Assessment & Evaluation in Higher Education
Vol. 30, No. 2, April 2005, pp. 123–136

Student evaluations of course and teacher: factor analysis and SSA approaches

Erik H. Cohen*
Bar Ilan University and Ashkelon Academic College, Israel

The present study attempts to discover organizing principles in college students' evaluations of their formal and informal educational programs. The study was conducted during the 1998–99 academic year. The data consisted of 14,344 assessment sheets completed by students at the Jordan Valley Regional College in Israel. Each assessment sheet contained 12 variables. Factor analysis revealed two main factors: the course (seven variables) and the teacher (five variables). A Smallest Space Analysis (SSA) revealed an additional distinction. Some of the evaluation items relate directly to the course or teacher and some relate to the interaction between the course or teacher and the students. Distinguishing between these categories could be helpful in identifying or correcting bias in student evaluations.

Introduction

Student evaluations of their education have gained importance as competition to attract and keep students has become more intense among the ever-proliferating number of institutes of higher education. For the past century, students' relationship to their schools and teachers has been once of the most widespread topics of educational research. Cashin (1988) cites over 1300 separate studies on student evaluations of teachers alone. Many studies have focused on comparative relations according to group differences such as gender, socioeconomic status, type of school, and achievements (Amin, 1994; Centra & Gaubatz, 2000; Radmacher & Martin, 2001; Schmidt, 2003). Student evaluations have been the subject of much controversy concerning whether they indicate actual quality in learning and teaching, or if students favor teachers and courses with little work and easy grading (Felder, 1992; Greenwald & Gillmore, 1997; Marsh & Roche, 2000; Kaplan *et al.*, 2002; Johnson, 2003; among many others).

*38 Bethlehem Road, Jerusalem 93504, Israel. Email: msaharon@mscc.huji.ac.il

ISSN 0260-2938 (print)/ISSN 1469-297X (online)/05/020123–14
© 2005 Taylor & Francis Ltd
DOI: 10.1080/0260293042000264235

In his comprehensive survey, Lewy (1985) noted the meagerness and non-cumulative character of studies regarding the students' relation to their schools as well as the absence of a theoretical framework that would allow for interpretation of the research data. There has been little attention given to the structural nature of the evaluation. The few noteworthy studies that have done so (Yee & Runkel, 1969; Darom *et al.*, 1980) have had little impact on subsequent research. While the relationship to schools has been treated as a multivariate concept, this has generally been through the setting of scales (one or more) based on a number of items that do not necessarily cover the totality of the content of the items. Recently, Greimel-Fuhrmann and Geyer (2003) published a structural analysis of students' evaluations of their teachers.

For guidance in formulating a hypothesis about the organizing principles underlying student evaluations of courses and teachers, we may look to studies in which organizing principles underlying quality of life for students were identified. Williams and Batten, in their 1981 study of quality of life among high school students identified six areas of evaluation: positive affect, negative affect, status, identity, opportunity and professors. Roberts and Clifton, in their 1992 study of quality of life among university students, suggest that the six areas may be re-conceptualized as four dimensions in college and university milieus: positive affect, interaction with professors, interaction with students, and negative affect (pp. 131–133). Two of the four Roberts and Clifton dimensions are distinctly interpersonal: interaction with professors and interaction with other students. Positive affect and negative affect also suggest an interaction between the student and the course (i.e. interest vs. disinterest) or between the student and the teacher (i.e. encouragement vs. fear).

The widely used Students' Evaluation of Educational Quality (SEEQ) also focuses on relationships between teacher, course and students.[1] The SEEQ measures nine factors of teaching effectiveness: learning/value, enthusiasm, organization, group interaction, individual rapport, breadth of coverage, exams, assignments, and workload/difficulty. These factors have been verified in over 30 published studies (e.g. Marsh, 1982, 1984, 1987, 1991; Marsh *et al.*, 1997; Marsh & Hocevar, 1991; Marsh & Kunkin, 1992; Marsh & Roche, 1993). Based on these previous studies, it may be hypothesized that one of the organizing principles concerning evaluation of classes will relate to this concept of interaction.

In addition to the controversy over lenient grading affecting evaluations, an overemphasis on students' emotional response to entertaining teachers, regardless of educational content of the class has been the basis of criticisms of the legitimacy of student evaluations (Naftulin *et al.*, 1973; Ware & Williams, 1975). Therefore, the interactive elements of the evaluation, i.e. the interpersonal dynamic between teacher and students or among students, are not sufficient in evaluating a course. The course content and the teacher's effectiveness may be considered separate categories of evaluation. For quality of classroom experience, then, we can propose five categories: interaction with professors, interaction with other students, interaction with/interest in course material, course quality (i.e. organization, depth of content) and teaching quality.

This study uses a number of multidimensional data analysis techniques to attempt to identify organizing principles of the way students assess their education. Such organizing principles could be used to more effectively design and analyze the results of student evaluations. It is a continuation of previous studies, which looked at how students assess their educational programs (Levy, 1991; Cohen 1995, 1999a,b; Cohen & Sabbagh, 1999). While these other studies have asked students to evaluate life on campus in a broader sense, including social and personal aspects, this study focuses on the classroom experience.

In this article I am concerned with how students conceptually organize elements of evaluation. Student evaluations, like consumer surveys, may reveal as much about the students (consumers) as about the instructors or courses (Hinton, 1993; Bligh, 1998). Thus, student evaluation sheets may be used to understand the logic behind *how* students evaluate classes, and not just how well they rated a specific class. A typology of how principles underlying evaluation may then be used to improve evaluation forms, potentially minimizing, or at least identifying, potential biases.

Methodology

Data collection

The data used for this study was collected in the 1998–1999 academic year. It consists of 14,344 assessment sheets completed by students at the Jordan Valley Regional College (JVRC) in Israel. The questionnaire was designed by faculty members of the college and was distributed and collected by the JVRC administration.

Questionnaires were anonymous and contain no demographic data on the students. Therefore, results can only be evaluated as a whole. Differences in evaluation of the school along gender, ethnic, religious or socioeconomic lines cannot be evaluated. In addition, we do not know how many students from each teacher or course are represented in the sample. It is possible that some teachers' students are over-represented since teachers at JVRC generally teach three or four courses, and students complete assessment sheets at the end of each course. The data provided by the college for this analysis do not include how many assessment sheets were completed for a given teacher. However, given the nature of the college, in which each student studies with many teachers, this is unlikely to affect the data.

Additionally, while the declared main goal of this study from the perspective of the college was to evaluate the teachers and courses themselves, our purpose here is to identify the conceptual structure that students use in evaluating their education. The analytic tools used in this study have proven useful in the interpretation of other studies, such as evaluations of educational tours, feelings of well-being in life, social indicators of quality in life in various nations, and more (Levy & Guttman, 1975; Levy 1986a,b, 1990; Cohen, 2000a,b; Cohen *et al.*, 2001). The uncovering of a structure used as a basis for evaluation is helpful in verifying the instrument of measurement itself. The structure may indicate whether or not categories of evaluation have been left out or repeated. If the structure is useful, it will be applicable to respondents from

Table 1. Means and standard deviations of the various evaluation items

Item	Mean	Standard Deviation
1. the course was well organized	4.00	0.96
2. the goals described in the course syllabus were clear and helped to achieve the aims defined by the teacher	3.95	1.01
3. the bibliography helped me understand the issues	3.90	1.04
4. the course delivered knowledge in the subject	4.00	1.04
5. the course awakened interest in the material	3.86	1.14
6. the course gave me tools for analysis and independent reflection	3.79	1.12
7. general assessment of course quality	3.92	1.05
8. the teacher presented material in a clear and interesting manner	3.98	1.12
9. the teacher was well prepared for the course	4.27	0.94
10. the teacher answered questions, remarks and criticisms	4.22	1.02
11. the teacher had a respectful relationship with the student	4.25	1.03
12. general assessment of teacher's teaching ability	4.13	1.06

a wide variety of situations. Therefore, though over- or under-representation of students in a particular class may be important in the college's evaluation of itself, it does not alter the usefulness of this analysis, which is essentially an assessment of the evaluation tool.

The questionnaires contained 12 variables, listed in Table 1. Each variable could be rated on a scale from 1 (unsatisfactory) to 5 (excellent).

Several categories of evaluation *not* included in this evaluation form should be noted. The twelve questions used as variables in this study represent some, but not all of the nine factors in the SEEQ. Of particular interest is the absence in the JVRC questionnaire of the two factors related to difficulty and grading. Also missing from this questionnaire are items related to group dynamics among students.

Data analysis

Two types of data analysis were performed on this data set. Each enables us to examine different aspects of the data and together present us with a rich and complex view of how students at the college evaluated their educational experience. The first analysis was a classic factor analysis using the Principal Component Analysis extraction method. It was conducted in two dimensions.

The second analysis was a Smallest Space Analysis (SSA), a procedure developed by Louis Guttman (1968, 1982). It is one tool in his comprehensive Facet Theory. Facet Theory is a systematic approach in which first the issues relevant to the research concept are defined and specified. These are the facets which will help the researcher to generate the research variables. Each of the facets should correspond to an aspect of the empirical data (Levy, 1994). This articulation of the research

variables may be used in designing the questionnaire and in formulating hypothesis about the research topic. The SSA is then used to produce a graphic representation of the data.

The differences between the more widely recognized factor analysis and the SSA seem subtle at first glance, but they are important. Both techniques seek to understand the inter-relationships among a large number of variables and to find common underlying structural elements. Both are used for data reduction. The factor analysis is a statistical approach that derives a set of uncorrelated variables, the factors. There is always a primary factor and subsequent factors. In contrast, the SSA, by providing a spatial representation of the data, which shows regions of correlated data, not only indicates various semantic 'factors', but also provides their structural organization. The structure itself reveals much about the data. And there may be more than one interpretation of the spatial representation, depending on the theoretical basis of the study.

As with factor analysis, the first step in conducting an SSA is the calculation of a correlation matrix. It may be noted that SSA allows for more flexibility in types of data analyzed, such as non-metric data (Thapalia, 2004). In this case we used the Pearson Coefficient. The data from the correlation matrix are then plotted as points in a Euclidean space. Strongly correlated items appear close to each other in the space, and those that are weakly correlated appear far apart. This cognitive map reveals regions of related items. Regions of the SSA map are not interpreted solely by proximity, but on the theoretical basis of the study. For example, items at opposite ends of the space may be grouped together in a structure of concentric circles representing a spectrum of central-peripheral regions. A series of vertical regions may group together items at opposite ends of the map (horizontally) in regions representing a progression of theoretical concepts. While the factor analysis provides the researcher with a set of underlying factors summarizing the information, the structure of the links between these factors is not indicated. The contribution of the SSA is to provide the researcher with information about the type of structure by which the variables are organized (Maraun, 1997). For a general presentation of the theoretical works of Guttman and their numerous applications, see Canter (1985), Borg and Lingoes (1987), Levy (1985, 1994), Young (1987), Borg and Groenen (1997), and Cohen (2003).

It is important to note that the borders dividing regions in SSA maps are subjective. A variety of configurations of regions can be recognized, with equal validity, as long as the specificity and continuity of the regions is respected. This may be understood with the metaphor of geographic maps. If one were mapping natural ecosystems, the Israeli city of Eilat, on the Gulf of Aqabah, would be placed in a region together with the Egyptian city of Taba and the Jordanian city of Al Aqabah, while it would be in a different region than Haifa, which is on the Mediterranean coast. However, if one is considering political boundaries, Eilat and Haifa would be in the same region, while Eilat, Taba and Al Aqabah would be in different regions, despite their proximity. Division of regions in the map depends on the theoretical or conceptual basis for drawing the borders.

Results

Distribution table

The first observation that can be made from these data is that the students at the JVRC gave overwhelmingly positive evaluations to their education along all twelve variables. As shown in Table 1, the mean for all twelve variables was well above the mid-point of the scale, the nearest being 3.79. The standard deviations are all between .94 and 1.14, indicating that these means closely reflect the responses of most students, and not a wide variety of answers (i.e. many ratings of 1 and many ratings of 5) disguised by a mean close to the mid-point. The item receiving the most positive evaluations was 'The teacher was well prepared for the course.' The item receiving the least positive evaluations was 'The course gave me tools for analysis and independent reflection.'

Factor analysis

The distribution table shows specifics about the students' attitudes towards their school. However, factor and multidimensional analysis techniques allow us to gain from this data set knowledge that is much more widely applicable in the field of educational assessment. The relationships between these variables reveal several important points about how college students evaluate their education.

Factor analysis revealed two main factors, clearly recognizable in two dimensions: the *course* (7 variables) and the *teacher* (5 variables) (Table 2 and Figure 1). For each factor we included only items with a communality index of greater than 0.6.

Smallest Space Analysis

We then re-analyzed the results from these twelve variables using the Smallest Space Analysis technique. Correlation matrices were calculated using both the Pearson coefficient and a regression free, non-linear coefficient of correlation known as MONCO. In this specific case the structural results from these two coefficients were similar. We therefore chose to present only the better-known coefficient of Pearson.

Several observations can be made directly from the correlation matrix. First, all of the correlations are positive. According to Guttman's First Law of Attitude, if the items have a positive or zero correlation, they may be considered part of the same set of attitude items (Gratch, 1973; Guttman & Levy, 1982). This is the case in relation to these 12 items. This is significant in terms of an attempt to understand the ways in which students evaluate their teachers and courses. If one or more of the items had strongly a negative correlation with several of the other items, we could suggest that the applicability of that item to the evaluation be re-considered.

As we can see in Table 3, the 12 items are strongly inter-correlated, the lowest correlation being 50. For the seven items related to the course, the weakest correlation (though still relatively strong) was with the item: the teacher had a positive relationship with the students (#11). The fact that this item was less strongly correlated

Table 2. Factor analysis bi-dimensional components matrix Extraction method: principal component analysis

	component 1	component 2
1. the course was well organized	.608	.552
2. the goals described in the course syllabus were clear and helped to achieve the aims defined by the teacher	.698	.362
3. the bibliography helped me understand the issues	.731	.130
4. the course delivered knowledge in the subject	.762	.438
5. the course awakened interest in the material	.795	.336
6. the course gave me tools for analysis and independent reflection	.790	.320
7. general assessment of course quality	.740	.516
8. the teacher presented material in a clear and interesting manner	.519	.727
9. the teacher was well prepared for the course	.377	.796
10. the teacher answered questions, remarks and criticisms	.275	.869
11. the teacher had a respectful relationship with the student	.220	.863
12. general assessment of teacher's teaching ability	.438	.813

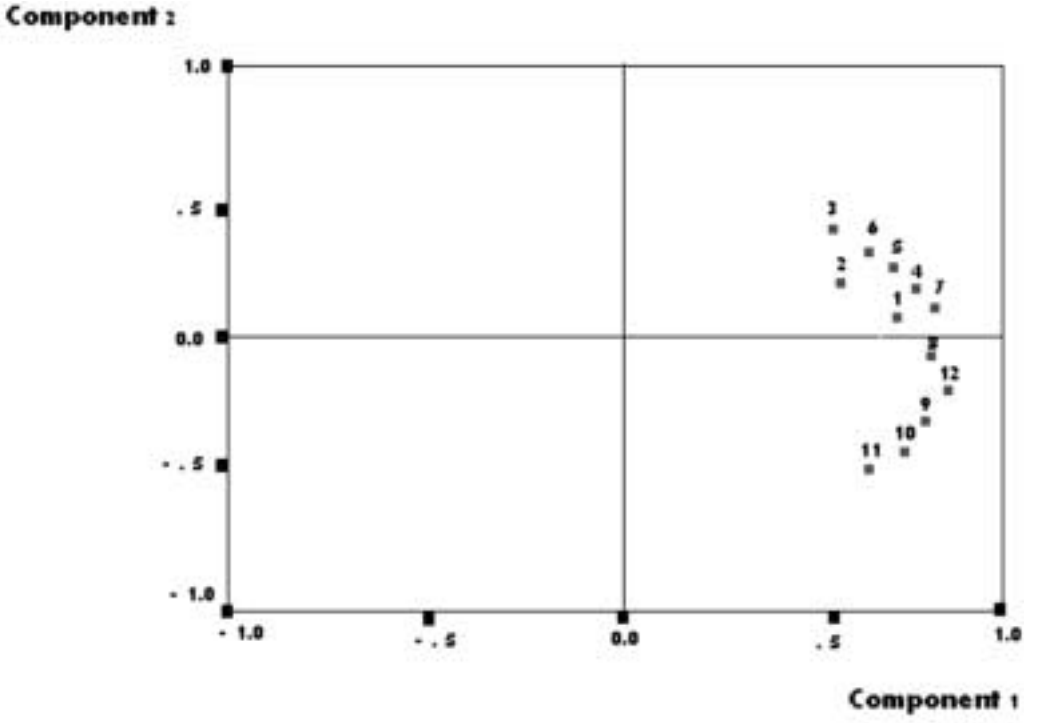
does not mean either that the students at JVRC had poor relationships with their teacher or that the teacher–student relationship is unimportant in the students' evaluation of their education. It means that the students' evaluation of the teacher–student relationship is less strongly related to their evaluation of the course.

The two items 'The teacher presented material in a clear and interesting manner' (#8) and 'General assessment of teacher's teaching ability' (#12) were the most strongly inter-correlated with the other items overall, their *lowest* correlations being 64 and 63 respectively, in both cases with the item 'The bibliography helped me understand the issues' (#3). These two items, the teacher's clear presentation of the material and general teaching ability are tightly bound with the students' evaluations of other aspects of the course.

The relationships represented in the matrix can be more clearly understood by plotting them on a schematic map. Figure 2 shows the results of the SSA, with distinct regions of correlated data. As outlined in the methodology section, the borders between the regions are not necessarily determined by proximity. We are not looking for clusters of points, but rather for specific and continuous regions of conceptually related data.

In this case, I distinguished between four regions in the map, in a grid formation. Along the x-axis, the left-hand side of the map contains the items related to the course. This corresponds to factor 1 in the factor analysis. The right-hand side of the map contains items related to the teacher, and corresponds to factor 2 in the factor analysis.

An additional distinction can be made along the y-axis. In the lower region of the map are evaluational items in which the student is an active participant. I have called this region 'in relation to the student'. By this I mean that these aspects of evaluating the educational experience have to do with the ability of the teacher or the course to



1. The course was well organized
2. The goals described in the course syllabus were clear and helped achieve the aims defined by the teacher
3. The bibliography helped me understand the issues
4. The course delivered knowledge in the subject
5. The course awakened interest in the material
6. The course gave me tools for analysis and independent reflection
7. General assessment of course quality
8. The teacher presented material in a clear and interesting manner
9. The teacher was well prepared for the course
10. The teacher answered questions, remarks and criticisms
11. The teacher had a positive relationship with the students
12. General assessment of teacher's teaching ability

Figure 1. The component plot of the bi-dimensional components matrix

involve the student. The course awakened in the student interest in the material, the course gave the student tools for independent analysis, or the teacher had a positive relationship with the students. Each of these indicates an evaluation of the course or teacher as they relate to the student. The items in the upper half of the map relate to an evaluation of the course or the teacher independent of any involvement on the part of the student. I have called this region ‘related directly to the teacher or course’. The course was well organized, the goals were clear, the teacher was well prepared; each of these depends only on the teacher or the course curriculum, and does not involve the participation of the student or a relationship between the student and the teacher or material.

Table 3. Pearson coefficient correlation matrix of the twelve evaluation items * Number of cases = 14344

	1	2	3	4	5	6	7	8	9	10	11	12
1 |	100	79	68	70	61	63	75	70	71	60	57	69
2 |	79	100	72	71	64	67	75	70	66	60	57	68
3 |	68	72	100	72	65	64	71	64	60	57	55	63
4 |	70	71	72	100	78	75	80	71	65	59	56	69
5 |	61	64	65	78	100	78	79	68	56	54	50	63
6 |	63	67	64	75	78	100	80	68	58	55	52	66
7 |	75	75	71	80	79	80	100	78	69	65	62	77
8 |	70	70	64	71	68	68	78	100	77	71	66	84
9 |	71	66	60	65	56	58	69	77	100	73	67	77
10 |	60	60	57	59	54	55	65	71	73	100	82	77
11 |	57	57	55	56	50	52	62	66	67	82	100	75
12 |	69	68	63	69	63	66	77	84	77	77	75	100

*the original coefficients were multiplied by 100 and rounded into integer numbers

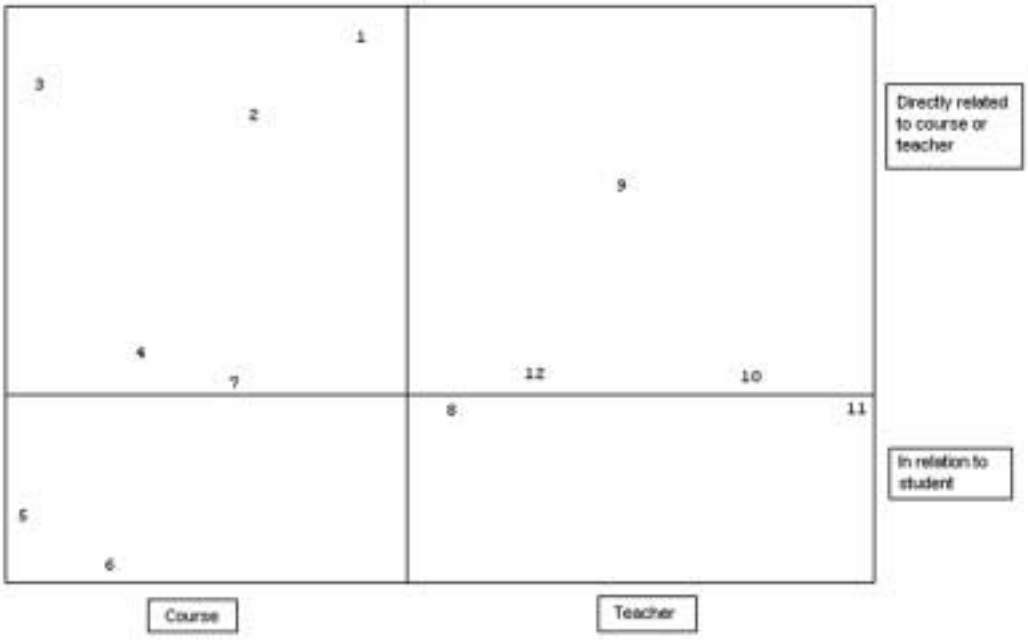
Conclusion

The factor analysis found a clear distinction between the teacher and the course. The nine factors in the SEEQ, cited above, emphasize only the teacher. As many of the criticisms of student evaluations relate to interactions with the teacher (ability to entertain, leniency), the importance seen here of the course as a distinct factor is significant.

Additionally, the distinction uncovered in the SSA analysis, between the inherent quality of a teacher or course and quality of teacher or material in relationship to the students has not been addressed in the wide body of literature on student evaluations. If, as seen in this analysis, the teacher and the course comprise distinct elements independent of their relationship with the student, then these items have not been given sufficient attention. In his in-depth analysis of lectures as an educational tool, Bligh (1998, p. 174) critiques the practice of student evaluations and concludes, 'The truth is, there are lecturers, students and interactions between them. Students' opinions will reflect all three.' The findings of this study expand on this observation: there are lectures (courses), teachers, students and interactions between all of them.

Yet another study with findings similar to those presented here is Cohen's (1981) meta-analysis of student evaluations. Ten factors were found in that factor analysis. Two of the factors are the same as the two found in the factor analysis presented in this study: course and instructor. Three others approximate to categories in the SSA: skill (involving the teacher), rapport (teacher in relation to students), and course structure (involving the course). Another factor with a high coefficient was student progress. This can be seen either related to (though not identical with) course in relation to the student.

Making a clear distinction in the wording of the evaluation sheets between the quality of the course or teacher and their relationship with the students may help in



1. The course was well organized
2. The goals described in the course syllabus were clear and helped achieve the aims defined by the teacher
3. The bibliography helped me understand the issues
4. The course delivered knowledge in the subject
5. The course awakened interest in the material
6. The course gave me tools for analysis and independent reflection
7. General assessment of course quality
8. The teacher presented material in a clear and interesting manner
9. The teacher was well prepared for the course
10. The teacher answered questions, remarks and criticisms
11. The teacher had a positive relationship with the students
12. General assessment of teacher's teaching ability

Figure 2. Bi-dimensional smallest space analysis map

identifying biases towards lenient or entertaining teachers. This factor could possibly yield a fifth region or sub-region in the map: *students in relation to other students* (group dynamics within the classroom).

Standardized evaluation sheets, such as the SEEQ, are adapted by individual institutions. The factors discussed in this survey may be included in future surveys and in the analysis of student evaluations. Survey design may take care to include items related to each of the map regions, since each represents a distinct aspect of the students' evaluations. Evaluations designed using the typology presented here may help clear up questions concerning student bias towards lenient or entertaining teachers.

The applicability of these four categories could be verified by applying the same methodology to other lists of variables, particularly those used in evaluation of

different educational settings (i.e. elementary schools, informal educational programs). Considering these categories when designing questionnaires would not preclude organizing the questionnaire in a way appropriate to a specific situation. For example, a study conducted by Forrester-Jones (2003) and published in this journal included a questionnaire organized logically based on the structure at a particular college (lecture/seminar/module). A re-reading of the list of questions, bearing in mind the results of the present study, finds that each of the four categories in the typology is represented by at least one question, and that questions related to 'teacher in relation to the student' are the most numerous. Thus the specific questionnaire can be organized for the convenience of a particular setting while incorporating the concepts of this typology. If a subsequent study were to be conducted, it could be decided, for example, to examine more closely issues related to the course or to the teachers independent of their relationship with students.

Future studies based on this typology may further investigate this issue in several other directions. If more than one college were involved in a study, it could provide a basis for inter-institutional comparisons. If some basic socio-demographic data on the students were provided (age, gender, year of study, etc.) it could be useful in a comparison of sub-populations in the institution. Demographic differences within or between student populations may influence how courses and teachers are evaluated. An analysis of sub-populations based on various demographic characteristics could help in understanding what evaluations reveal about yet another possible category: *related directly to the student him/herself*, as suggested by Bligh (1998).

The typology presented here should be verified in a number of different student populations. The four categories in the typology are broad enough to allow for a wide range of student concerns and priorities regarding education. Other possible categories suggested here, such as related directly to the student him/herself (i.e. what the survey reveals about the 'consumer') and group dynamics (students in relation to other students) may be examined in more detail.

Acknowledgements

Thanks to The Jordan Valley Regional College for allowing me to use their data and to Allison Ofanansky for editing the manuscript. A preliminary version of the article was presented at the *Fifth International Conference on Logic and Methodology*, Cologne, October 2000. Thanks also to the anonymous reviewers for their useful comments and suggestions.

Note

1. Since its creation in the early 1980s, responses to the SEEQ have been collected from over 1 million undergraduate and graduate students in 50,000 courses. Though the SEEQ has been translated into a number of other languages, and a Chinese version was tested with positive results (Marsh et al. 1997), to the best of our knowledge, it has not yet been translated into Hebrew or used by universities in Israel.

Notes on contributor

Erik H. Cohen, Ph.D., teaches in the School of Education at Bar-Ilan University and Ashkelon Academic College in Israel. In 2002–2003, he served as the Secretary of the Facet Theory Association. He is the director of Research & Evaluation, which conducts international surveys in the fields of evaluation, education and identity and he has been widely published on these subjects.

References

Amin, M. (1994) Gender as a discriminating factor in the evaluation of teaching, *Assessment & Evaluation in Higher Education,* 19, 135–143.

Bligh, D. (1998) *What's the use of lectures?* (5th edn) (Exeter, Intellect).

Borg, I. & Groenen, P. (1997) *Modern multidimensional scaling: theory and applications* (New York, Springer Verlag).

Borg, I. & Lingoes, J. (1987) *Multidimensional similarity structure analysis* (New York, Springer Verlag).

Canter, D. (Ed.) (1985) *Facet theory: approaches to social research* (New York, Springer Verlag).

Cashin, W. E. (1988) *Student ratings of teaching: a summary of the research* (Manhattan, KS, Kansas State University Center for Faculty Evaluation and Development).

Centra, J. & Gaubatz, N. (2000) Is there gender bias in student evaluations of teaching? *Journal of Higher Education,* 71(1), 17–33.

Cohen, E. H. (1995) Towards a strategy of excellence, a systemic analysis and policy research based on external variables in SSA, in: J. J. Hox, G. J. Mellenbergh & P. G. Swanborn (Eds) *Facet theory analysis and design* (Amsterdam, University of Amsterdam Press), 55–62.

Cohen, E. H. (1999a) Prior community involvement and 'Israel Experience' educational tours, *Evaluation and Research in Education,* 13(2), 76–91.

Cohen, E. H. (1999b) Informal marketing of 'Israel Experience' educational tours, *Journal of Travel Research,* 37(3), 238–243.

Cohen, E. H. (2000a) Multidimensional analysis of international social indicators, *Social Indicators Research,* 50(1), 83–106.

Cohen, E. H. (2000b) A facet theory approach to examining overall and life facet satisfaction relationships, *Social Indicators Research,* 51(2), 223–237.

Cohen, E. H. (2003) *Facet theory bibliography, Ljubljana–Jerusalem: Facet Theory Association.* Available online at: www.psy.mq.edu.au/FTA (accessed 19 October 2004).

Cohen, E. H., Clifton, R. A. & Roberts, L. W. (2001) The cognitive domain of the quality of life of university students, *Social Indicators Research,* 53(1), 63–77.

Cohen, E. H. & Sabbagh, C. (1999) Justice judgments among adolescents: Israel and Germany, in: R. M Schweizer (Ed.) *Seventh international facet theory conference, design and analysis* (Berne, Facet Theory Association), 111–126.

Cohen, P. A. (1981) Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies, *Review of Educational Research,* 51, 281–309.

Darom, A., Ritz, Y. & Avrahami, A. (1980) *The development of an instrument for measuring quality of life in the Israeli school* (Ramat Gan, Bar Ilan University Press) (in Hebrew).

Felder, R. (1992) What do they know anyway? *Chemical Engineering Education,* 26(3), 134–135.

Forrester-Jones, R. (2003) Students' perceptions of teaching: the research is alive and well, *Assessment & Evaluation in Higher Education,* 28(1), 59–69.

Gratch, H. (Ed.) (1973) *Twenty-five years of social research in Israel* (Jerusalem, Jerusalem Academic Press).

Greenwald, A. G. & Gilmore, G. M. (1997) No pain, no gain? The importance of measuring course workload in student ratings of instruction, *Journal of Educational Psychology,* 89(4), 740–743.

Greimel-Fuhrmann, B. & Geyer, A. (2003) Students' evaluation of teachers and instructional quality-analysis of relevant factors based on empirical evaluation research, *Assessment & Evaluation in Higher Education,* 28(3), 229–238.

Guttman, L. (1968) A general nonmetric technique for finding the smallest coordinate space for a configuration of points, *Psychometrika,* 33, 469–506.

Guttman, L. (1982) Facet theory, smallest space analysis and factor analysis, *Perceptual and Motor Skills,* 54, 491–493.

Guttman, L. & Levy, S. (1982) On the definition and varieties of attitude and wellbeing, *Social Indicators Research,* 10, 159–174.

Hinton, H. (1993) Reliability and validity of student evaluations: testing models versus survey research models, *PS, Political Science & Politics,* 26(3), 562–570.

Johnson, V. E. (2003) *Grade inflation: a crisis in college education* (New York, Springer Verlag).

Kaplan, M., Mets, L. & Cook, C. (2002) *Questions frequently asked about student ratings forms: summary of research findings* (Ann Arbor, MI, University of Michigan, Center for Research on Learning and Teaching).

Levy, S. (1985) Lawful roles of facets in social theories, in: D. Canter (Ed.) *Facet theory: approaches to social research* (New York, Springer Verlag), 59–96.

Levy, S. (1986a) Psycho-economic well-being: the case of Israel, in: A. J. MacFadyen & H. W. MacFadyen (Eds) *Economic psychology: intersections in theory and application* (Amsterdam, Elsevier), 585–617.

Levy, S. (1986b) *The structure of social values* (Jerusalem, Israel Institute of Applied Social Research), 44–66.

Levy, S. (1990) The mapping sentence in cumulative theory construction: well-being as an example, in: J. J. Hox & J. deJong-Gierveld (Eds) *Operationalization and research strategy* (Amsterdam, Swets & Zeitlinger), 155–177.

Levy, S. (1991) What is a good school in the eyes of the students? *Megamot,* 33(3–4), 535–548 (in Hebrew).

Levy, S. (Ed.) (1994) *Louis Guttman on theory and methodology: selected writings* (Aldershot, Dartmouth Publishing Company).

Levy, S. & Guttman, L. (1975) On the multivariate structure of well-being, *Social Indicators Research,* 2, 361–388.

Lewy, A. (1985) The student's relation to the school, in: S. Zachs (Ed.) *Thought and practice in education and treatment* (Tel Aviv, Papyrus), 157–168 (in Hebrew).

Maraun, M. D. (1997) Appearance and reality: is the Big Five the structure of trait descriptors? *Personality and Individual Differences,* 22(5), 629–647.

Marsh, H. W. (1982) SEEQ: a reliable, valid and useful instrument for collecting students' evaluations of university teaching, *British Journal of Educational Psychology,* 52, 77–95.

Marsh, H. W. (1984) Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility, *Journal of Educational Psychology,* 76, 707–754.

Marsh, H. W. (1987) Students' evaluations of university teaching: research findings, methodological issues, and directions for future research, *International Journal of Educational Research,* 11, 253–388.

Marsh, H. W. (1991) A multidimensional perspective on students' evaluations of teaching effectiveness: a test of alternative higher-order structures, *Journal of Educational Psychology,* 83, 285–296.

Marsh, H. W. & Hocevar, D. (1991) Students' evaluations of teaching effectiveness: the stability of mean ratings of the same teachers over a 13-year period, *Teaching and Teacher Education,* 7, 300–314.

Marsh, H. W. & Kunkin, M. (1992) Students' evaluations of university teaching: a multidimensional perspective, in: J. C. Smart (Ed.) *Higher education: handbook of theory and research* (New York, Agathon), 143–233.

Marsh, H. W. & Roche, L. A. (1993) The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness, *American Educational Research Journal,* 30, 217–251.

Marsh, H. W. & Roche, L. A. (2000) Effects of grading leniency and low workloads on students' evaluations of teaching: popular myth, bias, validity or innocent bystanders? *Journal of Educational Psychology,* 92, 202–228.

Marsh, H. W., Choi-Man, K., Hau, C. & Sui, T. (1997) Students' evaluations of university teaching: Chinese version of the students' evaluations of educational quality instrument, *Journal of Educational Psychology,* 89(3), 568–572.

Naftulin, D. H., Ware, J. E. & Donnelly, F. A. (1973) The Doctor Fox lecture: a paradigm of educational seduction, *Journal of Medical Education,* 48, 630–635.

Radmacher, S. & Martin, D. (2001) Identifying significant predictors of students evaluations of faculty through hierarchical regression analysis, *Journal of Psychology,* 135(3), 259–269.

Roberts, L. W. & Clifton, R. A. (1992) Measuring the affective quality of life of university students: the validation of an instrument, *Social Indicators Research,* 27, 113–137.

Schmidt, R. (2003) Comparing the student evaluation of 60 rural and urban type high schools in Southern California, *Education,* 123(3), 450.

Thapalia, C. (2004) *Multivariate statistics: an introduction.* Available online at: http://trochim.human.cornell.edu/tutorial/flynn/multivar.htm (accessed 26 January 2004).

Ware, J. E. & Williams, R. E. (1975) The Doctor Fox effect: a study of lecturer effectiveness and ratings of instruction, *Journal of Medical Education,* 50, 149–156.

Williams, T. & Batten, M. (1981) *The quality of school life* (Hawthorn, Australia, The Australian Council for Educational Research).

Yee, A. H. & Runkel, P. J. (1969) Simplical structures of middle class and lower class pupils' attitudes towards teachers, *Developmental Psychology,* 1, 646–652.

Young, F. W. (1987) *Multidimensional scaling: history, theory and applications* (Hillsdale, NJ, Erlbaum Associates).

Copyright of Assessment & Evaluation in Higher Education is the property of Carfax Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.