The Irish Journal of Education, 1969, III, 2, pp 117-122

VALIDITY STUDIES WITH A TEST OF HIGH-LEVEL REASONING

GEORGE SHOUKSMITH
Queen's University, Belfast

A measure of high-level reasoning, the Canterbury Reasoning Test (CRT) was employed in two studies, aimed at establishing its construct and predictive validity Data from the first investigation a factorial study with second-year university students, is used to provide evidence for the construct validity of the test as a measure of two basic aspects of high-level reasoning, (i) logical reasoning and (ii) pure associative reasoning The second investigation, a practical study carried out for the New Zealand Forest Service, suggests that the CRT has some validity as a predictor of success in university studies

The shortage of general reasoning tests suitable for use with subjects of high educational standing led the author to design a measure of a predominantly non-verbal nature, the Canterbury Reasoning Test (4) The Canterbury Reasoning Test (CRT) is a timed test of thirty minutes duration and consists of fifty items, mainly of the number and letter series variety It has been found to have a test-retest (one month interval) reliability co-efficient of 85 and a split-half reliability co-efficient of 90 Early studies with the test, with a group of 78 first year university students, produced correlations of r= 61 with the Moray House Test No 10, and of r= 72, with the Numerical Ability Test of the Differential Aptitude Test (DAT) battery These findings with their implications for at least a 'g' and 'n' factor content, suggested that a more detailed analysis of the factorial nature of the test might be worth while It was hoped, however, that as well as providing a research tool for the investigation of high-level reasoning, the CRT would also be useful for the practical task of selecting university students A correlation of r= 53 between the CRT and the class examination results of a group of 94 first year students taking a course in experimental psychology (including statistics) suggested that the test did have potential in the practical field

The present paper reports two studies which provide evidence for the construct and predictive validity of the test in the directions indicated by the foregoing

CONSTRUCT VALIDITY

The first study consists of a factorial investigation of a number of tests,

aimed at establishing the underlying constructs measured by the CRT Together with the CRT the following tests were administered Terman's Concept Mastery Test (an established test of high-level ability), the Mill-Hill Synonyms Test, for comparison with the Concept Mastery Test, and the complete Differential Aptitude Test battery, since the author had previously carried out a factor analysis of the battery for classroom demonstration purposes and this analysis could act as a guide for the present study

Subjects

The subjects were a group of 50 second year students (28 males, 22 females) of the University of Canterbury, New Zealand

Results

The matrix of correlations was factored by the centroid method and independently by the principal components technique Both methods produced qualitatively similar results, but the presence of negative latent roots of considerable size suggested that the latter analysis was unstable, so we returned to the centroid analysis for interpretation Three significant centroid factors were extracted and rotated graphically to a simple structure solution, modified to fit the findings of the previous analysis of the DAT battery The rotated factor matrix is shown in Table 1, where all loadings of $\pm$ 20 or less have been excluded, being regarded as zero projections (6)

TABLE 1

ROTATED FACTOR MATRIX

	Tests	Factors I	II	III
1	Canterbury Reasoning Test	—	77	63
2	Concept Mastery, Pt I Synonyms	66	—	37
3	Concept Mastery, Pt II Analogies	56	—	70
4	Mill-Hill Vocabulary Pt I Synonyms	75	—	—
5	DAT Verbal Ability	62	—	42
6	DAT Abstract Reasoning	—	64	52
7	DAT Numerical Ability	34	59	41
8	DAT Space Relations	—	70	—
9	DAT Mechanical Comprehension	—	37	45
10	DAT Clerical Speed and Accuracy	—	—	40
11	DAT Language Usage, Pt I Spelling	70	—	—
12	DAT Language Usage, Pt II Sentences	76	—	—

Interpretation of Factors

Factor I This is easily and uniquely interpreted as the traditional v ed factor of the British school Its major loadings are in the verbal tests but it is not purely a verbal factor since it has a small but significant loading in the DAT Numerical Ability Test This suggests that in Factor I 'learned material' is represented, or in other words, there is an educational or schooling component Therefore the wider interpretation as v ed seems better

Factor II This is a complex factor and Vernon's suggestion (7) that at high ability levels an alteration in the structure of 'k' and 'n' occurs, so that they are no longer opposed, seems relevant to its interpretation Non-verbal 'g' tests, Vernon concludes, link up with 'k' and 'n', the latter then becoming detached from the v ed cluster Thus Factor II appears to represent the formal logical aspect of reasoning isolated by Holzinger and Harman (3), it occurs frequently in the literature where solutions of problems in arithmetic, syllogisms and the like are concerned We can, therefore, define it as 'logical reasoning'

Factor III The highest loading for this factor is in the Analogies sub-test of the Concept Mastery Test, where completion of an item by the use of an appropriate concept, the calling to mind of pertinent information, is involved Part I of the Concept Mastery Test, which does not require the production of a concept in the same way as Part II has a lower loading It would seem, therefore, that Factor II is a second reasoning factor which is concerned with the less formal aspects of reasoning—reasoning which is linked with the production of new ideas or the relating of two or more ideas to form a new concept, in much the same way as Hebb (2) suggests later learning works

The factorial analysis confirms Davis's (1) division of reasoning into two factors, the first being concerned with logical analysis of a problem situation and the second involving the production of meaningful associative responses This second factor Davis calls 'pure judgment and reasoning' At first sight, it seems somewhat surprising that this factor, which Davis suggested had something to do with fluency and which suggests the creative aspect of reasoning, should enter to such a high degree into the CRT, a seemingly traditional 'convergent thinking' test Two explanations for the findings may be put forward Firstly, the aspect of fluency which operates in Factor III may be that which involves the calling to mind of pertinent information, which in turn could well be a major component in success in the CRT Secondly, the response style of the CRT may be regarded as being open, as on many items the length of the information required and

in some cases the number of 'bits' of information, is not specified A further study involving the CRT, reported elsewhere (5), showed that changing the answer format of this test affected both patterning of obtained response to the test items and *Ss*' overall performance Performance with the 'open' answer format used here, appears to involve a wider, more open reasoning ability

PREDICTIVE VALIDITY

The New Zealand Forest Service operates a cadet scheme, training suitably qualified entrants to be either rangers or foresters Some of the latter group, providing they hold the NZ University Entrance qualification, may be chosen on the basis of their work in initial training courses to undertake university studies Recently, a further development was suggested, that of offering special cadetships giving direct entry to university to suitably qualified applicants The present data was collected primarily to assess the usefulness of standard tests in selecting candidates whose potential for university study qualified them for 'direct entry cadetships' The results are used here to provide information about the predictive validity of the CRT which was included in the proposed selection battery

Subjects

The test battery was given to the existing group of forester cadets (N=39) Of this group, 10 were attending university, having been chosen on the basis of their initial work in the Forest Service The remaining 29 cadets were continuing with normal in-service training at one of the Forest Service's own schools These two groups were used as the criterion groups

Results

The CRT was administered to the complete sample, together with the following measures, Vernon Graded Mathematics Test, DAT Abstract Reasoning Test, Terman Concept Mastery Test and Raven Matrices Test Mean scores on all variables for university and non-university groups were compared and t-tests for significance were computed These results appear in Table 2 and indicate which variables were capable of distinguishing between the two groups

TABLE 2

COMPARISON OF MEAN SCORES OF UNIVERSITY AND NON-UNIVERSITY GROUPS OF FORESTER TRAINEES

Test	Diff between Means $(\bar{X}_1 - \bar{X}_2)$	t (n=39)	Level of significance
1 Canterbury Reasoning Test	8 7	2 278	$p < 05$
2 Vernon Graded Maths	6 88	2 099	$p = 05$
3 Concept Mastery Test	16 53	1 784	NS
4 DAT Abstract Reasoning Test	3 1	1 55	NS
5 Raven Progressive Matrices Test	0 98	0 73	NS

From Table 2, it can be seen that both the Vernon Mathematics Test and the CRT distinguish between the university and non-university cadets, these being the only tests from the initial battery having that capacity There is some further evidence, not revealed in the table, that the CRT has powers for relatively finer discrimination between university students of differing ability levels Four of the ten foresters attending university obtained 2 × B-grade passes in each of two consecutive years in the units they were taking A further four failed to reach a C-grade pass m two subjects in the same two years When the test scores of these two sub-groups were compared, it was found that the mean standard score on the CRT for the group of 'good' students differed significantly from the mean standard score of the 'poor' students Since the size of the sub-groups is extremely small, these findings are not conclusive Taken with the general findings, however, they do suggest that in the CRT we have a measure which is capable of making relatively precise predictions about a student's potential for success in university studies in science

CONCLUSIONS

The data on which the current studies were based, were collected in New Zealand The test itself, however, has been used extensively in the United Kingdom and throughout Ireland Currently, the CRT is being used to assess the potential of candidates for training in computer work In this it appears to be reasonably successful, as one would anticipate from the predictive validity study reported here The two studies reported indicate that the CRT offers a new approach to the assessment of reasoning with many practical, potential uses, which are extended by virtue of the fact

that the test's ceiling is very high It may be used with university students or with 'A' stream pupils in senior forms of secondary schools Its measurement is predominantly non-verbal, which in the many assessment problems in scientific and technical areas, is an added advantage The factorial analysis shows that it has a variance equally distributed between both reasoning factors, the CRT does not contain a significant verbal or educational component This factorial nature of the test suggests a number of uses for it in the research area, particularly in the newer areas of study concerning 'open' thinking or even 'creative' reasoning

REFERENCES

1 DAVIS F B The AAF qualifying examination *Army Air Force's aviation psychology program* Research Report No 6 Washington, D C US Government Printing Office, 1947

2 HEBB, D O *The organisation of behavior* New York Wiley, 1949

3 HOLZINGER, K J, and HARMAN, H H Comparison of two factorial analyses *Psychometrika*, 1938, *3*, 45-60

4 SHOUKSMITH, G A, *Manual for the Canterbury Reasoning Test* Christchurch, N Z The Caxton Press for the Carnegie Social Sciences Research Fund, 1960

5 SHOUKSMITH, G A Response bias produced by changing the answer format of a test *Psychological Reports*, 1968, *22*, 1310

6 THURSTONE, L L *Primary mental abilities* Chicago University of Chicago Press, 1938

7 VERNON, P E *The structure of human abilities* London Methuen, 1950