group.* University of London Press: London.

Starr, J. M., Whalley, L. J., Inch, S. & Shering, P. A. (1992). The quantification of the relative effects of age and NART-predicted IQ on cognitive function in healthy old people. *International Journal of Geriatric Psychiatry* **7**, 153–157.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin* **87**, 245–251.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale.* Psychological Corporation: New York.

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

Williams, E. J. (1959). The comparison of regression variables. *Journals of the Royal Statistical Society (Series B)* **21**, 396–399.