The Irish Journal of Education 1984, *xviii* 1 pp 3 22

THE VALIDITY OF THE DAT AS A MEASURE OF SCHOLASTIC APTITUDE IN IRISH POST–PRIMARY SCHOOLS

Michael O Martin and Bernard O'Rourke*
Educational Research Centre
St Patrick's College, Dublin

This study examines the criterion-related validity of the Differential Aptitude Tests (DAT) (Verbal Reasoning and Numerical Ability) using as criterion variables performance on objective tests of achievement (Drumcondra Attainment Tests) and on public examinations – the Intermediate Certificate Examination (ICE) and the Leaving Certificate Examination (LCE) The aptitude and objective achievement tests were administered to students in 39 post primary schools in Ireland at the beginning of their second year At the beginning of the students' third year the same aptitude test and a higher-level objective achievement test were administered At the end of that year, the students sat for the ICE At the end of their fourth year the students again took the DAT and at the end of their fifth year they sat for the LCE The DAT was found to correlate highest with scores on the objective tests of achievement next highest with performance on the ICE and lowest with LCE performance Correlations with indices of overall LCE performance were higher than correlations with performance on individual subjects DAT scores were found to be stable over time Gender differences favouring boys existed on both the Verbal Reasoning and Numerical Ability tests

A growth in the use of objective tests in the Irish educational system in recent years has not been matched by the publication of empirical studies of the validity of the tests being used While this situation has important implications for users of test information (1), the absence of validation data has the additional effect of limiting greatly the effectiveness of debate among those contemplating using test information to make decisions of various kinds The ongoing debate on third level selection procedures, for example, has been hampered by the fact that little empirical evidence is available relating to the criterion related (predictive) validity of objective tests of scholastic aptitude Prompted by this situation the study

* Requests for off prints should be sent to Bernard O Rourke, Educational Research Centre, St Patrick's College Dublin 9

reported in this paper was undertaken to provide validation data on a popular measure of scholastic aptitude

More specifically, the objective of the study was to relate the performance of students on two tests from the battery of Differential Aptitude Tests (DAT) – Verbal Reasoning (VR) and Numerical Ability (NA) – to their performance on other second level assessment instruments The DAT battery, which was developed in the United States (3), was adapted and standardized for use in Ireland (8) and has been used extensively in post primary schools While scores on the separate VR and NA tests of the DAT were considered, the study focused primarily on the combined score for the two tests, a VR-NA composite score The practice of using the combined score on the VR and NA tests as a measure of scholastic aptitude was encouraged by the authors of the American DAT test manual who viewed the combined VR NA test score as tapping 'the same general area as most tests of scholastic aptitude' (3, p 6), norms were presented (3, Ch 3) to facilitate use of the VR NA composite score as a measure of scholastic aptitude The extensive validity data (based on over 200 groups of second level pupils) reported in the American test manual (3, Ch 5) suggest that the VR NA composite score was a better predictor of course performance than scores on any of the eight individual DAT tests, although the NA test score was the best individual predictor in most cases reported, even for English and Literature

Data are presented in this paper relating specifically to criterion related validity A test has criterion related validity if scores on the test are correlated with scores on a criterion or outcome measure that we are interested in estimating (6) In evaluating the validity of the DAT as a measure of scholastic aptitude, the evidence of criterion related validity is taken as indicative of construct validity (5), a test is regarded as construct valid if it measures the concept or construct – in this study, scholastic aptitude – that it purports or is intended to measure In a general sense, construct validity is the more fundamental (27), for a test of scholastic aptitude to be judged construct valid, it should exhibit criterion related validity against nominated criteria (e g , performance on examinations) The data presented in this paper relate directly to the criterion related validity of the test, and indirectly therefore, to construct validity Criterion related validity is evaluated against two sets of criterion variables, performance on a series of objective tests of achievement and on two public examinations, the Intermediate Certificate Examination (ICE) and the Leaving Certificate Examination (LCE) A consideration of criterion

related validity is preceded by the presentation of data on the stability of test scores over time and on gender differences

METHOD

Sample

The data used in this study were drawn from a database established on a stratified random sample of Irish post primary schools in the 1970s All second level schools in the Republic of Ireland were stratified by type (secondary, vocational, comprehensive), by gender of student served (male, female, mixed), and by size – large (greater than 350 students) or small (equal to or less than 350 students) Within strata, schools were selected randomly Table 1 presents data on the number of schools for which test scores and examination results were available The students who partici pated in the present study were in their second year in post primary school in autumn 1974

TABLE 1

NUMBER OF SCHOOLS IN SAMPLE CLASSIFIED BY TYPE SIZE AND GENDER OF STUDENT SERVED

	Secondary		Vocational		Comprehensive
	Large	Small	Large	Small	
Male	7	4	–	–	–
Female	6	4	–	–	–
Mixed	–	5	2	8	1
Total	13	13	2	8	1
	26		10		1

Measure of scholastic aptitude

Differential Aptitude Tests, Form T Verbal Reasoning and Numerical Ability tests The Verbal Reasoning test was designed as 'a measure of ability to understand concepts framed in words' (3, p 6) Fifty items make up the test all are analogies Skills of abstraction or generalization and constructive thinking rather than verbal fluency or vocabulary recogni tion are assessed The 40 item Numerical Ability test assesses skills in reasoning with numbers and manipulation of numerical relationships

Separate VR and NA scores were used in analyses as well as a composite VR NA score, which was simply the sum of the two individual test scores (i e , VR + NA)

Criterion variables

Three independent sets of criterion variables were used All were measures of achievement The first set was composed of scores on a series of objective achievement tests, the second was composed of scores on the ICE and the third of scores on the LCE

Objective achievement tests Drumcondra Mathematics Test, Level V, Form A (9), Drumcondra Mathematics Test, Level VI, Form A (10) A Mathematics score was computed for each test by summing the Computation, Concepts, and Problem Solving subtest scores
Drumcondra Irish Test, Level V, Form A (9), Drumcondra Irish Test, Level VI, Form A (10) An Irish Reading score was computed by summing the Reading Vocabulary and Reading Comprehension subtest scores Drumcondra English Test, Level V, Form A (9), Drumcondra English Test, Level VI, Form A (10) Two of the five English subtests scores were used to compute scores for English Reading (Reading Vocabulary and Reading Comprehension) and two to compute scores for English Language (Capitalization/Punctuation and Usage/Parts of Speech)

Intermediate Certificate Examination The ICE may be taken by students who have followed an approved course of not less than three years' duration in a recognised school (18) Normally, students are about 15 years of age when they take the examination In 1975 76, courses leading to examinations were offered in 21 subjects For three of these subjects (Irish, English, and Mathematics), an option on Higher and Lower level courses and examinations was offered As well as these three subjects, History, Geography, Science, French, and Commerce were the next most popular subject choices in 1975 76 (16) The majority of students sit for examinations in between six and eight subjects Since the Department of Education reports the results of public examinations in the form of letter grades rather than numerical scores, some numerical trans formation was required to facilitate statistical analysis For our study, the ICE letter grades were assigned numerical values as follows A = 7, B = 6, C = 5, D = 4, E = 3, F = 2, and G ('no grade') = 1

Leaving Certificate Examination The LCE is normally taken two years after the ICE, when students are about 17 years of age (18) In

1977 78, courses leading to examinations were offered in 31 subjects An option on separate Higher and Ordinary level courses and examinations was offered in most subjects The pattern of subject popularity was broadly similar to that of the ICE, although there were marked gender differences for some subjects (e g , Biology, Physics, Chemistry, ,and common level technical subjects such as Engineering Workshop Theory and Practice, Technical Drawing, and Building Construction) (17) The majority of students sit for examinations in six or seven subjects The numerical scaling of LCE letter grades used the same transformation procedure as described for the ICE grades Correlations between DAT scores and LCE subject grades are presented for each subject for both examination levels (Higher and Ordinary) As well as using students' grades for individual subjects in the LCE, several indices of overall performance on the examination were calculated Procedures for computing indices of overall LCE performance are commonly known as 'points schemes', overall performance being the arithmetic sum of points earned from a maximum of six subjects While there are several such systems in use, they vary only slightly in design (14, 31) and performance (11, 28) One of the most common scoring systems makes use of the following grade to points transformations A = 5 points, B = 4, C = 3, and D = 2 for grades on Higher level examinations and A = 2 and B = 1 for grades on Ordinary level examinations, all other grades at both levels being assigned 0 points, this is referred to as the 543221 scheme Other corres ponding schemes which have been the subject of empirical study are 742110, 853210, 842110, and 632110 (28) All these points schemes were used as criterion variables in the present study, as was an additional scheme designed to give credit for the full range of possible grades (i e , including grades E and F on Higher level examinations and grades C to F on Ordinary level exammations) The ordering of the grade weights (points) in the extended scheme was derived from an examination of the ordering of the mean DAT score equivalents of the letter grades for the most common LCE subjects at both Higher and Ordinary levels The weights for the Higher level grades were A = 11 points, B = 10, C = 8, D = 7, E = 4, and F = 2, the weights for the Ordinary level grades were A = 9 points, B = 6, C = 5, D = 3, E = 1, and F = 0 One would have expected correlations between scores based on this scheme and DAT scores to have been higher than correlations between scores based on the other schemes and DAT scores since the analyses were based on the complete ability range of those in the sample who sat for the LCE

Procedure

The VR and NA tests of the DAT were administered to students at the beginning of their second year in post primary school in autumn 1974 The tests were readministered to the same students in autumn 1975, when students were commencing their third year, and again in summer 1977, at the end of their fourth year

The Drumcondra Attainment Tests, Level V, in Irish, English, and Mathematics were also administered in autumn 1974 (beginning of second year) A year later, in autumn 1975 (beginnmg of third year), Level VI of the same test series was administered

The students sat for the ICE in summer 1976 at the end of their third year and for the LCE in summer 1978 at the end of their fifth year Examination results were obtained from the Department of Education

Both the DAT and the Drumcondra Attainment Tests were administered in the students' own school, usually by a guidance teacher, and test answer sheets were returned to the Educational Research Centre for scoring

RESULTS

Stability over time and gender differences

Intercorrelations for three administrations of the DAT over a two and a half year period are presented in Table 2, coefficients are presented for the VR and NA tests and for the VR NA composite (Relevant distributional statistics are presented in Table 3) The lowest correlation (74) is between the autumn 1974 and summer 1977 administrations of the NA test (a period of some 30 months), while the highest correlation (87) is between the autumn 1974 and autumn 1975 VR NA composite scores

On all three testing occasions, boys scored higher than girls on the VR NA composite measure (Table 3) The column headed Eta^2 in Table 3 shows the percentage of the total variance in DAT scores attributable to gender differences, this statistic provides a measure of the size of those differences Most of the gender effect on VR NA composite scores was due to differences on the NA test, although the difference between boys and girls on VR was statistically significant also There is evidence from these data that the difference between boys and girls increased with age, however, this finding should be interpreted in the context of the fall off

TABLE 2

CORRELATIONS BETWEEN DAT SCORES OVER A THREE YEAR PERIOD

	Autumn 1974	Autumn 1975	Summer 1977
Verbal Reasoning			
Autumn 1974	–	85	81
Autumn 1975	–	–	84
Numerical Ability			
Autumn 1974	–	79	74
Autumn 1975	–	–	80
VR NA			
Autumn 1974	–	87	81
Autumn 1975	–	–	84

in numbers tested in the final year of data collection

When computed separately for boys and girls, intercorrelations between repeated administrations of the DAT comparable to those presented in Table 2 showed no systematic gender differences in stability of test scores over time

DAT and objective test performance

Correlations between the DAT scores and Drumcondra Attainment Test scores for the autumn 1974 and autumn 1975 administrations are presented in Table 4 as evidence of criterion related validity of the concurrent type Similar patterns in the magnitude of the coefficients were discernable in the correlation matrix for each year In general, the coefficients based on the autumn 1975 administration tended to be slightly larger

The coefficients for criterion related validity of the predictive type (autumn 1974 DAT and autumn 1975 attainment tests) (Table 5) were comparable to the concurrent coefficients for autumn 1974, but invariably smaller than the concurrent coefficients for autumn 1975

TABLE 3

MEAN DAT SCORES OVER A THREE YEAR PERIOD BY GENDER

	Total sample			Boys			Girls				
	M	SD	N	M	SD	N	M	SD	N	Eta^2	p
Verbal Reasoning											
Autumn 1974	19 6	8 5	1469	20 6	8 7	789	18 5	8 2	671	1 5	01
Autumn 1975	23 8	10 2	2022	25 5	10 2	1093	22 0	9 7	927	2 9	01
Summer 1977	32 4	9 6	608	34 5	8 9	235	31 1	9 7	373	3 0	01
Numerical Ability											
Autumn 1974	14 7	7 0	1468	16 1	7 5	795	13 1	6 0	673	4 6	01
Autumn 1975	18 1	7 8	1992	19 5	7 9	1089	16 5	7 3	903	3 6	01
Summer 1977	23 6	8 7	605	26 4	8 2	229	21 9	8 5	376	6 5	01
VR NA											
Autumn 1974	34 3	14 0	1462	36 6	14 5	793	31 5	12 8	669	3 3	01
Autumn 1975	42 0	16 6	1978	44 9	16 7	1086	38 5	15 7	892	3 8	01
Summer 1977	55 9	16 4	594	60 9	15 2	228	52 9	16 3	366	5 6	01

TABLE 4

CORRELATIONS BETWEEN DAT SCORES AND SCORES ON THE DRUMCONDRA ATTAINMENT TESTS LEVELS V AND VI

Measure	1	2	3	4	5	6	7
1 Verbal Reasoning		69	94	73	63	78	72
2 Numerical Ability	61		89	84	63	61	66
3 VR NA	92	86		85	69	76	75
4 Mathematics	67	80	81		65	66	68
5 Irish Reading	52	56	60	65		62	68
6 English Reading	70	53	69	63	64		72
7 English Language	68	61	72	71	67	73	

Note Coefficients below the diagonal refer to the autumn 1974 administrations of the DAT and Level V attainment tests while those above the diagonal refer to the autumn 1975 administrations of the DAT and Level VI attainment tests

TABLE 5

CORRELATIONS BETWEEN AUTUMN 1974 DAT SCORES AND SCORES ON THE DRUMCONDRA ATTAINMENT TESTS LEVEL VI (AUTUMN 1975)

DAT	Mathematics	Irish Reading	English Reading	English Language
Verbal Reasoning	66	53	75	65
Numerical Ability	77	54	56	58
VR NA	79	59	74	69

Compared with the correlations between the respective VR and NA tests and the attainment tests, the correlations between the VR NA composite and the attainment tests were marginally, though not consistently, better (Tables 4 and 5) The correlations between the VR NA composite and the English Reading attainment test scores were marginally lower than the correlations between the individual VR test and the English Reading attainment test scores, this occurred for both sets of concurrent coefficients as well as for the predictive coefficients

Across the four subject areas examined, the DAT VR and NA tests correlated differentially with the attainment test scores The correlations between the DAT NA scores and the Mathematics attainment scores were the highest In general, the DAT test scores correlated less well with the Irish attainment test scores than with the scores on the other attainment measures

DAT and ICE performance

Criterion related validity of the DAT against performance on the 1976 ICE was examined using the DAT scores from the autumn 1975 administration Correlations between the autumn 1975 DAT scores and the summer 1976 ICE results were, in general, substantial (Table 6) For the five most popular subjects (i e , History, Geography, Science, French, and Commerce) with common level examinations (as distinct from separate Higher- and Lower level examinations), the median correlation with VR NA was 56 Other less frequently taken subjects showed significant correlations too, for example, Latin (51), Home Economics (50), and Mechanical Drawing (49)

Since the provision of Higher and Lower level (in effect, streamed) courses for Irish, English, and Mathematics in the ICE inevitably leads to a restriction in the ability range of students taking a particular course (and sitting for the examination) and since correlations based on truncated distributions are often attenuated, it might be expected that the correlations between DAT scores and results on these subjects would be less than for subjects in which a single examination was taken by all students This seems to have happened for Irish and English, but not for Mathematics, the correlations between the VR NA composite score and the Higher and Lower level Mathematics examination results were 50 and 60 respectively (Table 6)

To resolve the problem of truncated distributions, transformations to a common numerical scale of both Higher and Lower level examination letter grades were made Regression analyses suggested that the relationship between the (quantified) letter grades and VR NA could be safely assumed to be linear for each subject at both levels and that the estimated slopes of the regression lines for Higher and Lower level examination results were very similar for each subject The results of these analyses confirmed that a satisfactory transformation to a single scale of results from both examinations was possible

TABLE 6

CORRELATIONS BETWEEN
AUTUMN 1975 DAT SCORES AND 1976 ICE RESULTS

ICE subject	N	DAT scores Verbal Reasoning	Numerical Ability	VR NA
Irish – Higher	501	18	27	24
Irish – Lower	1 174	32	32	35
English – Higher	818	38	32	40
English – Lower	915	34	31	40
Mathematics – Higher	453	40	48	50
Mathematics – Lower	1 278	50	57	60
History	1 689	56	48	57
Geography	1 687	58	55	62
Science	1 155	47	48	52
Music & Musicianship A*	259	35	32	38
Music & Musicianship B	6	–	–	–
French	1 151	51	48	54
Latin	338	47	43	51
Greek	20	15	47	36
Hebrew	0	–	–	–
German	45	28	54	46
Italian	0	–	–	–
Spanish	32	19	39	37
Home Economics	618	47	42	50
Commerce	1 053	48	56	56
Art	738	30	25	30
Woodwork	403	20	16	22
Metalwork	178	41	30	41
Mechanical Drawing	368	42	47	49

Note Correlations based on 15 students or less were not computed

*Syllabus A in Music and Musicianship contains written and aural components while Syllabus B contains in addition a practical component

The transformation was effected by replacing each letter grade for both examination levels with the mean VR NA score of all students receiving that grade This procedure, known as criterion scaling (2), is designed to yield a single scale for results from both examination levels such that the linear relationship between the derived scale and the criterion (VR NA) is maximized Following this procedure, correlation coefficients were computed between VR NA scores and the scaled scores In each case, the

coefficient was considerably greater than when computed separately for Higher and Lower level results For example, while the obtained correlation coefficients between VR NA scores and Irish scores were 24 and 35 for the Higher and Lower level examinations respectively, the correlation between VR-NA scores and the scaled Irish scores was 58 The coefficient was 63 for the scaled English score (compared with 40 for both levels separately) and 72 for the scaled Mathematics score (compared with 50 and 60 for the Higher and Lower level results respectively)

DAT and LCE performance

Correlations between the autumn 1975 DAT scores and summer 1978 LCE grades for Higher and Ordinary courses are presented in Table 7 In most cases, as expected, the correlations were appreciably less than their counterparts for the ICE results Only in the case of Irish was the size of the correlation maintained

To achieve an overall comparison between ICE and LCE correlations, the criterion scaling procedure was applied to LCE Irish, English, and Mathematics grades The resulting scores were correlated with the VR NA score In general, the correlations held their values well, given that a further two years had elapsed before the LCE was taken and that approximately 40% of the students were lost to the sample over this period (Table 8)

The correlations between the DAT VR NA composite and the measures of overall LCE performance are presented in Table 9 In addition to the DAT scores from the autumn 1975 administration, data are also included in Table 9 for the summer 1977 administration of the DAT Given the level of attrition in the sample over the period, it was decided not to make use of the summer 1977 DAT scores in analyses of individual LCE subject performance For the full sample, the correlations between the autumn 1975 DAT VR-NA composite and the measures of overall LCE performance were almost identical to those obtained when the summer 1977 DAT VR NA composite scores were used In both cases, the median correlation was 54 over the five schemes examined by Moran and Crowley (28) Use of the points scheme designed to give credit to the full range of grades on both Higher- and Ordinary-level examinations (viz , the extended scheme) improved the correlation between the DAT VR NA composite and the points scores, again to almost the same degree for both administrations of the DAT

Given the observed gender differences on both the component tests, the

TABLE 7

CORRELATIONS BETWEEN AUTUMN 1975 DAT SCORES AND 1978 ORDINARY AND HIGHER LEVEL LCE RESULTS

	Examination level							
	Ordinary				Higher			
LCE subject	N	VR	NA	VR NA	N	VR	NA	VR NA
Irish	723	27	28	30	314	24	21	26
English	643	27	24	30	430	13	12	15
Mathematics	952	46	56	57	97	22	21	24
Latin	30	37	37	42	93	36	08	26
Greek	0	–	–	–	5	–	–	–
Hebrew	0	–	–	–	0	–	–	–
French	324	28	26	31	262	32	23	33
German	16	55	50	59	18	28	12	21
Italian	0	–	–	–	0	–	–	–
Spanish	9	–	–	–	8	–	–	–
History	209	38	25	38	210	24	26	29
Geography	361	29	23	30	334	22	16	23
Applied Mathematics	4	–	–	–	21	19	46	39
Physics	91	07	13	11	103	40	33	43
Chemistry	80	17	33	28	95	16	27	31
Physics & Chemistry	14	–	–	–	24	−08	15	02
Biology	203	30	26	32	169	37	32	41
Agricultural Science	22	03	12	08	26	05	20	13
Agricultural Economics	3	–	–	–	1	–	–	–
Mechanics	0	–	–	–	0	–	–	–
Home Economics (SS)*	57	31	05	22	91	34	13	30
Home Economics (General)	74	18	19	23	78	46	36	49
Accounting	114	20	25	28	131	24	27	28
Business Organization	186	18	06	13	137	21	23	26
Economics	144	16	−03	07	70	00	02	00
Economic History	6	–	–	–	12	–	–	–
Art (including Crafts)	99	39	25	36	76	23	37	34
Music & Musicianship A†	0	–	–	–	7	–	–	–
Music & Musicianship B	8	–	–	–	5	–	–	–
Engineering Workshop	1	–	–	–	52	13	19	19
Technical Drawing	1	–	–	–	108	21	21	23
Building Construction	2	–	–	–	43	−05	15	06

Note Correlations based on 15 students or less were not computed

* The abbreviation SS refers to the Social and Scientific Home Economics course

† Syllabus A and Syllabus B in Music and Musicianship both include written and aural components but differ in the requirements for the instrumental component

TABLE 8

CORRELATIONS BETWEEN AUTUMN 1975 DAT VR NA SCORES AND 1976 ICE AND 1978 LCE CRITERION SCALED IRISH ENGLISH AND MATHEMATICS RESULTS

Subject	ICE N	ICE r	LCE N	LCE r
Irish				
Scaled	1 675	58	1 037	53
Higher	501	24	314	26
Lower/Ordinary	1,174	35	723	30
English				
Scaled	1 733	63	1 073	54
Higher	818	40	430	15
Lower/Ordinary	915	40	643	30
Mathematics				
Scaled	1 731	72	1 049	65
Higher	453	50	97	24
Lower/Ordinary	1 278	60	952	57

TABLE 9

CORRLLATIONS BETWELN DAT VR NA SCORES (AUTUMN 1975 AND SUMMER 1977) AND 1978 LCE BASED POINTS SCHEMES

Points scheme	Total sample N	Total sample r	Boys N	Boys r	Girls N	Girls r
			DAT AUTUMN 1975			
543221	846	55	445	55	401	58
742110	779	53	415	52	364	55
853210	779	55	415	55	364	56
842110	779	53	415	51	364	55
632110	779	54	415	53	364	56
Extended	1 075	64	530	61	545	66
			DAT SUMMLR 1977			
543221	382	54	139	53	243	55
742110	351	53	133	50	218	56
853210	351	54	133	51	218	57
842110	351	52	133	49	218	56
632110	351	54	133	50	218	57
Extended	487	63	167	56	320	65

correlations between the VR NA composite and measures of overall LCE performance were computed separately for boys and girls (Table 9) The results suggest that the DAT VR NA composite was marginally better at predicting the overall LCE performance of girls than of boys It is noteworthy that no gender differences were observed in scores on any of the points schemes with one exception on the extended scheme, boys (M = 36 0, SD = 14 1) scored higher than girls (M = 33 1, SD = 13 7) (F = 17 17, df = 1, 1656, $p < 0001$)

DISCUSSION

Information on stability of DAT test scores of the kind presented in this paper complements the test statistics reported traditionally in test manuals, which are usually based on a single test administration Our findings indicate a high level of test stability (equivalent to test retest reliability) for a period of two and a half years, suggesting that inter individual differences remain relatively consistent during the middle post-primary years Yet, there is an obvious increase in the raw scores on the test As Hilton points out, high relational stability indices 'do not mean that students do not change, they mean that the students' relative standing on the measures in question remains very nearly the same from one grade level to the next' (13, p 10)

The observed gender differences in DAT scores parallel those found elsewhere for mathematical ability but not those for verbal ability Gender differences in standardized test performance in mathematics (aptitude and achievement) are normally found to favour boys (at least in the post primary age range), the pattern being generalizable over countries, time, and instruments (21, 23, 29, 30) The differences we found on the verbal reasonmg test in favour of boys run contrary to expectations induced by the balance of research on gender related differences in measured verbal abilities which points to 'female superiority on verbal tasks being one of the more solidly established generalizations in the field of sex differences' (23, p 25) It is of note that the American DAT test manual reports no gender differences on the VR test (nor indeed on the NA test) (3) Research carried out in Ireland using verbally loaded tests presents mixed results No gender differences have been found in several studies using a test of reading comprehension (NS6 Reading Attainment Test) with primary school pupils of age 10 and 11 (26, 33) and post primary school first year pupils (32) A more recent study using the same test with a sample of primary school pupils, aged 10 and 11 years found differences

favouring boys (34) Martin and Kellaghan (25) observed no gender differences on tests of English attainment at third (Drumcondra English Test, Level II, Form A) and fifth (Drumcondra English Test, Level III, Form A) standards in primary school or on a test of Irish attainment (Drumcondra Irish Test, Level III, Form A) at fifth standard, although girls in third standard were found to score higher on an Irish attainment test (Drumcondra Irish Test, Level II, Form A) Using the Drumcondra Verbal Reasoning Test (DVRT) with a national sample of 500 11-year-old primary school pupils, Kellaghan and Macnamara (19) found no gender differences either However, in a longitudinal study of the same sample of pupils, Greaney and Kellaghan (12) found that, among students who were still attending school, differences on the DVRT favouring boys had emerged by the time they took the ICE and were still evident at LCE level That the gender differences observed in this and in the present study might be attributable to the dropout of less able boys is suggested by evidence both of a higher rate of male attrition (at least in the early second-level years) and of ability linked persistence through second level (12) Since different types of verbally loaded tests and samples have been used, it would seem that generalization from the Irish findings needs to be qualified in terms of the type of 'verbal' test (e g , reasoning, comprehension) and the age and school level of the sample Further, cultural factors may influence the direction of gender differences (cf 7), suggesting the need for caution in generalizing findings across countries

In a review of the full DAT test battery, Bouchard (4) argues that the strong intercorrelations between the component tests indicate that they are not independent measures Similarly, Linn notes that 'differential predictions and profile interpretations are generally not defensible in terms of the existing evidence of differential validity' (22, p 659) While the correlations between the VR and NA tests observed in this paper were lower than those reported in the American test manual (3, Ch 7), they were still high, and partially account for the small differences in the relationships between the respective VR and NA tests and performance in a number of subject areas on both objective achievement measures and public examinations In addition, the high intercorrelations help explain the sometimes negligible improvement in prediction of performance on both types of criterion measure achieved by summing the scores on the VR and NA tests to produce the VR NA composite score

In our study, we found that the criterion related validity of the DAT VR NA test was highest for the objective achievement tests, next highest

for the Intermediate Certificate Examination, and lowest for the Leaving Certificate Examination Three factors may be considered in attempting to provide partial explanations of these findings predictor-criterion time lag, range restriction in both predictor and criterion variables, and a method effect (i e , objective test vs essay type examination)

Predictor-criterion time lag might be used to explain partly the lower criterion related validity of the DAT against individual subject performance on the LCE, the LCE was taken some 30 months after the DAT was administered However, this explanation seems less appropriate in relation to the prediction of ICE performance, with a comparable one year time lag, the DAT was found to correlate higher with the objective achievement tests than with subject performance on the ICE In addition, the comparable predictive validities of the autumn 1975 and summer 1977 DATs against overall LCE performance suggest that predictor criterion time lag is less important than the analysis of individual subject performance might indicate

Several factors combine to reduce the effective sample size for any analyses of individual subject performance on both public examinations Population and sample attrition contribute to range restriction in the predictor variable (12) and also, by implication, in the criterion variables More obviously, the existence of two course/examination levels for some subjects at ICE level and for most subjects at LCE level probably contributes greatly to range restriction in criterion scores Further, analyses of individual LCE subject performance can be expected to yield more fragmented and less consistent results than would be the case for individual ICE subject performance since there are fewer examinees and more subject options, as well as the choice of course/examination levels For the ICE and LCE subjects in which Higher- and Lower-level examinations are taken, criterion related validity of the VR NA test against individual subject performance was poor except when some account was taken of range restriction in the examination grades For example, the weak correlation with performance on LCE Higher-level Mathematics can be attributed partly to the effects of range restriction in the comparatively small number of candidates taking Higher level Mathematics (cf 11) The correlation between the VR NA composite scores and criterion-scaled ICE and LCE subject performance illustrates the effects of range restriction on the variance in Irish, English, and Mathematics examination performance attributable to the DAT VR NA composite There are also suggestions that the grading system used in the public examinations may

itself contribute to range restriction in examination results (cf 15)

A method effect might provide a partial explanation of lower criterion related validity of the DAT VR NA for the public examinations than for the objective achievement tests Like the DAT, the achievement tests used in this study are objective, multiple choice tests In contrast, the method of assessment characteristically used in the ICE and LCE is open ended and of the essay type Given comparable content (e g , mathematics) over assessment procedures, a difference in criterion related validity might be attributable to the difference in the method of assessment

In terms of criterion related validity against the criterion of overall LCE performance, the data presented in this paper indicate that, depending on the points scheme and on student gender, from 26% to 44% of criterion variance can be accounted for by the scores on the VR NA composite measure of scholastic aptitude administered some 30 months before the LCE was taken Our findings regarding the criterion related validity of the DAT VR NA composite measure are supported by the findings of an earlier study in which the Primary Mental Abilities (PMA) test was administered one year before the participating students took the LCE, a correlation of 53 (N = 825) was observed between the PMA test and a measure of overall LCE performance quite similar to the ones examined in this study (24)

Until quite recently, standardized objective educational tests received little use in Ireland and this kind of testing (or examining) was thought to be relatively unfamiliar to teachers and students alike (20) With the growth of testing within the school system, the traditional psychometric emphasis on correct (i e , valid) interpretation and use of test information amounts, in practical terms, to the continuous monitoring of tests and how they respond to or are affected by, for example, changes in curricula, subject provision, and population (e g , the apparent emergence of gender differences on verbal tests), and the passage of time Validation studies of the type presented here are an integral feature of this kind of process and facilitate optimal use of test information in an educational setting

REFERENCES

1 AMERICAN PSYCHOLOGICAL ASSOCIATION *Standards for educational and psychological tests* Washington DC Author 1974

2 BEATON A E Scaling criterion of questionnaire items *Socio Economic*

Planning Sciences 1969, *2* 355 362

3 BENNETT, G K , SEASHORE, H G , & WESMAN, A G *Manual for the Differential Aptitude Tests Forms S and T* (5th ed) New York Psychological Corporation, 1974

4 BOUCHARD, T J Review of Differential Aptitude Tests In O K Buros (Ed) *The eighth mental measurements yearbook Volume 1* Highland Park, NJ Gryphon Press, 1978

5 BROWN, F G *Principles of educational and psychological testing* (3rd ed) New York Holt, Rinehart, & Winston, 1983

6 CRONBACH, L J Test validation In R L Thorndike (Ed), *Educational measurement* (2nd ed) Washington, DC American Council on Education, 1971

7 DYWER, C A Sex differences in reading An evaluation and a critique of current theories *Review of Educational Research* 1973, *43* 455-467

8 EDUCATIONAL RESEARCH CENTRE *Differential Aptitude Tests Adapted version* Dublin Educational Research Centre, St Patrick s College, 1975

9 EDUCATIONAL RESEARCH CENTRE *Drumcondra Attainment Tests (Irish English Mathematics) Level V Form A* Dublin Educational Research Centre, St Patrick s College 1976

10 EDUCATIONAL RESEARCH CENTRE *Drumcondra Attainment Tests (Irish English Mathematics) Level VI, Form A* Dublin Educational Research Centre St Patrick's College 1978

11 GREANEY V The predictive validity of the Irish Leaving Certificate Examination In J Coolahan (Ed) *University entrance requirements and their effect on second level curricula* Dublin Irish Federation of University Teachers 1979

12 GREANEY V & KELLAGHAN T *Equality of opportunity in Irish schools A longitudinal study of 500 students* Dublin Educational Company 1984

13 HILTON, T L Predictability and intellectual growth – some comments on the degree and interpretation of growth correlations Research Memorandum RM 72 1 Princeton NJ Educational Testing Service, 1972

14 HUMPHREYS A Cognitive non-cognitive biographic and demographic correlates of students performance in first university examinations Unpublished Master s Thesis, University College Cork, 1977

15 *ICE report* Final report of the Committee on the form and function of the Intermediate Certificate Examination Dublin Stationery Office, 1975

16 IRELAND DEPARTMENT OF EDUCATION *Tuarascail staitistiuil (Statistical report) 1974/75 1975/76* Dublin Stationery Office, 1977

17 IRELAND DEPARTMENT OF EDUCATION *Tuarascail staitistiuil (Statistical report) 1977 78* Dublin Stationery Office 1979

18 IRELAND DEPARTMENT OF EDUCATION *Rules and programme for secondary schools 1982 83* Dublin Stationery Office 1982

19 KELLAGHAN T , & MACNAMARA, J Family correlates of verbal reasoning ability *Developmental Psychology* 1972, 7 49 53

20 KELLAGHAN, T , MADAUS, G F , & AIRASIAN P W *The effects of standardized testing* Boston Kluwer Nijhoff, 1982

21 KELLAGHAN, T , MADAUS, G F , AIRASIAN, P W , & FONTES P J The mathematical attainments of post primary school entrants *Irish Journal of Education* 1976, *10* 3 17

22 LINN, R L Review of Differential Aptitude Tests In O K Buros (Ed) *The eighth mental measurements yearbook Volume 1* Highland Park NJ Gryphon

Press, 1978

23 MACCOBY, E E, & JACKLIN C N *The psychology of sex differences* Stanford, CA Stanford University Press, 1974

24 MADAUS, G F, KELLAGHAN, T, & RAKOW, E *A study of the sensitivity of measures of school effectiveness* Report submitted to the Carnegie Corporation, New York, December, 1975 Dublin Educational Research Centre, St Patrick s College, Chestnut Hill MA School of Education, Boston College, 1975

25 MARTIN, M, & KELLAGHAN, T Factors affecting reading attainment in Irish primary schools In V Greaney (Ed), *Studies in reading* Dublin Educational Company, 1977

26 McDONAGH, D A survey of reading comprehension in Dublin city schools *Irish Journal of Education* 1973, *7* 5 10

27 MESSICK, S Constructs and their vicissitudes in educational and psychological measurement *Psychological Bulletin* 1981, *89* 575 588

28 MORAN, M A, & CROWLEY, M J The Leaving Certificate and first year university performance *Journal of the Statistical and Social Inquiry Society of Ireland* 1979, *24* 231 266

29 PALLAS, A M & ALEXANDLR, K L Sex differences in quantitative SAT performance New evidence on the differential coursework hypothesis *American Educational Research Journal* 1983, *20* 165 182

30 PETERSEN A C CROCKETT L & TOBIN RICHARDS, M H Sex differences In H E Mitzel (Ed), *Encyclopedia of educational research Volume 4* (5th ed) New York Free Press, 1982

31 RAFTERY, A *University selection procedures in Ireland* Paper read at Union of Students in Ireland Seminar Belfast, March 1977

32 SWAN D Sex differences in reading achievement In V Greaney (Ed), *Studies in reading* Dublin Educational Company, 1977

33 TRAVERS, M A second replication of a survey of reading comprehension in Dublin city schools *Irish Journal of Education* 1976, *10* 18 22

34 WARD, N A fourth survey of reading comprehension in Dublin city schools *Irish Journal of Education* 1982 *16* 56 61