



ISSN: 2349-9141

Available online at <http://www.ijrr.com>

International Journal of Information Research and Review
Vol. 2, Issue, 08, pp. 979-984, August, 2015



OPEN ACCESS JOURNAL

Full Length Research Paper

THE BIASES THAT AFFECT THE EVALUATION OF UNIVERSITY TEACHING BY STUDENTS FROM THE FACULTIES OF SCIENCE IN MOROCCO

^{1,*}Jaouad Alem, ²Ines Bouguerra, ³Amina En-Nya and ⁴Khalid Taghzouti

¹School of Physical Activity, Laurentian University, 935 Ramsey Lake Road, Sudbury (Ontario) Canada

²Higher Institute of Sport and Physical Education of Sfax, Laurentian University, Tunisia

³Faculty of Rabat, Mohamed V University

⁴Faculty of Rabat, University Mohammed V

***Corresponding Author**

Received 23rd July 2015; Published 31st August 2015

Abstract

This article aimed to identify biases that could affect items that assessed university education by students (SET); the 8 items that assess SET are those who have been used by Laurentian University. The questionnaire was administered to Moroccan university students at the Faculty of Sciences Rabat in 2013. The reliability is checked by calculating Cronbach's alpha. T-tests and ANOVA were used to verify the influence of potential biases. The regression analysis on the average for SET was used to estimate the relative weight of some moderating variables. Results have not revealed the influence of the students or teachers genre. The study showed, however, that certain variables were likely to influence the results of SET.

Keywords: SET, validity, bias, Morocco.

Copyright © Jaouad Alem et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To cite this paper: Jaouad Alem, Ines Bouguerra, Amina En-Nya and Khalid Taghzouti, 2015. The biases that affect the evaluation of university teaching by students from the faculties of science in Morocco, *International Journal of Information Research and Review*. Vol. 2, Issue, 08, pp. 979-984. August, 2015.

INTRODUCTION

Empirically, Marsh (1983, 1984; Marsh and Bailey, 1993) argues that SET are multidimensional, reliable, stable and relatively valid in front of a variety of quality educational indicators. Harvey and Hebert (2013) investigated the psychometric properties of the EEA in terms of response rates, satisfaction rates, reliability and dimensionality factors according to two conditions for the award (online versus paper) in students from the University of Québec at Rimouski. A first sample of 20,245 students completed the paper version and a second sample of 16,432 students completed the online version reported the results of a reliability coefficient greater than a threshold of 0.80 (almost the same, with a deviation of 0.01). Chemsî, Radid, Sidiqet and Talbi (2011) have designed and validated a computerized tool for evaluating the teaching and training in distance by students of the faculty of Ben M'sik Morocco. 50 students completed survey assessing by 24 items and 105 students responded to the training evaluation 13 items. Cronbach's alpha estimated of 0.92 while the items that assess training is of the order of 0.85. Marsh (1983, 1984) also supports the SET are not affected by hypothetical bias. However, there is no consensus among authors in dimensions of the construct to be measured by the SET.

Marsh and Dunkin (1992) also indicate that most instruments that measure teaching skills are based on a mixture of logical and pragmatic considerations rather than on a specific theoretical construct. According to Poissant (1996), there would be several thousand questionnaires with many and varied dimensions, including the construction process and the quality of the resulting of evaluation will vary.

According to certain hypothesis, the enthusiasm of teachers, and their use of positive evaluations of the work and examinations ("evaluation leniency") would be factors that skew the SET and invalidate the interpretation in terms of teaching's skills. However, authors such Marsh and Dunkin (1992) consider the SET as one of the best ways to evaluate the education offered in the context of a course by an academic resource and it is affected by few potential biases. In this research, we focus on factors possibly related to the SET such as the gender of students and the teachers, the age of the students and teachers, the academic level of students, the academic performance of students and the type of assessed courses (science and engineering versus humanities). The study tend so to identify the biases that can affect the validity of the SET, as measured by the instrument used in the Laurentian University (eight items).

Literature review

The variable « gender »

In general, the relationship between "gender" and SET is such as negligible. For example, the study of Costin and others (1971) shows no significant difference from the students who evaluate or from teachers who are evaluated. They both receive the same evaluation despite their gender. Furthermore, Basow and Distenfeld, (1985); Bennett, (1982); Elmore and Lapointe, (1974) indicate that women teachers are evaluated less severely than their male counterparts. The study of Aleamoni and others (1980) reported that girls evaluate positively certain aspects of education. Also, the performance of the female teachers in class is often assessed more critically than men teachers (Sandler, 1991). Female teachers receive lower ratings than their male counterparts especially male students, however, the evaluations of female students are not affected by the gender of the teacher (Basow and Howe, 1987; Feldman, 1993; Hancock, Shanon and Trentath 1993). According to Kaschak, (1978); Lombardo and Tocci (1979), there is no bias related to gender.

The variable "field of study"

The study of Centra (1979) on thousands of classes in various fields of study reveals that humanities professors get slightly better results than the social sciences and natural sciences. Feldman (1978) arrived at the same results on this subject. He noted that although research findings are not satisfactory, we should not neglect the variable field of study when interpreting results for administrative purposes. Cashin and Clegg (1987) report that the ratings are higher in the SET in applied arts and nursing. In addition, Cashin and Clegg (1987) add that the humanities and the arts get better evaluations compared to the social sciences, which are better evaluated than math and science courses.

The variable "age"

Richardson (1994) wanted to test the course evaluation questionnaire (CEQ) of Ramsden (1991) at universities in England. The CEQ is composed of 30 items that measured the following six dimensions: quality of teaching (8 items); clear objectives and standards (5 items); the appropriate workload (5 items); the appropriate evaluation (6 items), the emphasis on the independence of students and choice (6 items). 256 questionnaires has been sent to students from various disciplines in the social sciences, 95 completed and returned the completed questionnaires. Internal consistency as measured by Cronbach is 0.79 for quality evaluation; 0.77 for clear objectives; 0.71 for the appropriate workload of 0.47 for appropriate assessments, 0.55 for emphasis on the independence of students, 0.42 for the subscale memory. These coefficients are generally lower than the corresponding figures obtained by Ramsden (1991), but they tend to follow the same pattern.

Also, variance analyzes were performed on scores on the five factors of the first order and the second order factor, using the independent variables whose year of study, the curriculum, gender, as well as the age covariate.

These analyzes showed that only one variable is significant. Indeed, age was negatively associated with scores associated with the appropriate workload ($\beta = -0.28$; $F = 5.11$; $df = 1, 62$, $p < 0.05$). It is not surprising that adults with household responsibilities students said their workload as (even) less appropriate than did the younger. According to Marsh (1984, 1983, 1982, 1980), the variable "workload" has an influence on other variables (students, courses), which has repercussions on the evaluation of teaching. For the author, this variable is not a source of bias since it is inherent to learning and it is desirable to teaching. Some studies like Costin, William, Greenouch and Menges (1971) and those of Doyle (1983) show that more advanced students in their studies' end of the 1st, 2nd and 3rd cycles tend to a more positive evaluation of the teaching the less advanced students. According to the review of Feldman (1978), some studies show a positive correlation between the academic year and the evaluation of teaching while other studies have found no correlation.

Griffin, Coates, Mcinnis and Richard (2003) extended course evaluation questionnaire (CEQ) of Ramsden (1991). In addition to the dimensions of the original CEQ, the new instrument has the following dimensions: "student support", "learning resources", "the organization of the course", "learning community", "the qualities of the graduate (s)" and "intellectual motivation". 3691 students from 15 universities in Australia completed the CEQ including three versions of the questionnaire were used. The purpose of this study was not necessarily to determine overall satisfaction, each of the six extra dimensions were analyzed separately.

The Rasch reliability index was used to indicate to what extent each of the dimensions are separated from their respective variables. The index has a range from 0.0 to 1.0 with a value of 1.0 indicating that each dimension only contributes to the clarification of the variable. The measuring the reliability of separation can be understood as a measure of the validity of construct; and the results indicated that all dimensions have measurement characteristics coherent compared with latent traits, because their values ranged from 0.85 to 0.93. It seems that the degree to which students feel motivated and intellectually stimulated by their university experience increases with age. However, the data associated with the "learning community" dimension show an opposite trend. Older students feel they are least part of a "university community" that involves working collaboratively with other students. The students' impressions of the generic skills obtained during their university experience increases with grade. The scores for the dimension "graduate qualities" also show a significant increase over the years, indicating a change in the attitudes of students towards their courses over the years. However, the scores of the dimension "learning community" decrease with the years of study.

Variable "teaching dimension"

The results of many studies (Hildebrand, Wilson and Dienst, 1971), reported by Doyle (1983); Kulik and McKeachie, (1975); Feldman (1976); Centra (1979) and Marsh (1984, 1987) show that the organization and clarity of the course, the enthusiasm and the dynamism of the teacher, his interaction with his students and the feedback that it gives to his students are the most frequently mentioned dimensions in research on SET.

A study of Feldman (1988) shows that the importance put by students and teachers, is granted to certain dimensions such as the concern of the teacher student progress, organization and clarity of course, control of matter, the dynamism of the teacher, availability, compliance of the student and his assessment of student learning (justice, fairness and quality of examinations).

Downie and Moller (2002) have used the course evaluation questionnaire (CEQ) of Ramsden (1991) to evaluate the student experience for an entire curriculum. The CEQ consisted of questions related to the quality of teaching and learning, student support, and finally administrative services. Minor changes were made to sections of the student support and administrative services while Section of the quality of teaching and learning has been kept intact. The CEQ is composed of 56 items that measured the following six dimensions: the academic environment, education, skills development, appropriate evaluation, appropriate workload standards and clear objectives. The last item was evaluating the overall quality of post-secondary studies.

The pilot study was conducted among students in their final year in four disciplines including tourism, leisure, hotel trades and sports. 1100 questionnaires sent to students attending 25 programs, 634 were completed and returned. On average, students are more satisfied with the dimension "skills development" (the average is 3.8 and 3.9) followed by the academic environment (average of 3.3 and 3.4), teaching and adequate evaluation (average 3.2 and 3.3). Students are less satisfied with the dimensions "clear standards and goals" (average of 3.1) and "appropriate workload" (average of 2.9 and 3.0). According to the students, the best aspects of their studies course then monitoring methods and teaching and staff skills. However, students feel that some aspects should be improved especially communication including timely feedback on their work, more information on what is expected of them and more time with tutors.

Ajouter une phrase telle que: cette recension nous permet de libeller et de tester les quelques hypothèses de recherche suivantes:

Seven moderating variables were selected and seven hypotheses follow:

Hypothesis 1: The age of students has an effect on their SET.

Hypothesis 2: The Student gender has no effect on their SET.

Hypothesis 3: The gender of teachers has no effect on the SET of students

Hypothesis 4: The experience of the teacher has no effect on the SET of students.

Hypothesis 5: The sample of students has no effect on the SET of students.

Hypothesis 6: The year of study has an impact on the SET of Moroccan students; SET of 1st year students are more severe.

Hypothesis 7: Academic average student is positively correlated with their SET.

MATERIALS AND METHODS

The questionnaire, sample and statistical techniques are specified in the following lines

Questionnaire

The students evaluation questionnaire of teaching's quality is the one that is currently used at Laurentian University. This instrument consists of eight items, it had been previously analyzed its reliability, its degree of discrimination and its construct validity for a population of student physical education (Alem, Radid, Ahami; Hariti and Hamrouni; 2013; Alem, Bouguerra, Hamrouni and Hariti, 2015). Indeed, Alem *et al.* (2013 and 2015) had experienced this form of university students Maghreb, it appeared that it was possible to eliminate redundant items and offer academic course programs in the Maghreb SET a simple grid, composed of only 4 items with the following metric qualities:

- Items are reliable in terms of internal consistency,
- The questionnaire is fairly easy to administer in terms of number of items: 4 items retained 'capture' more than two-thirds of the total variance.
- The items are valid only in terms of their ability to evaluate two distinct dimensions of quality of education: the climate and learning objectives.

Sample

The study was conducted on a group of 1123 students in Rabat Sciences faculty who agreed to answer an anonymous questionnaire consisting of eight moderators and eight items that measure the SET. The eight moderating variables are age, students gender, teachers gender, the type of courses evaluated (science versus humanities), the number of students, the year (1st, 2nd, 3rd or 4th year) and the academic average for the student.

The statistical techniques used

Before checking the presence of potential biases, reliability of 8 items was analyzed with Cronbach's alpha is a good measure of the internal consistency of a latent variable or a psychological construct, acceptable values normally being at or above .70 (Nunnally, 1978). However, values of .60 are also acceptable (George and Mallery, 2003; Hair, Black, Babin, Anderson and Tatham, 2006), especially if the construct is measured with few items; these authors propose the following interpretations alpha coefficients unacceptable <.60; poor 0.60 to 0.69; acceptable from 0.70 to 0.79; good 0.80 to 0.89; excellent > .89. Thereafter, t-tests and ANOVA were used to verify the influence of the students and teachers gender on SET. Finally, a correlation analysis between the other moderators and SET has estimated the respective weight of each of moderator variables.

RESULTS

The main results of the study are reported below. They come in two distinct sections: the analysis of the reliability and the influence of potential biases.

It appears that the alpha is acceptable because it is well above the acceptable threshold of 0.80. In addition, Table 2 shows it does not seem possible to optimize the alpha coefficient.

Table 1. Analysis of the statistical reliability of the questionnaire

Cronbach's alpha	Number of items
,85	8

The influence of potential biases

T-tests and ANOVA were used to verify the influence of students and professors gender:

Table 2. Statistics for each item

	Medium scale if the item is deleted	Variance scale if the item is deleted	The correlation between the item and the corrected total	Cronbach's alpha if item deleted
it1	20,49	57,63	,51	,84
it2	20,76	56,51	,62	,83
it3	20,89	56,43	,55	,84
it4	21,02	54,74	,65	,83
it5	20,84	55,15	,63	,83
it6	21,35	57,42	,51	,84
it7	20,94	55,23	,65	,83
it8	21,10	54,07	,61	,83

- SET according to gender of students: female students do not value their courses more hardest than their male counterparts: respectively M = 2.98, SD = 1.04, N = 822 versus M = 3.03, SD = 1 03, N = 285; t = -0.7, df = 1,105, p = 0.48.
- SET accoding to gender of teachers: courses taught by women (M = 2.96, SD = 1.04, N = 449) were not evaluated more hardest than the courses taught by men (M = 3.00, SD = 3.04, N = 627): t = -0.72, df = 1074, p = 0.47.

The correlation matrix between the other moderators and SET has estimated the respective weight of each of moderating variables:

The correlations are all significant but rather weak, except for the two variables of quality educational support 'and self academic average reported by students.

Table 3. Correlations between moderators and other SET

Moderator variables	Correlation with SET	N
age	0,11**	1103
The sample of student?	-0,11**	1043
Duration of the final exam	0,07*	1055
Time allowed for the examination	0,20**	1091
The quality of the didactic support	0,53**	1071
Students skills in French	0,29**	1066
The number of years experience of teachers	0,12**	938
academic average reported students	0,34***	1079

** The correlation is significant at the 0.01 level (2-tailed).

* The correlation is significant at the 0.05 level (2-tailed).

The higher the quality of instructional support is perceived, the more students evaluate positively their courses (0.53 **). When students perform in their courses, the SET is high (R = 0.34 ***).

DISCUSSION

The results show that the alpha is acceptable because it is well above the acceptable threshold of 0.80 (0,85).This result is supported by the findings of Marsh (1983, 1984) and Marsh and Bailey (1993).

It agrees, too, with the results of Harvey and Hébert (2013), where, respectivement the reliabilities of the two modalities, ie online and paper was above the threshold of 0.80. Chemsî *et al.* (2011) reported an even higher reliability of the order of 0.922. The results of T-tests and ANOVA to verify the influence of students and professors gender show thatfemale students do not value their courses more hardest than their male counterparts. This result is in accordance with the study of Costin and co, (1971).

In this study, we note that courses taught by women were not evaluated more hardest than the courses taught by men what is in divergence with studies of Basow and Distenfeld, (1985); Bennett, (1982); Elmore and Lapointe, (1974); Basow and Howe, (1987); Feldman, (1993); Hanock, Shanon and Trentath, (1993). The higher the quality of instructional support is perceived, the more students evaluatepositively their courses (0.53**).

Indeed, the literature reviewshows that students give importance to certain dimensions as: the organization and clarity of the course, the enthusiasm and dynamism of the teacher, his interaction with his students and the feedback it gives to his students (Hildebrand, Wilson and Dienst, 1971; reported by Doyle (1983); and McKeachie Kulik, 1975; Feldman, 1976; Centra, 1979 and Marsh, 1984, 1987), the concern of the teacher 's progress student, organization and clarity of course, mastery of the material, the dynamism of the teacher, his availability, his respect for the student and his assessment of student learning (justice, impartiality and quality of reviews) (Feldman, 1988), course content monitoring methods and teaching and staff skills (Downie and Moller, 2002). More students perform in the courses evaluated over their own evaluations are high (R = 0.34 ***). Indeed, the literature shows that the evaluation of teaching depends on the field of study (Centra, 1979; Feldman, 1978; Cashin and Clegg, 1987). It also depends on the student's motivation degrees, which in turn increases with age, so a tendency to evaluate teaching standards more positively (Griffin, Coates, Mcinnis and Richards, 2003).

Conclusion

This research aimed to identify certain biases that can affect the validity evaluations of teaching by Moroccan students. Analysis of the internal consistency of the items by calculating the Cronbach's alpha indicates that it is acceptable. Indeed, the coefficient exceeds the threshold of 0.8. T-tests and ANOVA did not demonstrate that the gender of students or teachers influenced the SET. Significant but weak correlations between the SET and the age of students, the number of teacher's years of experience, the sample of students, the time allocated for the review and the time to respond to reviews.

As against three variables seem to affect more SET: quality didactic support, the French language skills of students and their average academic.

This research is the result of the study Alem *et al* (2013) had already tested before the same questionnaire in his three North African Universities (Morocco, Algeria and Tunisia). Their study did suggested that the SET are actually multidimensional, reliable and discriminating, but his validity remains to be demonstrated. All these results can be used to qualify the findings of the recent study of Harvey and Hebert (2013) who argue that SET should not be made for summative purposes related to the teaching career but could be considered as a process valid and reliable formative framework to stimulate reflection.

REFERENCES

- Alem, J., Bouguerra, I., Hamrouni, S. and Hariti, H. 2015. The Student Evaluations of Teaching by Maghreb Students in Physical Education. *International Journal of Innovation and Research in Educational Sciences*. Volume 2, Issue 2, ISSN (Online): 2349–5219.
- Alem, J., Radid, M., Ahami, A., Hariti, H. and Hamrouni, S. 2013. *Les qualités métriques de l'évaluation de l'enseignement par les étudiants maghrébins*. 5ème Colloque en mesure et évaluation: la communication des résultats d'évaluation. Axe de la mesure. ACFAS 2013.
- Basow, S.A. and Distenfeld, M.S. 1985. Teacher Expressiveness: More Important for Males than Females? *Journal of Educational Psychology*, 77, 45-52.
- Basow, S.A. and Howe, K.G. 1987. *Evaluations of College Professors: Effects of Professors, Sex- Type and Sex, and Students' Sex*. Psychological Reports, 60, 671- 678.
- Bennett, S.K. 1982. Student Perceptions of and Expectations for Male and Female Instructors: Evidence Relating to the Question of gender Bias in Teaching Evaluations. *Journal of Educational Psychology*, 74.
- Cashin, W.E. and Clegg, V.L. 1987. *Are student ratings of different academic fields different?* Paper presented at the annual meeting of the American Educational Research Association. Washington, DC. - See more at: <http://teaching.berkeley.edu/blog/evaluating-evaluations-part-1#sthash.124JOqJL.dpuf>
- Chemsi, G., Radid, M., Sadiq, M., and Talbi, M. 2011. *Conception et validation d'un outil informatisé pour l'évaluation des enseignements et des formations à distance par les étudiants: EVAL-EFDE*. Radisma, Numero 7.
- Cortina, J.M. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98–104.
- D'Apollonia, S. and Abrami, P.C. 1997. Navigating student ratings of instruction. *American Psychologist*, 52(11), 1198-1208.
- Downie, M. and Moller, I. 2002. *The Ramsden Course Experience Questionnaire: A pilot study of final-year students on hospitality, leisure, sport and tourism degree courses*. Retrieved from: http://www.heacademy.ac.uk/assets/hlst/documents/johlste/0010_downie_vol1no1.pdf
- Elmore, P.B. and Lapointe, K. 1974. Effect of Teacher Sex and Student Sex in the Evaluation of College Instructors. *Journal of Educational Psychology*, 66, 386-389.
- Feldman, K.A. 1993. College Students' Views of Male and Female College Teachers: Part II- Evidence From Student Evaluations of Their Classroom Teachers. *Research in Higher Education*, 34(2), 151-211.
- George, D. and Mallery, P. 2003. SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon, p. 231
- Griffin, P., Coates, H., Mcinnis, C. and James, R. 2003. The development of an extended course experience questionnaire. *Quality in Higher Education*, 9(3), 259-266
- Hair, J., Black, B. Babin, B., Anderson, R. and Tatham, R. 2006. *Multivariate Data Analysis* (6th edition). Upper Saddle River, NJ: Prentice-Hall.
- Hair, J., Black, W., Babin, B., Anderson, R. and Tatham, R. 2006. *Multivariate Data Analysis* (6th ed.). New Jersey: Pearson Educational, Inc.
- Hancock, G.R., Shannon, D. and Trentham, L. 1993. Student and Teacher Gender in Ratings of University Faculty: Results from Five Colleges of Study. *Journal of Personal Evaluation in Education*, 6(3), 235-248.
- Harvey, L., and Hébert M. H. 2013. *Les qualités psychométriques de l'évaluation de l'enseignement par les étudiants et les étudiants: impacts des modalités pédagogiques et de passation*. Actes du 25ème colloque de l'ADMEE-Europe Fribourg 2013: Evaluation et autoévaluation, quels espaces de formation ?
- Kaschak, E. 1978. *Sex Bias in Student Evaluations of College Professors*. *Psychology of Women Quarterly*, 3, 235-242.
- Lombardo, J. and Tocci, M. 1979. Attribution of Positive and Negative Characteristics of Instructors as a Function of Attractiveness and Sex of Instructor and Sex of Object. *Perceptual and Motor Skills*, 48, 491-494.
- Marsh, H. W. 1983. Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. *Journal of Educational Psychology*, 75, 150-166.
- Marsh, H. W. 1984. Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-754.
- Marsh, H. W. and Bailey, M. 1993. Multidimensionality of students' evaluations of teaching effectiveness: A profile analysis. *Journal of Higher Education*, 64, 1 - 18.
- Marsh, H. W. and Dunkin, M. 1992. *Students' evaluations of university teaching: A multidimensional perspective*. In J.C. Smart (Ed.), *Higher education: Handbook on theory and research* (Vol. 8, pp. 143-234). New York: Agathon Press
- Nunnally, J. C. 1978. *Psychometric theory*. New York: McGraw-Hill Inc.
- Poissant, H. 1996. *L'évaluation de l'enseignement universitaire*. Montréal: Éditions Logiques.
- Ramsden, P. 1991. *A performance indicator of teaching Quality in Higher Education: the Course Experience Questionnaire*. *Studies in Higher Education* 16(2), 129-150.

Richardson, J.T.E. 1994. A British evaluation of the Course Experience Questionnaire. *Studies in higher education*, 19 (1), 59-68.

Sandler, B.R. 1991. *Women Faculty at Work in the Classroom, or, Why it Still Hurts to be a Women in Labor*. *Communication Education*, 40, 6-15.
