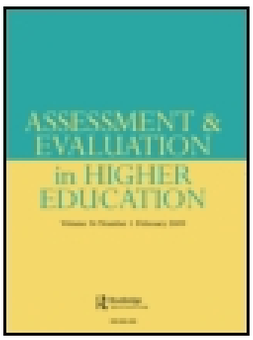

Routledge
Taylor & Francis Group

Assessment & Evaluation in Higher Education

ISSN: 0260-2938 (Print) 1469-297X (Online) Journal homepage: http://www.tandfonline.com/loi/caeh20

Assessing the Teaching Quality of Accounting Programmes: An evaluation of the Course Experience Questionnaire

MARANN BYRNE & BARBARA FLOOD

To cite this article: MARANN BYRNE & BARBARA FLOOD (2003) Assessing the Teaching Quality of Accounting Programmes: An evaluation of the Course Experience Questionnaire, Assessment & Evaluation in Higher Education, 28:2, 135-145, DOI: 10.1080/02602930301668

To link to this article: http://dx.doi.org/10.1080/02602930301668

Published online: 27 May 2010.

Assessment & Evaluation in Higher Education, Vol. 28, No. 2, 2003 Carfax Publishing Taylor & Francis Group

Assessing the Teaching Quality of Accounting Programmes: an evaluation of the Course Experience Questionnaire

MARANN BYRNE & BARBARA FLOOD, *Dublin City University Business School, Dublin City University, Ireland.*

ABSTRACT *In recent years, measuring the efficiency and effectiveness of higher education has become a major issue. National governments are now demanding greater public accountability for funds invested in the sector, resulting in the emergence of various performance indicators relating to both teaching and research. The Course Experience Questionnaire (CEQ) was developed to measure the perceived quality of teaching in degree programmes. It evolved from research that identified curriculum, teaching and assessment as key determinants of students' approaches to learning and, in turn, the quality of their learning outcomes. The CEQ data are intended for use in making comparisons within fields of study over time and/or across institutions. However, no European study has reported on its suitability to evaluate teaching within an accounting programme. This paper outlines the development of the CEQ and confirms its reliability and construct validity for use in the accounting discipline in an Irish context.*

Introduction

Higher education institutions world-wide are experiencing a wave of quality evaluation, as demand for public accountability escalates (Kwan, 1999; Romainville, 1999; Buckley & Hurley, 2001). Quality teaching and research are the primary objectives of higher education and, consequently, there is a growing interest in the development of performance indicators to measure the outputs of these two activities. To date, most performance measures used in higher education have focused on research activity. Designing performance indicators for teaching is fraught with difficulty (Bok, 1986) and many, such as student progression rates and dropout rates, are inappropriate for evaluating the quality of the teaching process (Romainville, 1999).

ISSN 0260-2938 print; ISSN 1469-297X online/03/020135-11 © 2003 Taylor & Francis Ltd
DOI: 10.1080/0260293032000044491

Quality teaching is that which facilitates quality learning among students (Marsh & Roche, 1994; Biggs, 1999). Research has shown that students are best placed to evaluate many aspects of teaching quality, and their ratings are valid, multidimensional and quite reliable (Marsh, 1987; Wachtel, 1998). Despite such research, many academics still feel threatened by students' ratings of teaching, fearing that they will lack objectivity and may be used inappropriately by institutions (Stringer & Finlay, 1993). Evaluating teaching at the course level (the term 'course' refers to a full course of study such as a degree programme) rather than at the individual module level is usually viewed more positively by staff, and for the purposes of maintaining and enhancing teaching quality within an institution it is argued that such a focus is more appropriate (Ramsden, 1991; Richardson, 1994). However, there appears to be a general shortage of robust instruments to evaluate teaching quality at the course level, and thus the Course Experience Questionnaire (CEQ) could have much to offer higher education institutions.

The CEQ is an instrument that was developed in Australia where it is widely used as a performance indicator of teaching quality. It was designed in the context of a theory of teaching and learning in which students' perceptions of curriculum, teaching and assessment are recognised as affecting their learning approaches and the quality of their learning outcomes (e.g., Marton & Saljo, 1976; Entwistle & Ramsden, 1983). The CEQ specifically aims to monitor teaching quality at the course level and to measure differences between academic units in those aspects of teaching that students can evaluate (Ramsden, 1991; Lizzio *et al*., 2002). There is little research examining the appropriateness of the CEQ outside Australia or in specific disciplinary settings and thus this paper examines the validity of the instrument in the context of accounting courses at an Irish university.

Within the accounting discipline, numerous reports have criticised accounting programmes for failing to encourage students to adopt deep approaches to learning and to develop as independent learners (AAA, 1986; IFAC, 1996; Albrecht & Sack, 2000). These reports have called upon accounting educators to re-evaluate their programmes and to instigate changes that enhance the quality of student learning. Accordingly, this study assesses the suitability of the CEQ for evaluating teaching quality in the accounting discipline.

The remainder of the paper is structured as follows. Firstly, the development of the CEQ and the nature of the instrument are outlined. Secondly, the validation of the instrument for use with accounting students at an Irish university is presented. Finally, the potential uses and limitations of the CEQ in the quality-conscious higher education sector are examined.

Development of the CEQ

The development of the CEQ took place at a time of considerable reform in the higher education sector in Australia. In the late 1980s, two discussion papers (Green and White Papers) on higher education demanded greater accountability for the public funds invested in the sector and called for the development of appropriate indicators of institutional performance (Dawkins, 1987, 1988). In particular, student satisfaction was identified as a measure of institutional performance that should be considered. In response, the Performance Indicators Research Group (PIRG) was commissioned to examine a number of quantitative indicators that could be used to evaluate the performance of individual institutions and the performance of the university sector as a

whole (Long, 1995). One measure included in this evaluation that provided evidence on student satisfaction was the CEQ.

The CEQ was devised by Ramsden (1991) with the objective of measuring students' perceptions of the teaching quality of their courses. The instrument does not capture every aspect of teaching effectiveness, rather it aims to gather reliable and valid data on those elements of teaching that students experience directly. Quality is a relative concept (Romainville, 1999), thus the instrument is suitable for measuring differences in the quality of teaching between comparable academic units, or within academic units over a period of time. The instrument evolved from research on student learning carried out at Lancaster University in the 1980s, where Ramsden and Entwistle (1981) devised the Course Perceptions Questionnaire (CPQ) to identify the factors in the learning environment that influenced students' approaches to learning. This questionnaire elicited students' views on the quality of their courses using the following scales: Good Teaching, Openness to Students, Freedom in Learning, Clear Goals and Standards, and Appropriate Workload. It was found that when academic departments were perceived to provide these characteristics, students were likely to learn more effectively. Further studies confirmed the importance of these characteristics in defining good teaching (Entwistle & Ramsden, 1983; Ramsden *et al*., 1989; Entwistle & Tait, 1990) and thus, in designing the CEQ scales, efforts were made to capture these factors (Ramsden, 1991). Preliminary investigations of the CEQ confirmed the internal consistency of the scales and demonstrated clear evidence of its ability to discriminate validly between courses (Ramsden, 1991). It was on the basis of this evidence that the PIRG recommended a national trial of a 30-item version CEQ to evaluate its suitability as a performance indicator of teaching in the higher education sector (Linke, 1991).

The 30-item CEQ consists of five scales: Good Teaching, Clear Goals, Appropriate Workload, Appropriate Assessment, and Emphasis on Independence. Students are required to evaluate the teaching they experience during their course of study by responding to the 30 items using a 5-point Likert scale ranging from 'definitely disagree' to 'definitely agree'. The instrument was trialled in 1990 among final-year students of different disciplines across a range of higher education institutions in Australia, with 3372 valid responses being received (Ramsden, 1991). To test the validity and reliability of the instrument, the structure of the five scales as envisaged by the development work was confirmed by factor analysis, with just three items not loading as expected conceptually (Ramsden, 1991). The internal consistency of the five scales was then examined by determining Cronbach alpha values and all were found to be satisfactory. The instrument was further validated by examining the relationships between the scale scores and various external criteria. One of the external criteria used was students' learning approaches as measured by the Approaches to Studying Inventory (ASI). The expected relationships between students' perceptions of teaching and learning approaches were confirmed, in that perceptions of good teaching were significantly positively related to a deep approach to learning (Ramsden, 1991). Subsequent analysis with the trial data, comparing the CEQ scores of different departments in different institutions, revealed that the instrument was satisfactorily able to discriminate between the teaching performance of different academic units (Ramsden, 1991). The initial validation of the instrument was enhanced by reference to another data set collected in the same year. The Mathews Report, reviewing the state of the accounting discipline in higher education in Australia, had, by arrangement with the CEQ developers, used the instrument to gather data from accounting students (Mathews *et al.*, 1990).

Separate analysis conducted on this data confirmed the satisfactory statistical properties of the instrument (Ramsden, 1991).

The highly satisfactory results from the CEQ trial resulted in the development of a shorter form of the instrument, which was then incorporated into the Graduate Careers Council of Australia's (GCCA) annual survey of graduates in 1993 (Eley, 1998). This 23-item version is made up of the items with the strongest loadings for four of the five scales in the original 30-item instrument. The Emphasis on Independence scale was eliminated due to its weaker scale structure and a new scale of Generic Skills was introduced (Wilson *et al.*, 1997). This new scale was included in light of increasing emphasis on transferable generic skills in higher education, and was designed to provide a brief uniform national measure of graduates' perceptions of the extent to which their courses had developed these skills. The alpha values of the five scales in the 23-item CEQ, while marginally lower than those for the original 30-item version, indicated acceptable levels of internal consistency. Additionally, the factor structure of this short form of the instrument was very satisfactory, leading Wilson *et al.* (1997, p. 41) to conclude that it "offers a stable factor structure equal to the … full form, with the advantage of cleaner relationships between items and scales". This 23-item short form of the CEQ is now the most commonly used version of the instrument (Wilson *et al.*, 1997).

Given the potential uses of the CEQ and its strong statistical properties, there is surprisingly little research reporting the replication of the original study (Ramsden, 1991) or the validation of the instrument for use in other contexts (Wilson *et al.*, 1997). Trigwell and Prosser (1991), using a sample of 55 final-year Australian nursing students, reported a scale structure for the CEQ broadly similar to that reported from the national trial. Richardson (1994, p. 60) validated the CEQ for use in the British higher education environment, as he contends that "it should not be assumed that an instrument developed in Australia would necessarily transfer to a British context". The alpha values calculated on the British data were lower than those reported in Ramsden (1991), and while the values for three of the scales were very acceptable, those for the Appropriate Assessment scale (0.47) and the Emphasis on Independence scale (0.55) were less satisfactory. First order factor analysis broadly confirmed the five-scale pattern of the original study. However, the composition of the Appropriate Assessment scale did not appear to be robust, and a number of other items loaded on different scales than would have been expected (Richardson, 1994). Further analysis identified one dominating second order factor, which Richardson concluded could be interpreted as the CEQ providing a "single global measure of perceived teaching quality" (p. 65). While the sample size used in this study was small (89 usable responses) and has been subsequently questioned, Richardson argues that the small sample size is only problematic if the findings of his study were substantially different to those reported elsewhere. Hence, he concludes that the CEQ can be employed to evaluate teaching in academic units in Britain.

In another British study, Broomfield and Bligh (1998) validated the CEQ specifically for use with medical students. While the study confirmed the basic scale structure of the instrument, it reported the splitting of the Good Teaching scale into two factors relating to, firstly, teacher interaction and presentation and, secondly, the quality of feedback given to students. Eley (1998) examined alternative ways of presenting the CEQ items. In a study with 352 business and engineering students in Australia, he prepared three parallel versions of the 23-item CEQ. One version was simply the regular form of the instrument where students respond to the items on the 'agree–disagree' Likert scale. The second version altered the regular CEQ question stems to devise Behavioural Observa-

TABLE 1. Population and sample response rate

Class Group	Population	Responses	% Response
BAAF 2000	110	71	64.5%
BAAF 2001	96	75	78.1%
MBSA 2000	33	26	78.8%
MBSA 2001	41	32	78.0%
Total	*280*	*204*	*72.9%*

tion Scale (BOS) questions, and the final version used Dimensional Rating Scale (DRS) questions. Both the internal reliabilities of the scales and the fit of the factor structure to those scales were better for the BOS and DRS versions of the instrument compared to the 'agree–disagree' standard version. Additionally, the alternative versions showed increased capabilities to distinguish between academic units, leading to the suggestion that alternative question formats could further enhance the psychometric properties of the CEQ. While such alternative formats may come into wide-scale use in the years ahead, it should be noted that the reliability and validity of the standard form of the instrument reported in the Eley study were very acceptable and in line with those reported in the studies described previously.

Objectives and Data Collection

The primary objective of the current study is to establish the reliability and validity of the CEQ for use with accounting students in an Irish university. The results of this process are compared to the findings from other validation studies. Additionally, the conditions for the effective application of the CEQ in this environment are outlined.

The population for the study consisted of students at Dublin City University (DCU) on the BA in Accounting and Finance (BAAF) and the MBS in Accounting (MBSA) programmes graduating in 2000 and 2001. The CEQ was distributed to the BAAF classes and the MBSA 2001 students at a lecture in the last week of their academic year. For the MBSA 2000 students, questionnaires were mailed to students' home addresses, approximately five months after their graduation. A follow-up letter was sent one month later to maximise the responses received. The purpose of the study was explained to all groups and they were reassured that their responses would not be used in any context other than for the purposes of this research. Table 1 presents a breakdown of the potential population and the response rate for each class.

The 23-item CEQ was the instrument used in this study. As previously outlined, it contains five scales: Good Teaching, Clear Goals, Appropriate Workload, Appropriate Assessment, and Generic Skills. The Good Teaching Scale (GTS) is made up of six items and is defined by teaching characteristics such as providing students with feedback, explaining issues and concepts, making the programme interesting, motivating students and understanding their problems. Four items constitute the Clear Goals and Standards Scale (CGSS) and embrace issues relating to the extent to which students are clear about what is expected of them. The Appropriate Workload Scale (AWS) comprises four items and assesses students' perceptions of the reasonableness of the workload. Three items make up the Appropriate Assessment Scale (AAS) and capture the extent to which assessment encourages understanding of material rather than rote memorisation. The Generic Skills Scale (GSS), comprising six items, attempts to evaluate the extent to

TABLE 2. Items of the Course Experience Questionnaire

Good Teaching Scale (6 items)	
3	The teaching staff of this course motivated me to do my best work
7	The staff put a lot of time into commenting on my work
15	The staff made a real effort to understand difficulties I might be having with my work
16	The teaching staff normally gave me helpful feedback on how I was going
17	My lecturers were extremely good at explaining things
19	The teaching staff worked hard to make their subjects interesting
Clear Goals and Standards Scale (4 items)	
1	It was always easy to know the standard of work expected
6	I usually had a clear idea of where I was going and what was expected of me in this course
13*	It was often hard to discover what was expected of me in that course
23	The staff made it clear right from the start what they expected from students
Appropriate Workload Scale (4 items)	
4*	The workload was too heavy
14	I was generally given enough time to understand things I had to learn
20*	There was a lot of pressure on me to do well in this course
22*	The sheer volume of work to be got through in this course meant it couldn't all be thoroughly comprehended
Appropriate Assessment Scale (3 items)	
8*	To do well in this course all you really needed was a good memory
12*	The staff seemed more interested in testing what I had memorised than what I had understood
18*	Too many staff asked me questions just about facts
Generic Skills Scale (6 items)	
2	The course developed my problem solving skills
5	The course sharpened my analytical skills
9	The course helped me develop my ability to work as a team member
10	As a result of my course I feel confident about tackling unfamiliar problems
11	The course improved my skills in written communication
21	My course helped me to develop the ability to plan my own work

* = Reverse coded items.

which students perceive that their courses developed generic skills. The specific items within each scale are set out in Table 2.

The collected data were entered into an SPSS file. Before undertaking the analysis, the coding of items that had a meaning opposite to that of the relevant scale was reversed, so that a high score on each item corresponded to a judgement of good teaching (Richardson, 1994). For example, where a student had responded 1 ('strongly disagree') to item 13 ('it was often hard to discover what was expected of me in this course'), this was recoded as 5 to reflect the student's perception of clear goals and standards. All items that were recoded are identified in Table 2 by an asterisk.

Results and Discussion

Reliability and Validity of CEQ

Cronbach's coefficient alpha was used to measure the internal consistency of each scale. The alpha can have a value ranging from 0 to + 1 and the higher the value the greater the internal consistency. However, as reliability is a function of the number of items in the scale, the alphas for scales of different lengths are not directly comparable (Ainley & Johnson, 2001). The alphas calculated in the current study are set out in Table 3.

TABLE 3. Alpha values and between scale correlations

	Alphas	CGSS	AWS	AAS	GSS
GTS	*0.76*	0.382	0.262	0.221	0.358
CGSS	*0.78*		0.316	0.084	0.239
AWS	*0.73*			0.118	0.068
AAS	*0.69*				0.138
GSS	*0.66*				

While the alpha values for three of the scales (GTS, AAS and GSS) were lower than those reported in Wilson *et al.* (1997) they are nevertheless satisfactory, and demonstrate moderate to high levels of internal consistency. The alpha values were then compared to the correlations between the scales and in each case they were greater than the associated correlations, providing further evidence of the internal reliability of the five scales (Eley, 1998).

Following the approach in previous validation studies, exploratory factor analysis was conducted using principal components extraction (Ramsden, 1991; Richardson, 1994; Wilson *et al*., 1997). Given the correlations between the scales, an oblique rotation was deemed most appropriate (Cattell, 1978). A combination of the scree test and the eigenvalue greater than one rule was used to determine the number of factors to be extracted. This resulted in the expected five-factor structure which explained 53.5% of the variance. Item 21 ('plan own work') was the only item that did not load as expected, loading negatively on the AWS rather than on the GSS. This suggests that students negatively associate the ability to plan their own work with their workload. Item 11 ('written communication') loaded on the GSS as envisaged conceptually, but it also cross-loaded less strongly onto the GTS and the AWS. This indicates that students associate the development of their communication skills with good teaching, but they perceive it as adversely affecting their workload. All other items loaded as expected and the full factor pattern is set out in Table 4.

The factor pattern illustrated in Table 4 confirms the clean relationships between the items and scales of the 23-item form of the CEQ as reported in Wilson *et al*. (1997) and validates the appropriateness of the instrument for capturing students' perceptions of five key aspects of teaching quality. Factor analysis on the scale scores yielded a one-factor solution and demonstrated that the CEQ can provide a single global measure of perceived teaching quality.

Further evidence of the validity of the CEQ is provided by examining the relationships between CEQ scores and external criteria (Wilson *et al*., 1997). One such criterion frequently used is overall course satisfaction (Ramsden, 1991; Wilson *et al*., 1997). The CEQ instrument distributed asked students to state the extent of their overall satisfaction with their course. All the CEQ scales showed a significant positive correlation with overall satisfaction, strengthening the instrument's validity for use with Irish accounting students.

Potential Applications and Limitations of the CEQ

This study has demonstrated that the CEQ is a suitable performance indicator that can be used by accounting educators to help them assess the teaching quality of accounting courses in Ireland. It offers reliable and useful feedback on accounting students'

TABLE 4. Factor analysis of CEQ

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
GTS					
7	0.812				
16	0.778				
15	0.696				
19	0.532				
3	0.506				
17	0.459				
AWS					
22		0.720			
20		0.707			
4		0.705			
14		0.597			
21		− 0.306			
AAS					
8			0.807		
12			0.770		
18			0.760		
GSS					
2				0.786	
5				0.744	
10				0.648	
9				0.610	
11	0.307	− 0.325		0.344	
CGSS					
1					0.793
23					0.768
13					0.724
6					0.702

Note: Loadings less than 0.30 are omitted.

perceptions of aspects of teaching that are central to the quality of their learning. Furthermore, within institutions the collection of CEQ data over time will facilitate the monitoring of changes in teaching quality at the individual course level and the department level.

The CEQ also provides a basis for inter-institutional comparisons within the accounting discipline. The sharing of CEQ data facilitates the identification of best practices of teaching that could be used collectively to enhance accounting education. However, care is needed in making comparisons between institutions, as observed differences may be caused by extraneous factors, such as variations in student characteristics and levels of funding. Ainley and Long (1994) contend that informed judgement is necessary in making all CEQ comparisons, as scores should be regarded as indicative rather than conclusive.

CEQ scores provide reliable indicators of teaching strengths and weaknesses but do not generate a complete measure of teaching quality. As teaching is a highly complex cognitive activity, multiple sources of data are required to provide a comprehensive evaluation (McKenzie *et al*., 1998; Saroyan & Amundsen, 2001). Therefore, the CEQ should not be used in isolation in designing intervention strategies regarding teaching

practices, rather it should be used in conjunction with other performance indicators. Wilson *et al*. (1997, p. 48) recommend that CEQ results for a course should be viewed as a starting point rather than an end point in any evaluation process, and survey results should be used "to indicate productive directions for investigations into underlying factors affecting teaching quality". Furthermore, as the CEQ is a course-level performance indicator, it should not be used to provide diagnostic feedback regarding individual lecturers.

Conclusions

Evaluating teaching quality is a topic of major interest in the higher education sector. This study has highlighted the potential of the CEQ to offer reliable and useful feedback on teaching effectiveness. The conceptual foundations of the instrument and its development in Australia were described and studies validating the instrument in other contexts were examined.

This study validated the CEQ for use with Irish accounting students. Factor analysis confirmed the five-scale structure of the instrument and reliability analysis supported the internal consistency of each scale. The instrument offers Irish accounting educators the opportunity to gather a valuable set of data that can be used to assess the quality of teaching and to identify areas for improvement. It also has the potential to facilitate benchmarking for best teaching practices in the discipline of accounting.

Acknowledgement

The authors wish to acknowledge the assistance of Jim Morrow (MBSA graduate) in gathering the data for this study.

Notes on Contributors

MARANN BYRNE is a senior lecturer in Accounting at Dublin City University. She is the course director for the BA in Accounting and Finance and her current research interests include students' approaches to learning, second level accounting and the history of accounting education in Ireland.

BARBARA FLOOD is a lecturer in Accounting at Dublin City University. She is the course director for the postgraduate Professional Diploma in Accounting. Her current research interests focus on the learning experiences of students within professional accounting education, the history of accounting education in Ireland and conceptions of learning. *Correspondence*: Dublin City University Business School, Dublin City University, Dublin 9, Ireland. E-mail: barbara.flood@dcu.ie

REFERENCES

AINLEY, J. & JOHNSON, T. (2001) *Course Experience Questionnaire 2000: an interim report* (Melbourne, Australian Council for Educational Research).

AINLEY, J. & LONG, M. (1994) *The 1994 Course Experience Questionnaire Graduates* (Melbourne, Graduate Careers Council of Australia).

ALBRECHT, W. S. & SACK, R. J. (2000) *Accounting Education: charting the course through a perilous future* (Sarasota FL, American Accounting Association).

AMERICAN ACCOUNTING ASSOCIATION (AAA) (1986) Committee on the Future Structure, Content and Scope of Accounting Education (The Bedford Committee) Future accounting education: preparing for the expanded profession, *Issues in Accounting Education*, 1, pp. 168–195.

BIGGS, J. (1999) *Teaching for Quality Learning at University* (Buckingham, SRHE/Open University Press).

BOK, D. (1986) *Higher Learning* (Cambridge, MA, Harvard University Press).

BROOMFIELD, D. & BLIGH, J. (1998) An evaluation of the 'short form' course experience questionnaire with medical students, *Medical Education*, 32, pp. 367–369.

BUCKLEY, F. & HURLEY, J. (2001) Change in the nature of the university: the quality regime and its effects on the university teaching environment, *Social Science Information*, 40 (4), pp. 546–576.

CATTELL, R. B. (1978) *The Scientific Use of Factor Analysis in Behavioural and Life Sciences* (New York, Plenum Press).

DAWKINS, J. S. (1987) *Higher Education: A Policy Discussion Paper* (Green Paper) (Canberra, Australian Government Publishing Service).

DAWKINS, J. S. (1988) *Higher Education: A Policy Statement* (White Paper) (Canberra, Australian Government Publishing Service).

ELEY, M. G. (1998) The course experience questionnaire: some effects of altered question formats on the CEQ's effectiveness as an instrument. Paper presented at the 24th Annual Conference of HERDSA, Auckland, New Zealand, 7–10 July.

ENTWISTLE, N. J. & RAMSDEN, P. (1983) *Understanding Student Learning* (London, Croom Helm).

ENTWISTLE, N. J. & TAIT, H. (1990) Approaches to learning, evaluations of teaching and preferences for contrasting academic environments, *Higher Education*, 19, pp. 169–194.

INTERNATIONAL FEDERATION OF ACCOUNTANTS (IFAC) (1996) *International Education Guideline 9—Pre-qualification Education, Assessment of Professional Competence and Experience Requirements of Professional Accountants* (New York, IFAC).

KWAN, K. (1999) How fair are student ratings in assessing the teaching performance of university teachers?, *Assessment and Evaluation in Higher Education*, 24 (2), pp. 181–195.

LINKE, R. D. (1991) *Report of the Research Group on Performance Indicators in Higher Education* (Canberra, Australian Government Publishing Service).

LIZZIO, A., WILSON, K. & SIMONS, R. (2002) University students' perceptions of the learning environment and academic outcomes: implications for theory and practice, *Studies in Higher Education*, 27 (1), pp. 27–52.

LONG, M. (1995) The course experience questionnaire and Australian universities, *Unicorn*, 21 (3), pp. 27–35.

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

MARSH, H. W. & ROCHE, L. A. (1994) *The Use of Students' Evaluations of University Teaching to Improve Teaching Effectiveness* (Canberra, Australian Government Publishing Service).

MARTON, F. & SALJO, R. (1976) On qualitative differences in learning I—outcome and process, *British Journal of Educational Psychology*, 46, pp. 4–11.

MATHEWS, R. L., BROWN, P. R. & JACKSON, M. A. (1990) *Accounting in Higher Education: Report of the Review Group of the Accounting Discipline in Higher Education* (Canberra, Australian Government Publishing Service).

MCKENZIE, J., SHEELY, S. & TRIGWELL, K. (1998) Drawing on experience: an holistic approach to student evaluation of courses, *Assessment and Evaluation in Higher Education*, 23 (2), pp. 153–163.

RAMSDEN, P. (1991) A performance indicator of teaching quality in higher education: the course experience questionnaire, *Studies in Higher Education*, 16 (2), pp. 129–150.

RAMSDEN, P. & ENTWISTLE, N. (1981) Effects of academic departments on students' approaches to studying, *British Journal of Educational Psychology*, 51, pp. 368–383.

RAMSDEN, P., MARTIN, E. & BOWDEN, J. (1989) Social environment and sixth form pupils' approaches to learning, *British Journal of Educational Psychology*, 59 (2), pp. 129–142.

RICHARDSON, J. T. E. (1994) A British evaluation of the course experience questionnaire, *Studies in Higher Education*, 19 (1), pp. 59–68.

ROMAINVILLE, M. (1999) Quality evaluation of teaching in higher education, *Higher Education in Europe*, 24, pp. 415–424.

SAROYAN, A. & AMUNDSEN, C. (2001) Evaluating university teaching: time to take stock, *Assessment and Evaluation in Higher Education*, 26 (4), pp. 341–353.

STRINGER, M. & FINLAY, C. (1993) Assuring quality through student evaluation, in: C. ELLIS (Ed.) *Quality Assurance for University Teaching* (Buckingham, SRHE/Open University Press).

TRIGWELL, K. & PROSSER, M. (1991) Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes, *Higher Education*, 22, pp. 251–266.

WACHTEL, H. (1998) Student evaluation of college teaching effectiveness: a brief review, *Assessment and Evaluation in Higher Education*, 23 (2), pp. 191–211.

WILSON, K. L., LIZZIO, A. & RAMSDEN, P. (1997) The development, validation and application of the course experience questionnaire, *Studies in Higher Education*, 22 (1), pp. 33–53.