

Students' Evaluation of Tertiary Educational Goals: The Impact of Lecturer and Student Characteristics

Nerina J. Caltabiano¹ and Marie L. Caltabiano²

Abstract: This study is concerned with how tertiary students evaluate educational goals within their degree programs. Crooks' (1988) classification of educational practices in terms of short-term and medium-term consequences is used. This study assesses the viability of his classificatory system within a university student sample. The current study extends previous research by considering both lecturer characteristics (e.g., lecturer supportiveness, teaching quality, approachability of lecturing staff and availability of lecturing staff) and student characteristics (e.g., age and student's year level) which may predict how educational practices are evaluated. Evaluation data are provided by 164 volunteer students (31.5% males and 68.5% females). The principal component analysis was able to establish a list of short-term and medium-term goals appropriate for a university student sample. Lecturer supportiveness predicts short-term goals while teaching quality and student's year level predicted medium-term goals.

I. Introduction.

Student evaluation of teaching effectiveness (SETE) has been used in Australian universities for some time now as part of the quality assurance process. Student evaluation data are used by individual departments in making personnel decisions, in the allocation of teaching resources, as well as in decisions about whether or not to offer a subject. Student ratings influence faculty decisions regarding promotion and tenure of lecturing staff, as well as in the award of teaching merit grants. Students also use evaluation data in selection of degree courses and specific subjects. Such use of student rating data has been termed summative evaluation, in contrast to the use of teaching effectiveness ratings by individual instructors for the purpose of improving teaching, otherwise referred to as formative evaluation (Theall & Franklin, 2001).

Ratings of teaching effectiveness are typically made on a teacher rating form (TRF), with considerable similarity in the types of questions asked across tertiary institutions in the United States, Australia and the United Kingdom. These questions generally ask about the lecturer's knowledge of the subject area, the clarity of the lecturer's explanations, willingness to answer questions, fairness in grading assessment, and punctuality. Instructors can also choose to include in their subject evaluations, items from an additional list of questions. These questions evaluate features specific to the discipline e.g. laboratory sessions, field trips; the use of technology such as computer-generated slide shows, the Internet, and electronic mail to communicate with students;

¹ Senior Lecturer, James Cook University – Cairns Campus, School of Psychology, P.O. Box 6811, Cairns, Qld. 4870, Australia. Nerina.Caltabiano@jcu.edu.au.

² Senior Lecturer, James Cook University – Cairns Campus, School of Psychology, P.O. Box 6811, Cairns, Qld. 4870, Australia. Marie.Caltabiano@jcu.edu.au.

and in the case of cross-campus teaching, questions might evaluate the learning experience via video-lecturing. Evaluations of teaching effectiveness are important because they give insight into the quality of the learning experience for the student, and subsequently how degree programs are evaluated in terms of the attainment of their educational goals.

The last 30 years have seen a proliferation of research studies addressing the validity of student evaluations of teaching effectiveness. Concerns about validity have centred on a number of issues: students not being qualified to assess the lecturer's competence, that ratings may reflect popularity, the possibility of lecturer ratings being influenced by grades received, that students rate highly those lecturers who are lenient and from whom they learn the least, and that gender and other situational variables (such as class size, year level and required versus elective subjects) may influence ratings (Marsh, 1984; Ory, 2001; Theall & Franklin, 2001). The underlying rationale of research addressing the validity of student evaluations of teaching effectiveness, is that if student ratings are valid, they should predict criteria of effective teaching such as performance on examinations and achievement in the subject. In a meta-analysis by Cohen (1981) of 41 studies across 68 courses, the correlation between student achievement and their ratings of the lecturer was 0.43, indicating that students rate highly those lecturers from whom they learn the most. More recent research indicates that student expectations about their grade in a subject (Greenswald & Gillmore, 1997), and actual grades received (D'Apollonia, Lou & Abrami, 1998) affect their evaluation of the lecturer. The question of rating validity has largely been addressed in relation to evaluations of teaching effectiveness. The literature is scant on whether these same variables (e.g. grading, class size, year level) and indeed others, affect how the educational goals of degree programs are evaluated.

Less attention has been given to instructor characteristics which may influence ratings of teaching effectiveness. Williams and Ceci (1997) found a difference of 2 scale-points on student ratings when the instructor varied presentation style so as to be more enthusiastic. Williams (2001) reported that student evaluations of instructional style were strongly correlated with overall evaluations of the course. Other research (Shevlin, Banyard, Davies & Griffiths, 2000) has found that teacher effectiveness ratings are accounted for by perceptions of the lecturer's personality or 'charisma,' lecturer ability and module attributes. The quality of the learning experience depends on more than the actual instruction received. Instructors can influence students in a number of ways such as in motivating them to read further on a topic, to reflect critically on issues, and in the application of ideas to novel situations. It would be expected that instructor characteristics such as warmth, attitude and supportiveness would influence perceptions of the learning experience, and ultimately how the degree program is evaluated.

Given that the learning experience involves student and instructor in a dynamic interaction, it is surprising that little research has been focussed on the interactive influence of student and instructor characteristics on how degree programs are evaluated. From earlier research, we know that student ratings are higher for elective rather than required courses (Marsh, 1984) and in higher level courses rather than in lower level courses (Feldman, 1978) indicating that the student's year level should be taken into consideration. Given that a large proportion of university students are mature aged involved in study for the first time or changing career paths, age of the student should be

considered in student evaluations. In a study by Stringer and Irwing (1998) which used structural modelling to examine the effects of course, student and instructor characteristics on students' evaluations of teaching, it was found that teaching quality exerted a direct influence on evaluations of course integration, and had an indirect influence on stimulation/learning and overall evaluations. Interestingly, student characteristics (such as motivation, study habits and prior knowledge) and course characteristics explained a substantial amount of the variance in overall evaluations independent of teaching quality.

With minor exceptions (Stringer & Irwing, 1998) most of the research into student evaluations of teaching effectiveness has centred on subjects or modules rather than degree programs/courses. In addition, where teaching has been evaluated in specific modules, TRFs of the individual universities have been used. While these rating forms evaluate a number of worthwhile attributes of the module and the teaching quality, they tend to be global rather than specific in their evaluation, and often do not provide any detailed information on whether the educational goals and objectives of the degree program were satisfied. Crooks (1988) provides a list of consequences associated with student classroom evaluation practices which might be relevant to the educational experience of university students. These educational consequences, referred to as goals by McInerney and McInerney (1994), are classified in terms of duration of impact (short-term, medium and long-term) on learning strategies, and skill development. In this study, we sought to obtain more detailed information on university students' evaluation of the educational goals of their degree programs, rather than focussing on specific modules. Of interest also was whether student and lecturer characteristics affect how degree programs are evaluated.

Accordingly, the aims of the research were to:

- 1) study how students evaluate the educational goals within their degree programs
- 2) examine the effect of lecturer characteristics (such as supportiveness, teaching quality, approachability, and availability) and student characteristics (age and year level) on student evaluations of educational practices.

II. Method.

A. Subjects.

The students (N=164) were a convenience sample of volunteers recruited from several degree programs within the university setting. No course instructors solicited student participation in this study. Whenever the results do not add up to 164 it is because the missing values have been omitted. The students were 31.5% (n = 51) males and 68.5% (n = 111) females aged between 18 and 55 years (M = 29.0, SD = 9.6), enrolled in a cross-section of undergraduate courses. This gender bias reflects the gender-mix within the university student population. More specifically the students can be categorized as follows with regards to the courses they were enrolled in: Bachelors of Social Sciences and Arts (n = 64, 41.0%), Bachelor of Psychology (n = 43, 27.6%), Bachelor of Education (n = 32, 20.5%), Bachelor of Science (n = 8, 5.1%) and Bachelors of Management, Law and Tourism (n = 9, 5.8%). Approximately one third of the sample came from each of the year levels of the students' degree programs. The student sample

was evenly split in terms of students who lived alone and those who lived with a partner. Only 22.1% (n = 36) of the students were not employed, while pursuing their studies, with the remaining 77.9% (n = 128) being involved in some type of work for pay. The majority of students (n = 107, 70.4%) were residing in rental accommodation and 29.6% (n = 45) either owned their own home or were paying it off. About half of the student respondents (n = 76, 46.9%) lived ten kilometres or less from the campus, 49.4% (n = 80) lived between ten and 50 kilometres from the university and the remaining 3.7% (n = 6) of the students lived 50 or more kilometres away from the university.

B. Questionnaire.

After obtaining ethical clearance a questionnaire specifically designed for this study was used to collect the anonymous data from a voluntary student sample enrolled in different degree programs and from different year levels. The questionnaire consisted of items assessing the educational goals of the student's course of study, lecturer characteristics, and student characteristics.

Table 1: Percentage of students responding that the educational goals have been met in their university course of study.

Item	N	%
Helping students feel a sense of accomplishment.	117	71.3
Checking that students have adequate prerequisite skills and knowledge to effectively learn the material to be covered.	88	54.3
Communicating and reinforcing the instructor's or the curriculum's broad goals for students, including the desired standards of performance.	107	65.6
Focusing attention on important aspects of the subject.	128	78.5
Influencing students' choice of (and development of) learning strategies and study patterns.	76	46.6
Describing or certifying students' achievement in the course, thus influencing their future activities.	86	53.1
Reactivating or consolidating prerequisite skills or knowledge prior to introducing new material.	92	65.1
Encouraging active learning strategies.	92	56.1
Helping students to monitor their own progress and develop skills of self-evaluation.	74	45.1
Influencing students' motivation to study the subject and their perceptions of their capabilities in the subject.	89	54.3
Giving students opportunities to practise skills and consolidate learning.	107	65.2
Guiding the choice of further instructional or learning activities to increase mastery.	84	51.2
Providing knowledge of results and corrective feedback.	117	71.3

The educational goals presented here come from Crooks' (1988) review of the impact of classroom evaluation practices on students. These practices are classified in terms of eight short-term consequences and five medium-term consequences that can be used as a checklist of the qualities of good educational practices (McInerney & McInerney, 1994). The percentage of students reporting that the educational goals have been met in their chosen university course of study is presented in Table 1. An example of a short-term goal is: "guiding the choice of further instructional or learning activities to increase mastery". A medium-term goal from the list is "influencing students' motivation to study the subject and their perceptions of their capabilities in the subject". The goal that was seen as being met by most of the students was that of focusing attention on important aspects of the subject (n = 128, 78.5%). The goal that was least seen as having been met in their course of study was that of helping students to monitor their own progress and develop skills of self-evaluation (n = 74, 45.1%). The reliability coefficient (Cronbach's Alpha) for the short-term goals was 0.60, and for the medium-term goals was 0.58.

The items assessing lecturer characteristics focused on lecturer supportiveness, teaching quality, approachability and availability of the lecturing staff. The student characteristics included age, gender, course and year level information, marital status, employment status, whether they were or were not living in rental accommodation and the distance they resided from the university.

III. Results.

A principal components analysis was used to see whether similar clusters of educational goals (Crooks, 1988; McInerney & McInerney, 1994) could be identified within a university student sample. The analyses for this study were performed using SPSS 11 for Windows. Bartlett's test of sphericity (276.57) is large and the associated significance level is small (p = .000), consequently it appears unlikely that the population correlation matrix is an identity, that is, all diagonal terms are one and all off-diagonal terms are zero. The Kaiser-Meyer-Olkin measure of sampling adequacy found to be in the middling range (0.67) is an index for comparing the magnitudes of the observed correlation coefficients to the magnitude of the partial correlation coefficients. Both these measures suggest that the use of the factor model is appropriate. Table 2 presents the factor structure obtained.

The principal components factor analysis using varimax rotation revealed two factors which have been labelled 'short-term goals' and medium-term goals'. A factor loading exceeding 0.30 was used to determine high loading items. Two items that were originally medium-term goals in fact were found to load on the current short-term goals factor (ie., communicating and reinforcing the instructor's or the curriculum's broad goals for students, including the desired standards of performance; and checking that students have adequate prerequisite skills and knowledge to effectively learn the material to be covered). With regards to the second factor three items that originally loaded on the short-term goals factor were found to load on the medium-term goal factor (ie., helping students feel a sense of accomplishment; encouraging active learning strategies and focusing attention on important aspects of the subject). Using the scree criteria the eigenvalues for the factors were 2.85 and 1.41 respectively. Factor scores were computed

for each of the two factors where the two factor scores take into account the contribution of each scale item to each factor. These were then used as dependent variables in the regression analyses.

Table 2: Varimax rotated factor matrix for the evaluation goals (N = 157).

Item	Factor 1	Loadings 2	Communality h ²
Giving students opportunities to practise skills and consolidate learning	.747	.037	.560
Communicating and reinforcing the instructor's or the curriculum's broad goals for students, including the desired standards of performance	.702	.031	.494
Providing knowledge of results and corrective feedback	.594	.046	.355
Checking that students have adequate prerequisite skills and knowledge to effectively learn the material to be covered	.431	.281	.265
Guiding the choice of further instructional or learning activities to increase mastery	.430	.239	.242
Reactivating or consolidating prerequisite skills or knowledge prior to introducing new material	.390	-.010	.152
Focusing attention on important aspects of the subject	.314	.288	.182
Helping students to monitor their own progress and develop skills of self-evaluation	.282	.242	.138
Influencing students' choice of (and development of) learning strategies and study patterns	.080	.746	.563
Helping students feel a sense of accomplishment	.015	.574	.330
Influencing students' motivation to study the subject and their perceptions of their capabilities in the subject	.102	.560	.324
Encouraging active learning strategies	.246	.550	.363
Describing or certifying students' achievement in the course, thus influencing their future activities.	.003	.546	.298
Eigenvalues	2.853	1.413	4.266
% Common Variation	66.88	33.12	100.00
% Total Variation	21.95	10.87	32.82

Standard multiple regressions were used to analyse the data. The assumptions of sample size, multicollinearity and singularity, outliers, normality, linearity,

homoscedasticity and independence of residuals were not violated in these analyses. Educational goals were then regressed on predictor variables of lecturer supportiveness, teaching quality, approachability of lecturing staff, availability of lecturing staff, age, gender and year level. Table 3 presents the significant regression results for short-term goals. Of all the variables considered, lecturer supportiveness (that is, the greater the supportiveness) made the largest unique and significant contribution (Beta = 0.23, $p = .007$) in explaining short-term goals. The variables considered explain eight percent of the variance in short-term goals.

Table 3: Regression of short-term goals on lecturer and student characteristics (N=153).

Predictor	Beta	SE	t	p
Lecturer Supportiveness	0.23	0.12	2.76	0.007
Teaching Quality	0.07	0.20	0.90	0.37
Approachability of lecturing staff	0.05	0.19	0.51	0.61
Availability of lecturing staff	0.01	0.18	0.16	0.87
Age	-0.14	0.01	-1.64	0.10
Year level	0.12	0.09	1.50	0.13
Intercept			-1.03	
R²			0.08	
Adjusted R²			0.04	
F = 2.10, df = 6,152 p = .05				

Table 4 presents the significant results for the medium-term goals. Year level (Beta = -0.21, $p = .01$) and teaching quality (Beta = 0.21, $p = .01$) best predicted medium-term goals. These two significant variables were equal in their contribution in explaining the medium-term goals. The variables were found to account for 11 percent of the variance in this model.

Table 4: Regression of medium-term goals on lecturer and student characteristics (N=152).

Predictor	Beta	SE	t	p
Lecturer Supportiveness	0.13	0.11	1.58	0.11
Teaching Quality	0.21	0.20	2.53	0.01
Approachability of lecturing staff	0.02	0.18	0.20	0.84
Availability of lecturing staff	0.04	0.17	0.46	0.65
Age	-0.02	0.01	-0.21	0.84
Year level	-0.21	0.09	-2.58	0.01
Intercept			-0.20	
R²			0.11	
Adjusted R²			0.08	
F = 3.12, df = 6,152 p = .007				

IV. Discussion.

The focus of this study was on how students evaluate entire degree programs rather than with evaluation of individual subjects/modules and individual instructors, and

whether evaluations of the degree program are influenced by lecturer and student characteristics. While perceptions of subjects taken, and evaluations of the teaching quality within these subjects would to some extent influence overall ratings of the quality of the degree, the approach taken allowed for a more in-depth evaluation of whether students perceived the degree to have satisfied basic educational goals such as consolidating prerequisite skills, facilitating motivation to learn, and encouraging active learning strategies. These goals were adapted from Crooks' (1988) list of educational consequences associated with student classroom evaluation practices. As these practices were considered relevant to the educational experiences of university students, the data were factor analysed to assess the viability of using Crooks' classificatory system in a tertiary setting. Factor analysis confirmed Crooks' classification of short-term and medium-term influences, with only minor differences in the items loading on each factor.

While most universities routinely collect data on student evaluations of teaching effectiveness, this is often done at the subject and individual instructor level for either formative or summative purposes of evaluation. It is at the discretion of individual departments to have final year students evaluate their degree programs. This study is advocating that Staff Development units work on developing a standardised form to facilitate the monitoring of educational goals/practices of degree programs within university faculties. While the present study found Crook's (1988) classification to be a useful basis of such evaluation, it could be supplemented with discipline-specific questions relating to student placements and other applied learning approaches. As a means of compiling formative evaluative data, student evaluations of the degree program could be compared with evaluations of alumni working in the profession. If there is a deficit in skill development in a specific area, as identified by recent graduates, this information can be used in decisions about subject offerings and in staff selection decisions.

In this research we were interested in assessing the influence of lecturer characteristics and student characteristics on how degree programs were evaluated. Evaluation of short-term goals such as checking prerequisite skills, providing opportunities to consolidate learning, providing corrective feedback, and developing skills of self-evaluation were influenced by lecturer supportiveness. Students who perceived these short-term goals to have been adequately met in their degree program were more likely to also perceive lecturing staff to be supportive. While previous research has examined the influence of instructor characteristics on student ratings, the emphasis has been on stylistic and presentation variables rather than on more subtle person characteristics (Williams & Ceci, 1997). Lecturer personality was reported by Shevlin et al. (2000) to account for effectiveness ratings, as much as ability. The present finding would seem to substantiate the need to acknowledge the importance of person variables in the interactive learning experience.

In relation to the evaluation of medium-term educational goals of degree programs, it seems that other instructor characteristics relating to knowledge base and teaching quality are also important. Such medium-term goals include influencing choice of learning strategies and study patterns, motivating students, encouraging active learning strategies and focusing attention on important aspects of a subject. In Stringer and Irwing's (1998) study, teaching quality was found to have a direct effect on student's evaluation of course integration, while indirectly influencing stimulation/learning. There

are therefore consistencies across the two studies even though the level of analysis is different – degree program compared to individual subjects.

In regard to student characteristics, year level was negatively associated with evaluations of the extent to which medium-term goals had been met by the degree. Students at higher levels tended to be more critical than lower-level students possibly because by the final year of their course, students are more knowledgeable of a discipline and have higher expectations regarding training. Likewise, Sailor, Worthen and Shin (1997) found that upper-level students were more critical in their student evaluations. These findings are contrary to previous research by Feldman (1978) where higher ratings are given by higher level students. The difference in findings may be due to evaluations of subjects/individual instructors in the Feldman study as opposed to evaluation of a degree course in the present study.

The explanatory model in the present study did not account for much of the variance in student evaluations of their degree programs. A number of interpretations could be offered for this finding. Firstly, student evaluations may actually be robust, such that they vary little across student characteristics such as age or year level. Moreover, lecturer characteristics such as perceived competence, approachability and supportiveness may exert little influence in how tertiary educational goals are evaluated. Secondly, the explanatory variables may have been correctly specified for a different dependent variable. That is, the current study used Crook's short-term and medium-term educational goals to evaluate degree programs. Future research may wish to replicate the present findings using a different set of evaluative criteria to assess degree programs. Lastly, perhaps other variables may need to be considered in relation to tertiary educational goals. Some such variables may be whether or not the university is a regional one, with limited student resources, the quality/quantity of library resources with access to electronic databases and electronic journals, the use of internet teaching and video-lecturing limiting access to lecturing staff, and the availability of student placement opportunities for regional universities.

Student learning in university settings is multifaceted. Students may learn through intrinsically motivated reading and exploration of ideas, in classroom group situations through the sharing of ideas, and through lecturer imparted knowledge. The learning process can be considered to be dynamic. Even in typical lecture settings, students are not inactive recipients of knowledge. They actively work on the information, ask questions, debate and discuss aspects of the learnt material. As such evaluation of the learning experience should also take into account student factors and situational factors (setting, module characteristics), in addition to instructor characteristics. Where such variables are found to influence student evaluation ratings, this does not invalidate the evaluation process but rather gives a more realistic account of interactive influences in the dynamic of teaching.

References

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

Crooks, T.J. (1988). "The impact of classroom evaluation practices on students." *Review of Educational Research*, Vol. 58, pp. 438-481.

- D'Apollonia, S. Lou, Y. & Abrami, P.C. (1998). "**Making the grade: A meta-analysis on the influence of grade inflation on student ratings.**" Paper presented at the 79th Annual meeting of the American Educational Research Association, San Diego, April.
- Feldman, K. A. (1978). "Course characteristics and college students' ratings of their teachers and courses: What we know and what we don't." *Research in Higher Education*, Vol. 9, pp. 199-242.
- Greenswald, A. G. & Gillmore, G. M. (1997). "Grading leniency is a removable contaminant of student ratings." *American Psychologist*, Vol. 52, pp.1209-1217.
- Marsh, H. W. (1984). "Students' evaluation of university teaching: Dimensionality, reliability, validity, potential biases and utility." *Journal of Educational Psychology*, Vol. 76, pp.707-754.
- McInerney, D.M. & McInerney, V. (1994). "*Educational Psychology: Constructing Learning.*" Sydney: Prentice Hall.
- Ory, J. C. (2001). "Faculty thoughts and concerns about student ratings." *New Directions for Teaching and Learning*, Vol. 87, pp.3-15.
- Sailor, P., Worthen, B.R. & Shin, E. (1997). "Class level as a possible mediator of the relationship between grades and student ratings of teaching." *Assessment and Evaluation in Higher Education*, Vol. 22, No. 3, pp.261-270.
- Shevlin, M. Banyard, P., Davies, M. & Griffiths, M. (2000). "The validity of student evaluation of teaching in higher education: Love me, love my lectures?" *Assessment and Evaluation in Higher Education*, Vol. 25, No. 4, pp. 397-405.
- Stringer, M. & Irwing, P. (1998). "Students' evaluations of teaching effectiveness: A structural modelling approach." *The British Journal of Educational Psychology*, (<http://proquest.umi.com/pqdweb>)
- Theall, M. & Franklin, J. (2001). "Looking for bias in all the wrong places: A search for truth or a witch hunt in student rating of instruction?" *New Directions for Institutional Research*, Vol. 109, pp. 45-56.
- Williams, R. (2001). "*Course evaluations: A strategy for improving instruction*" (Report No. HE033764). Tennessee: Author. (ERIC Resources in Education ED449759).
- Williams, W. M. & Ceci, S. (1997). "How'm I doing? Problems with student ratings of instructors and courses." *Change*, Vol. 29, No. 5, pp.13-23.