Scandinavian Journal of Psychology, 1993, 34, 305–317

On the relation between demographic variables and neuropsychological test performance

ERIK LYKKE MORTENSEN[1] and ANDERS GADE[2]

[1]*Institute of Preventive Medicine, Kommunehospitalet, Copenhagen, Denmark*
[2]*Psychological Laboratory, University of Copenhagen and Department of Neurology, Rigshospitalet, Copenhagen, Denmark*

Mortensen, E. L. and Gade, A. (1992). On the relation between demographic variables and neuropsychological test performance. *Scandinavian Journal of Psychology, 34*, 305–317.

This paper reports a study of the relation between demographic variables and neuropsychological test performance in a sample of 141 normal subjects. A preliminary analysis demonstrated the importance of age, educational level, sex, and Verbal IQ for the prediction of neuropsychological test performance. The main study is a detailed analysis of the relation between these four variables and neuropsychological test performance. The results are presented in the form of standardized regression coefficients and are discussed in relation to the problem of predicting expected premorbid performance in neuropsychological tests.

Key words: Neuropsychological test performance, premorbid test performance, demographic variables, prediction of premorbid performance.

Erik Lykke Mortensen, Institute of Preventive Medicine, Kommunehospitalet, 1399 Copenhagen Denmark

A fundamental problem in diagnostic neuropsychological testing is the considerable variance in cognitive performance associated with demographic variables such as sex, age, and education. Parsons & Prigatano (1978) discussed the importance of these background variables in neuropsychological research, and subsequently the clinical relevance has been demonstrated and discussed mainly in relation to the Halstead-Reitan neuropsychological battery (e.g. Bornstein, 1985; Heaton *et al.*, 1986; Leckliter & Matarazzo, 1989). For subtests of the WAIS and WAIS-R a relation between test performance and these demographic variables has also been demonstrated (e.g. Heaton *et al.*, 1986); Kaufman *et al.*, 1988), and significant correlations between WAIS performance and performance in the Halstead-Reitan battery have been demonstrated in several studies (Leckliter & Matarazzo, 1989).

Premorbid intelligence may also be considered a demographic variable and performance in selected verbal WAIS subtests may be viewed as an index of premorbid intelligence. This was our point of view in a review of studies of the influence of age, education, and intelligence on neuropsychological test performance (Gade & Mortensen, 1984), and for normal subjects we demonstrated that Verbal IQ explained considerable variance in neuropsychological test performance beyond that explained by age and education.

The relatively high correlations between demographic variables and neuropsychological test performance in normal subjects may be a source of diagnostic error to the extent that test norms do not sufficiently take these correlations into account: Intellectual impairment may be overdiagnosed in old low-education patients and underdiagnosed in young high-education patients. Similarly, variation in premorbid intelligence may be a source of diagnostic errors, and it should be remembered that demographic characteristics may interact in unknown ways.

Gade *et al.* (1984) tentatively suggested the use of multiple regression equations to obtain a relatively precise estimate of the expected premorbid neuropsychological test performance of a person with specified demographic characteristics, and this possibility has been explored in various contexts (Gade *et al.*, 1985*a*; Gade *et al.*, 1985*b*; Gade *et al.*, 1988, Mortensen *et al.*, 1991). Our approach is similar to recent attempts to predict Halstead-Reitan performance from demographic variables (Alekoumbides *et al.*, 1987; Karzmark *et al.*, 1984; Ryan *et al.*, 1987) but is also similar to recent studies using test scores as an index of premorbid ability (Schlosser & Ivison, 1989; Grober & Sliwinski, 1991).

The purpose of this paper is to present a more detailed description of the relation between demographic variables and neuropsychological test performance, and to explore in detail the problems in using demographic characteristics to estimate the expected premorbid test performance. Our neuropsychological battery is the standard battery of Rigshospitalet, Copenhagen. This battery was described by Gade *et al.* (1988) and by Gade & Mortensen (1992) who presented raw score distributions for a sample of 141 normal subjects. The results of this sample are the basic for the present report. Although many detailed results will necessarily be sample specific, we believe that a detailed study and discussion will elucidate general methodological problems in the prediction of expected premorbid test results.

METHOD

Subjects

The sample consisted of 141 brain healthy subjects who were selected according to criteria described by Gade *et al.* (1988) and Gade & Mortensen (1992). The educational level of each subject was calculated as the sum of number of school years (range 7–12) and an occupational training index (range 1–5). Gade & Mortensen (1992) described these variables and the split of the sample into two educational levels and three age levels: For educational level a score in the range 8–12 was considered low and a score in the range 13–17 was considered high. The range of the three age levels was 20–40, 41–60, and 61–83 years, respectively.

Test battery

The sample of 141 subjects was tested with a battery that is described fully in a separate article (Gade & Mortensen, 1992). In this context we concentrate on a Verbal prorated WAIS IQ and the standard neuropsychological test battery from Copenhagen University Hospital (Rigshospitalet). The WAIS subtests were the Information, Similarities, Digit Span, and Vocabulary tests, and the standard neuropsychological battery consisted of Proverb Interpretation, Classification, Paired Associates Learning, List Learning (Buschke), Digit Span, Sentence Repetition, Symbol Digit Modalities Test (SDMT), Trail Making Test, Block Design, and Visual Gestalts. Two scores are generated from the Digit Span and the Trail Making tests, and thus a total of 12 test scores is derived from the battery. A factor analysis of the battery suggested four factors, and consequently the means of the corresponding four groups of tests are analyzed in this paper. Finally, we include the Total Mean of all 12 test scores because this mean is often used as an index of the general level of test performance.

All individual test scores were normalized and standardized to a mean of 50 and a standard deviation of 10 (see Mortensen & Gade (1992) for details). Similarly, all means of test scores were restandardized to a mean of 50 and a standard deviation of 10.

Data analysis

In a preliminary analysis, age and educational level were analyzed as qualitative variables with only 3 and 2 levels. This analysis is relatively uncomplicated, but clearly more precise information on the relation between demographic variables and test performance is needed for the prediction of expected premorbid performance of individual patients. Therefore, age and educational level were analyzed as quantitative independent variables together with Verbal IQ and sex in multiple regression analyses of the relation between demographic variables and neuropsychological test performance.

To analyze all tests in detail would be outside the scope of this paper, and therefore only the means of groups of test scores are analyzed with regression analysis. As these test means are means of highly correlated tests, it should still be possible to demonstrate major types of relations between demographic variables and neuropsychological test performance.

The relations between the four predictor variables and test performance were analyzed in a large number of alternative hierarchical regression analyses. In these analyses the order of the predictor variables was arranged with main effects first, then lower order interactions, and finally higher order interactions. Non-linear or quadratic main effects were tested, and when significant the relevant variables were included in the hierarchical regression analyses. Non-significant variables were discarded from further analysis, and in this way more simple equations were gradually developed.

In both clinical practice and research it may be difficult to obtain an unbiased estimate of the premorbid Verbal IQ of brain damaged patients (Mortensen *et al.*, 1991). Alternative regression equations were therefore obtained for age, sex and educational level without Verbal IQ.

Because the age range is so wide in the sample, means of residuals were calculated separately for the three age subgroups and the low and high educational groups. The sample was also divided into a below average Verbal IQ group ($N = 37$), an average Verbal IQ group ($N = 68$), and an above average Verbal IQ group ($N = 36$), and the means of residuals were calculated for each of these groups.

RESULTS

Sample background

Table 1 shows background data for the three age groups, two educational groups and females and males, respectively.

In the present context it is important to observe that the mean Verbal IQ is significantly lower in the youngest age group than in the two older age groups, and this result should be borne in mind in the interpretation results of analyses without IQ as a covariate. These IQ differences between age groups may be related to both subject selection procedures and IQ test standardization factors (see Gade & Mortensen, 1992).

An increasingly large proportion of the Danish population receives higher education, and it is therefore understandable that an interaction between education and age can be demonstrated for Verbal IQ ($p = 0.01$). Independent studies using a military draft board

Table 1. *Background information in 7 subgroups*

	Age Groups			Education		Sex	
Variable	Young	Middle	Old	Low	High	F	M
No. of Subjects	45	58	38	86	55	63	78
Age							
Mean	31.62	49.59	69.18	50.51	46.98	51.43	47.28
SD	6.26	5.99	6.88	16.00	15.17	15.45	15.79
Educational Level							
Mean	12.20	11.16	11.11	9.74	14.18	11.27	11.64
SD	2.54	2.51	2.64	1.36	1.47	2.76	2.44
School Years							
Mean	9.36	8.52	8.58	7.67	10.56	8.78	8.82
SD	1.64	1.66	1.77	0.99	0.94	1.79	1.66
Occupational Training							
Mean	2.84	2.64	2.53	2.07	3.62	2.49	2.82
SD	1.19	1.13	1.08	0.92	0.73	1.18	1.09
Prorated Verbal IQ							
Mean	104.38	110.41	111.61	101.76	119.84	107.98	109.47
SD	13.54	13.64	19.62	12.86	13.08	14.09	16.85

Table 2. *T score means for three age groups, two educational groups, and sex*

	Age Groups			Education		Sex	
Variable	Young	Middle	Old	Low	High	F	M
Total Mean	54.78	51.79	41.37	45.64	56.65	49.52	50.27
Abstraction Mean	53.56	50.95	43.47	46.52	54.84	48.94	50.44
Proverb Interpretation	50.98	52.10	45.92	47.38	54.35	50.55	49.69
Classification Test	55.51	50.09	43.24	47.07	54.51	47.94	51.62
Verbal Learning Mean	52.82	50.71	44.76	46.95	54.20	51.63	48.28
Paired Associates	53.09	50.53	45.53	46.87	54.96	50.82	49.33
List Learning	51.87	50.46	46.64	48.51	52.47	53.03	47.48
Span Mean	51.31	52.41	45.05	46.62	55.49	48.62	51.26
Digit Span Forward	51.18	51.62	46.11	47.24	54.29	48.59	51.13
Digit Span Backward	51.44	51.87	44.79	47.62	53.27	47.81	51.45
Sentence Repetition	50.27	52.24	46.32	46.35	55.75	49.81	50.18
Visuo-Motor Mean	55.96	51.36	40.92	46.10	56.13	49.51	50.42
SDMT	55.09	51.01	42.39	46.08	55.95	50.92	49.19
Trail Making A	55.24	50.31	43.18	47.27	54.18	50.59	49.46
Trail Making B	54.33	51.28	42.82	46.10	56.02	49.79	50.12
Block Design	54.44	51.64	42.47	46.92	54.98	48.60	51.24
Visual Gestalts, Learning and Retention	55.93	51.44	40.50	46.92	54.60	47.89	51.55

group test have confirmed that intelligence differences between educational levels have decreased in Denmark (Teasdale & Owen, 1992).

Preliminary analysis

Table 2 presents the mean test results for the three age groups, the two educational groups, and females and males, respectively. Only main effects were significant in analysis of variance, and when Verbal IQ was introduced as a covariate, these analyses of covariance also showed significant main effects only, although a few p values for interactions did approach significance.

For the age factor a non-linear relation between age and test performance is suggested by small differences between the youngest and the middle age groups and by the dramatically lower performance of the oldest age group. This low level of performance is even more remarkable because the mean Verbal IQ is highest in the oldest age group (cf. Table 1).

In the analysis of variance the effects of educational level appear stronger than the effects of age in most tests, but in the analysis of covariance the effect of education is in many cases non-significant. This result is understandable because educational level and verbal IQ are highly correlated (Table 1).

The analysis of variance demonstrates that females show better performance than males in the List Learning Test and that males show better performance than females in the Digit Span and Visual Gestalts tests. The male group has a higher Verbal IQ mean than the female group (Table 1), and with Verbal IQ as a covariate, only the sex difference in List Learning remains significant (the analysis of covariance also showed a significant sex difference favoring females for the SDMT).

Table 3. *The correlations of the four predictor variables with the five test means*

Test mean	Sex	Age	Education	VIQ
Abstraction Mean	−0.08	−0.39**	0.53**	0.52**
Span Mean	−0.13	−0.20*	0.49**	0.67**
Verbal Learning Mean	0.17*	−0.31**	0.43**	0.42**
Visuo-motor Means	−0.05	−0.63**	0.55**	0.39**
Total Mean	−0.04	−0.53**	0.63**	0.60**

Note: Correlations marked with * are significant at the five percent level, and correlations marked with ** are significant at least at the one percent level of significance.

Regression analysis

Table 3 shows the linear correlations between the four predictor variables and the five means of subgroups of tests. Age, educational level, and Verbal IQ correlate significantly with all test means. For sex the correlation is significant only in the case of verbal learning (Table 2 shows that females obtain better results than males).

In regression analysis alternative orderings of both main effects and interactions were tried, but in general hierarchical order was only important for Verbal IQ and educational level (the correlation between these two variables was 0.61). For the Verbal Learning Mean, the Visuo-Motor Mean, and the Total Mean educational level did explain significant variance beyond that explained by Verbal IQ, but this was not the case for the Abstraction and Span Means.

Table 4 shows the multiple correlations and the amounts of variance explained for the analyses with and without Verbal IQ included. The final regression equations explain most variance for the Total Mean and the mean of the Visuo-Motor tests. Less variance is explained for the Abstraction and Verbal Learning means, and the mean of the Span tests is a special case because a digit span score is part of the Verbal IQ.

Table 5 shows the final standardized regression coefficients for the analysis with Verbal IQ included. The pattern of the results is complicated in the sense that a unique combination of predictor variables is obtained for each test mean: Verbal IQ and the linear age component are included in all five regression equations, but the quadratic age component is not included

Table 4. *The multiple correlations and the proportions of variance explained for models with and without Verbal IQ included*

	Abstraction Mean	Span Mean	Verbal Learn. Mean	Visuo-Motor Mean	Total Mean
VIQ Included					
Multiple R	0.732	0.782	0.668	0.841	0.914
Proportion of Variance Expl.	0.537	0.611	0.446	0.708	0.836
VIQ not Included					
Multiple R	0.640	0.558	0.577	0.780	0.786
Proportion of Variance Expl.	0.409	0.311	0.332	0.624	0.617

Note: All multiple correlations are highly significant.

Table 5. *Standardized regression coefficients for the final regression equations for the four background variables*

Independent Variables	Abstraction Mean	Span Mean	Verbal Learn. Mean	Visuo-Motor Mean	Total Mean
Sex	—	—	0.259**	—	0.086*
Age/Linear	−0.523**	1.061*	−0.498**	−0.069	−0.078
Age/Quadratic	—	−0.635*	—	−0.623*	−0.588**
Verbal IQ	0.631**	1.138**	0.429	0.411**	0.659**
Education	—	—	−1.468**	0.188*	0.124*
Age/Linear x IQ	—	−0.970*	—	—	—
Education x Verbal IQ	—	—	2.231**	—	—

Note: Coefficients marked with * are significant at the five percent level, and coefficients marked with ** are significant at the one percent level of significance or less.

for the means of the Abstraction and the Verbal Learning tests. Educational level does not explain significant variance beyond that explained by Verbal IQ for the means of the Abstraction and the Span tests, but does so for the other three test means. In the case of the Span Mean there is a significant interaction between age and Verbal IQ, and in the case of the Verbal Learning Mean the interaction between IQ and educational level is significant. Finally, the regression equation includes sex for the Verbal Learning mean, and again this indicates that females perform better than males. The result for the Total Mean also includes a sex factor, but in this case interpretation is simplified by the fact that no interactions are significant.

Table 6 presents a much simpler pattern for the analyses without verbal IQ: For four of the five test means performance is a linear function of educational level and non-linear function of age. For Verbal Learning the results are clearly different because performance is a linear function of age and because the interaction between age and educational level is significant. Furthermore, the regression equation again includes a term to take into account significant sex differences.

It may be added that the interaction between educational level and age for the Total Mean is almost significant in the analysis without IQ included ($p = 0.06$). This interaction is not

Table 6. *Standardized regression coefficients for the final regression equations without Verbal IQ included*

Independent Variables	Abstraction Mean	Span Mean	Verbal Learn. Mean	Visuo-Motor Mean	Total Mean
Sex	—	—	0.263**	—	—
Age/Linear	0.723	1.282**	−1.107**	0.419	0.732*
Age/Quadratic	−1.046**	−1.425**	—	−0.983**	−1.182**
Education	0.489**	0.475**	−0.186	0.464**	0.563**
Age/Linear × Education	—	—	0.956**	—	—

Note: Coefficients marked with * are significant at the five percent level, and coefficients marked with ** are significant at least at the one percent level of significance.

Table 7. *The correlations of IQ with residuals of the final regression equations without IQ. The means of the residuals are also shown for three IQ subgroups with Verbal IQ Means of 89.40, 108.75, and 128.86 respectively*

Test mean	VIQ Correlation	Mean Residuals Low IQ	Average IQ	High IQ
With VIQ				
Abstraction Mean	—	−0.55	0.36	−0.11
Span Mean	—	−0.31	−0.12	0.55
Verbal Learning Mean	—	−0.90	0.81	−0.61
Visuo-motor Means	—	0.63	−0.23	−0.21
Total Mean	—	0.05	−0.04	0.03
Without VIQ				
Abstraction Mean	0.34**	−4.11**	0.72	2.87*
Span Mean	0.46**	−5.72**	0.50	4.94**
Verbal Learning Mean	0.25**	−3.30*	0.58	2.30
Visuo-motor Means	0.33**	−2.10*	−0.05	2.24*
Total Mean	0.52**	−4.37**	0.33	3.86**

Note: Correlations and means marked with * are significant at the five percent level, and coefficients marked with ** are significant at least at the one percent level of significance.

included in the equation presented, but one may want to include the interaction term if the analysis is to be used for predictive purposes.

Because Verbal IQ is so strongly related to test performance, residuals can be expected to correlate with IQ if Verbal IQ is omitted from the regression equations.

That this is indeed so is demonstrated in Table 7, which shows the correlations between Verbal IQ and residuals, and also the mean residuals for three IQ subgroups. For the equations without Verbal IQ all correlations are significant, and the means of residuals deviate significantly from zero in most cases for both the low and high IQ subgroups.

We also calculated the means of residuals for the three age groups and the two educational groups of the preliminary analysis. However, none of the mean residuals deviate significantly from zero, and in this sense the data suggest that errors of prediction are not systematically related to age or education.

DISCUSSION

Our major purpose has been to describe the relation between background variables and performance in neuropsychological tests. Our aim is to develop a practical clinical method to predict the expected premorbid performance level of both individual patients and groups of patients. Therefore, methodological problems such as the question of cross-sectional versus longitudinal research designs are relatively unimportant in the present context. Furthermore, a complete review of relevant literature is beyond the scope of this empirical study. We will, however, compare our findings with related research.

Abstraction mean

The two Abstraction tests are a Proverb Interpretation test and a version of the Kasanin-Hanfmann Classification subtest (Lezak, 1983). Abstraction is often considered an aspect of intelligence and Proverb Interpretation is part of the WAIS Comprehension test. Previously, Gorham (1956) found that Proverb Interpretation test scores vary with educational level, and therefore our results for intelligence and educational level are not unexpected. Indeed, it can be argued that the only unexpected result for the Abstraction mean is the fact that the regression equation with Verbal IQ included showed a simple linear (rather than a non-linear) decline in performance with age. The means in Table 2 and the regression equation without Verbal IQ suggest that age-related decline is only substantial (five T score points or more) above the age of sixty (which is consistent with the findings of Albert *et al.*, 1990). Although the non-linear age component is not significant ($p = 0.16$), it may be included in the regression equation with IQ, but this equation also predicts a substantial decline before the age of 50.

Span mean

The Span Mean consists of Digit Forward, Digit Backward and Sentence Repetition (Spreen & Benton, 1969). Most data are of course available for the Digit Span tests, but in our sample Forward and Backward Digit Span correlate 0.61 and 0.55 with the Sentence Repetition test, and thus the latter test may be interpreted in the context of Digit Span research. Digit Span is one of the verbal WAIS subtests, and therefore the relatively high correlation of the Span Mean with Verbal IQ is understandable (cf. Table 3). A correlation between Span Mean and educational level may be predicted from the correlation with IQ, but still a correlation of 0.49 may seem surprisingly high (Vargo & Black, 1984, has demonstrated a relation between intelligence and the Sentence Repetition test).

Birren & Morrison (1961) similarly found a relatively high correlation between education and WAIS digit span (0.43), but they found essentially no correlation with age when education was kept constant. For our data regression analysis shows a significant non-linear relation with age. The equation with educational level included (Table 6) predicts peak performance in the late forties and the early fifties (cf. Table 2), but the equation with Verbal IQ (Table 5) predicts peak performance in the twenties (this difference can probably be explained by the positive sample correlation between age and Verbal IQ). Finally, it should be noted that other studies have also demonstrated a moderate decline in Memory Span with increasing age (e.g. Botwinick & Storandt, 1974; Craik, 1977; Kaufman *et al.*, 1988).

In our sample a significant interaction was observed between age and Verbal IQ: The difference in Span performance between low and high IQ subjects is much larger for younger than older subjects, and this means that the low IQ subjects showed much less age-related decline in performance than the high IQ subjects. This interaction may be caused by measurement artefacts (floor effects?), but it may also reflect the fact that attention and related functions tend not to decline below a certain level in normal subjects. The latter suggestion may be supported by the fact that Mortensen & Kleven (1992) in a longitudinal study found significant Digit Span decline in the decade between ages 50 and 60, but not in the decade between ages 60 and 70.

Verbal learning mean

An age-related decline in verbal learning ability has been demonstrated by the results from the Baltimore Longitudinal Study (Arenberg & Robertson-Tchabo, 1977) and has been corroborated by many other studies (see for instance the norms of the Wechsler Memory Scale presented by Lezak, 1983). Much of this research has suggested a non-linear relation between age and verbal memory with marked performance decline limited to the later decades of life. This pattern is also suggested by the group means in Table 2, but still the non-linear age component never approached significance in regression analysis.

In hierarchical regression analysis educational level was significant when it was entered after Verbal IQ, and also the interaction between these two factors was clearly significant in predicting verbal learning. Even though Verbal IQ was significant when it was entered after educational level, it proved insignificant in the equation with the interaction term included (Table 5). The interaction can be illustrated by the fact that the correlation between educational level and verbal learning performance is 0.06, 0.23, and 0.58 in the low IQ, average IQ, and high IQ groups defined in Table 7. Thus, in this sample educational level appears to be much more important at relatively high IQ levels. This rather unexpected result may be a sampling artefact and should be confirmed before interpretation is attempted.

The interaction between educational level and age is significant in the analysis without IQ: The subjects with relatively low education show a marked age-related performance decline and the subjects with relatively high education show little or no decline. This is illustrated by the fact that the linear correlation between age and verbal learning performance was −0.43 and −0.07 in the low and high educational groups in Table 1. Analysis with Verbal IQ included shows no interaction between age and IQ nor between age and educational level, which suggests that preservation of verbal learning ability is not necessarily related to general intellectual functioning. The interaction between age and educational level in the analysis without IQ may in fact reflect the correlation between age and IQ in the sample.

Finally, the sex factor is clearly significant for both the equations with and without Verbal IQ which both predict differences of about 5 T score units. Many previous studies have shown that females outperform males in verbal ability, but even so the above figures may seem surprisingly high (see for example Hyde, 1981). Andersen (1976) has previously demonstrated a significant sex difference in the Paired Associate Learning test. In our study the difference was greater in the List Learning test.

Visuo-motor mean

The effect of educational level is significant with Verbal IQ included for the Visuo-Motor Means and this mean also shows a marked age related performance decline. The magnitude of this decline may be explained by the fact that speed is an important aspect of most of the tests combined in this mean (cf. Hertzog, 1989).

The norms presented in Lezak (1983) for the Symbol Digit Modalities Test (SDMT) obviously demonstrate age related decline, and this result has been confirmed for Danish subjects by Nielsen *et al.* (1989). Similarly, numerous studies found age related decline in both parts of the Trail Making Test (e.g. Bornstein, 1985; Long & Klein, 1990), and Nielsen *et al.* (1989) also confirmed this result for Danes. The review by Leckliter & Matarazzo (1989) confirms a negative relation between age and Trail Making performance and a positive relation between performance in this test and education and IQ.

Birren & Morrison (1961) found that WAIS Block Design correlated −0.32 with age and 0.44 with education, and these relations have been confirmed by more recent studies (e.g. Heaton *et al.*, 1986; Kaufman *et al.*, 1988). For the Visual Gestalt test a negative relation between age and performance has been demonstrated by both Andersen (1976) and Nielsen *et al.* (1989).

It may be concluded that the results for the Visuo-Motor Mean in Tables 5 and 6 are consistent with previous studies of individual tests.

Total mean

All 12 test scores are combined to obtain the Total Mean, and a consequence of this procedure is that the four cognitive domains are weighted according to the number of tests within each area (e.g. there are five Visuo-Motor test scores, but only two Abstraction and two Verbal Learning scores). It can be argued that the four cognitive factors represented in the test battery should have equal weighting in the Total Mean by simply computing the average of the four test score means. The reason that we have averaged individual test scores is that our factor analysis of the battery can only be considered tentative and that the four test score means do not discriminate equally well between normal and brain damaged subjects (Mortensen & Gade, 1992). By averaging individual test scores the Total Mean will behave more like the Visuo-Motor Tests that discriminate well between these two groups of subjects.

Actually, the regression equations for the Total Mean (as for the Visuo-Motor Mean) have relatively strong non-linear relations to age (see Tables 5 and 6). Educational level is included in the equations with IQ included (as is the case also for the Visuo-Motor Mean and the Verbal Learning Mean). On the average the equation for the Total Mean predicts that females will score 1.73 better than males in T score units. The large sex difference in the Verbal Learning Tests makes this result understandable.

Table 4 showed that the multiple correlations are highest for the Visuo-Motor Mean and the Total Mean. The size of the multiple correlations may to a certain extent be a function of the specific cognitive functions measured, but it is likely that they also depend on the number of tests averaged in each mean. The more test scores included in a mean the more likely it will behave like a measure of general intellectual ability and therefore be more closely related to Verbal IQ and educational level. Finally, the data of Gade *et al.* (1985*a*) suggest that test means composed of many test scores tend to be more reliable than means composed of few test scores, and the more reliable test means are likely to have higher correlations with background variables.

The issue of Verbal IQ

Table 4 showed that the equations with Verbal IQ included explain more variance than the equations without Verbal IQ. In this sense omission of Verbal IQ from the equations can be considered a case of "specification error". Verbal IQ is significantly correlated with both age and educational level, and therefore the estimates of the effects of age and education will be biased in the equations without IQ (see Berry & Feldman, 1985).

One of the reasons that the equation without Verbal IQ are still of interest is that Verbal IQ may be affected by brain disease. Thus, Mortensen *et al.* (1991) found that the mean observed Verbal IQ in a group of neurological patients with diffuse cerebral atrophy was about 7 IQ points lower than expected from the age and educational level of the group (see also Willshire *et al.*, 1991). Use of equations with Verbal IQ included may therefore result in biased estimates of expected premorbid test scores and consequently in biased estimates of intellectual impairment. This problem can be avoided by using only demographic variables like sex, age, and education, and this procedure may in fact be preferable in some settings. However, Table 7 shows that the residuals of the regression equations without Verbal IQ will be highly correlated with verbal IQ. In practice this means that for a given educational level patients with relatively low IQ will tend to perform at a lower level than expected, and relatively high IQ patients will tend to perform at a higher level than expected. This may of course have serious consequences for the evaluation of individual patients, even though it may be less problematic in group comparisons if no systematic premorbid IQ group differences can be expected.

Table 7 shows the correlations between Verbal IQ and residuals for the complete sample, but the question may be asked whether the correlation is of the same size for educational levels. Formal statistical tests of the interaction between Verbal IQ and the other three predictor variables show a significant interaction between educational level and IQ only for Verbal Learning residuals, and this result is understandable from the results of the analysis with IQ included. In general Table 7 can therefore be interpreted as showing the average bias in evaluating premorbid level from the equations without Verbal IQ.

Wilson *et al.* (1978, 1979) suggested an estimate of premorbid IQ based on age, sex, education, occupation, and race, and the review by Klesges & Troster (1987) shows a continued interest in estimating WAIS IQ from demographic variables. In this study we have used a prorated Verbal IQ to predict the expected performance in neuropsychological tests. Verbal IQ has of course been chosen instead of the Full Scale IQ because it is supposed to be less affected by brain damage than the performance subtests of the WAIS. It should, however, be observed that the verbal WAIS subtests are not necessarily the cognitive tests that are least affected by brain damage, and a possible research strategy is to develop a test score mean that is a composite of test scores that appear least affected by brain damage. Any cognitive test that intellectually impaired patients perform at the level expected from their scores on demographic variables may be of potential value in the prediction of expected premorbid performance in neuropsychological tests.

Range of expected scores

To illustrate the influence of background variables on the expected performance of normal subjects and on the expected premorbid performance of patients with possible brain damage, Table 8 shows the minimum and maximum predicted *T* scores observed in the normal sample. The range of the predicted scores partly depends on the proportion of explained variance, and for the equations with Verbal IQ included it varies from about three times the *T* score standard deviation of 10 to more than four times this standard deviation. For the equations without Verbal IQ included the range is narrower, varying from about two and a half times to almost four times the *T* score standard deviation.

For the Total Mean the minimum expected *T* score was predicted for a 76-year-old female with an educational score of 9 and a Verbal IQ of 75. The maximum expected *T* score was predicted for a 35-year-old male with an educational score of 14 and a Verbal IQ of 135. Table 8 and these examples demonstrate that the use of an overall mean of a normal sample as the basis for clinical decisions may lead to either over- or underdiagnosis of intellectual impairment. In clinical practice and research the predominant type of diagnostic error will of course depend on the sample distribution of demographic variables.

A traditional way of making test norms available to clinicians is in printed tables. However, if such tables were to provide the details actually available, both their construction and use would be very time-consuming. Instead, it is possible to use the regression equations in computer programs that for each test or test mean can evaluate the expected level of performance for any given combination of the demographic variables. A preliminary version of such a program has been developed by the first author, and this program is generally available in Denmark.

Table 8. *Minimum and maximum predicted T scores for the sample of 141 normal subjects (based on final equations with and without Verbal IQ included)*

	VIQ included		VIQ not included	
Test Means	Minimum score	Maximum score	Minimum score	Maximum score
Abstraction Mean	27	65	32	63
Span Mean	29	69	34	61
Verbal Learning Mean	35	66	37	62
Visuo-Motor Means	26	66	28	66
Total Mean	23	67	28	66

Generalization

The basic principle discussed in this paper is that the relation between demographic variables and neuropsychological test performance should be systematically analyzed by multiple regression or any other statistical technique appropriate for a particular set of data. The establishment of relevant equations to predict expected test performance should therefore be a relevant goal for any neuropsychologist. It is, however, important to note that our equations are only appropriate for Danes, and this is indeed one of the reasons that we have decided to present standardized regression coefficients only (the non-standardized regression equations can be obtained from the first author).

It should also be observed that a given set of equations can be expected to be fully valid only within limited time periods because relations between demographic variables and neuropsychological test performance depend on changing life circumstances such as educational opportunities for females and males respectively. Similarly, equations including Verbal IQ or any other test will only remain valid as long as test norms remain valid, and much evidence suggests that IQ norms are only valid for relatively short time periods (Flynn, 1984).

This study was supported by grants from the Danish Medical Research Council to Anders Gade. We thank Agnete Jønsson and Hanne Udesen for assistance in data collection.

REFERENCES

Albert, M. S., Wolfe, J. & Lafleche, G. (1990). Differences in abstraction ability with age. *Psychology and Aging*, *5*, 94–100.

Alekoumbides, A., Charter, R. A., Adkins, T. G. & Seacat, G. F. (1987). The diagnosis of brain damage by the WAIS, WMS, and Reitan Battery utilizing standardized scores corrected for age and education. *International Journal of Clinical Neuropsychology*, *9*, 11–28.

Andersen, R. (1976). Verbal and visuo-spatial memory. Two clinical tests administered to a group of normal subjects. *Scandinavian Journal of Psychology*, *17*, 198–204.

Arenberg, D. & Robertson-Tchabo, E. A. (1977). Learning and Aging. In J. E. Birren & K. W. Schai (Eds). *Handbook of the psychology of aging*, pp. 421–449. New York: Van Nostrand.

Berry, W. D. & Feldman, S. (1985). *Multiple regression in practice*. London: Sage Publications.

Birren, J. E. & Morrison, D. F. (1961). Analysis of the WAIS subtests in relation to age and education. *Journal of Gerontology*, *16*, 363–369.

Bornstein, R. A. (1985). Normative data on selected neuropsychological measures from a nonclinical sample. *Journal of Clinical Psychology*, *41*, 651–659.

Botwinick, J. & Storandt, M. (1974). *Memory, related functions, and age*. Springfield: Charles C. Thomas.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schai (Eds), *Handbook of the psychology of aging*, pp. 384–420. New York: Van Nostrand.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.

Gade, A. & Mortensen, E. L. (1984). The influence of age, education, and intelligence on neuropsychological test performance. Paper presented at the 3rd Nordic Conference in Behavioral Toxicoloy, Århus, 1984.

Gade, A. & Mortensen, E. L. (1992). Standardization of a neuropsychological test battery: Methods and basic results. (in preparation).

Gade, A., Mortensen, E. L. & Bruhn, P. (1988). "Chronic painter's syndrome". A reanalysis of psychological test data in a group of diagnosed cases, based on comparisons with matched controls. *Acta Neurologica Scandinavica*, *77*, 293–306.

Gade, A., Mortensen, E. L. & Udesen, H. (1984). The pattern of intellectual impairment in toxic encephalopathy. International conference on organic solvent toxicity, Stockholm, 1984.

Gade, A., Mortensen, E. L., Udesen, H. & Bruhn, P. (1985*a*) On the importance of control data and background variables in the evaluation of neuropsychological aspects of brain functioning. In: *Neurobehavioural methods in occupational and environmental health*. Environ. Health: Doc 3. Copenhagen: WHO, 91–96.

Gade, A., Mortensen, E. L., Udesen, H. & Jønsson, A. (1985*b*). Predictors of cognitive performance: age, education, and intelligence. Presented at the 8th INS European conference, Copenhagen, 1985. *Journal of Clinical and Experimental Neuropsychology, 7,* 144, (abstract).

Gorham, D. R. (1956). A proverb test for clinical and experimental use. *Psychological Reports, Monograph Supplement, No. 1, 2,* 1–12.

Grober, E. & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology, 13,* 933–949.

Heaton, R. K., Grant, I. & Matthews, C. G. (1986). Difference in neuropsychological test performance associated with age, education, and sex. In I. Grant & K. M. Adams (Eds), *Neuropsychological assessment of neuropsychiatric disorders,* pp. 100–120. New York & Oxford: Oxford University Press.

Hess, G. (1974). *Wais anvendt på 698 50-årige.* København: Akademisk Forlag.

Hertzog, C. (1989). Influences of cognitive slowing on age differences in intelligence. *Developmental Psychology, 25,* 636–651.

Hyde, J. S. (1981). How large are cognitive gender differences? *American Psychologist, 36,* 892–901.

Karzmark, P., Heaton, R. K., Grant, I. & Matthews, C. G. (1984). Use of demographic variables to predict overall level of performance on the Halstead-Reitan Battery. *Journal of Consulting and Clinical Psychology, 52,* 663–665.

Kaufman, A. S., McLean, J. E. & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS-R subtests. *Journal of Clinical Psychology, 44,* 231–248, 1988.

Klesges, R. C. & Troster, A. C. (1987). A review of premorbid indices of intellectual and neuropsychological functioning: What have we learned in the past five years? *International Journal of Clinical Neuropsychology, 9,* 1–11.

Leckliter, I. N. & Matarazzo, J. D. (1989). The influence of age, education, IQ, gender, and alcohol abuse on Halstead-Reitan neuropsychological test battery performance. *Journal of Clinical Psychology, 45,* 484–511.

Lezak, M. D. (1983). *Neuropsychological assessment.* New York & Oxford: Oxford University Press.

Long, C. J. & Klein, K. (1990). Decision strategies in neuropsychology II: Determination of age effects on neuropsychological performance. *Archives of Clinical Neuropsychology, 5,* 335–345.

Mortensen, E. L. & Gade, A. (1992). Linear versus normalized T scores as standardized neuropsychological test scores. *Scandinavian Journal of Psychology, 33,* 230–237.

Mortensen, E. L., Gade, A. & Reinisch, J. M. (1991). A critical note on Lezak's 'Best Performance Method' in clinical neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 13,* 361–371.

Mortensen, E. L. & Kleven, M. (1992). A WAIS longitudinal study of cognitive development during the life span from ages 50 to 70. (Submitted).

Nielsen, H., Knudsen, L. & Daugbjerg, O. (1989). Normative data for eight neuropsychological tests based on a Danish sample. *Scandinavian Journal of Psychology, 30,* 37–45.

Parson, O. A. & Prigatano, G. P. (1978). Methodological considerations in clinical neuropsychological research. *Journal of Consulting and Clinical Psychology, 46,* 608–619.

Ryan, C. M., Morrow, L. A., Bromet, E. J. & Parkinson, D. K. (1987). Assessment of neuropsychological dysfunction in the workplace: Normative data from the Pittsburgh Occupational Exposures Test Battery. *Journal of Clinical and Experimental Neuropsychology, 9,* 665–679.

Schlosser, D. & Ivison, D. (1989). Assessing memory deterioration with the Wechsler Memory Scale, the National Reading Test, and the Schonell Graded Word Reading Test. *Journal of Clinical and Experimental Neuropsychology, 11,* 785–792.

Spreen, O. & Benton, A. L. (1969). *Neurosensory Center Comprehensive Examination for Aphasia.* Victoria, B. C.: Neuropsychological Laboratory Dept. of Psychology, University of Victoria.

Teasdale, T. W. & Owen, D. R. (1992). 30-year secular trends in the cognitive abilities of Danish male school-leavers at a high educational level. (Submitted).

Vargo, M. E. & Black, F. W. (1984). Normative data for the Spreen-Benton Sentence Repetition Test: Its relationship to age, intelligence, and memory. *Cortex, 20,* 585–590.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence.* Baltimore: Williams & Wilkins Company.

Willshire, D., Kinsella, G. & Prior, M. (1991). Estimating WAIS-R IQ from the National Adult Reading Test: A cross-validation. *Journal of Clinical and Experimental Neuropsychology, 13,* 204–216.

Wilson, R. S., Rosenbaum, G. & Brown, G. (1979) The problem of premorbid intelligence in neuropsychological assessment. *Journal of Clinical Psychology, 1,* 49–53.

Wilson, R. S., Rosenbaum, G., Brown, G., Rourke, D., Whitman, D. & Grisell, J. (1978). An index of premorbid intelligence. *Journal of Consulting and Clinical Psychology, 46,* 1554–1555.

Received 30 March 1992