

Examination of an Extension of Guttman's Model of Ability Tests

Aharon Tziner,
Tel-Aviv University

Avigdor Rimmer,
Pilat, Ltd., Tel-Aviv

An extension of Guttman's structural model of ability tests was devised and investigated with two samples consisting, respectively, of 335 and 225 males. The examinees in the first sample came for vocational guidance after their military service and were administered a 17-test battery. The second sample consisted of applicants for various jobs in an organization and were administered a 14-test battery. For each sample, a matrix of intercorrelations between scores was obtained based on the number of correct responses. The matrices were submitted to Guttman-Lingoes Smallest Space Analysis. The two-dimensional structure found was a radex in which (1) the facet of the language of presentation radially divided the space and (2) the facet of mental operation formed concentric rings. The significance of these findings for theoretical and applied problems relating to ability tests is discussed.

The underlying structure of ability tests is a topic of considerable interest (cf., Cunningham, 1981; Kelderman, Mellenbergh, & Elshout, 1981). There appear to be two primary reasons for this. First, from a general perspective of ability testing it is important to have a coherent structural model that reasonably organizes the tests into meaningful facets. Such a model could aid in deciding, in a systematic and rational manner, what new tests should be developed in order to measure facets previously only partially assessed, or not assessed at all. Second, researchers are interested in determining

whether there is a universal structural model of general applicability that does not vary even when used for different testing objectives (e.g., vocational counseling or selection for job openings in organizations) or when used on individuals with varying levels of education, age, and so forth.

Previous studies that were aimed at examining the foregoing questions were mainly based on factor analysis (see, for instance, Guilford, 1980, 1981). This method, however, has been sharply criticized on several grounds. Basic assumptions of its statistical model, as, for instance, linearity, are often violated (Kelderman et al., 1981). Apart from this, factors with only one or two salient loadings are frequently found when using factor analysis to reveal the structure of ability tests (Cunningham, 1981). In such a context, the decision regarding the appropriate number of factors becomes unusually difficult, thus leading to solutions that are not parsimonious and giving rise to an inconsistency across investigators (Schlesinger & Guttman, 1969). Furthermore, although the factor analytic strategy of research has often been used to test hypotheses of structure, it has also been shown to be inadequate for this purpose (Guttman, 1958; Shapira & Zevulun, 1979; Shye, 1978).

A better approach to the structural configuration issue might emerge from Guttman's facet theory (for a detailed exposition see Shirom, 1982; Shye, 1978). Relying on this approach, Schlesinger and Guttman (1969) classified ability tests into two fac-

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 8, No. 1, Winter 1984, pp. 59-69
© Copyright 1984 Applied Psychological Measurement Inc.
0146-6216/84/010059-11\$1.80

ets: (1) the language of presentation of the test items, which may be verbal, figural, or numerical; and (2) the mental operation required by the test, namely, rule inference or rule application. They then defined a mapping sentence to corroborate the structural configuration based on the foregoing facets. The empirical investigation, using high school students as subjects, firmly supported the hypothesized structural configuration: a radex yielded by a radial arrangement of tests according to the language of presentation facet, together with a pattern of concentric circles formed by the mental operation facet.

Nevertheless, it is important to outline a weakness inherent in one of the facets—that of mental operation. This facet did not take into account the important differentiation in mental operations indicated by Piaget (1956; and also in Brainerd, 1978). According to Piaget, the rule of application operation entails three levels of cognitive complexity organized hierarchically from the lowest to the highest:

1. Ability to successfully handle concrete items (designated here *concrete rule application*),
2. Ability to successfully handle signs (*clerical rule application*), and
3. Ability to successfully handle cognitive concepts (*abstract-cognitive rule of application*).

For this reason the mental operation was restructured in the present study so that it would be composed of four elements instead of two:

1. Rule inference,
2. Abstract-cognitive rule application,
3. Clerical rule application, and
4. Concrete rule application.

Using this extension, a mapping sentence was devised similar to that formulated and examined by Schlesinger and Guttman (1969) in order to extract the maximum from the operation facet. The aim was to find out empirically whether or not the radex structural configuration is still preserved after the introduction of the above-mentioned extension. Two replications were carried out: one on a sample of individuals applying for vocational counseling, the other on a sample of selected applicants for job positions in organizations. It was then decided to

test the mapping sentence, employing two replications, in order to ascertain whether the objectives of testing in any way affected the structural configuration of ability tests that emerged.

Research Mapping Sentence

The mapping sentence tested in the current study is shown in Figure 1. The aim was to examine empirically whether the radex would still be an appropriate structural configuration to portray geometrically the interrelationships among the elements of facets, even when using this elaborate mapping sentence.

Method

Subjects and Procedure

The first sample consisted of 335 males, mostly aged 21 to 22, selected at random from applicants for vocational guidance after completion of their military service. The second sample consisted of 225 males aged 21 to 35, also selected at random, from applicants for job positions in an organization. Both groups completed ability tests that were administered and scored by professional psychologists—17 tests for those in the first sample and 14 for those in the second. Examinees in the first sample were tested at the Vocational Branch of the Israeli Ministry of Labor and Welfare in Tel Aviv, and those in the second sample were tested at a private psychological testing company providing selection and placement services.

Variables

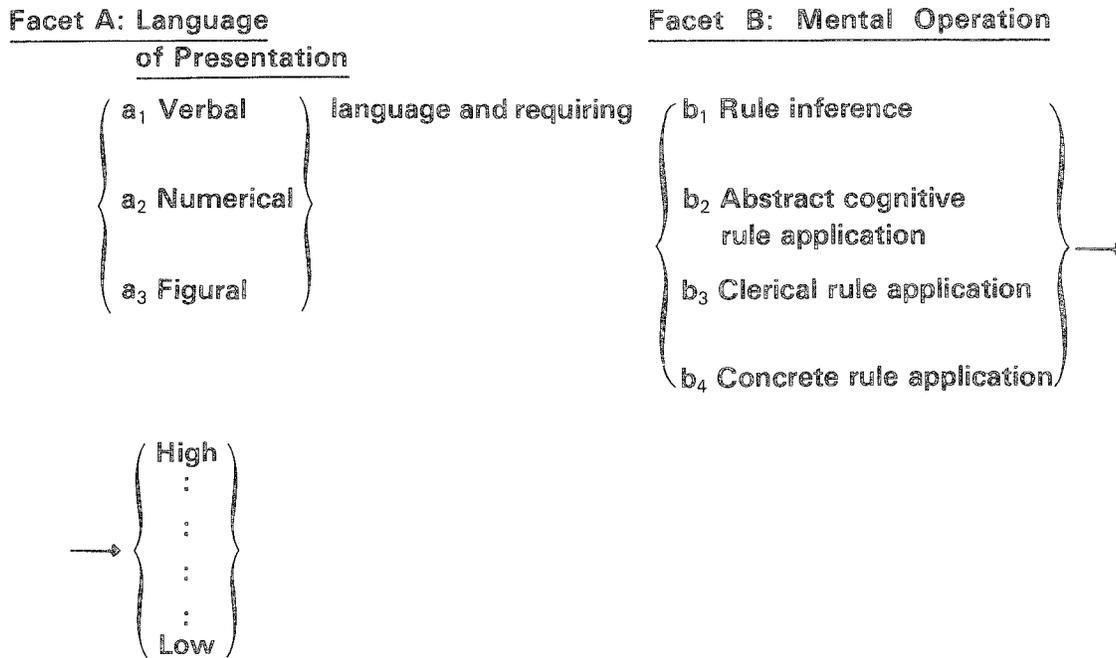
The combination of elements of the two facets, A and B, forms profiles (*structuples* in facet theory terminology) represented here by the ability tests. For instance, the Verbal Analogies test represents the a_1b_1 structuple. The detailed list of tests employed, along with the structuples to which they correspond, is presented in Table 1.

Data Analysis Procedure

The investigation of the empirical structure of the ability tests was carried out using Guttman-

Figure 1
The Mapping Sentence for Ability Tests

The performance of subject (x) on a psychometric ability test item presented in



Lingoes Smallest Space Analysis (SSA; see, e.g., Guttman, 1968; Lingoes, 1973). This is a non-metric multidimensional scaling procedure that geometrically represents the correlation matrix based on the order of magnitude of the intercorrelations among the variables—in this case, the tests. SSA portrays the tests in the current study as points in an Euclidean space. The more the correlation between the corresponding tests increases, the nearer the two points are to each other. More explicitly, if r_{ij} is the observed correlation between tests i and j , and if d_{ij} is the calculated distance between these two tests for the smallest space, then SSA attempts to find out the space with the minimal number of dimensions in which the following inequalities are preserved:

$$d_{ij} < d_{kl} \text{ whenever } r_{ij} > r_{kl} \quad (1)$$

$$(i, j, k, l = 1, 2, \dots, n)$$

Order among the correlations is thus preserved in

the smallest possible Euclidean space, and the absolute sizes of the input correlation coefficients can be reproduced from the scattergram of the r_{ij} on the d_{ij} (Shepard diagram) printed by the computer. Such a chart indicates how well $r_{ij} = f(d_{ij})$, where f is a monotonically decreasing function.

The goodness of fit between the smallest possible Euclidean space and the correlation matrix is numerically measured by a coefficient of alienation $(1 - r^2)^{1/2}$, where r is a rank-order correlation between the intercorrelations of the tests and their corresponding geometric distances. The smaller the coefficient of alienation, the better the fit. Zero represents a perfect fit. In practice, however, a coefficient of alienation smaller than .15 is considered a good fit (see Guttman, 1968).

To find evidence regarding the structural configuration in the SSA output, it is necessary to apply the principles of contiguity regions (Elizur, 1981; Shapira & Zevulun, 1978). In this connection, and

Table 1
Psychometric Ability Tests and the Structuples they Represent
in the Two Replications

First replication, n=335	Second replication, n=225	Structuple
1. Verbal analogies	---	a ₁ b ₁
2. Finding the odd word		
3. Vocabulary test	1. Rewriting scrambled sentences	a ₁ b ₂
4. Classifying and adding expenses		a ₂ b ₂
5. Arithmetic problems	2. Arithmetic problems	
6. Raven's progressive matrices	3. Raven's progressive matrices	
	4. Form perception	
7. Space perception (identifying a body by its envelope)	5. Bennett's mechanical comprehension	a ₃ b ₂
	6. Space perception (identifying a body by envelope)	
8. Word comparison	---	a ₁ b ₃
9. Identifying words in small printing		
10. Number comprison	7. Addition exercises	a ₂ b ₃
11. Number code substitution	8. Mixed exercises	
12. Arithmetic exercises		a ₃ b ₃
13. Identifying figures in small printing	9. Figure acuity	
14. Minnesota form board	10. Koh's cubes	a ₃ b ₄
15. Hand and leg coordination	11. Hand and leg coordination	
16. Fixing different sorts of pins in form board	12. Pin insertion	
17. Marking three points in a circle	13. Gilbert's test	
---	---	a ₂ b ₄
	14. Mathematical progressions	a ₂ b ₁

in order to detect a radex structure, one of the facets in the mapping sentence of this study (A) should have acted as a *polarizing* facet, while the other (B) should have acted as a *modulating* facet. A facet is designated as modulating when it divides the space into concentric circles and as polarizing when three or more lines originating at the same point (or pole) separate the space by diverging into different directions (Schlesinger & Guttman, 1969). Additional relevant details pertaining to the principles of contiguity in facet analysis can be found elsewhere (Gratch, 1973; Shye, 1978).

Results

First computed among the number of correct responses to the tests were Pearson intercorrelations. Tables 2 and 3 display the results for the first and second replication, respectively. The first point of interest that can be seen in both tables is that all the correlations are zero or positive, fulfilling one of the requirements set up by Guttman for facet analysis (for more details refer to Gratch, 1973). In addition, the correlations in the first application are much higher than in the second application, a

Table 2
Correlations Between Ability Tests in the First Replication (n=335)

Ability tests	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Verbal analogies	—																
2. Odd word	.80	—															
3. Vocabulary	.78	.82	—														
4. Arithmetic problems	.70	.64	.57	—													
5. Expenses problems	.67	.68	.64	.61	—												
6. Raven's matrices	.69	.67	.60	.63	.63	—											
7. Space perception	.54	.58	.49	.51	.52	.68	—										
8. Words comparison	.61	.62	.62	.51	.66	.50	.36	—									
9. Words identification	.67	.68	.71	.54	.69	.53	.40	.72	—								
10. Numbers comparison	.52	.50	.46	.53	.64	.47	.38	.69	.63	—							
11. Number-code substitution	.54	.53	.49	.51	.61	.53	.38	.64	.59	.73	—						
12. Arithmetic exercises	.53	.54	.49	.63	.60	.51	.44	.57	.53	.65	.60	—					
13. Figures identification	.42	.48	.37	.40	.51	.56	.42	.41	.44	.40	.42	.42	—				
14. Marking three points in a circle	.33	.30	.37	.40	.37	.33	.29	.32	.32	.42	.45	.41	.28	—			
15. Minnesota form board	.43	.46	.38	.43	.45	.57	.54	.52	.32	.42	.46	.37	.54	.27	—		
16. Hand/leg coordination	.25	.28	.21	.18	.29	.36	.36	.24	.19	.28	.26	.18	.27	.16	.38	—	
17. Fixing different sorts of pins in formboard	.27	.28	.19	.28	.35	.37	.33	.30	.28	.32	.36	.28	.39	.26	.49	.38	—

Note: Intercorrelations of $r = .11$ and above are significant at least at $p < .05$, two-tailed.

Table 3
Correlations Between Ability Tests in the Second Replication (n = 225)

Ability Tests	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Mathematical progressions	—													
2. Arithmetic problems	.87	—												
3. Raven's progressive matrices	.67	.61	—											
4. Space perception	.52	.48	.59	—										
5. Form perception	.46	.41	.54	.54	—									
6. Bennett's mechanical comprehension	.51	.45	.68	.53	.55	—								
7. Rewriting scrambled sentences	.34	.32	.39	.27	.18	.32	—							
8. Addition exercises	.43	.47	.24	.24	.22	.17	.05	—						
9. Mixed exercises	.41	.43	.22	.16	.26	.13	.02	.59	—					
10. Figure acuity	.37	.33	.45	.48	.49	.46	.18	.21	.17	—				
11. Koh's cubes	.33	.29	.48	.46	.47	.55	.27	.15	.08	.39	—			
12. Hand/leg coordination	.20	.18	.20	.30	.25	.34	.17	.12	.03	.36	.29	—		
13. Pin insertion	.13	.10	.19	.21	.26	.29	.05	.12	.07	.29	.22	.43	—	
14. Gilbert's test	.08	.06	.12	.11	.19	.27	.04	.02	.00	.26	.24	.61	.45	—

Note: Intercorrelations of $r = .13$ and above are significant at least at $p < .05$, two-tailed

fact that can be explained by the first sample being more heterogeneous than the second, as dictated by the objective of testing (vocational guidance versus selection and placement for a particular organization).

These matrices were then subjected to the SSA-I computer program, which maps the tests as points on the Euclidean space of two dimensions. These test configurations are shown in Figures 2 and 3, corresponding to the matrices in Tables 2 and 3, respectively. The coefficients of alienation obtained were .11 and .095, respectively, for the configurations in Figures 2 and 3, which can be con-

sidered as indicating good and very good fit of the two-dimensional plots to the original intercorrelation matrices (as references for goodness of fit, see Guttman, 1968; Kenny & Canter, 1981). Both figures show that every test is contiguous with others that share the same facet elements. This is consistent with the principle of contiguity in facet theory (see Shye, 1978), which explicitly states that items having the same facet elements should be more highly correlated and thus closer together in the configurative plot than items that do not. This is illustrated here, for example, by Verbal Analysis and Finding-the-Odd-Word tests pertaining to the

Figure 2
SSA Map for the 17 Tests in the First Group
(Vocational Guidance Situation)

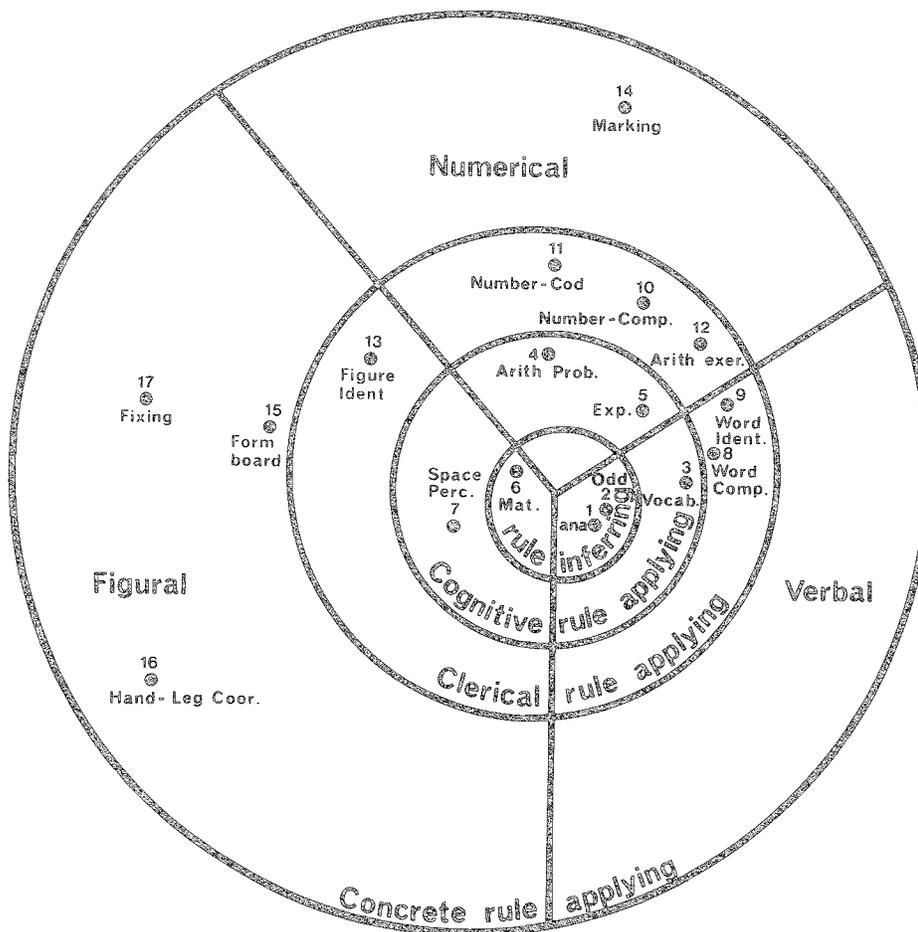
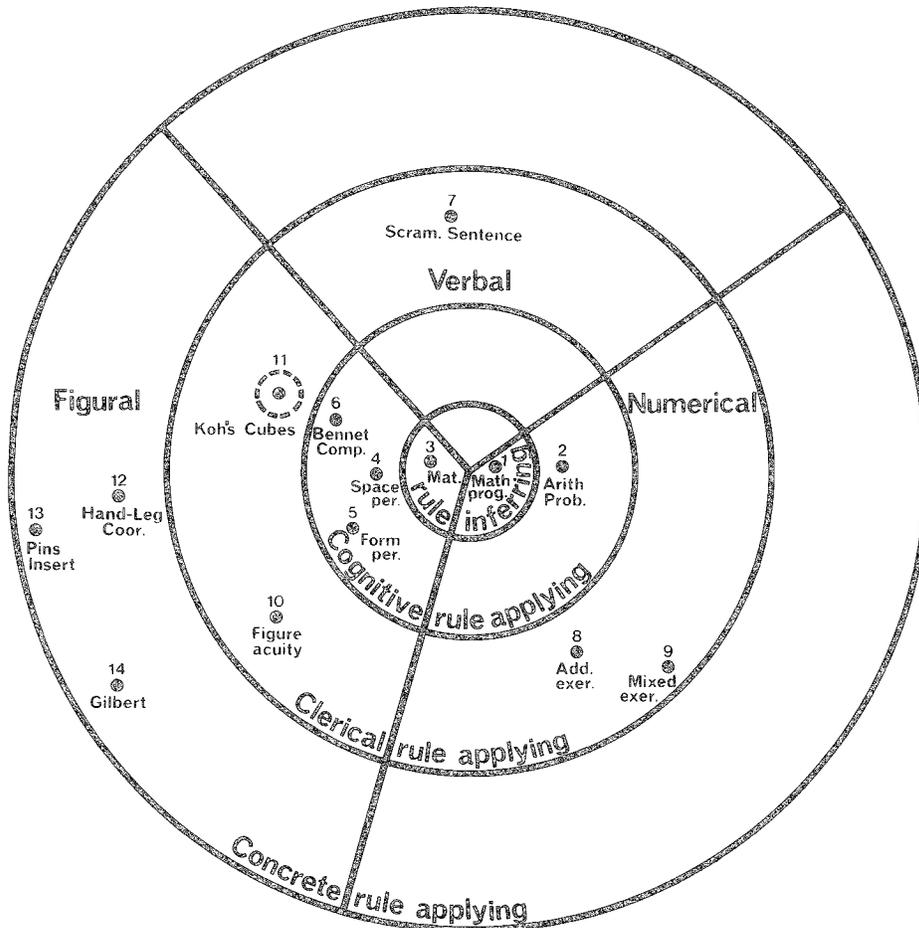


Figure 3
SSA Map for the 14 Tests in the Second Group
(Selection Situation)



same structuple a_1b_1 in Figure 2. These are clearly closer to each other than to the Identifying Figures in Small Print test, which represents structuples a_3b_3 .

Furthermore, both figures have a similar configurative structure, as is expected according to the mapping sentence: a radex configuration formed by joint action of the language presentation facet (A), which divides the space radially (polarizing facet), and the mental operation facet (B), which divides the space from the center to the outside in concentric rings (modulating facet). Certain implications

are evident concerning the order of elements in Facet B. This facet divides the space so that (1) the rule-inferring ability tests occupy the central region, (2) encircled contiguously by the cognitive-abstract rule applying test region, (3) next by the clerical rule-applying tests regions, and (4) furthest by the concrete rule-applying tests region. Hence, the tests in each region are much closer to each other than to tests outside the region. For instance, Mathematical Progressions tests and Raven Progressive Matrices tests are much closer to each other than to Addition Exercises tests (Figure 3),

since both pertain to the rule-inference region, while the latter pertains to the clerical rule-applying region.

There is, however, one noticeable exception to the above conclusion: In general, the ability tests representing structuple a_3b_4 in Figure 3 cluster together in the figural area within the concrete rule circle. This is correct for all of them, except for Koh's cubes test, which is unexpectedly situated in the clerical circle. It may be reasonable to suppose that this seemingly odd finding could be accounted for by the fact that it presents stimuli in a concrete form, although the type of mental operation required by this test is the rule of inference. Thus, this test is positioned somewhere in the mid area, as is indeed shown by its place in Figure 3.

Discussion

The implications of the results in this study are two-fold. First, although only a relatively small number of tests were involved, of which not all the structuples were equally represented (if at all) by tests, and although a new elaborated mental operation facet (instead of the original one generated by Schlesinger & Guttman, 1969) was used, the radex proved again to be an appropriate configuration of the structure of ability tests. This conclusion is further supported by the fact that the radex structure did not evolve by chance, but rather that the empirical SSA results confirmed what had previously been predicted by the mapping sentence of the current study. Second, the configuration of points representing the ability tests exhibited the same structure (the radex) in both replications, albeit practically the tests administered to the testees in the two samples differed greatly, as dictated by the objective of testing (i.e., vocational guidance versus selection and placement). Consequently, it appears plausible to conclude that the radex structure can be expected to reproduce itself even if other tests are utilized, given that the two facets—mental operation and language of presentation—are represented.

Of additional interest is the fact that those few tests employed in both replications (e.g., Raven's Progressive Matrices) retained location in the same

regions of the radex in Figure 2 as in Figure 3, thus indicating that the placement of a test in the radex is apparently invariant with respect to other tests in the same battery; rather, the location is predictable by the mapping sentence. In contrast, the structural invariance is clearly opposed to the salient difference in patterns of intercorrelation between tests, notably upon scrutiny of Tables 1 and 2. This probably demonstrates that intercorrelations are to a large extent affected by various extraneous factors (e.g., characteristics of examinees) unlike the remarkable stability of the configurative structure—the radex. In this regard the authors are, in essence, consistent with Elizur (1982), who arrived at a similar conclusion in his investigation of work values structure. In order to further support the conclusions concerning this structure, more studies across various sociodemographic characteristics of examinees (e.g., age, level of education, type of occupation) are certainly needed; they should, in addition, deal with the objectives of testing (e.g., assessment of managerial potential, assessment for promotion and for placement in organizations).

A further implication emerged from the findings concerning the radex structure. As mentioned previously, the mental operation facet divided the space into four concentric regions. These regions may be conceived as being arranged in a gradient order formed according to mental abstractness when the concrete rule-application is the lowest level and rule inference is the highest. This implies that a taxonomy of ability tests might be devised according to the above-mentioned four levels of abstractness. Further, this taxonomy would be maintained, no matter in what language test items are presented—verbal, numerical, or figural.

If the veracity of this conclusion is proved in further similar replications, a practical contribution, in addition to the theoretical contribution, could be made to the understanding of the underlying structure of ability tests. For instance, in order to design a battery of ability tests for selection, it would only be necessary to obtain information—via a job analysis procedure—about the mental operation and the language of presentation that predominantly accounts for successful performance in

that job. Then, all that would be required would be to select or to construct tests corresponding to the type of mental operation and the chosen language of presentation revealed in the job analysis process.

Some readers may consider this suggestion to be oversimplified. The authors trust, however, that as it is theoretically reasonable, it offers an approach for devising test batteries from a minimal number of components that would still contain all the relevant core elements. Some redundancy among predictors would exist, even if the proposed test battery procedure is followed, since, as has been pointed out, the tests tend to relate to one another. Yet, this redundancy would be substantially smaller than that expected in a test battery structured in a different manner. Some overlap among predictors, however, would not be detrimental as long as it is not too large and it increases the overall validity of the test battery as a whole (cf. Tziner & Dolan, 1982).

To conclude, the authors agree with those researchers who maintain that the ultimate test for the validity of findings is their recurrence in numerous studies (cf. James, Mulaik, & Brett, 1982). For this reason additional research is suggested in order to accurately assess the extent of recurrence of the findings, in this study. This is particularly the case when comparing results from facet analysis, employed here with those emanating from other methods of structure exposure, such as Kruskal's multidimensional scaling analysis (Kruskal & Wish, 1978).

References

- Brainerd, C. J. (1978). *Piaget's theory of intelligence*. Englewood Cliffs NJ: Prentice-Hall.
- Cunningham, W. R. (1981). Ability factor structure differences in adulthood and old age. *Multivariate Behavioral Research*, 16, 3-22.
- Elizur, D. (1981). *The importance and structure of job outcomes*. Unpublished manuscript, Bar-Ilan University.
- Elizur, D. (1982). *Facets of work values: A structural analysis of work outcomes*. Unpublished manuscript, Bar-Ilan University.
- Gratch, H. (1973). *Twenty-five years of social research: The story of the Israel Institute of Applied Social Research*. Jerusalem: Jerusalem Academic Press.
- Guilford, J. P. (1980). Fluid and crystallized intelligences: Two fanciful concepts. *Psychological Bulletin*, 88, 406-412.
- Guilford, J. P. (1981). Higher-order structure-of-intellect abilities. *Multivariate Behavioral Research*, 16, 411-435.
- Guttman, L. (1958). What lies ahead for factor analysis? *Educational and Psychological Measurement*, 18, 497-515.
- Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for configuration of points. *Psychometrika*, 33, 465-506.
- James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Conditions for confirmatory analysis and causal inference*. San Francisco: Sage Publications.
- Kelderman, H., Mellenbergh, G. J., & Elshout, J. J. (1981). Guilford's facet theory of intelligence: An empirical comparison of models. *Multivariate Behavioral Research*, 16, 37-61.
- Kenny, C., & Canter, D. (1981). A facet structure for nurses' evaluation of ward designs. *Journal of Occupational Psychology*, 54, 93-108.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. London/Beverly Hills CA: Sage Publications.
- Lingoes, J. C. (1973). *The Guttman-Lingoes nonmetric program series*. Ann Arbor: Mathesis Press.
- Piaget, J. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.
- Shapira, Z., & Zevulun, E. (1979). On the use of facet analysis in organizational behavior research: Some conceptual considerations and an example. *Organizational Behavior and Human Performance*, 23, 411-428.
- Shirom, A. (1982). What is organizational stress? A facet analytic conceptualization. *Journal of Occupational Behavior*, 3, 21-38.
- Schlesinger, I. M., & Guttman, L. (1969). Smallest space analysis of intelligence and achievement tests. *Psychological Bulletin*, 71, 95-100.
- Shye, S. (Ed.) (1978). *Theory construction and data analysis in the behavioral sciences*. San Francisco: Jossey-Bass.
- Tziner, A., & Dolan, S. (1982). Validity of an assessment center for identifying future female officers in the military. *Journal of Applied Psychology*, 67, 728-736.

Acknowledgments

Both authors contributed equally to this article. The authors thank L. Guttman for his helpful comments on an earlier draft of this article.

Author's Address

Send requests for reprints or further information to Aharon Tziner, Department of Labor Studies, Faculty of Social Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel.