USE OF SERVICES QUALITY THEORY AS THE BASIS FOR STUDENT EVALUATIONS OF TEACHERS: A PRELIMINARY STUDY

Joseph L. Orsini, California State University, Sacramento

ABSTRACT

This study is a preliminary investigation of the feasibility of using a general theory of services quality in an application of measuring college teaching quality. The results are generally supportive, with the predictiveness of the test instrument being comparable to an existing instrument, which was developed in the college education industry.

INTRODUCTION

Teaching effectiveness is one of the most important aspects of faculty development and promotion in higher educational institutions (Hildebrand, Wilson and Dienst 1971). While various means of evaluating teaching are used, the most popular continues to be those performed by students (McCallum 1984). The education literature is rich in discussions of the use of student ratings of faculty (Cohen 1981), where concerns generally center on the composition of the particular instrument being used (Lammers and Kirchner 1985).

The purpose of this paper is to perform an exploratory study regarding the application of emerging theory in services quality to college teaching services. Characteristics of general services quality are applied specifically to the college teaching industry, and an instrument designed to test the predictiveness of these characteristics. A comparison is made between this services quality instrument and an existing teaching quality instrument, which was developed within the college education industry in accordance with accepted procedures.

HYPOTHESES

Within services industries generally the employment of user ratings of services quality is firmly established, both theoretically (Lewis and Booms 1983) and in application (Center for the Study of Services 1981). As with teaching evaluation, services generally are concerned with the selection of service quality attributes, and the investigation of consumer and service characteristics which may affect perceptions of service quality. A typical approach in the services industries is to determine the relevant characteristics which consumers use in their service selection decision, then survey a group of consumers in order to see the predictiveness of the respective characteristics.

The emergence of the importance of services theory in marketing has led to the proposal of a series of service quality characteristics which are thought to be generally applicable to all service industries. The study by Parasuraman, Zeithaml and Berry (1985) proposes ten universally applicable service quality characteristics, which are all evaluated in this study: Reliability, Responsiveness, Competence, Access, Courtesy, Communication, Credibility, Security, Understanding, and Tangibles. The hypotheses investigated in this study are that these ten items are significantly related to overall teaching quality.

METHODOLOGY

A descriptive research design was used in this study, where the ten characteristics proposed by Parasuraman, et al. (1985), were included in a questionnaire for college students' evaluation (see Figure 1). A comparison was made between the predictiveness of these quality measures and an existing instrument, widely noted in the literature, devised by Hildebrand, et al. (1971).

The ten basic characteristics were evaluated as complete concepts in this preliminary study rather than attempting to differentiate any subcomponents that may exist. This is a practice also undertaken in the college education industry (March 1984). The general concepts indicated in Parasuraman, et al. (1985), were subjectively reformulated to apply to the specific industry, in accordance with the terminology currently used in the industry literature.

The subjects utilized were a convenience sample of 136 undergraduate business majors at an urban campus of the California State University. Industry practice was utilized (Marsh 1984), where half of the students were randomly selected to use both instruments to evaluate a "good" teacher of their choice; the other half were to evaluate a "poor" teacher of their choice. In addition, half the students were randomly assigned to first utilize the test instrument prior to using the existing instrument, while the reverse was true for the other half.

Several kinds of analysis were performed. Distributions of the characteristics were examined to assure a variety of responses. The association of each characteristic with the dependent variable (overall evaluation of teacher quality) was analyzed to determine the predictiveness of each characteristic. Correlations among the service quality characteristics was performed both by simple bivariate correlation and factor analysis. The overall predictiveness of the characteristics was analyzed by regular stepwise regression and by principle components regression. Analysis of both

the test instrument and the existing instrument was performed for comparison purposes.

RESULTS

The results of the examination of service quality characteristic means and standard deviations, and their individual predictiveness of overall quality, is indicated in Table 1. Generally, the mean level of the ratings was high, averaging about 3.5 on the six-point scale (where 0 indicated the absence of the characteristic and 5 the full presence of the characteristic). The standard deviation averaged just under one and one-half scale points. These would appear to be reasonable results, given that faculty generally are viewed as capable of performing their job (Hildebrand, et al. 1971). The bivariate correlations are all highly significant ($p < .001$), thus supporting the hypotheses. Generally, a one-point increase on the characteristic quality scale corresponded to a one-point increase on the overall teaching quality scale.

As would be anticipated from the correlations indicated in Table 1, the quality ratings of the test instrument are moderately correlated with each other. Correlations among the ten items range from .48 to .79 ($p < .01$; not shown), averaging about .60. Application of a factor analysis (varimax rotation) to the ten measures resulted in the production of one factor, further indicating the strength of the associations. For comparison purposes, the Hildebrand, et al. (1971) items resulted in four factors, although the first contained 88 percent of the variance. The bivariate correlations between the latter instrument's individual items and the measure of overall quality were slightly higher than those of the service quality instrument items; this will be discussed further below.

Stepwise regression was utilized to assess the overall predictiveness of the measures. As indicated in Table 2, the characteristics Credibility, Communication, Reliability and Courtesy produced the best model (defined as that model where no additional variable is able to increase the R-squared more than .01). This model accounted for 80 percent of the variance (adjusted R-squared). A principal components regression reflected the above findings (not shown), with the one resulting component having an R-squared of .82. For comparison purposes, the principal components analysis of the Hildebrand, et al. (1971) instrument resulted in the three components (not shown) with an R-squared of .91. However, the R-squared of the first component alone was .90 (the BMDP4R program automatically enters components in stepwise fashion). Further, when the individual items were combined into their theoretical constructs (Hildebrand, et al. 1971), and entered into a stepwise regression, the two variables in the best model yielded an adjusted R-square of .72. Thus, when both instruments were tested in their combined constructs form, their levels of predictiveness were quite comparable. Stepwise regression of the 35 individuals items of the Hildebrand, et al. (1971), instrument also resulted in the production of a model with 4 items only, and an adjusted R-square of .91.

DISCUSSION AND CONCLUSIONS

The purpose of this study was to serve as a preliminary investigation of the hypothesis of a general theory of services quality, applied as a measure of college teaching quality. Insofar as the theory developed by Parasuraman, et al. (1985), was based on four industries unrelated to teaching (i.e., the retail banking, credit card service, securities brokerage, and product repair and maintenance industries), this was a substantial test.

The results of this test as an application of theory were generally favorable. The approach used was to translate the ten service quality concepts into language relevant to the college teaching industry, then use these concepts as predictors of overall teaching quality. The resulting predictiveness was comparable to a carefully prepared instrument developed in the industry and tested under similar conditions.

While the results of this preliminary investigation were favorable to the theory, the findings should be interpreted with caution. In addition to the non-random and comparatively small sample used, subjective judgment was employed in the translation of concept to applied measure. Possibly an excessively "liberal" interpretation of the concepts allowed the resulting favorable predictiveness to occur where it was not warranted. Particularly troublesome in this regard are the results of the factor analysis, which showed only one factor instead of the hypothesized ten. While the combination of small sample and generally adequate product quality could produce this result, as indicated by the heavy loading on the first factor of the Hildebrand, et al. (1971) instrument, certainly further investigation is needed.

Further research would be warranted even if these issues were not present. This investigation used the ten theoretical constructs as complete items, where frequently they appear to be reasonably composed of items which should be tested separately and formed into a scale, if necessary. The responsiveness construct, for example, may better be divided into two items, as it would appear plausible to find faculty who answer questions quickly, but are slow to return exams and papers.

Further research should also be undertaken as an enhancement of theory development. In true interdisciplinary fashion, the findings of one discipline may enhance theory development in another. For example, some items in the Hildebrand, et al. (1971) instrument regarding teaching style (energy, enthusiasm, and a sense of humor) do not seem to be represented among the ten service quality constructs, thus may be a relevant addition. Quite possibly further research would reveal others. Given the importance of services and their quality, further research in this area would seem to be strongly warranted.

FIGURE 1

Test Instrument

CALIFORNIA STATE UNIVERSITY, SACRAMENTO
School of Business and Public Administration

SDOT II

	Not at all Descriptive					Very Descriptive
RELIABILITY--instructor is consistent and dependable in performance with adequate justification for schedule changes; promises are honored.	0	1	2	3	4	5
RESPONSIVENESS--instructor provides timely responses to questions; exams and other materials are returned in a timely manner.	0	1	2	3	4	5
COMPETENCE--instructor possesses the required skill and knowledge in the subject of the course.	0	1	2	3	4	5
ACCESS--instructor is reasonably available outside of class, either in person or by telephone.	0	1	2	3	4	5
COURTESY--instructor shows politeness, respect, consideration and friendliness toward students.	0	1	2	3	4	5
COMMUNICATION--instructor uses a language level understandable to students; concepts are clearly explained and questions fully answered.	0	1	2	3	4	5
CREDIBILITY--instructor is trustworthy and believable; has the students' best interests at heart.	0	1	2	3	4	5
SECURITY--instructor keeps students informed, as much as feasible, regarding their grade standing in the course.	0	1	2	3	4	5
UNDERSTANDING--instructor makes the effort to: understand student needs, provide individualized attention as feasible, and recognize students outside of class.	0	1	2	3	4	5
TANGIBLES--instructor makes effective use of available visual media (chalkboard, overhead projector, etc.); personal appearance is clean and neat.	0	1	2	3	4	5
OVERALL RATING --instructor is among the best I have had at CSUS.	0	1	2	3	4	5
--instructor is among the best I have had at the School of Business and Public Administration.	0	1	2	3	4	5
Other Criteria--_____	0	1	2	3	4	5
_____	0	1	2	3	4	5

Comments--_____

_____

TABLE 1

Test Instrument Statistics

Variable	Mean	Deviation	Bivariate Coefficients Corr.	Regr.
Reliability	3.64	1.32	.78	1.12
Responsiveness	3.55	1.45	.68	.89
Competence	3.99	1.21	.71	1.10
Access	3.65	1.13	.60	1.01
Courtesy	3.47	1.46	.74	.96
Communication	3.50	1.41	.79	1.07
Credibility	3.50	1.44	.83	1.08
Security	3.27	1.49	.64	.82
Understanding	3.31	1.44	.79	1.04
Tangibles	3.72	1.29	.60	.90
Overall Rating	3.07	1.88	1.00	--

TABLE 2

Results of Stepwise Regression of Quality Characteristics

Variable	Coef.	S. E. Coef.	F	Sig.
Credibility	.31	.11	8.6	.004
Communication	.38	.09	19.4	.000
Reliability	.42	.09	22.0	.000
Courtesy	.27	.08	11.4	.001
Constant	-1.82	.24	56.2	.000

Adjusted R-square = .80
Overall F = 120.4 @ d.f. = 4 and 116

REFERENCES

Center for the Study of Services (1981), Service Evaluation, 1, Washington, D.C.: Center for the Study of Services.

Cohen, Peter A. (1981), "Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisection Validity Studies," Review of Educational Research, 51 (Fall), 281-309.

Hildebrand, Milton, Robert C. Wilson, and Evelyn R. Dienst (1971), Evaluating University Teaching, Berkeley: Center for Research and Development in Higher Education.

Lammers, H. Bruce, and Don F. Kirchner (1985), "The Construct Validity of Teaching Behavior Evaluation Methods: A Multitrait-Multimethod Analysis," Journal of Marketing Education, 7 (Summer), 35-44.

Lewis, Robert C., and Bernard H. Booms (1983), "The Marketing Aspects of Service Quality," in L. Berry, G. Shostack, and G. Upah (eds.), Emerging Perspectives on Services Marketing, Chicago: American Marketing Association.

Marsh, Herbert W. (1984), "Students' Evaluations of Teaching: Dimensionality, Reliability, Validity, Potential Biases, and Utility," Journal of Educational Psychology, 76, 707-754.

McCallum, L. W. (1984), "A Meta-Analysis of Course Evaluation Data and its Use in the Use in the Tenure Decision," Research in Higher Education, 21, 150-158.

Parasuraman, A., Valerie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and its Implications for Future Research," Journal of Marketing, 49 (Fall), 41-50.