Educational Research (ISSN: 2141-5161) Vol. 2(8) pp. 1417-1430 August 2011
Available online@ http://www.interesjournals.org/ER

Full Length Research Paper

Methods behaving differently: The effects of method of data analysis on understanding student satisfaction with their educational experience

Walid El Ansari[1]* and Reza Oskrochi[2]

[1]Faculty of Applied Sciences, University of Gloucestershire, Gloucester United Kingdom.
[2]School of Technology, Department of Mathematical Sciences, Oxford Brookes University, Oxford, United Kingdom.

Accepted 20 June, 2011

This is a large-scale survey (138 modules, 2650 questionnaires) into student satisfaction at the School of Health and Social Care of a British University in the UK. We assessed the extent and influence of the statistical method employed in the analysis of satisfaction data on the actual understanding of student satisfaction. Satisfaction was computed by four commonly used statistical methods of analyses of satisfaction data selected through literature review. After scrutiny, one was chosen to act as the 'preferred method'. Each method was used to individually analyse the dataset twice: once without controlling for the effects of variables and clusters that were under investigation; and again after controlling for such effects. Findings of the analyses were compared to those of the 'preferred method'. After controlling, some initially observed effects of some variables on satisfaction were subsequently lost. Compared to the findings of the 'preferred method', different methods exhibited over- or underestimation of satisfaction. Satisfied students were post-qualifying students who attended term one, academic level 2 modules in smaller classes (in terms of student numbers) and whose assessments were 50/50 combinations of coursework/ exams.

Keywords: Satisfaction with education, learning and teaching, student evaluations, statistical methods, evidence base

INTRODUCTION

The education sector is witnessing three transformations. Firstly, higher education institutions are increasingly attentive to monitoring the quality of the learning and teaching experiences (Taylor and Hill, 1993; Chapple and Murphy, 1996), in order to prepare students for a multifaceted, rapidly changing world (Lowe, 1996). Second, there is more emphasis for practice to be premised on valid evidence. This focus on quality that is evidence-based has contributed to a third transformation: the monitoring of students' satisfaction with their learning experiences as a source of course evaluation (El Ansari, 2002a and 2002b). Student satisfaction acts as guidance for students, decision-makers, and institutional performance (Kerridge and Mathews, 1998; White et al., 1999), and to assure learning and teaching quality standards (Murphy and Harris, 1995).

These aspirations about the delivery of quality educational experiences for Health and Social Care students, as well as the monitoring of quality by evaluations of student satisfaction have gained many voices. The result is more studies of evaluations based on student ratings than of all the other means used to evaluate college teaching combined (Cashin, 1988). Some of these evaluations have explored satisfaction using qualitative interviews and focus groups (Kapborg, 2000; Latter et al., 2000). Less inquiries employed mixed methods comprising both qualitative and quantitative paradigms (Saksomboon, 2002; El Ansari, 2004). More

*Corresponding author E-mail: walidansari@glos.ac.uk

traditionally, student ratings of their teaching and learning experiences have been assessed using quantitative surveys and questionnaires (El Ansari and Oskrochi, 2004; Brown and Edelmann, 2000; Espeland and Indrehus, 2003; El Ansari and Oskrochi, 2006). The scope of this paper is the methodological limitations of this latter 'quantitative' group; particularly the influences of the type and inherent characteristics of the statistical method employed in the analysis of satisfaction data on the actual understanding of student satisfaction.

Satisfaction is a multidimensional construct. Four terms are used in this paper to reflect the aspects that need consideration in satisfaction studies. The term 'variable/s' indicates the range of demographic or educational-related variables that influence student satisfaction. These include gender, ethnicity, disability, age, academic term and level, student's entry qualifications, study mode (full/ part time), qualification aim and class size (number of students) (Ofori, 2000; El Ansari, 2002a). The term 'question/s' designates the actual items that are investigated in questionnaires of student satisfaction with their educational experience. These include the areas that comprise the learning and teaching 'climate' (El Ansari, 2003): module administration and organisation, module team, feedback mechanisms, assessment procedures, course characteristics, and university resources (Lee et al., 1999; El Ansari, 2002b). Due to the multiplicity of these 'climate' areas, the number of questions employed in satisfaction questionnaires is often large and requires reduction into a smaller number of constructs. Hence the term 'cluster/s' indicates the grouping/s of question into meaningful constructs of the learning and teaching experience. These clusters are then investigated individually or collectively in relation to the variables under study in order to assess their individual/ collective influences on satisfaction. The term 'method/s' specifies the statistical methods that are used in analysing the data.

Quantitative studies of satisfaction: some methodological challenges

In spite of the body of literature on student surveys with their health and social care learning experiences, a feature of the published studies is the conceptual and technical variations in the range of the definitions, methodologies and analysis techniques that are employed by different authors. For instance, at the *conceptual* level of definitions, in Wales, Glossop (2001) highlighted the lack of common definitions to make valid comparisons between teaching institutions. At the *technical* level of response scales for the questions of student satisfaction surveys, there seems relative agreement in the use of response scales to indicate the degree of agreement/ disagreement with a group of statements, or to rate the importance of each question. Nevertheless, in Australia, Eley and Stecher (1997) reported concerns about the use of different response scales. They compared two types of response scale formats used in students' questionnaires of teaching evaluation, and found that some formats yielded measurable improvements in reliability and in the capability to distinguish amongst levels of teaching quality. Likewise, at the level of the *analysis* of satisfaction surveys, the variations of the statistical methods that are employed to analyse the data are more pronounced. These exhibited many dissimilarities; most published studies on student satisfaction are in disagreement in the methods employed in the analysis of responses.

Variations in analyses techniques: some examples

Method 1: analysis based on generating the *extent and the index of satisfaction* using the proportion of satisfied students– in the UK, Hayden and Thompson's (1996) satisfaction survey employed equations on the data to compute three different/ parallel indicators of student satisfaction: the extent of and the index of satisfaction. They calculated overall satisfaction as the product of these two indicators.

Method 2: analysis based on using the *proportion of satisfied students*– in a survey of nurses in the UK, Eaton et al. (2000) identified differences between student satisfaction levels by calculating the proportion of students who agreed in their rating with the statements that the questionnaire investigated.

Method 3: analysis based on using a *mean score of satisfied students* that is then employed in a *univariate* manner– in Wales, Kinsella et al. (1999) surveyed and compared student nurses based on the students' mean rating of the satisfaction questions. In Jordan, Nahas et al. (1999) compared the mean score of satisfaction for each question, and for clusters that comprised groups of questions. Kniveton (1996) used mean satisfaction scores to compare students' perceptions of assessment methods and Cowman (1996) to study the course experiences of student nurses. In Cameroon, Amin (1994) computed the mean student ratings to examine teacher's characteristics that explained the students' overall ratings of their courses. In England, Hoskins et al. (1997) examined the degree performance as a function of students' demographic variables employing similar methods.

Method 4: analysis based on using a *mean score of satisfied students* that is then employed in a *multivariate* manner. In Hong Kong, Kwan (1999) examined the effects of the course characteristics on student satisfaction ratings of university teaching using a multivariate analysis of variance (MANOVA). Similarly, in

the UK, El Ansari and Oskrochi (2004) employed MANOVA to examine the individual and collective effects of many educational/ demographic variables on the satisfaction of health and social care students.

Some investigations have highlighted the effects of variations in definitions or response formats (Glossop, 2001; Eley and Stecher, 1997). Similarly, El Ansari and Moseley (2011) examined five common choices of satisfaction summary measures that are commonly used in the literature and highlighted that the five measures were correlated, but levels of student satisfaction varied widely according to the summary measure that was used. Fewer inquiries addressed the influence of the type of statistical method employed in the analysis of student satisfaction ratings on the actual findings, conclusions and subsequent implications. Thus the question that this paper examined was: when methods of analyses differ, do differences matter?

Aim of the paper

This study explored the extent and influence of the type of statistical method employed in the analysis of satisfaction data on the actual understanding of student satisfaction with their educational experience. The inquiry formed part of a wider survey into student satisfaction at the School of Health and Social Care of a British University in the UK. The six aims were to:

- Illustrate some published examples of the variations in the statistical methods employed in the analyses of student satisfaction
- Select four different statistical methods that are commonly reported in the analysis of satisfaction, compare their properties and choose one of them as a 'preferred method' for subsequent comparisons
- Employ each of the four statistical methods individually to analyse a survey (2650 questionnaires) of student satisfaction. For each method, two separate analyses were undertaken: an initial one that did not control for the effects of other variables and clusters that were under investigation; and a second analysis that controlled for such effects
- Compare the student satisfaction findings of each of the three statistical methods with those of the 'preferred method' on two occasions: once without controlling for the variables and clusters that were under investigation and the other after controlling for such effects
- Assess the extent and manner of influence of the type of statistical method on understanding student satisfaction
- Discuss the satisfaction findings according to the preferred method

In addition, the study sought to answer the questions: "What explains satisfaction?" and, "What explains students' achieved grades?"

METHODS AND FINDINGS

Four methods of analyses of survey satisfaction data: selection criteria and examples

The selection of the methods of analyses that this paper examined was premised on a review of the literature on student satisfaction (particularly in Health and Social Care education). Studies that employed quantitative methods were noted. The methods below were considered because they fulfilled the selection criteria: 1) were commonly used by published authors from different countries; 2) illustrated a contrast from one another (assumptions, rationale, and how satisfaction was measured); 3) represented different degrees of 'rigour' (confounding effects, possibility to assess the individual effects of each variable per se); 4) represented different requirements for the analysis (level of statistical expertise and time required to compute the method); and 5) offered analytical advantages/ disadvantages over one another (use in the analyses of all the available information that the dataset offers, control of associations between the study variables as well as correlations between the teaching and learning clusters). Four methods emerged:

- Method 1: analysis based on generating three-dimensional indicators: the *'extent'*, *'index'* and *'overall'* satisfaction using the proportion of satisfied students. Each indicator is applied to each cluster independently (Hayden and Thompson, 1996). Mathematical equations were employed to generate the 'extent' and 'index' of satisfaction premised on whether respondents expressed either strong satisfaction/ satisfaction with each variable. Overall satisfaction was computed as the multiplication of the 'extent' and 'index'.
- Method 2: analysis based on using the *proportion of satisfied students*. The dependent variable is the dichotomous response of satisfaction/ dissatisfaction with each learning/ teaching cluster. The independent variables are the demographic/ educational variables. The statistical tools include tests of proportion or logistic regression for the binary outcomes (Kleinbaum and Klein, 2002). Each variable's effect is assessed by including it as an independent variable into the logistic regression model.
- Method 3: analysis based on the *mean score of satisfied students*. The dependent variable is the mean score of satisfaction in each cluster. The independent variables are the demographic/ educational variables. The statistical tools include one or two sample t-test;

analysis of variance; or linear regression for the mean score of satisfaction (Pawitan, 2001) applied to each cluster. Each variable's effect is assessed by including that variable as an independent variable into the linear regression model.

- Method 4: analysis based on the multivariate realisation of *mean score of satisfied students*. The dependent variables are the mean scores of satisfaction in every cluster. The independent variables are the demographic/ educational variables. The statistical tools include multivariate test of hypothesis or multivariate analysis of variance (MANOVA) (Tabachnick and Fidell, 1996). Before controlling, each variable's effect is assessed by including that variable in the model. After controlling, each variable's effect is assessed by including all demographic and educational variables in the model. In this method, the associations between the clusters are also controlled for.

Table 1 scrutinised five features of the four methods described above. The similarities and differences suggested that Method 4 would act as the 'preferred method': it was the only method that allowed for correlations between the variables to be considered and also for the associations between clusters of teaching and learning to be controlled. Hence, this method was selected to act as the 'benchmark' that the findings of the other methods would subsequently be compared with.

The dataset

In the academic year 2000/ 2001, after ethical approval, the first author undertook a student satisfaction survey at the School of Health and Social Care of a British University in the UK. The one-page questionnaire (18 close ended items scored on 5-point scales, 1= 'Positive Perception' and 5 = 'Negative Perception') was adopted from that developed and validated in England (Kerridge and Mathews, 1998). The questions addressed educational-related factors: module administration, academic level and term, module team, assessment procedures, course characteristics, and university resources (Lee et al., 1999; El Ansari, 2002a). It also explored student-related and demographic variables: gender, ethnicity, disability, age, entry qualifications, study mode (full/ part time), and qualification aim (Rhodes et al.1999; Ofori, 2000; Chevannes, 2000; El Ansari, 2002b and 2003). The sampling frame was all the modules in the School. Participation in the study was voluntary; those not wishing to participate could leave or stay in the class when the questionnaires were administered. Depending on the class size (student numbers), the response rates varied from 75% to 100%. Smaller classes had higher response rates. The questionnaire was first piloted, data protection and confidentiality were observed, and the survey yielded 2662 questionnaires. Table 2 depicts the survey's 18 questions by the mean percent of students who were satisfied and by the mean score of satisfaction.

Data reduction: three clusters of teaching and learning

The dataset was prepared for the analysis by categorising the 18 questions into meaningful constructs. Hierarchical cluster analysis reveals the natural groupings within the data so the degree of association is strong between members of the same cluster and weak between members of different clusters (Clustan, 2005; Statsoft Electronic Textbook, 2005). Using SPSS, cluster analysis determined the number and composition of underlying constructs to be used in subsequent analyses. The survey's 18 questions generated three learning and teaching clusters. Figure 1 shows the Dendrogram that was generated by the process, the cluster's name; the questions that comprised each cluster; their average linkage; and their Cronbach's α (internal consistency) (Cronbach, 1951). Pearson's coefficients showed positive, high to fair correlations between the three clusters (0.81 to 0.58, $P = 0.01$).

The survey: modules and students

Configurations of teaching teams delivered the 138 modules that were surveyed. These Health and Social Care modules included life sciences, pharmacology, performance and movement, physiotherapy, occupational therapy, management and organisation, and professional development modules. Others were public health, research, psychology, and psychosocial care modules. Many nursing modules were also represented: general practice, school, community, learning disability, public health, mental health, cardiac surgical, medical, and adult and children's nursing. Other modules adressed perioperative or palliative care, rehabilitation, care of the child, collaborative practice, quality and clinical governance, planning nursing care, chronic illness, play, pain, and relationships and the dying.

The contribution of each of 138 modules to the sample was varied: 78% of the modules contributed less than 1% of the total questionnaires each; 19% of the modules contributed 1-3% each; only 3% of the modules contributed 3-6% each; and no modules contributed more than 6% of the total questionnaires each. The sample comprised 89% females, 99% of the students did not report disability and 77% were full-timers. 'White' ethnicity comprised most (92%) of the sample, and mean student age was 28.9 years (range 18 - 59 years). About 28% of

Table 1: Researching student satisfaction: a comparison of 5 main features of 4 analytical methods

	Method 1	Method 2	Method 3	Method 4
1. Overview				
Description	Indices of satisfaction**:** extent, index, and overall student satisfaction	Percentage of satisfied students	Mean satisfaction score of students across questions or clusters of questions	
Example/s of use	Hayden & Thompson (1996)	Kinsella et al. (1999); Eaton (2000)	Nahas et al. (1999)	Kwan (1999)
Assumptions, rationale	Measuring 3 indices of satisfaction could provide a holistic 3D view of satisfaction	Face value: if more students report satisfaction, then the module is satisfying i.e. a binary response of satisfied/dissatisfied	If mean satisfaction scores are computed for each question (Method 3) or cluster (Method 4), then 'missing responses' bias might be avoided	
Measuring satisfaction	How satisfaction is measured is very important in selecting the appropriate method for analysis. But how satisfaction is considered has effects on the comparability of findings across studies and generalisability. Considering what constitutes satisfaction, studies employ different cut offs of the 5-point Likert scales where 1= 'strong positive perception', 5= 'strong negative perception'. Some studies consider satisfaction when participants report either the 1st or 2nd option; others consider it if either the 1st, 2nd, or 3rd options are reported (see El Ansari and Moseley, 2011). Method (2) considers it if average satisfaction with all questions of a cluster is less than 3			
2. Analysis				
Unit of analysis	Students' rating of satisfaction to each learning & teaching cluster, each comprising several individual questions		Students' mean scores of satisfaction for each learning & teaching cluster, each comprising several individual questions	
Initial Computation	Equation for each of the 3 satisfaction indices	% of students reporting satisfaction	Computation of mean satisfaction scores across each cluster	
Comparisons potential	All methods allow comparisons across categories of variables (demographic: males/females, disabled/ not disabled etc. or educational: study mode, class size etc.)			
3. Confounding				
Confounding effects	High as controlling cannot be done	Low potential for confounding effects of demographic or educational variables only if controlling is undertaken		
Controlling (variables)	Method does not allow controlling generally	Methods allow for controlling the effects of other variables (although not always undertaken in published studies)		
Controlling (clusters)	Method does not allow controlling generally	Methods do not allow for controlling for the effects of other clusters		Possible to control for effect of clusters and overall effect of all clusters
Did our analysis control?	No, method does not allow controlling	Yes for variables	Yes for variables	Yes for variables and for clusters
Individual net effects	Not possible to assess	Possible to assess the net effects of each of the variables per se on satisfaction		

Table 1 cont. Researching student satisfaction: a comparison of 5 main features of 4 analytical methods

	Method 1	Method 2	Method 3	Method 4
4. Requirements				
Computational demand	High (3 different indices need to be generated)	Low with most of statistical packages		Intermediate with most statistical packages
Statistical expertise	Low	Intermediate		High
Time required	High, no package for calculating the indices	Less with most statistical packages		Intermediate with most statistical packages
5. Pros and cons				
Advantages	Provides insights; Does not ignore depth of satisfaction (strong satisfaction / satisfaction are not treated equally)	Possible to control for effects of other variables Uses more of the available information in the data than 1	Does not ignore depth of satisfaction. Uses more info than methods 1&2	Most robust as it controls for associations between variables and is the only method that also controls for the associations between clusters
Limitations	Least robust; uses minimal of available information; Does not allow for controlling; Requires many calculations	Not possible to control for effects of other clusters Ignores depth of satisfaction (strong satisfaction/ satisfaction treated equally)	Ignores associations between clusters	Almost no limitation
Recommended?	Not recommended (controlling for effects of variables and clusters is unlikely); if used then only for categorical data	Controlling for variables must be undertaken Only for binary response data (yes/ no)	For quantitative univariate responses (mean score); no cluster associations	For multivariate quantitative response data (e.g. mean score); as it allows for associations between clusters; controlling for variables is essential
Preferred method?	No	For binary responses	For univariate quantitative responses	For multivariate quantitative responses

Table 2: Descriptive statistics for questions employed in the survey

	Question	% of students who are Satisfied (C. I.)	Mean score of satisfaction* (C. I.)
Q1	Module ran smoothly	57 (55 –59)	2.42 (2.38 - 2.46)
Q2	Module increased my interest in the subject	59 (57 – 61)	2.43 (2.38 - 2.47)
Q3	Module team provided opportunity to ask questions	65 (63 – 67)	2.26 (2.21 - 2.30)
Q4	Module material was well presented	58 (56 – 60)	2.40 (2.36 - 2.44)
Q5	Module was thought provoking	60 (59 – 62)	2.36 (2.32 - 2.40)
Q6	Module assessment methods were appropriate	54 (51 – 55)	2.49 (2.44 - 2.54)
Q7	Module team displayed good knowledge	77 (75 – 78)	1.91 (1.87 - 1.95)
Q8	Module team correctly assumed level of skills I had	48 (46 – 50)	2.60 (2.56 - 2.64)
Q9	Module information available at beginning of module	71 (70 – 73)	1.99 (1.95 - 2.03)
Q10	Received helpful feedback	46 (44 – 48)	2.60 (2.56 - 2.64)
Q11	Seminar group sizes were small enough	59 (58 – 61)	2.33(2.29 - 2.38)
Q12	References needed for module available in library	38 (36 – 40)	2.88 (2.84 - 2.93)
Q13	Work required for module was appropriate	46 (45 – 48)	2.68 (2.64 - 2.72)
Q14	Module elements integrated into meaningful whole	55 (53 – 57)	2.44 (2.40 - 2.47)
Q15	Module was intellectually stimulating	61 (59 – 63)	2.35 (2.31 - 2.40)
Q16	Module is expected to be of direct use in my career	80 (79 – 82)	1.76 (1.73 - 1.80)
Q17	Module made me look at my profession differently	51 (50 – 53)	2.52 (2.48 - 2.56)
Q18	Module team styles were clear/ informative/ stimulating	58 (56 – 60)	2.40 (2.36 - 2.44)

C.I.: 95% Confidence Interval; *smaller numbers indicate more positive perception (i.e. higher satisfaction)

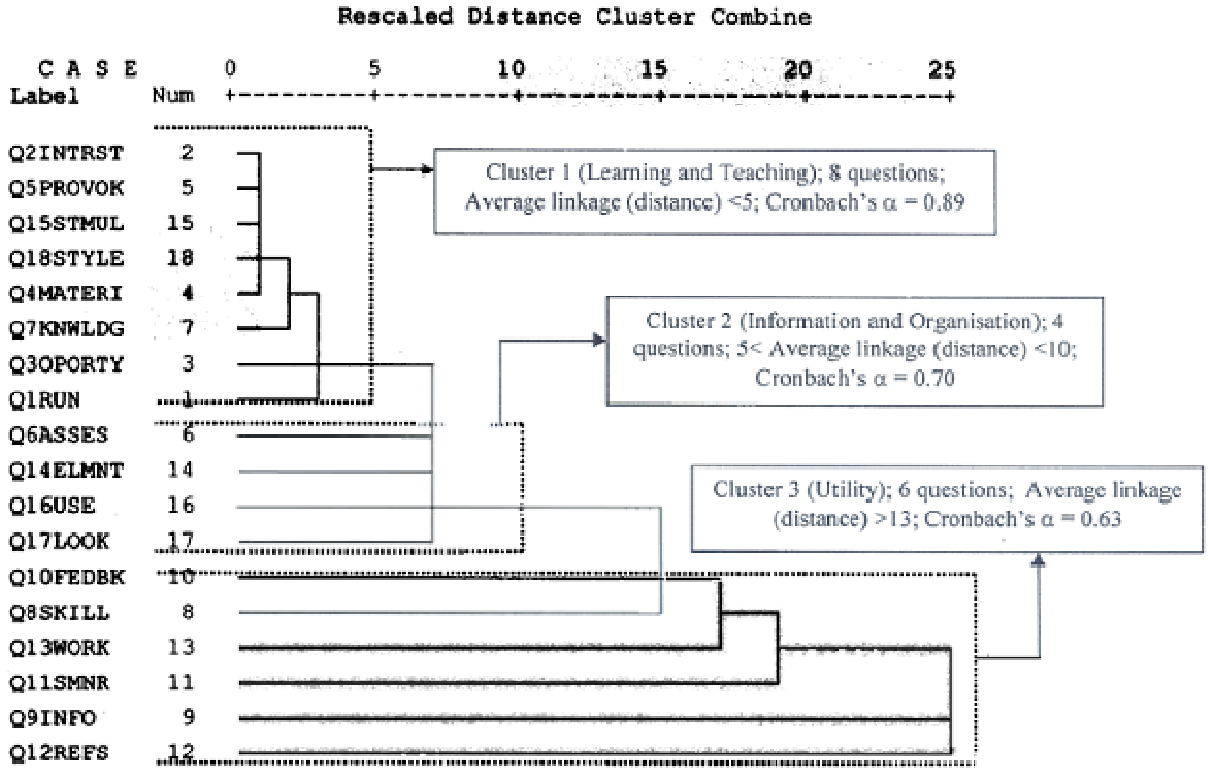

Figure 1: Dendogram: clustering of eighteen questionnaire items using average linkage

Table 3: Effects of demographic and educational variables (before controlling) on student satisfaction with individual clusters of learning and teaching by analysis method

	Effect of variables on clusters according to Preferred Method (Method 4)	Whether significant* effect of variables on clusters of learning and teaching is detected by			
Variable		Method 1	Method 2	Method 3	Method 4
	Cluster[#]	L I U	L I U	L I U	L I U
Demographic					
1. Gender	No effect	N N N	N N N	N N N	N N N
2. Disability	No effect	N N N	N N N	N N N	N N N
3. Ethnicity	No effect	N N N	N N N	N N N	N N N
4. Age bracket	Older students more satisfied	Y **N** Y	Y **N** Y	Y Y Y	Y Y Y
Educational					
5. Academic Term	Term one students more satisfied	**Y** Y Y	N **N** **N**	N Y Y	N Y Y
6. Academic Level	Level two students more satisfied	**Y** Y Y	N Y Y	**Y** Y Y	N Y Y
7. Study Mode	Part-timers more satisfied	Y **N** **N**	Y Y Y	Y Y Y	Y Y Y
8. Entry Qualification	Professional/ other postgraduate entry qualifications more satisfied	Y Y Y	**N** Y **N**	Y Y Y	Y Y Y
9. Nature of module	Post-qualifying more satisfied	Y Y Y	Y Y Y	Y Y Y	Y Y Y
10. Qualification aim	Diplomas more satisfied than others	Y Y Y	Y Y Y	Y Y Y	Y Y Y
11. Class size	Smaller class size more satisfied	Y Y Y	Y Y Y	Y Y Y	Y Y Y
12. Assessment Strategy	More satisfaction with 50/50 coursework and exam assessment	Y Y Y	Y Y Y	Y Y Y	Y Y Y

*: significant at $P<0.05$; [#] The three letter configuration indicates whether significant effect of variables is detected on each of the three clusters of Learning and Teaching (L), Information and Organisation (I), and Utility (U); N: no; Y: yes; Boxed bolded letters e.g. **Y** or **N** indicate the particular variable and cluster that are showing differences in their findings by method when compared with the same findings of the Preferred Method

the sample were traditionally aged students (<21 years), 20% were mature (21 – 25 years) and 52% were older mature students (>25 years). Almost 20% of respondents attended term one modules, while 43% and 37% attended term two and three modules of the academic year 2000/2001. The class size (student numbers) also varied, where 46% were of a large size (≥60 students), while classes with fewer than 20 students contributed 19% of the sample. Medium sized classes (either 20-40 or 41-60 students) contributed about 17.5% of the sample each.

As regards academic levels, Level 3 comprised 59% of the sample (defined as levels of performance expected from notional third year undergraduates). Level 4 (MSc) students were 6%. Roughly 29% of students attended post qualifying modules, and 9% attended postgraduate modules. For entry qualifications, 68% of respondents had an A Level or equivalent (taken at the end of students' first year of the sixth form).

About 3% had GCSE/O-level qualifications (taken by 15/16 year olds at secondary school), and 9% had sub-degree qualification (e.g. HNC/ HND). Approximately 12% had a degree qualification: Bachelor of Arts or Science. Postgraduate entry qualifications were about 1% of the sample, while 'other' qualifications contributed about 4%. In terms of qualification aim, approximately 86% of the students aimed either for BSc or BA, while Diplomas and MScs comprised 3% and 6%. Three quarters of the modules were assessed by coursework (2500 - 3500 word essay), while the rest were assessed by various combinations of coursework and exams.

Effect of statistical method on satisfaction - before controlling

Table 3 shows the findings of the four methods: the effects of twelve demographic/ educational variables (before controlling) on student satisfaction with three individual clusters. The three-letter configuration of each cell of the table indicates whether a significant effect of each of the variables (rows) is detected, by the particular statistical method, on each of the three clusters (columns) of Learning and Teaching, Information and Organisation, and Utility. When the findings of any of the methods disagreed with those of the 'preferred method', this was highlighted in bold and boxed in the three-letter configuration in the intersect between the particular variable and the cluster. The table suggested that when controlling was not undertaken for the confounding effects of the variables and clusters, in about 83% of the time, the four methods agreed in their findings (e.g. the influences of gender, disability, ethnicity, nature of module, qualification aim, class size and assessment strategy on satisfaction). However, in other instances (five variables - age bracket, academic term, academic level, study mode, entry qualification), some of the findings of the three methods disagreed with the 'preferred method' (method 4) across one/ more clusters. For instance, the findings of each of methods 1 and 2 disagreed with the 'preferred method' in five instances across one or more clusters (Teaching and Learning, Information and Organisation, and Utility) as regards five variables (age bracket, academic term and level, study

Table 4: Effects of demographic and educational variables (after controlling) on student satisfaction with individual clusters of learning and teaching by analysis method

	Effect of variables on clusters according to Preferred Method (Method 4)	Whether significant* effect of variables on clusters of learning and teaching is detected by			
Variable		Method 1	Method 2	Method 3	Method 4
	Cluster[#]	L I U	L I U	L I U	L I U
Demographic					
1. Gender	No effect	—	N N N	N N N	N N N
2. Disability	No effect	—	N N N	N N N	N N N
3. Ethnicity	No effect	—	N N N	N N N	N N N
4. Age bracket	No effect[1]	—	N N N	N **Y** **Y**	N N N
Educational					
5. Academic Term	Term one students more satisfied	—	Y **N** Y	Y Y Y	Y Y Y
6. Academic Level	Level two students more satisfied	—	Y Y Y	Y Y Y	Y Y Y
7. Study Mode	No effect[1]	—	N N N	N N N	N N N
8. Entry Qualification	No effect[1]	—	N N N	N **Y** **Y**	N N N
9. Nature of module	Post-qualifying more satisfied	—	Y **N** Y	Y Y Y	Y Y Y
10. Qualification aim	No effect[1]	—	N N N	N N N	N N N
11. Class size	Smaller class size more satisfied	—	Y **N** Y	Y Y Y	Y Y Y
12. Assessment Strategy	More satisfaction with 50/50 coursework and exam assessment	—	Y Y Y	Y Y Y	Y Y Y

*: significant at $P<0.05$; [#]:The three letter configuration indicates whether significant effect of variables is detected on the three clusters of Learning and Teaching (L), Information and Organisation (I), and Utility (U); — : not applicable as method does not allow for controlling; N: no; Y: yes; Boxed bolded letters e.g. **Y** or **N** indicate the particular variable and cluster that are showing differences in their findings by method when compared with the same findings of the Preferred Method; [1] : variables that exhibited effects before controlling but (Table 3) lost those effects after controlling

mode and entry qualifications). The findings of method 3 showed the least disagreement (1 instance). There seemed no apparent tendency for certain methods to systematically influence the effects of particular variables on particular learning and teaching clusters.

Effect of statistical method on satisfaction - after controlling

Table 4 depicts the findings of the four methods: the effects of twelve demographic/ educational variables on student satisfaction with individual clusters after controlling for correlations between the variables (Methods 2, 3 and 4) and also after controlling for the associations between the clusters (only method 4). Method 1 did not allow for controlling (hence was not included in this analysis). The table suggested three pertinent issues.

- *Initial effects that were subsequently lost*: according to the findings of 'preferred method', the four variables of age bracket, study mode, entry qualification, and qualification aim, initially showed significant effects on satisfaction before controlling (Table 3), but lost those effects after controlling (Table 4). Thus controlling is mandatory and sometimes radically alters the findings. Uncontrolled analyses provided an erroneous impression that these variables critically influenced satisfaction, when truly in this dataset they did not. Their initial observed effects were probably due to correlation with other variables, and in the presence of other important variables, their net effect was ignorable. Such knowledge helps to focus researchers' attention and resources on those variables that truly influenced satisfaction.
- *'False positive' satisfaction findings (Type 1 error)*: a serious finding is related to method 3, where in two instances this method reported findings that grossly disagreed with the 'preferred method'. Table (4) suggested that although the two variables of age bracket and entry qualification (after controlling) exhibited no effects ('preferred method'), method 3 still reported that these variables significantly influenced satisfaction as regards the Information and Organisation, and Utility clusters (denoted by **Y**). Hence caution is required when employing models similar to method 3 to analyse satisfaction data, as these could generate 'false positive' (an overestimation of) satisfaction findings, particularly in relation to the Information and Organisation, and the Utility aspects of the educational experience.
- *'False negative' satisfaction findings (Type 2 error)*: another serious finding is related to method 2, where in three instances this method reported findings that partially disagreed with the 'preferred method'. Table (4) suggested that although the three variables of academic term, nature of module and class size (after controlling) exhibited significant effects across the three clusters ('preferred method'), method 2 still reported that these variables did not influence satisfaction as regards the second cluster of Information and Organisation. Hence caution is required when using method 2 to analyse satisfaction data, as it could generate 'false negative' (an

Table 5: Effects of demographic variables, educational variables and satisfaction with each cluster of teaching and learning on students' attained grade

Variables	Effect of variables or clusters on Grade	Controlling Undertaken?*	
		No	Yes
Demographic			
1. Gender	No effect	**Y**	N
2. Disability	No effect	N	N
3. Ethnicity	Students of 'White' ethnicity have better grades	Y	Y
4. Age bracket	Older students have better grades	Y	Y
Educational			
5. Academic Term	No effect	**Y**	N
6. Academic Level	Students on modules of higher academic levels have better grades	Y	Y
7. Study Mode	No effect	N	N
8. Entry Qualification	Students with degree entry qualification perform better	Y	Y
9. Nature of module	No effect	**Y**	N
10. Qualification aim	No effect	N	N
11. Class size	Smaller classes (student numbers) have better grades	**N**	Y
12. Assessment Strategy	Students on modules with assessments comprising more coursework than exams have better grades	Y	Y
Clusters: satisfaction with each cluster of teaching and learning			
Satisfaction with L/T cluster	No effect	**Y**	N
Satisfaction with I/O cluster	Students satisfied with I/O cluster have better grades	Y	Y
Satisfaction with U cluster	No effect	**Y**	N

* Controlling for the effects of other variables and clusters; L/T: Learning and Teaching cluster; I/O: Information and Organisation cluster; U: Utility cluster; Boxed bolded letters e.g. **Y** or **N** indicate instances where controlling for the effects of other variables and clusters influenced the findings

underestimation of) satisfaction findings, particularly in relation to the Information and Organisation aspects of the educational experience.

Effects of variables and clusters on students' attained grade

In order to analyse the effects of the independent confounders (i.e. demographic/ educational variables; and level of satisfaction with each learning and teaching cluster) on grade, Table 1 suggested that method 3 was appropriate. As grade is a univariate quantitative measure, hence Methods 1, 2, and 4 were not appropriate (suited for categorical data, binary data and multivariate data respectively). Analysis was undertaken twice (Table 5): without controlling for the effects of the variables and clusters (a series of t-tests/ ANOVAs); and with controlling for the effects of the variables and clusters (linear regression model which considers their correlations).

Before controlling, gender exhibited a significant effect on grade. However this influence was not a genuine effect of gender, but rather it was the effect of other variables that were correlated with gender on grade [e.g. ethnicity (0.13), entry qualification (0.1) and academic level (0.1), P = 0.001]. The initial observed effect of gender disappeared when the effects of those variables were controlled for. Similarly, academic term was correlated with assessment strategy (0.26) and with academic level (0.13); and nature of the module was correlated with age bracket (0.42) and academic level (0.39). Hence the initial observed effects of academic term and nature of the module on grade disappeared when these correlations were controlled for. In the same manner, the three satisfaction clusters were correlated with each other (0.61 to 0.39, P = 0.001). Hence, after controlling, only one of these three clusters (Information and Organisation cluster) exhibited genuine effects. Thus for some variables or clusters, their initial influences were not truly their own, rather these effects were 'mediated' through other variables/ clusters.

Conversely, class size showed no significant effect on grade before controlling but it had a significant influence after controlling. However, class size was correlated with academic level and assessment strategy (0.43 and 0.32 respectively, P = 0.001). Thus without controlling for the effect of these two variables, the initial observed effect of class size on grade was 'diluted' or 'concealed' by the effect of the other variables. Only after controlling for academic level and assessment strategy did the true (genuine) effect of class size on grade appear.

Table 6: Variables and clusters that explain student satisfaction

Variable[1]	% of satisfaction explained by variable	
	Before controlling	After controlling*
Academic Term	2	2
Academic Level	2	11
Nature of module	7	7
Class size	6	5
Assessment Strategy	6	6
Ability	3	4
Total explanatory power of the variables combined	Not appropriate to sum up the effects of individual variables	35

[1] all significant at $P<0.05$; * : based on the Preferred Method (Method 4 - Wilk's λ)

Table 7: Variables and clusters that explain students' achieved grades

Variable/ cluster[1]	% of grade explained by variable/ cluster	
	Before controlling	After controlling*
Ethnicity	2	Not appropriate to calculate the effects of individual variables/ clusters
Age	2	
Academic Level	3	
Entry qualification	2	
Class size	0	
Assessment Strategy	4	
Satisfaction with the I cluster[a]	1	
Total explanatory power of the variables/ clusters combined	Not appropriate to sum up the effects of individual variables/ clusters	13

[1] all significant at $P<0.05$; * : based on the univariate analysis (Method 3, employing R^2); [a] : Information and Organisation cluster of the learning and teaching experience

What explains satisfaction?

Table 6 depicts the explanatory powers of the final five variables that were significant in explaining satisfaction after controlling for all other variables. Each variable's explanatory power changed due to controlling. At times controlling decreased the explanatory power (class size), in other instances the explanatory power was increased (academic level). Academic level exhibited the greatest explanatory power (11%) followed by nature of module (7%). Assessment strategy, class size and academic term each explained 6%, 5% and 2% of the reported satisfaction. Collectively, these five variables explained 35% of the satisfaction levels. None of the clusters contributed to explaining satisfaction.

What explains achieved grade?

Table 7 shows the explanatory powers of the final variables and cluster that displayed significant effects in explaining the students' resultant knowledge as measured by their achieved grades, after controlling for all other variables and clusters. Collectively, the six variables of ethnicity, age, academic level, entry qualification, class size, assessment strategy, as well as one cluster (satisfaction with the Information and Organisation cluster) explained 13% of students' grades that were achieved in their modules.

DISCUSSION

There is emphasis on student evaluations in monitoring the quality of learning and teaching in higher education (El Ansari et al., 2002). This paper considered some issues that have hampered the quality, comparisons and generalizations of satisfaction research. Such research needs to be adequate, in an effort to generate valid evidence for researchers, practitioners and policy makers.

Within higher education, studies have reported the effects of variations in definitions or response formats (Glossop, 2001; Eley and Stecher,1997). In an era where methodological variations need to be minimised so

that findings are subsequently comparable, this state of matters is not unique to satisfaction studies. Concerns about common understandings and clarity of the concepts, terms, response formats and analytical variations have been voiced in partnership (El Ansari and Weiss, 2005), poverty (Mowafi and Khawaja, 2005), social network analysis (Hawe et al., 2004) public health (Rychetnik et al., 2004) and educational (El Ansari and Moseley, 2011) research.

Further, statistical methods are a challenging topic to teach and learn and there is evidence that it is often faultily applied in medicine (Altman, 1982, 1991 and 2002; Altman and Bland, 1991; Altman et al., 1983; O'Fallon et al., 1978) and other scientific disciplines. Such errors range from aspects of design and analysis, to reporting and interpretation (Garcia-Berthou and Alcaraz, 2004), to misuse (Dickinson, 2002).

In relation to the first aim of this study, we illustrated the variations in the statistical methods employed in the analyses of student satisfaction data citing some examples. To this end the paper has provided international examples of published papers of student satisfaction that illustrate such variations.

As regards to the second aim, we selected and contrasted four statistical methods that were commonly reported in the analysis of satisfaction data, in order to generate a 'preferred method' for subsequent comparisons. In relation to this aim the paper examined a variety of studies and statistical techniques employed in satisfaction studies, compared the features and aspects of four methods that fulfilled the selection criteria and voted for a 'preferred method'.

In connection with the third and fourth aims, we employed each of the four methods individually to analyse a dataset of student satisfaction and to compare the findings of each of the four techniques with those of the 'preferred method'. To this end, tables 3 and 4 depicted the analysis of the dataset using each of statistical methods in turn, once before controlling and once after. This is consistent with others (Pearson, 2004). For example, Garcia-Berthou and Alcaraz (2004) checked, employing three different statistical packages to undertake the same tests, all the statistical results reported in all the papers of particular volumes of *Nature* published in 2001, and also 12 randomly selected papers from the *BMJ* (2001). They found that 11.6% of the computations in *Nature* and 11.1% of the computations in the *BMJ* were incongruent. However, this paper is concerned with a parallel issue: the effects of employing different statistical methods on understanding student satisfaction.

For the fifth aim, we assessed the extent and type of influence of statistical method on understanding student satisfaction. Before controlling, the four methods exhibited considerable agreements in their findings across seven of the twelve demographic and educational variables. There were no tendencies for certain methods to systematically influence particular clusters of the learning and teaching experience. However, after controlling, three important trends emerged: initial significant effects of four variables on satisfaction were subsequently lost; and methods 3 and 2 were prone to generate 'false positive' and 'false negative' satisfaction findings respectively, with possible over- or underestimation of satisfaction.

Finally in relation to the sixth aim, we discussed the satisfaction findings according to the preferred method. In this study, respondents who were more satisfied were post-qualifying students who undertook academic level 2 modules that ran in term one, attended classes that were of smaller sizes (in terms of student numbers) and whose final assessment was a 50/50 combination of coursework/ exam assessments.

These findings are in agreement with others. As regards post-qualifying students and academic level, Feldman (1978) and Marsh (1987) found that higher level courses received higher ratings. However, the associations between course level and ratings were diminished when other variables (class size) was controlled for (Feldman, 1978). This highlights the importance of controlling in satisfaction studies.

This investigation also found class size effects: students in medium-sized classes were the least satisfied. Smaller and larger classes were more satisfied. Wachtel (1998) similarly found that smaller classes received higher ratings. The findings confirm that the relationship between class size and student ratings in not linear, but a U-shaped curvilinear relationship, with small and large classes receiving higher ratings than medium-sized ones (Koushki, 1982; Feldman, 1984). Class size effect may be specific to teaching and learning dimensions (e.g. group interaction or instructional rapport) (Wachtel, 1998).

In connection with the association between the final assessment type of the modules and student satisfaction, there is sparse literature that addressed this relationship. However, other parallel variables have been described in terms of the course workload, but even such variables attracted controversy. Ryan et al. (1980) reported that the introduction of mandatory student ratings led faculty to reduce course workloads and make examinations easier. Marsh (1987) found a positive correlation between 'workload' and student ratings, where more difficult courses were rated more favourably. However, Wachtel (1998) reported that the course level and student age might be confounding factors. Hence it is essential to control for the variables and clusters in studies of student satisfaction.

Finally, in terms of the question: "What explains satisfaction?, this study found that collectively, five variables (academic level, nature of module, Assessment strategy, class size and academic term) explained 35% of

the satisfaction levels. This suggested that the understanding of academic, educational and other variables that could contribute to explaining student satisfaction still needs to be uncovered. Similarly, in terms of the question: "What explains students' achieved grades?", collectively, six variables (ethnicity, age, academic level, entry qualification, class size, assessment strategy, as well as one cluster - satisfaction with the Information and Organisation) explained 13% of grades that students actually achieved in their modules. Again, this suggested that the understanding of academic, educational and other variables that could contribute to explaining student grades is yet to be revealed.

A limitation of this inquiry is that it examined one student cohort in one institution in one academic year. Although the response rates were high, the sample was self-selecting. Due to confidentiality, data could not be retrieved to judge how respondents compared with the total student body.

CONCLUSION: RESEARCH AND PRACTICE IMPLICATIONS

Module teams and educational administrators need to achieve quality standards that enhance satisfying educational experiences for their learners. But the comparability, interpretations and generalisation of findings across studies, institutions and countries is hampered. Hence the development of standardised, robust methods for the analysis of satisfaction surveys is critical for a valid evidence base for policy and practice. Controlling is necessary in student satisfaction studies where many variables collectively explain the satisfaction with the experience. Researchers need to consider the effects that different statistical methods of analysis have on the findings and conclusions of satisfaction studies, resulting in an over- or under-estimation of satisfaction. The need for standardized measurement and analysis of satisfaction data is not limited to education but also for satisfaction research generally e.g. patient satisfaction, where demographic (patient), condition- and setting-(unit/ward/health worker) related variables contribute to satisfaction with clusters of the patient's care experience. Until robustness is achieved, researchers need to realise the effects of methods on results. This inquiry highlighted research and practice considerations that require attention if learner satisfaction is to contribute to educational developments for a competent health and social care workforce.

ACKNOWLEDGEMENTS

The authors thank the staff and the students of the School of Health and Social Care at the University where the study was implemented. Special thanks go to Ms. Julie Poulton for the data retrieval and linkage, and to Prof. Terence Davis and Dr Diane Pearson for constructive comments on earlier drafts of the paper.

REFERENCES

Altman DG, Bland JM (1991). Improving doctors' understanding of statistics. J R Stat Soc. 154: 223-267.
Altman DG, Gore SM, Gardner MJ, Pocock SJ (1983). Statistical guidelines for contributors to medical journals. BMJ. 286: 1489-1493.
Altman DG (2002). Poor-quality medical research: what can journals do? JAMA. 287: 2765-2767.
Altman DG (1982). Statistics in medical journals. Stat Med. 1: 59-71.
Altman DG (1991). Statistics in medical journals: developments in the 1980s. Stat Med. 10, 1897-1913.
Amin ME (1994). Correlates of course evaluation at the Faculty of Letters and Social Sciences of the University of Yaounde. Assessment and Evaluation in Higher Education. 18(2): 135-141.
Bird D, Roberts PM (1998). The role of library and information services in the modular curriculum. Nurse Educ. Today. 18: 583-591.
Brown H, Edelmann R (2000). Project 2000: a study of expected and experienced stressors and support reported by students and qualified nurses. J. Adv. Nurs. 31(4): 857-864.
Cashin WE (1988). *Student ratings of teaching: a summary of the research*. Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.
Chapple M, Murphy R (1996). The nominal group technique: extending the evaluation of students' teaching and learning experiences. Assess. Eval. High. Educ. 21: 147-158
Chevannes M (2001). An evaluation of the recruitment of black and minority ethnic students to pre-registration nursing. Nurs. Times Res. 6(2): 626-635.
Clustan. *Cluster analysis*. Retrieved November 12, 2010 from http://www.clustan.com/what_is_cluster_analysis.html
Cowman S (1996). Student evaluation: a performance indicator of quality in nurse education. J. Adv. Nurs. 24: 625-632.
Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. Psychometrica. 16: 297-334.
Dickinson HO (2002). Use and abuse of statistics in nursing research. J. Adv. Nurs. 39(5): 405-407.
Eaton N, Williams R, Green B (2000) Degree and diploma student nurse satisfaction levels. Nurs Stand. 14(4): 34-39.
El Ansari W, Oskrochi R (2006). What Matters Most? Predictors of Student Satisfaction in Public Health Educational Courses. Public Health. 120(5): 462-73.
El Ansari W, Oskrochi R (2004). What 'really' affects health professions students' satisfaction with their educational experience? Implications for practice and research. Nurse Educ. Today. 24(8): 644-55.
El Ansari W, Weiss ES (2005). Quality of research on community partnerships: developing the evidence base. Health Educ. Res. 21(2): 175-80.
El Ansari W (2004). Appraisal skills as a public health competency for evidence-based care: Student satisfaction with their research education. J. Public Health Manage. Pract. 10(4): 354-365.
El Ansari W (2003). Satisfaction Trends in Undergraduate Physiotherapy Education. Physiother. 89(3): 171-185.
El Ansari W (2002a). Student Nurse Satisfaction Levels With Their Courses: Part II – Effects of Academic Variables. Nurse Educ. Today. 22(2): 171-180.
El Ansari W (2002b). Student Nurse Satisfaction Levels With Their Courses: Part I – Effects of Demographic Variables. Nurse Educ. Today. 22(2): 159-170.
El Ansari W, Moseley L (2011). You get what you measure: Assessing and reporting student satisfaction with their health and social care educational experience. Nurse Educ Today. 31(2):173-8.

El Ansari W, Pearson D, Davis T (2002). Satisfaction with interprofessional multidisciplinary education? Influences of learners' demographic and academic characteristics. J. Interprof Care. 16(2): 174-175.
Eley MG, Stecher EJ (1997). A comparison of two response scale formats used in teaching evaluation questionnaires. Assess. Eval. High. Educ. 22(1): 65-79.
Espeland V, Indrehus O (2003) Evaluation of students' satisfaction with nursing education in Norway. J. Adv. Nurs. 42(3): 226-36.
Feldman KA (1978). Course characteristics and college students' ratings of their teachers: what we know and what we don't. Res. High. Educ. 9: 199-242
Feldman KA (1984). Class size and college students' evaluations of teachers and courses: a closer look. Res. High. Educ. 21: 45-116.
Garcia-Berthou E, Alcaraz C (2004). Incongruence between test statistics and P values in medical papers. BMC Med Res Methodol. May 28;4:13. doi:10.1186/1471-2288-4-13
Glossop C (2001). Student nurse attrition from pre-registration courses: investigating methodological issues. Nurse Educ. Today. 21(3): 170-80.
Hawe P, Webster C, Shiell A (2004). A glossary of terms for navigating the field of social network analysis. J. Epidemiol. Community Health. 58(12), 971-5.
Hayden M, Thompson J (1996). An investigation of the effectiveness of an evaluation instrument designed for use with modular masters courses in a range of international contexts. Assess. Eval. High. Educ. 21(1): 55-68.
Hoskins SL, Newstead SE, Dennis I (1997). Degree performance as a function of age, gender, prior qualifications and discipline studied. Assess. Eval. High. Educ. 22: 317-328.
Kapborg I (2000). The nursing education programme in Lithuania: voices of student nurses. Adv Nurs. 32(4): 857-63.
Kerridge JB, Mathews BP (1998) Student rating of courses in HE: further challenges and opportunities. Assess. Eval. High. Educ. 23: 71-82.
Kinsella FE, Willaims WR, Green BF (1999). Student nurse satisfaction: implications for the common foundation programme. Nurse Educ. Today. 19: 323-333.
Kleinbaum D, Klein M (2002). *Logistic Regression. A Self-Learning Text*, New York: Springer-Verlag. (2nd Edition).
Kniveton BH (1996). Student perceptions of assessment methods. Assess. Eval. High. Educ. 21(3): 229-237.
Koushki PA, Kuhn HAJ (1982). How reliable are student evaluations of teachers? Eng. Educ. 72: 362-367.
Kwan K-P (1999). How fair are students ratings in assessing the teaching performance of university teachers? Assess. Eval. High. Educ. 24(2): 181-195.
Latter S, Rycroft-Malone J, Yerrell P, Shaw D (2000). Evaluating educational preparation for health education role in practice: the case of medication education. J. Adv. Nurs. 32(5): 1282-1290.
Lee T, Mawdsley JM, Rangeley H (1999). Students' part-time work: towards an understanding of the implications for nurse education. Nurse Educ. Today. 19: 443-451.
Lowe I (1995). The changing nature of universities. In: L. Hewson S. Toohey (Eds) *The changing university: Papers presented at the Biennial UNSW Education Conference 1995, pp. 9-16*. Sydney: Professional Development Center. The University of New South Wales.
Marsh HW (1987). Students' evaluations of university teaching: research findings, methodological issues etc. Int. J. Educ. Res. 11: 255-378.
Mowafi M, Khawaja M (2005). Poverty. J Epidemiol Community Health. 59(4): 260-264.
Murphy RJL, Harris B (1995). Improving the quality of teaching in universities. Psychol. Teach. Rev. 4: 28-30.
Nahas VL, Nour V, Al-Nobani M (1999). Jordanian undergraduate nursing students' perceptions of effective clinical teachers. Nurse Educ. Today. 19: 639-648.
O'Fallon JR, Dubey SD, Salsburg DS, Edmonson JH, Soffer A, Colton T (1978). Should there be statistical guidelines for medical research papers? Biometrics. 34: 687-695.
Ofori R (2000). Age and 'type' of domain specific entry qualifications as predictors of student nurses' performance in biological, social and behavioural sciences in nursing assessments. Nurse Educ. Today. 20: 298–310.
Pawitan Y (2001). *In all likelihood: statistical modelling and inference using likelihood*. Oxford: Oxford University Press.
Pearson H (2004). Double check casts doubt on statistics in published papers. Nature. 429(6991): 490.
Rhodes RS, Davis DC, Odom BC (1999). Challenges and rewards of educating a profoundly deaf student. Nurse Educ. 24(3): 48-51.
Ryan JJ, Anderson JA, Birchler AB (1980). Student evaluation: the faculty responds. Res. High. Educ. 12: 317-333.
Rychetnik L, Hawe P, Waters E, Barratt A, Frommer M (2004). A glossary for evidence based public health. J. Epidemiol. Community Health. 58(7): 538-45.
Saksomboon K, McMillan M, Cholowski K (2002). An evaluation of the baccalaureate-nursing program within the (MOPH), Thailand. Nurse Educ. Today. 22(8): 641-7.
Snell J (1995). It's tough at the bottom. Nurs Times. 91: 58-59
Statsoft Electronic Textbook. Cluster Analysis. General purpose. Retreived January 11, 2007 from http://www.statsoft.com/textbook/stathome.html
Tabachnick BG, Fidell LS (1996). *Using multivariate statistics.* 3rd edition. New York: Harper Collins.
Taylor A, Hill F (1993). Quality Management in Education. Quality Assurance in Education. 1: 21-28.
Wachtel HK (1998). Student evaluation of college teaching effectiveness: a brief review. Assess. Eval. High. Educ. 23(2): 191-211.
White J, Williams R, Green BF (1999). Discontinuation, leaving reasons and course evaluation comments of students on the common foundation programme. Nurse Educ. Today. 19: 124-150.