Person. individ. Diff. Vol. 8, No. 4, pp. 585–586, 1987
Printed in Great Britain.

0191-8869/87 $3.00 + 0.00

Stability of factor loadings

ROBERT L. THORNDIKE
Teachers College, Columbia University 525 W 120th Street, New York, NY 10027, U.S.A.

(*Received 3 July 1986*)

Summary—Each of 17 tests was embedded in six different eight-test matrices. Stability of general factor loadings were determined over the six contexts. Correlations of factor loadings ranged from 0.52 to 0.94 with an average value of 0.83.

There has been renewed enthusiasm in the past decade for *g* as an overall indicator of cognitive functioning. Much of Jensen's recent research related the size of the *g*-loading of different item types to outside variables such as information-processing speed or ethnic-group differences. Hunter reports that a broad general factor accounts for most of the prediction of training and job performance that is possible from the Armed Services Vocational Aptitude Battery (ASVAB) (Hunter, 1983). We have found in several cross-validation analyses that most of the prediction obtainable with such batteries as the Differential Aptitude Tests and the USES General Aptitude Test Battery can be achieved with a uniform common-factor score (Thorndike, 1985). With this renewed interest in 'psychometric *g*', it becomes of interest to find out to what extent the *g*-loading of a specific test is inherent in the nature of the test and remains stable as the test is inserted in different sets of other tests.

We have made some exploration of this issue. The basic data are found in a table of correlations among a set of 65 variables reported in the Appendix (p. 902) of Report No. 3 of the Army Air Forces Aviation Psychology Program Research Reports (Guilford, 1947). The battery was composed of 45 research tests and the 20 tests of the Aircrew Classification Battery.

For the purpose of this study, the first 48 variables were divided into six sets of eight variables that provided the matrix into which each of the remaining tests were inserted, one variable at a time. The *g*-loading of each of the 17 was determined six times, each time in the context of a different set of reference tests. To provide some sense of the composition of the six batteries, the tests are listed by name in Table 1. Any reader who is sufficiently curious can find a full description of each test in the Army Air Force report.

The 17 tests that were inserted into the six research batteries were all tests from the AAF Aircrew Classification Battery. They are listed together with their six *g*-loadings in Table 2.

Correlations among the *g*-loadings are shown in Table 3. The values above the diagonal are based on the tests as administered. To provide for the possibility that some of the correlation might have arisen from differences among the tests in reliability, *g*-loadings were also calculated after correcting for the unreliability of the inserted test. The correlations among these loadings are shown below the diagonal of the table.

Table 1. Test Variables in AAF Research Battery

1–8	25–32
Map memory	Object identification I
Figure analogies	Object identification II
Spatial visualization	Plane position memory
Planning air maneuvers	Decoding
Vocabulary	Route planning
Map distance	Flight formation
Estimation of length	Aerial landmarks
Speed of identification	Pattern assembly
9–16	**33–40**
Memory for plane silhouettes	Block counting
Directional orientation	Discrimination reaction time I and II
Visualization of maneuvers	Discrimination reaction time III and IV
Planning a circuit	Plane name memory
Path tracing	Planning a course
Maze tracing	Compass orientation
Formation visualization	Competitive planning
Objectivity of perception	Camouflaged outlines
17–24	**41–48**
Visual memory	Angle estimation
Figure classification	Spatial reasoning
Spatial visualization	Tool function
Map planning	Word knowledge
Object recognition	Dial and table reading
Object of perception	Biographical data—pilot
Position orientation	Biographical data–navigator
Aerial orientation	Spatial orientation I

Table 2. Factor loadings of 17 classification tests when inserted in six different matrices

Test	Matrix 1	2	3	4	5	6
1. Spatial orientation II	63	65	63	58	51	62
2. Reading comprehension	62	47	54	53	52	68
3. Instrument comprehension	48	56	63	51	49	58
4. Mechanical principles	43	61	59	47	33	57
5. Speed of identification	52	48	48	51	59	53
6. Numerical operations I	48	26	40	40	50	50
7. Numerical operations II	52	32	46	46	53	55
8. Mechanical information	20	30	26	18	08	49
9. General information	30	39	35	27	18	48
10. Judgment	43	35	39	37	39	51
11. Arithmetic reasoning	61	48	56	53	51	62
12. Rotary pursuit	21	30	33	24	24	28
13. Rudder control	12	28	28	15	09	28
14. Finger dexterity	34	25	38	35	33	37
15. Complex coordination	46	53	57	51	48	54
16. Two-hand coordination	25	35	37	35	33	39
17. Discrimination reaction time	52	55	61	59	60	61

Table 3. Correlations of g-loadings when put in different test matrices*

Set of six research tests	1	2	3	4	5	6
1	—	0.63	0.83	0.93	0.92	0.88
2	0.66	—	0.89	0.77	0.52	0.71
3	0.85	0.89	—	0.94	0.80	0.79
4	0.94	0.78	0.95	—	0.94	0.83
5	0.93	0.55	0.83	0.94	—	0.73
6	0.91	0.72	0.81	0.85	0.78	

*Values below diagonal corrected for test unreliability.

The median correlation for the uncorrected values (above the diagonal) is 0.83, while the median correlation for values corrected for unreliability (below the diagonal) is 0.85. Thus, it makes little difference whether one uses the uncorrected or the corrected values. The correlations cover a considerable range—from 0.52 to 0.94 in the uncorrected data—presumably reflecting the degree of similarity of the tests in the matrices into which the 17 tests were inserted. It is difficult to judge whether the similarity was more or less than would be encountered in different aptitude batteries appearing in the test literature. But the data do provide support for a view that the g-loading of a type of test task has substantial stability and is to a considerable extent determined by the characteristics of the test itself, rather than the context in which it appears.

REFERENCES

Guilford J. P. (1947) *Printed Classification Tests*. Army Air Forces Aviation Psychology Program Research Reports. Rep. No. 3. U.S. Government Printing Office, Washington, D.C.

Hunter J. E. (1983) Validity generalization of the ASVAB second report. Unpublished research report.

Thorndike R. L. (1985) The central role of general ability in prediction. *Multivar. behav. Res.* **20,** 241–254.