ABSTRACT
         Three attributes of faculty competence--teaching,
research, and service--were each measured by three different
instruments--department chair evaluation, self-evaluation, and one
other measure (student evaluations, number of publications, or number
of committees). The data were analyzed through a
multitrait-multimethod matrix (MTMMM) to ascertain convergent and
divergent validity. Data were obtained from 57 faculty members, all
of whom held the doctorate. Because the required conditions were not
fully met for each criterion, applying the MTMMM approach led to
somewhat ambiguous results. While the limitations (e.g., small sample
size, questionable properties of the instrument, and uni-dimensional
assumption of traits) should be considered in interpreting the
results, the fact that confidential data are difficult to obtain
suggests both that the tentative conclusions may be of value to other
researchers and that further investigations are warranted. The
results indicate that the evidence for validity is relatively weak;
it was concluded that the results from these measures should be
interpreted with considerable caution. (Author/PN)

Convergent and Discriminant Validation of Measures

of Faculty Competence

by the Multitrait-Multimethod Matrix

Richard B. Hansen

Department of Economics

Bruce G. Rogers

Department of Educational Psychology and Foundations

University of Northern Iowa

Running head: CONVERGENT AND DISCRIMINANT VALIDATION

Abstract

In this study, three attributes of faculty competence--teaching, research,
and service--were each measured by three different instruments--department
chair evaluation, self-evaluation, and one other measure (student evaluations,
number of publications, and number of committees). The data were analyzed
through a multitrait-multimethod matrix to ascertain convergent and divergent
validity. The results indicate that the evidence for validity is relatively
weak; it was concluded that the results from these measures should be
interpreted with considerable caution.

Convergent and Discriminant Validation of Measures of Faculty

Competence by the Multitrait-Multimethod Matrix

The study of measures of faculty competence has received considerable attention in the higher education literature, particularly during the past two decades, with the strengths and limitations of numerous criteria of performance being discussed (Miller, 1979). From a national survey of department chairmen, Centra (1979) reported that the four major sources of information on faculty competence were chairman evaluations, colleague opinions, student ratings, and committee deliberations. While some published studies have reported correlations among measures from these sources (Blackburn & Clark, 1975), the generalization of such results requires evidence of validity; namely, that the various sources of information can consistently distinguish among different aspects of faculty competence and can agree when addressing the same aspect.

In their attempt to develop a methodological approach to validity which would lead to a more sophisticated understanding of psychological traits, Campbell and Fiske (1959), in a now classic article, proposed the concepts of convergent and discriminant validity. They described the multitrait-multimethod matrix (MTMMM) as an efficient tool for displaying correlations, and developed criteria for evaluating the MTMMM in terms of the proposed validity concepts. To apply this technique, several traits are each measured by several methods, with the intent of showing the distinctiveness of each trait, regardless of the method of measurement used. Convergent validity is evidenced when different methods of measuring the same trait show high correlations, while discriminant validity is evidenced when different traits measured by the same method produce low correlations.

The MTMMM has been widely used in the literature, where it has made notable contributions toward the investigation of human traits (Bajtelsmit, 1979; Lawler, 1967). Recently, the technique was used to investigate factors of teacher effectiveness, as measured by student ratings and instructor self-evaluations (Marsh, 1982).

When employing the MTMMM, the researcher is first confronted with the question of what traits are to be assessed. Numerous traits have been proposed as components of faculty competence; e.g., Centra (1979) reported ten, Miller (1974) reported nine, and Marsh (1982) reported nine. Since considerable overlap existed among many of these, it appeared appropriate to conceptualize them into the traditional "three-legged stool" of teaching, research, and service. It thus became the purpose of this study to investigate the convergent and discriminant validity of several common measures of these three traditional attributes using the logical analysis of the MTMMM approach.

Method

Subjects

The data for this research were obtained from eight departments in a

midwestern university (Hansen, 1979). Within each department, a random sample of faculty members was drawn proportional to the department size, and the resulting 11. faculty members were invited to participate in the study. Of the 73 who agreed to do so, complete information was obtained from 57 (4 females and 53 males). There were 22 full professors, 22 associate professors, and 13 assistant professors, all of whom held the doctorate.

Procedure

The MTMMM approach requires that the same multiple methods be used to evaluate each of the traits. For this study, three sources of evidence were sought. First, the appropriate department chair rated, on a scale of 1-7, the professor's performance in the areas of teaching, research, and service for two time periods; namely, for that current academic year and for the composite of all prior years. Second, each professor rated, on a scale of 1-7, his or her own performance in the same three areas for both time periods. Finally, one other measure of performance was collected for each of the three traits. Current teaching performance was measured by the mean student response to a question concerning the professor's overall teaching ability. No measure of previous teaching ability was available. Using data from a faculty questionnaire covering both time periods, research performance was measured by total publications (including books, textbooks, journal articles, and other publications), and service performance was measured by the total number of department, college, and university committees on which the faculty member served. Since these last three measures did not employ the same method, the analysis of the MTMMM was appropriately modified.

Results

Pearson product moment correlation coefficients were computed among all of the variables and cast into a MTMMM (see Table 1). The parenthesized values in the main diagonal represent reliability estimates, and the underlined values represent validity coefficients. Each area of the matrix will be explained below.

The reliability values were obtained by correlating the results representing the two time periods. It would have been preferable to have obtained these coefficients by gathering data at two different times. However, since both ratings were, in fact, obtained at the same time, the first ratings represent the assessor's memory of how the faculty member would have been rated previously. This approach assumes a perfect memory on the part of the rater, a highly questionable assumption (Kenny, 1979, p. 233). Although these were the most stable reliability estimates that were available for this study, the resulting values should be interpreted with caution.

As expected, the reliability coefficients in the MTMMM tended to be the largest values within the matrix; indeed, they appeared sufficiently large to warrant further examination of the convergent and divergent validity of the measures.

Convergent validity was investigated by inspecting the validity diagonals;

namely, those correlations between different methods of measuring the same
trait (the monotrait-heteromethod correlations). Of the nine validity
correlations, only five exceeded the critical value of .27 (for a sample size
of 57 at the .05 significance level). Although the evidence for convergent
validity is somewhat weaker than might be desired, it is not completely
unexpected. Marsh (1982, p. 275) reported a median correlation of only .30
between instructor self-ratings and ratings by their students. When the data
of Blackburn and Clark (1975) are conceived of as a MTMMM, the validity
diagonals show a median value of .45 with a range from .10 to .63. The results
of the present study are not inconsistent with those values.

Discriminant validity was investigated with three criteria. First, two
methods of measuring the same trait should correlate higher than two measures
having neither trait nor method in common. That is, a value in a validity
diagonal should be higher than the other values in its column and row in the
heterotrait-heteromethod triangles. Table 1 shows that the three validity
values for Teaching (.20, .36, and .34) were each larger than the values in
the corresponding rows and columns within the dashed triangles. The variable
Research showed similar results, but Service exhibited weaker evidence.

The second criterion states that two methods of assessing a particular
variable should show a higher correlation than should the same method when
applied to that particular variable and some other variable. That is, the
validity values in a particular column (or row) should be higher than the
heterotrait-monomethod values in the same column (or row), if method effects
are to be eliminated. From the data in Table 1, there was an indication of
strong method effects. For example, when the validity values of .20 and .36
for Teaching were compared with the values of .23 and .56 in the same column,
a method effect for department chair ratings could not be ruled out. While
the validity coefficients, on the average, tended to be higher than the
corresponding heterotrait-monomethod values, there were still enough exceptions
to give cause for concern.

The third criterion suggests that a similar pattern of trait inter-relation-
ships should be shown regardless of the method used. Thus, a similar pattern
of correlations should be found within each of the triangles. However, from
the data in Table 1, it was difficult to detect any consistent pattern within
the triangles.

Discussion

Applying the MTMMM approach to the measures of faculty competence in this
study led to somewhat ambiguous results, because the required conditions were
not fully met for each criterion. Some writers have suggested that, in certain
cases, path analysis can provide less ambiguous results than the Campbell-Fiske
criteria (Schmitt, 1978). However, since path analysis was primarily developed
to handle situations involving large numbers of traits and methods, the
Campbell-Fiske analysis appeared more appropriate for the present study,
considering the few traits and variables therein.

There are several factors that should be considered in interpreting the

findings of this study. First, the sample size was small. No doubt, more stable results would have appeared with a larger number of subjects, but research using confidential type data often must be restricted to relatively small sample sizes. Second, certain properties of the instruments might be questioned. Were the chair evaluations given for this study the same that the chair actually used when making official reports? Was the investigator-made student evaluation instrument as reliable and valid as instruments used elsewhere? Was the number of publications and the number of committees synonomous, respectively, with the quality of research and service? Again, the nature of the study imposed certain limitations on the types of operational definitions that could be used for the constructs. Third, the three traits may each have been a composite of other traits. When Marsh (1982) factor analyzed student rating scales, the results indicated several distinct factors, and he suggested that assuming student ratings to be uni-dimensional may be one factor leading to ambiguous results.

While the above limitations should be considered in interpreting the results, the fact that confidential data of this type are difficult to obtain suggests both that these tentative conclusions may be of value to other researchers, and that further investigations are warranted.

In summary, the MTMMM analysis indicated that the three hypothesized components of faculty competence could be distinguished one from another, but that the evidence for their convergent and discriminant validity was relatively weak. Therefore, while having some value, data purporting to measure such traits should be interpreted with considerable caution.

Table 1

Multitrait - Multimethod Matrix of Measures of Faculty Competence (N = 57)

Traits	Dept. Chair			Self			Other		
	T1	R1	S1	T2	R2	S2	T3	R3	S3
1. Dept. Chair									
Teaching (T1)	(88)								
Research (R1)	23	(83)							
Service (S1)	56	24	(72)						
2. Self									
Teaching (T2)	20	04	17	(73)					
Research (R2)	-28	43	-27	27	(60)				
Service (S2)	05	21	38	33	21	(66)			
3. Other									
Teaching (T3)	36	02	-02	34	01	05	(NA)		
Research (R3)	-01	35	05	-22	23	17	-15	(24)	
Service (S3)	-04	00	12	-13	00	20	-15	-25	(46)

Note. Parentheses are used to identify the values in the reliability diagonal. (The reliability of the Teaching variable was not available for these data.) Underlines are used to identify the values in the validity diagonals. Solid line triangles enclose the heterotrait-monomethod values. Broken line triangles enclose the heterotrait heteromethod values. The intercorrelations among the Other methods are not enclosed by solid lines because they are not mono-method. Decimal points are omitted.

References

Bajtelsmit, J. W. (1979). Convergent and discriminant validation of chartered life underwriter (CLU) examinations by the multitrait_multimethod matrix. Educational and Psychological Measurement, 39, 891-896.

Blackburn, R. T., & Clark, M. J. (1975). An assessment of faculty performance: Some correlates between administrator, colleague, student, and self-ratings. Sociology of Education, 48, 242-256.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-195.

Centra, J. A. (1979). Determining faculty effectiveness. San Francisco: Josey-Bass.

Hansen, R. B. (1979). The determinants of faculty salaries in selected depart-ments at the University of Nebraska-Lincoln, Dissertation Abstracts, 40, 1603.

Kenny, D. A. (1979). Correlation and Causality. New York: John Wiley & Sons.

Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. Journal of Applied Psychology, 51, 369-381.

Marsh, Herbert W. (1982). Validity of students' evaluation of college teaching: A multitrait-multimethod analysis. Journal of Educational Psychology, 74, 264-279.

Miller, R. I. (1979). The Assessment of College Performance. San Francisco: Josey-Bass.

Schmitt, N. (1978). Path analysis of multitrait-multimethod matrices. Applied Psychological Measurement, 2, 147-173.