Evaluation of Effective Teaching Using Student Ratings of Instruction: By Factor Analysis Approach

Nidhi Joshi*
Nitish Ghune**

Abstract

The broad objective of the study is to understand the factors considered by students in evaluating their Guides. Means, we are interested in indentifying the determinants (or factors), on the basis of which Students may judge their Guides behavior towards them, knowledge about their Subject and ability of their guidance. The study uses factor analysis to determine the validity and reliability of the evaluation instrument in assessing instructor or course effectiveness in principles of economics courses. Results show positive associations between student perception of teaching effectiveness, learning facilitation, effective communication, and clarity of course elements, and course evaluation and feedback.

Keywords: Student Rating, Effective Teaching, Factor Analysis.

Introduction

In recent years, student evaluations of Guide performance have become increasingly common in college campuses across the nation. In many college campuses, administrators use this result as a major determiner for making critical retention, promotion, and merit pay decisions about individual faculty members. For administrators, the attractiveness of student evaluations of faculty is that they provide an easy, seemingly objective assessment of teaching that does not require justification (Stone, 1995). The most common type of commercial student evaluation form utilizes a Likert-type scale for students to rate faculty related to a series of statements about the Behavior, Knowledge of topic, course and instruction. Among the three major approaches to faculty evaluation- Student rating, peer rating, and self assessment- student rating are most widely used because they provide a structured, systematic and economical way to obtain feedback on students' reactions to instructor and course. A large body of research shows that students rating, if obtained properly, are a reliable and moderately valid way of measuring teaching effectiveness. Using factor analysis we investigated the underlying factors related to the items on the survey which revealed factors relating to personal attributes of the professor facilitator, learning facilitation and quality of feedback. On the basis of these two factors:

1. Friendliness towards students

2. Knowledge and Guiding Ability on the subject

The goal of this study was to gather objective data regarding guide's attitudes and knowledge towards the conduct of research and scholarly activity in order to develop a plan for increasing productivity.

Review of Literature

A vast literature argues that teaching is a multidimensional process comprising a number of separable dimensions or guide attributes. For a sample of more recent research in the field, see Glynn, Sauer and Wood, 2005, and Dennis, 2009. There continues to be robust debate and discussion about the findings of this extensive body of research and what conclusions can be drawn about student evaluations of teaching and their use. According to Algozzine et al, 2004, student evaluation of teaching is a very complex and controversial issue with inconsistent research findings (p.138), while Kulik, 2001 argues that some studies on student evaluation of teaching SETs are “conflicting, confusing, and inconclusive” (p.10). Nevertheless, Kulik agrees with other studies that show that these evaluations are reliable, and valid measures of teaching effectiveness (Centra, 2003; Marsh, 1987; Penny, 2003; Spoorens and Martelman, 2006).

Research shows that students tend to take teaching evaluations more seriously than faculty and institutional members commonly believe. Students are

*Assistant Professor, Pioneer institute of Professional Studies, Indore (M.P)
**Research Fellow, Indian Institute of Technology, Roorkee (UK)

more willing to participate and offer meaningful feedback when they believe and can see that their input is being considered and incorporated by their instructors and the institution. In general, however, students do not perceive that their feedback is often used. Some studies show that students place most value on evaluations for formative purposes, but research also indicates that students believe their input should be considered for summative purposes. Students would like to see more specific items related to teaching effectiveness on student evaluation of teaching instruments (Sojka & Deeter-Schmetz, 2002; Chen & Hoshower, 2003).

Research also shows that faculty often believes that students do not take evaluations seriously and that ratings encourage grade leniency. Nonetheless, most faculties do pay attention to student feedback. Further, when evaluations are used for formative purposes, instructors show a high degree of motivation to improve their teaching based on student input. Studies have emerged showing how institutions and individual faculty members have begun using evaluations, consultations, and portfolios to improve instruction qualitatively. When faculty are well informed about the purposes of evaluation, much of their anxiety dissipates and willingness to learn from student feedback increases (Sojka, Gupta, & Deeter-Schmetz, 2002; Hativa, 1995; Gallagher, 2000; Bain, 2004). Teaching evaluations are commonly considered for summative purposes, including tenure, merit increase, retention for non-tenured faculty, promotion, and course assignment decisions. While research generally agrees that teaching evaluations offer an effective and meaningful way to inform these decisions, often such data are misused, misinterpreted, or overused. Some institutions use student ratings data as the sole criterion for evaluating teaching effectiveness, and these institutions often use only global items on student ratings forms to construct their evaluation. Such misuse can breed distrust between faculty and administrators, resentment on the part of instructors for evaluations, and hinder other formative uses of these data.

Instead, researchers recommend that the focus of student evaluation of teaching should be on desired educational outcomes and whether these outcomes are being achieved or not. In this regard, if student ratings forms are to be used, the instruments must be subjected to rigorous validity tests and analysis. In addition, the student-rating data is best used in combination with other criteria in order to provide a better assessment of teaching, which is a multidimensional construct.

Using evaluations to inform instructors of their teaching effectiveness and to aid them in improving or enhancing their teaching constitute the formative purposes of teaching evaluations. When used to inform teaching practices, specific dimensions of teaching must be identified and focused upon in order to bring about change. Research indicates that evaluations are most effective in improving teaching when faculty members understand and value the importance of such processes, and an institutional and departmental culture that supports and respects teaching is evident. Evaluation systems for formative purposes often encompass more than just student ratings of Guide effectiveness. Several colleges and universities have begun using portfolios, peer observation, self-review, and more qualitative approaches to improve teaching. Similarly, recent establishment of faculty development centers on many campuses reveals a trend toward investing in the formative uses of evaluations. See, Hobson & Talbot, 2001; Hoyt & Pallett, 1999; Theall & Franklin, 2001; Kulik, 2001; Gallagher, 2000; Johnson & Ryan, 2000; Hativa, 1995; Bain, 2004. Journal of College Teaching & Learning – May 2010 Volume 7, Number 559 Some of the myths about the usefulness of student ratings begin from faulty research studies, conflicting findings within the research literature, or reluctance on the part of some administrators and faculty to evaluate and be evaluated, respectively. Some common myths include students are not able to make informed and consistent judgments about their instructors; student ratings are essentially a popularity contest; students cannot make accurate judgments unless they have been away from the course for a while; student ratings are negatively related to student learning; student ratings are based upon expected grade in course.

Methodology

Data Collection

For Data collection we did survey of 64 students to find out the factors from which they evaluate their research Guide. For that we prepared questionnaire in which we covered 14 questions and asked individually that how they evaluate their Guide.

14 questions reflecting the students' attitude towards Guides personality, competence, knowledge and about

their prescribed notes or assignment, each measured on a scale ranging 1-5.

Respondents (Students) are typically asked to express their level of agreement on a scale (e.g.:5-point Likert scale: 1."strongly agree ", 2."agree", 3."neutral", 4. "disagree", 5. "strongly disagree").

We can say: This was a quick study done to understand how students evaluate their Guides. A survey was conducted using a questionnaire. The responses were analyzed using principal component (factor) analysis. From the results, variables were clubbed together for two factors:

1. Friendliness (towards students): which had questions on Manner, Interpersonal behaviour, respect for student, interest in student's work, confidence in the student, and freedom given to student to ask questions.

2. Knowledge and Guiding Ability (on topic of research): questions on manner, well trained and competent, good examine skills, great command on his research topic etc.

- Friendliness towards students: includes Q1, Q3, Q6, Q9, Q13, Q14.
- Knowledge and Guiding Ability over the Topic of research: includes Q2, Q4, Q5, Q7, Q8, Q10, Q11, and Q12.

Further, some questions are such that together they test one factor but in opposite directions. For example, Q2 and Q8. Both these questions test Knowledge but in opposite direction. Which means the scores which the respondents gave is higher for one and lower for other. These may be more such questions which can be grouped.

Method Employed

To understand the factors considered by students in evaluating their research Guides we have used **Principal Component (Factor) Analysis.**

Factor Analysis

- Factor analysis was first used in 1904 by Charles Spearman, a UK Psychologist.
- Spearman used factor analysis in his models of human intelligence. With factor analysis he developed theory that variety of cognitive test could be explained by one single factor.
- Factor analysis is quiet different to other statistical analysis method. Many statistical analysis methods are used to investigate the relation between independent and dependent variables.
- In contrast, factor analysis is used to study the patterns of relationship among many dependent variables, with the goal of discovering something about the nature of independent variables that affect them.
- Use of Factor analysis in psychology is most often associated with intelligence research. The use of factor analysis can also be observed in other domains of psychology such as personality, attitudes, etc.
- Factor analysis is used widely in other social sciences (investigating human behavior), education, business fields, biological science, etc.

Objectives of Factor Analysis

The main applications of factor analytic techniques are:

(1) To reduce the number of variables

(2) To detect structure in the relationships between 2 variables, that is to classify variables. Hence, FA is applied as a data reduction or structure detection method (Statsoft, 2003).

Reducing the number of variables means reducing complexity of data, thus it will be easier to analyze the data.

Empirical Analysis

The word Empirical denotes information acquired by means of observation or experimentation. Empirical Analysis is a way of gaining knowledge by means of direct and indirect observation or experience. Empirical evidence (the record of one's direct observations or experiences) can be analyzed quantitatively or qualitatively. In scientific use the term empirical refers to the gathering of data using only evidence that is observable by the senses or in some cases using calibrated scientific instruments. What early philosophers described as empiricist and empirical research have in common is the dependence on observable data to formulate and test theories and come to conclusions.

After applying factor analysis method on the survey data of Research Scholars' evaluations of their guides, we get following results with its interpretation.

Results With Interpretation

Kaiser-Meyer-Olkin (KMO) and Bartlett's Test

This table shows two tests that indicates the suitability of your data for Structure Detection.

KMO =.881 indicates the proportion of variance in the variables that might be caused by underlying factors. KMO value is very close to 1 which indicates that factor analysis may be useful with our data.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.881
Bartlett's Test of Sphericity	Approx. Chi-Square	608.062
	Df	91
	Sig.	.000

Bartlett's test measure the significance level, 0.001 indicates factor analysis may be appropriate with our data.

The KMO measures the sampling adequacy of data means it tell us that the data which you have taken are sufficient to meet a need satisfactory or not. Large values for the KMO measure indicate that a factor analysis of the variables is a good idea. KMO statistic varies between 0 and 1. A value 0 indicates that the factor analysis is likely inappropriate. And value 1 indicates that patterns of correlation are relatively compact and factor analysis should yield distinct and reliable factor. For these data the value is 0.881, which is seems to be good; so, we should be confident that factor analysis is appropriate for this data.

Bartlett' Test measures null hypothesis that the original correlation matrix is an identity matrix. For factor analysis to work we need some relationships between variables and if R-matrix were an Identity matrix then all correlation coefficients would be zero. Therefore we want this test to be significant (i.e. have a significance value less than 0.05). A significant test tells us that R-matrix is not an identity matrix; therefore there are some relationships between the variables. For this data Bartlett Test is highly significant ($p<0.001$), and therefore factor analysis is appropriate. Another indicator of the strength of the relationship among variables is Bartlett's test of sphericity. Bartlett's test of sphericity is used to test the null hypothesis that the variables in the population correlation matrix are uncorrelated. The observed significance level is .0000. It is small enough to reject the hypothesis. It is concluded that the strength of the relationship among variables is strong. It is a good idea to proceed a factor analysis for the data.

Communalities

Communalities indicate the amount of variance in each variable that is accounted for.

Initial communalities are the estimates of the variance in each variable accounted for by all the components or factors. For principal components extraction, this is always equal to 1.0 for correlation analysis.

Extraction communalities are the estimates of the variance in each variable accounted for by the components. The communalities in this table are almost high, which indicates that the extracted component represent the variables well. If any communality is very low in a principal components extraction, we may need to extract another component.

Total Variance Explained

This output table explained the variance by the initial solution, extracted component, and rotated component.

The first section of the table shows the Initial Eigen values.

The Total column gives the eigenvalue, or amount of variance in the original variables accounted for by each component.

The % of Variance column gives the ratio, expressed as a percentage, of the variance accounted for by each component to the total variance in all of the variables. Factor 1 explained 55.424% of total variance.

The Cumulative % column gives the percentage of variance accounted for by the first n components. For example, the cumulative percentage for the second component is the sum of the percentage of variance for the first and second components.

For the initial solution, there are as many components as variables, and in a correlations analysis, the sum of the eigenvalues equals the number of components. You have requested that eigenvalues greater than 1 be extracted, so the first three principal components form the extracted solution.

The second section of the table shows the extracted components. They explain nearly 63% of the variability in the original 14 variables, so you can considerably reduce the complexity of the data set by using these components, with only a 37% loss of information.

The rotation maintains the cumulative percentage of

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.759	55.424	55.424	7.759	55.424	55.424	4.849	34.635	34.635
2	1.007	7.192	62.616	1.007	7.192	62.616	3.917	27.982	62.616
3	.968	6.917	69.534						
4	.852	6.087	75.620						
5	.649	4.635	80.255						
6	.541	3.867	84.122						
7	.502	3.588	87.710						
8	.426	3.042	90.752						
9	.403	2.879	93.631						
10	.302	2.155	95.786						
11	.181	1.295	97.081						
12	.159	1.138	98.219						
13	.148	1.055	99.274						
14	.102	.726	100.000						

Extraction Method: Principal Component Analysis.

variation explained by the extracted components, but that variation is now spread more evenly over the components. The large changes in the individual totals suggest that the rotated component matrix will be easier to interpret than the unrotated matrix.

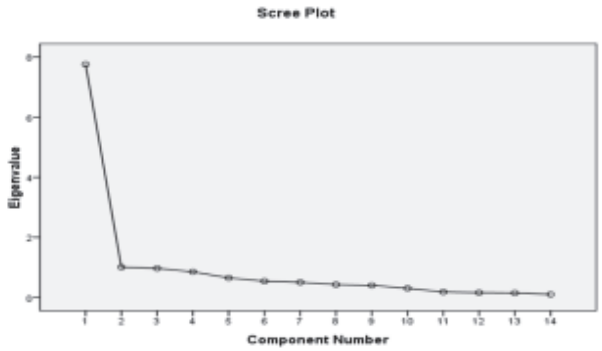
The cumulative variability explained by these two factors in the extracted solution is about 63%.

Before extraction SPSS has identified 14 linear components within the data set. The eigenvalues associated with each factor represents the variance explained by that particular linear component and SPSS also displays the eigen value in terms of percentage of variance explained.

Screen Plot

The Screen plot helps you to determine the optimal number of components. The eigen value of each component in the initial solution is plotted.

Generally, we want to extract the components on the steep slope. The components on the shallow slope contribute little to the solution. The last big drop occurs between the second and third components, so using the first two components is an easy choice.



Rotated Component Matrix

The rotated component matrix helps you to determine what the components represent.

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.The first component is most highly correlated with interest in me, cold & interpersonal, friendly manner, confide problem.

The second component is most highly correlated with doubt treatment, unanswered questions.

Factor 1 is responsible for understanding friendliness which comes from questions:

Q9 explained 85%

Q3 explained 83%

Q1 explained 74%

Q13 explained 70%

Factor 2 is responsible for understanding knowledge and ability of guidance on the topic which comes from questions:

Q7 explained 89%

Q10 explained 75%

We can reduce the size of the data file from 14 variables to 2 factors by using Factor Analysis with a principal components extraction. We have uncovered two latent factors that describe relationships between our variables. These factors suggest various patterns of evaluations, which we can be used more efficiently to find out students viewpoints towards their guides.

Conclusion

As per the survey questions q1,q3, q6,q9,q13,q14 used to explain friendliness of guide towards students and q2,q4,q5,q7,q8,q10,q11,q12 explain guide's knowledge and guiding ability over the topic. But the result shows that q1,q3,q9,q13 variables explain guide friendliness and q7,q10 explain guide knowledge and guiding ability because all of these are greater than 0.70.

Our model explained nearly 63% of variability in 14 variables. Means, variation explained by 2 factors friendliness, knowledge and guiding ability is 63%. There is about 63% chances that students can evaluate their guide through this model which is good explained.

So, by rotated component matrix table we can conclude that our hypothesis 1 i.e. friendliness exists towards students & hypothesis 2 i.e. guide has excellent knowledge and good guiding ability over the topic both are accepted which means guides are friendly with the students and having excellent knowledge and good guiding ability over the topic.

On an average factor 1 explained 77% of guide friendliness and factor 2 explained 81% of guide's knowledge and guiding ability.

References

1. Journal of College Teaching and Learning – May 2010. Evaluating Effective Teaching in College Level Economics Using Student Ratings Of Instruction: A Factor Analytic Approach.
2. Research and Scholarly Activity: Faculty Attitudes and Awareness. National forum of applied educational research journal volume 20, number 3, 2006.
3. A Multilevel Factor Analysis of Students' Evaluations of Teaching. Michael D. Toland and R. J. De Ayala, Educational and Psychological Measurement 2005.
4. Student Evaluation of Faculty: An Analysis of Survey Results, U21Global Working Paper Series, No. 003/2008.
5. Abrami, P.C., D'Apollonia, S. & Cohen, P.A. (1990) "Validity of Student Ratings of Instruction: What We Know and What We Do Not," Journal of Education Psychology, 82(2), pp. 219-231.
6. Aleamoni, L. (1987) "Student Rating: Myths Versus Research Facts," Journal of Personnel Evaluation in Education, (1) pp.111-119.
7. Chen, Y. & Hoshower, L.B. 2003. "Student Evaluation of Teaching Effectiveness: an Assessment of Student Perception and Motivation." Assessment and Evaluation in Higher Education 28(1): 71-88.
8. Glynn, Joseph. C., P. Sauer, and G. Wood, (2005), "Evaluating Evaluations: A Review of Student.
9. Marsh, H. W. & Roche, L.A. (1997). "Making Students' Evaluation of Teaching Effectiveness: The Critical Issues of Validity, Bias, and Utility." American Psychologist 52(11): 1887-1197.

Appendix

14 questions reflecting the students' attitude towards guide's personality, competence, knowledge and about their prescribed notes or assignment, each measured on a scale ranging 1-5.

Respondents (Students) are typically asked to express their level of agreement on a scale (e.g.:5-point Likert scale: 1."strongly agree ", 2."agree", 3."neutral", 4. "disagree", 5. "strongly disagree").

Students' evaluation of guide: A case

- 14 questions reflecting the students' attitude towards guides personality, competence, knowledge and about their prescribed notes or assignment, each measured on a scale ranging 1-5

- Respondents (Students) are typically asked to express their level of agreement on a scale (e.g.: 5-point Likert scale: 1. "strongly agree ", 2. "agree", 3. "neutral", 4. "disagree", 5. "strongly disagree").

Questions which get asked for Evaluation are:

Q1	My Guide treats me in a friendly manner	Friendly manner
Q2	I have some doubts about the ability of Guide	Doubts ability
Q3	My Guide seems cold and impersonal	Cold and interpersonal
Q4	My Guide does his/her best to keep me from worrying	Reassurance
Q5	My Guide examines me carefully as necessary	Careful examination
Q6	My Guide should treat me with some respect	Not enough respect
Q7	I have some doubts about the problem suggested	Doubts treatment
Q8	My guide seems very competent and well trained	Competent and trained
Q9	My guide seems to have a genuine interest in me as a person	Interest in me
Q10	My guide leaves me with many unanswered questions about my queries and doubts.	Unanswered questions
Q11	My guide uses words that I do not understand	Jargon
Q12	I have a great deal of confidence in my guide	Confidence
Q13	I feel I can tell my guide about my personal problems	Confide problems
Q14	Not free to ask questions	Unfree asking questions