The Student Evaluations of Teaching by Maghreb Students in Physical Education

Jaouad, Alem
School of Human Kinetic, Laurentienne University, 936 Chemin du Lac Ramsey, P3E 2C6 Sudbury, Ontario, Canada, email: jalem@ laurentienne.ca

Saber, Hamrouni
Institut Supérieur du Sport et de l’Éducation Physique de Tunis, Campus Universitaire de La Manouba 2010, Tunisie, email: sabeurhamrouni@yahoo.fr

Ines, Bouguerra
Institut Supérieur du Sport et de l’Éducation Physique de Sfax. Route de l'Aéroport, Km 3.5, BP 384, Sfax 3000, Tunisie, email: inessbouguerra@yahoo.fr

Hakim, Hariti
Institut d’éducation physique et sportive de l’Université d’Alger, Ville nouvelle de Mehelma, Zeralda Alger, Algérie, email: hhari@ yahoo.fr

Abstract: The aim of our research consisted in checking the conclusions of Marsh (2007) about the psychometric properties of the “EEE” (Assessment of the quality of High School Teaching). For the purpose of our research eight items of a questionnaire used by Laurentian University have been administered to Algerian (N1 = 83) and Tunisian students (N2 = 56) in physical education and sport in September 2012. Statistical analysis made it possible the eliminating of redundant items. Furthermore exploratory factor analysis confirmed just two of the theoretical dimensions of the questionnaire, i.e. the learning enviroment and consistency in the pursuit of teaching and learning. The reliability was checked by calculating the Crombach alpha which was more than 0.83. The method ‘Educational Testing Service’ was used to estimate the degree of discriminating items and has thus shown that the eight items possesses a discriminating effect. Our findings offer to the academic training programs in the Maghreb a grid of a simple multidimensional teaching evaluation by students, consisting of only 4 reliable and available items but their validity needs to be demonstrated.

In the light of these results we can retain the conclusion of some studies like those of Harvey and Hebert (2013) which suggest that teaching evaluations by students can be considered as a valid and reliable process in the form of a global interpretation that is made from the assessment instrument which fits into an overall result. That means that we should take in account some other factors to deliver a reliable, discriminating and valid questionnaire. These factors could be related to the gender of students and teachers, to their ages, to the personality of the teacher, to the academic level of students, to their academic performance, to the quality of education as the enthusiasm (March, 1982), and finally to the kind of the evaluated courses (science and engineering versus humanities).

Keywords: Item Analysis, Validity of the Construct, Evaluation of Teaching by Students, Maghreb.

I. INTRODUCTION AND RESEARCH PROBLEM

The issue of student evaluations of teaching (SETs) is now a concern for the different actors of education. This is a growing practice through the world. It is the cornerstone of improving the quality of teaching and its valuation based on the conviction that the assessment made quality. The majority of American universities, North American and European use the student evaluations of teaching to regularly assess the quality of university education[21]...
it. Reference [53] argues that there is no consensus, no model what a good teaching. Reference [4] note that there are definitions focus on the product of teaching, the processes related to teaching and a mixture of the two definitions. Empirically, [40]-[41]-[43] argues that student evaluations of teaching are multidimensional, reliable and stable and relatively valid to a variety of quality education indicators; the student evaluations of teaching are not affected by hypothetical bias.

The literature concerning the validity of the student evaluations of teaching is partially contradictory [30]. It describes the use of a number of authors to more or less complex methodologies, which were strongly criticized. This complexity is emphasized by [2], which identified five main methodologies: the multitrait-multimethod approach, multisection approach, bias analysis, laboratory methods and the multidimensional approach. Studies on the dimensionality of the evaluation of teaching by students is relevant and fit into the whole process of verifying the validity of the instruments used. Validity is a unitary concept. It is a judgment that is focused on all evidence about an instrument and determines whether the interpretations made from it are consistent with the intention [6] - [36]. The concept of construct validity [67] refers to the degree to which an operationalization (instrument) allows to measure the concept that it is supposed to represent. Reference [36] also use the concept of validity term as synonym and considers the associated procedures give meaning to test scores in connection with the theoretical model that is to be evaluated. Reference [39] specify that it is to determine the degree to which the test performance can be interpreted in terms of one or more constructed. This degree is established from the adequacy between the theoretical and empirical structure of the measured object [21]-[38].

There is no consensus in dimensions of the construct being measured. This could be a problem when the results of evaluations are used for decision making purposes [20]. Reference [44] precise, moreover, that most of the instruments are based on a mixture of logical and pragmatic considerations rather than on a specific theoretical construct. According to [49], there would be several thousand questionnaires to the many and varied dimensions, including the construction process and the quality of the evaluation rating will vary. As part of this research we focus on two North African countries: Tunisia and Algeria. The student evaluations of teaching are not implemented systematically. We question the psychometric qualities of student evaluations of university teaching.

II. THE THEORETICAL BACKGROUND

2.1. Metric Qualities of Student Evaluations of Teaching as Reported by Researchers from Different Universities:

2.1.1 Student Evaluations of Teaching in the United States:

One of the standard tools of the most quoted and most complex is the Students’ Evaluation of Educational Quality (SEEQ). This instrument consists of nine dimensions and 35 items [40]-[41]. In contrast, [46] report that a university (Yale Law) has optimized its assessment process by developing an abridged version of his instrument, composed of 8 items, has enhanced the rate of participation of students with the combined use of a coercive measure. Between these extremes, several instruments [11] - [34] - [37] - [60] - [61] - [64] measure between 2 and 5 dimensions and are formed from 12 to 25 items. Therefore there is some variability in the measured construct. The most of These instruments measure the three roles of the teacher identified by [26] and included in the work of Abrami [3] - [4] - [20] the teaching of the course material (delivery of instruction), ease of interacting with students (facilitation of interactions), and evaluation of learning (evaluation of student learning).

Reference [55] was used the questionnaire of evaluation of the quality of education (QEQE), to assess the relationship between the evaluation of educational scores by students and the number of courses being complemented by their instructors. The evaluation of the quality of education questionnaire consists of 32 items that measured the following nine dimensions: learning / academic value, the enthusiasm of the staff, organization / clarity, group interaction, individual report, subject to cover, review / evaluation, course sessions / work overall evaluation. Of the 21 schools that have been recruited to participate, 19 returned the demographic questionnaire while 201 physical education students have completed the evaluation of the quality of education questionnaire. The internal consistencies of the following dimensions: 0.87 learning / academic value; 0.91 for the enthusiasm of the staff; 0.87 for organization / clarity; 0.88 for group interaction, 0.91 individual report; 0.85 for the material to be covered; 0.93 examination / evaluation; 0.89 for the coursesessions / works; 0.91 overall evaluation. The Cronbach coefficient for the entire questionnaire is 0.97. Thus, the scores for each dimension and for the entire questionnaire were considered very reliable. The study found a fairly strong correlation (r = 0.65, P <.05) between the amount of formal courses completed and the value of the dimension work / reading. Otherwise, this means that teachers who have completed more formal courses tend to have higher scores on the dimension reads / work. Despite the low statistical power, the learning dimensions / academic value and material to cover have respectively correlation coefficients of 0.48 and 0.49. These positive correlations are considered important and this means that teachers who have completed more courses tend to have higher scores on learning dimensions / academic value and material to cover. It is the same score on the whole questionnaire (r = 0.38). However, the correlation is considered average between teachers who have completed over the course and the scores given by the student to dimensions enthusiasm of the staff (r = 0.37), organization / clarity (r = 0.3), group interaction (r = 0.36), and review / evaluation (r = 0.27).
2.1.2 Student Evaluations of Teaching in Quebec:
Reference [33] investigated the psychometric evaluation of teaching by students in terms of response rates, satisfaction rates, reliability and dimensionality of the factors according to two conditions for the execution (online versus on hard paper) among students of the University of Quebec at Rimouski. An initial sample of 20,245 students from 868 classes completed the paper version in the Fall 2007, Winter and Summer’s quarter 2008. A second sample of 16, 432 students from 1094 classes who completed the online version, during the Fall 2010, Winter and Summer’s quarter 2011. The questionnaire consisted of 26 items that measured six dimensions: Course Context (items 1-5), Organization and Clarity (items 6-13), Dynamism, interest and teaching skills (items 14-17) Interaction with students (items 18-20), Evaluation and Feedback (items 21-24) and General Appreciation (25 to 26 items). The paper modality has had more participation (difference of 26.3%) among students compared to the online modality. Furthermore, satisfaction rates are for the paper modality. The reliabilities of the two modalities are above the threshold of 0.80 and are almost identical (difference of 0.01). Finally, the hierarchical model is the one that best describes the data set. The factor structure of the online version is similar to the paper. Indeed, in the online version of the Omega reliability coefficients specific dimensions less than 0.30 and less than 0.35 in the paper version for dimensions: clarity and organization, dynamism, interest and teaching skills, interaction with students and students, evaluation and feedback and general appreciation. The coefficient of reliability of the hierarchical dimension General skill in teaching is 0.97 for the paper version and 0.95 for the online version.

The table below shows the factorial combination of 26 items as used of the University of Quebec at Rimouski (UQAR) following factor analysis Performed by [33]

Table 1: The Grouping of Items and Dimensions Reported by [33]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Q6</td>
<td>0,45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>0,56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Saturation of the Items on Each Factor of the Model Set at Five Factors in Factor Analysis with Principal Component, as Reported by [15]

<table>
<thead>
<tr>
<th>Dimensions/items</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0,755</td>
</tr>
<tr>
<td>Objectives / teaching</td>
<td>0,803</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The content / teaching</td>
<td>0,803</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The content / presented and discussed at the first session</td>
<td>0,656</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The plan / was respected</td>
<td>0,751</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The content / structured</td>
<td>0,818</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This course / acquisition</td>
<td>0,737</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The unit of teaching / adaptable</td>
<td>0,682</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The scheduled time / respected</td>
<td>0,439</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The amount of work required / Realistic adaptable to your level</td>
<td>0,600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.1.3 Student Evaluations of Teaching in Morocco:
Reference [15] have developed and validated a computerized tool for the evaluation of teaching and distance learning by students of Ben M'sik faculty. 50 students responded to questionnaires of teaching evaluation with 24 items and 105 responded to the questionnaires evaluation questionnaires comprising 13 items. The evaluation of teachings was measured with the 5 following factors: the organization of teaching, the relationship with teachers, evaluation methods, pedagogy and evaluation results. Evaluation of training was measured with the 6 following factors: the organization of teaching, the logistics, the schedule, the content of training, acquisition, communication. In the 2 questionnaires, most factors do not have double saturations unequivocal. Cronbach’s alpha, which estimates the reliability of the items was 0.92 while the number of items that assess training is 0.85. Both questionnaires are considered valid because the commonality of each item is greater than 0, contributions are greater than 0.5. The index of KMO is greater than 0, 76 and Bartlett coefficient is significant.

The students evaluation teaching model experienced in Morocco by [15] was composed of five dimensions, Table 2 shows the grouping of items in each dimension.
2.1.4 Student Evaluations of Teaching in Greece:

Reference [59] investigated the psychometric properties of the course evaluation questionnaire (CEQ) that they have adapted. 283 Greek students in Tourism Management of two technological education institutes responded to the questionnaires during spring 2008. A Likert scale of 5 levels was used by students, 1 (strongly disagree) to 5 (strongly agree) for each item. The course evaluation questionnaire consisted of 31 items, however, the course evaluation questionnaire adapted consists of 19 items. Items 1 to 30 reflects the perception of the quality of teaching and were measured with the following 5 factors: quality teaching, clear standards and objectives, appropriate workload, appropriate evaluations, emphasis on the student's independence. The item 31 measures the level of overall satisfaction of students towards their programs. The five-factor model describes best the instrument used for the index of Tucket-Lewis (TLI) is 0.943 which is greater than 0.85 which is the significant value. Similarly, the RMSEA was 0.042 and the IFC was 0.954. Cronbach's alpha, which estimates the reliability of all the items of the questionnaire was 0.88 and that of each dimension of the items ranged between 0.82 and 0.88; these values are satisfactory. In addition, the correlation coefficients with each dimension are significant (p<0.01).

<table>
<thead>
<tr>
<th>2.1.5 Student Evaluations of Teaching in Australia:</th>
</tr>
</thead>
</table>
| Reference [50] has developed an evaluation tool that measures the academic performance of organizational units of teachings (Race Experience Questionnaire). The psychometric properties of the CEQ (Course Experience Questionnaire) are described in terms of reliability, validity. In 1989, a total of 3372 questionnaires were returned by 13 students attending post-secondary institutions across all departments (humanities, health sciences, etc.). The questionnaire composed of 26 items that measure six dimensions: quality teaching, clear objectives, appropriate workload, appropriate evaluations, the emphasis on the student's independence. Internal consistency as measured by the Cronbach coefficient of these dimensions is satisfactory, it is in the order of 0.87 for quality teaching; 0.80 for clear objectives; 0.77 for the appropriate workload, 0.71 to appropriate evaluations and 0.72 for emphasis on the student's independence. Correlation analyzes demonstrate that there is a positive correlation between overall satisfaction and perception of adequate teaching and adequate evaluation methods. The correlation is highest for the dimension teaching quality and the lowest for the dimension appropriate workload. Also in Australia, [29] revised the Course Evaluation Questionnaire (CEQ) of [51] and spent the modified release they named the student experience evaluation questionnaire (SEEQ) with students from the University of Sydney. 7632 students completed the evaluation questionnaire of the student experience between 2001 and 2002 that included 23 items and 6 following dimensions: quality teaching, clear standards and objectives, appropriate workload, appropriate evaluations, generic skills. These authors reported Cronbach alpha coefficients satisfactory: 0.83 to quality teaching; 0.80 for clear objectives; 0.76 for the appropriate workload, 0.72 for appropriate evaluations and 0.77 for generic skills. The five-factor model describes best the model predicts for the RMSEA is 0.049, value that is within the range conventionally accepted (0.049 to 0.051).

| [63] determined the psychometric properties of two versions of the course evaluation questionnaire (CEQ) of [50]. On one hand, the original form (also called completed version) of the course evaluation questionnaire composed of 36 items that measured the following six dimensions: quality teaching, clear objectives, appropriate workload, appropriate evaluations, the emphasis on the student's independence, generic skills. On the other hand, the short form of course evaluation questionnaire consists of 23 items, excluding the student's independence dimension and a new dimension was added to either generic skills in addition to 5 other sizes mentioned -High. 3 samples of students attending 14 different schools (business, humanities, law, health sciences etc.) completed |
the questionnaires. The first sample included 1362 responses graduated on 1992; the second sample of 2130 responses graduated on 1993 and the third sample of 7370 responses graduated on 1994. The completed version (Course Experience Questionnaire 36) was administered to samples graduated on 1993 and 1994, while the shorter version (Course Experience Questionnaire 23) was administered to the sample graduated on 1992. For samples of 1993 and 1994, the internal consistency of items as measured by Cronbach's 0.86 to 0.88 for quality teaching; 0.82 and 0.82 for clear objectives; 0.75 and 0.74 for the appropriate workload, 0.74 and 0.73 for the appropriate evaluations, 0.68 and 0.67 for the student's independence, 0.80 and 0.79 for generic skills. Finally for the sample of 1992, the internal consistency was 0.88 for quality education; 0.76 for clear objectives; 0.69 for the appropriate workload, 0.70 for appropriate assessments, 0.77 for generic skills. Finally for the sample of 1992, the internal consistency was 0.88 for quality teaching; 0.76 for clear objectives; 0.69 for the appropriate workload, 0.70 for appropriate evaluations, 0.77 for generic skills. Thus, we see that the alpha coefficients for the three samples indicate moderate to high levels of internal consistency for all scales. These reliability coefficients for the first five dimensions are consistent with those of pilot study of [50] and generally higher than those of study of [52], in particular the independence dimension. The alpha coefficients for the short form (Course Experience Questionnaire 23 samples of 1992), while those of certain dimensions are marginally lower compared to the full form (Course Experience Questionnaire 36, 1993 students and 1994 samples), nevertheless they demonstrate acceptable levels of internal consistency. The alpha coefficients for the short form (Course Experience Questionnaire 23 samples of 1992), while those of certain dimensions are marginally lower compared to the completed form (Course Experience Questionnaire 36, samples students 1993 and 1994), nevertheless they demonstrate acceptable levels of internal consistency. Thus, [63] argue that the two forms can be considered reliable instruments. The course evaluation questionnaire is able to discriminate between the explicit goals of different disciplines, at least between medicine programs versus psychology programs. Positive and significant correlations were found between all aspects of course evaluation questionnaire and the overall satisfaction of the course, academic achievement and generic skills; this was the case for all samples. Correlations in connection with the satisfaction in course are compatible and / or higher than those of the original study of [50]. The quality of teaching, standards and clear objectives have a stronger correlation with satisfaction and academic success. The appropriate workload demonstrates a lower correlation with satisfaction, academic achievement and generic skills. Positive association between the scores of the course evaluation questionnaire and the learning outcome measures such as satisfaction, academic achievement and generic skills strengthens the validity of the instrument as a measure of quality teaching. The long form of the structure of the instrument (Course Experience Questionnaire 36), incorporating the additional dimension "generic skills" reflects a high degree of stability and the short form (Course Experience Questionnaire 23) also provides a structure steady: so the course evaluation questionnaire is a valid, reliable and stable instrument according to [63]. Reference [31] extended the course evaluation questionnaire of [50]. In addition to the existing course evaluation questionnaire, the new instrument has the following dimensions: student support, learning resources, course organization, the learning community, the qualities of the graduate (s) and intellectual motivation. 3691 students from 15 universities Australians completed the course evaluation questionnaire whose three versions of the questionnaire were used. Because their goal was not necessarily to determine overall satisfaction, each of the six additional dimensions were analyzed separately. The Rasch reliability index was used to indicate how each dimension is independent of the others. 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 of the reliability of separation can be understood as a measure of the validity of built; and the results indicated that all dimensions have measurement characteristics coherent with latent traits, since 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 indicates an opposite trend. Older students feel they are least part of an academic 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 of the dimension of 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 learning community size decreased with the years of study. Reference [25] used the course evaluation questionnaire (CEQ) of [50] to evaluate the student experience for an entire curriculum. The CEQ was composed of questions related to the quality of teaching and learning, student support, and final 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 course evaluation questionnaire consists of 56 items that measured the following six dimensions: the academic environment, teaching, skills development, appropriate evaluation, appropriate workload, clear standards and 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, hospitality and sport. 1100 questionnaires sent to students attending 25 programs, 634 were completed and returned the completed questionnaires. On average, students are most satisfied with skills development dimension (the average is 3.8 and
3.9) followed by the academic environment (average of 3.3 and 3.4), education and adequate assessment (average of 3.2 and 3.3). Students are least satisfied with clear standards and objectives sizes (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 content monitoring methods, teaching skills and efficient staff. By contrast, students feel that some aspects should be improved including communication including timely feedback on their work, more informations on what is expected of them and more time with tutors.

2.1.6 Student Evaluations of Teaching in England:

Reference [52] tested the course evaluation questionnaire (CEQ) of [50] in several universities. The course evaluation questionnaire consists of 30 items that measure the following six dimensions: quality of teaching (eight items); clear standards and goals (five items); the appropriate workload (five items); the appropriate evaluation (six items), the emphasis on the independence of students and choices (six items). The 256 questionnaires sent to students from various disciplines in the social sciences, 95 completed and returned the questionnaires. The internal consistency was 0.79 for quality teaching; 0.77 for clear objectives; 0.71 for the appropriate workload of 0.47 for appropriate evaluations, 0.55 for emphasis on the independence of students, 0.42 for the subscale memory. These coefficients are generally lower than those reported by [50]. Analyses of variance were also performed on scores on the five factors of the first order and the second order factor, using the independent variables whose year of study, curriculum, gender and 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 younger.

2.1.6 Student Evaluations of Teaching in Niger:

Reference [47] applied a modified version of the paper being student course evaluation questionnaire (CEQ) of [50]. The sample consisted of 2221 students from 17 Nigerisan universities in science (biology, chemistry, mathematics, etc.). The questionnaire composed of 61 items corresponding to the following dimensions: quality teaching, clear objectives, course materials and resources, generic skills, appropriate workload, deep learning strategies, evaluations appropriate, the emphasis on the student's independence, superficial learning strategies. The last item measured the satisfaction of students compared to their course of degree. From July to April 2008, students completed the questionnaire, using a Likert scale to 5 modalities. The internal consistency of the dimensions is greater than 0.61. The correlations are significant, positive and strong between the perception of students and generic skills. Similarly, there is a strong correlation between student perceptions and satisfaction towards their courses; This was measured through the last item. Finally, correlation coefficients between learning strategies and perceived quality by students are: 0.43 for quality teaching, 0.41 for clear objectives; 0.04 for the appropriate workload, 0.07 for the courses of materials and resources, appropriate evaluations 0.01 and 0.05 for emphasis on the student's independence and 0.62 for generic skills. In addition, the correlation coefficients between generic skills and perception of the quality of teaching experiences and the learning environment by students are: 0.55 for quality teaching; 0.51 for clear objectives; 0.10 for the workload appropriate, of 0.14 for the course resources, 0.08 for appropriate evaluations and 0.12 for emphasis on the student's independence. Reference [58] evaluated the applicability of a modified version of the abbreviated form of the course experience evaluation questionnaire (CEQ) of [8]. The course experience evaluation questionnaire modified consists of 30 is composed of 25 items that measured the following six dimensions: generic skills; quality teaching; educational resources (materials and equipment); clear objectives and standards; the appropriate workload; appropriate evaluation I (Student reaction to evaluation); appropriate evaluation II (the impact on staff assessment). Five specific elements to medical care have been added to the original version. These were selected on the basis of the program's course content. For example, which refer to oral communication (presentations), the learning tools (video clips), and equipment (use of video cameras) are not part of the original version of the questionnaire during evaluation (Course Experience Questionnaire 25), but mainly characterized the program. For example, which refer to oral communication (presentations), the learning tools (video clips), and equipment (use of video cameras) are not part of the original version of the course evaluation questionnaire (Course Experience Questionnaire 25), but mainly characterized the program. The new items were numbered from 25 to 29. Also, in the original version, Question 25 evaluates the overall student satisfaction compared to the course whereas in the modified version, it corresponds to item 30. In 1997-1998, 199 students of the first-year in general medicine, veterinary medicine and pharmacy, 168 completed the questionnaire. Students of medicine followed a comm.-

Copyright © 2015 IJIRES, All right reserved
Reference [8] determined the psychometric characteristics of the abbreviated course evaluation questionnaire (CEQ) among medical students. The short course of the evaluation questionnaire composes of 25 items that measured the following six dimensions: quality teaching, goals and clear standards, appropriate evaluation, appropriate workload and generic skills. In addition, the 25 course evaluation questionnaire contains an item relative overall student satisfaction in relation to the course. 149 students completed the course evaluation questionnaire using a Likert scale of 5 points. The internal consistency was 0.78 for generic skills; 0.57 for the appropriate workload; 0.37 for light standards and objectives; 0.66 for the quality teaching I; II 0.61 to quality teaching; 0.55 for appropriate evaluation. A strong structure with six factors reflect how medical students perceive the course accounted for 60% of the variance: Generic skills (25%); the appropriate workload (10%); clear objectives and standards (9%); Quality Teaching 1 (6%); Quality Teaching 2 (5%) and the appropriate evaluation (5%). Five of these are very similar to the Course Experience Questionnaire original. The sixth factor comes from a split in the dimension of quality education into two factors described like this: the quality of teaching 1 refers to the way the teacher interacts and presents educational materials to students and quality of teaching 2 refers to feedback given to students. The inter-correlation dimension was examined using bivariate correlation analysis of Pearson. It follows an interesting point: the dimension generic skills is connected in some way to all other dimensions, the strongest correlation with the dimension clear objectives and standards. The degree of correlation between the quality of teaching 1 and the quality of teaching suggests that the two sub-dimensions probably reflect different aspects of teaching.

Table 3 presents some instruments of measurement of teaching evaluations by students as listed by [33].

<table>
<thead>
<tr>
<th>studies</th>
<th>ND</th>
<th>NI</th>
<th>Correlations (Dimensions)</th>
<th>Alpha (Dimensions)</th>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feldman (1976)</td>
<td>19</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>19 dimensions that fall into three teacher roles</td>
</tr>
<tr>
<td>Marsh (1983)</td>
<td>9</td>
<td>35</td>
<td>na</td>
<td>0.87 to 0.98</td>
<td>Learning, Enthusiasm, Organization, Interaction, Individual report, covered material, exams, works, workload</td>
</tr>
<tr>
<td>Hess and Thorpe (2005)</td>
<td>al. 3</td>
<td>15</td>
<td>0.52 to 0.63</td>
<td>0.84 to 0.87</td>
<td>Evaluation, Communication, Conception, Preparation, Quality, Student, course, Professor, Global</td>
</tr>
<tr>
<td>Toland and Ayala (2005)</td>
<td>De 3</td>
<td>25</td>
<td>na</td>
<td>nd</td>
<td>Nulles</td>
</tr>
<tr>
<td>Wong and Fitzsimmons (2008)</td>
<td>3</td>
<td>19</td>
<td>0.80 to 0.87</td>
<td>0.91 to 0.93</td>
<td>Personal attributes, Facilitation of learning, Feedbacks</td>
</tr>
<tr>
<td>Layne and Layne (1999)</td>
<td>al. 3</td>
<td>14</td>
<td>Nulles</td>
<td>0.95 to 0.99</td>
<td>Preparation and presentation of the course, Interactions, Evaluation, Personal attributes, Facilitation of learning, Feedbacks</td>
</tr>
<tr>
<td>Byrne and Flood (2003)</td>
<td>5</td>
<td>23</td>
<td>na</td>
<td>0.66 to 0.78</td>
<td>Teaching, Goals, workload, Evaluation, Global</td>
</tr>
<tr>
<td>UQAR</td>
<td>5(6)</td>
<td>21(26)</td>
<td>-----</td>
<td>-----</td>
<td>(context), Organization, Dynamism, Interaction, Evaluation, General appreciation</td>
</tr>
<tr>
<td>Yale Law</td>
<td>na</td>
<td>8</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
</tbody>
</table>

aND = Number of dimensions
bNI = Number of items
cna = not available

dWe would remind that [45] stated that evaluations of university teaching by students are multidimensional, reliable, stable and relatively valid.

2.1.8 The Purpose and the Social Relevance of Research:

The specific objective is to analyze the psychometric qualities of students teaching evaluations as used by Laurentian University in terms of their fidelity, their degree of discrimination and their construct validity. We will administrate eight items at two universities in the Maghreb: Algeria and Tunisia. It is not to see a particular teaching as effective or not, but to verify the psychometric properties of a questionnaire consisting of 8 items. The social relevance is to provide reliable feedback to teachers...
about the quality of their education and also provide administrators with a tool to take administrative decisions related to career progression for teachers.

**III. METHODOLOGY**

**3.1 Sample:**
With the help of two North African collaborators, we set up in September 2012 a set of two anonymous samples of students: 83 Algerian students who followed a movement analysis course at the University of Algiers and 56 Tunisian students who have completed a course psychophysiology at the University of Tunis. A sample of 139 students of the Maghreb was therefore established in September 2012, which allowed to perform a pilot study. The Students are everyone enrolled in various courses at the postsecondary level in one or other of the mentioned universities: the University of Algiers (Algeria) and the University of Tunis (Tunisia).

**3.2 Instrumentation:**
We used the questionnaire for the evaluation of teaching by students from Laurentian University because it only consists of eight items on the one hand and on the other it has never been a empirical verification.

<table>
<thead>
<tr>
<th>Question 1: The teacher meets the objectives that he (she) has set</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 2: The teacher communicates clearly when (she) gives Instructions, present the course material, anime discussions.</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 3: The reports of the teacher with me are:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 4: The teacher has created an atmosphere of openness mind and tolerance in which I feel comfortable:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 5: The teacher's ability to stimulate my interest in learning is:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 6: In my opinion, the overall evaluation of my performance by the teacher is:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 7: I find the effectiveness of this teacher in this course:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 8: I would recommend a course taught by the teacher to student or a friend:</th>
<th>Very few</th>
<th>few</th>
<th>moderately</th>
<th>good</th>
<th>very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

The first seven dimensions of the questionnaire evaluate different aspects of the quality of teaching. The student must show its degree of agreement (or disagreement) to the items using a Likert-type scale to 5 terms, ie 1 = very low; 5 = very good).

**3.3 Conduct of Research:**
The students were asked to complete the evaluation questionnaire of teaching by students during the month from April to October 2013. Initially, the researcher (or his assistant) distributed questionnaires to students and explained to them the purpose of the research (estimating the psychometric properties of the eight items) and stated that the questionnaires were anonymous (or the names of students or the names of teachers who provided the selected courses shall not be sought or known). Thereafter, courses to evaluate were defined. It should be noted that none of the evaluated courses were taught by the teacher who distributed the questionnaires. He read the eight teaching evaluation questions by the students and explained them aloud. Then he gave the students time to answer a question before moving on to the next question. The questionnaires handover protocol was not to last more
than 10 minutes. The students became aware of the questionnaire and delayed responses in an attached answer sheet by circling their answer choices according to Likert scale proposed; this, without indicating the name of the professor of the evaluated subject. After answering, the researcher or his assistant were recovering locally anonymous responses and thanked the students for their kind cooperation. Finally, the researcher compiled the anonymous responses in an Excel file and sent by email to the principal investigator.

3.4 Data Analysis Method:
3.4.1 Reliability (or the internal consistency of items)
This metric quality refers to the preciseness and internal consistency of an instrument, that is to say, the proportion of variance attributable to the true score of the total variance[9]-[10]-[51]. It determines the extent to which a score is affected by the presence of sources of error. In classical test theory, the Cronbach's alpha is used on items that compose each of the dimensions and all the items of an instrument. Although there are several other ways to calculate a fidelity index, this index is the one that is most regularly reported [17]. The Cronbach's alpha [18]-[19] is a good measure of the internal consistency of a latent variable; the acceptable values are normally above 0.70 [48]. The interpretation of the alpha coefficient depends on the number of items, most of items there, plus alpha coefficients can be raised with yet rather average inter-item correlations [16]-[65].

3.4.2 The Degree of Discrimination:
This metric quality is very little studied by researchers. The degree of discrimination of the items will be estimated through an analysis of the items according to the method Educational Testing Service (ETS). According to the Educational Testing Service method (an item whose degree of discrimination is less than 3 indicates that the item is unable to differentiate the evaluated subjects [32].

3.4.3 Construct Validity:
Studies on the dimensionality of the evaluation of teaching by students is relevant and fit into the whole process of verifying the validity of the instruments used. Validity is a unitary concept. It is a judgment that is focused on the accumulated evidences about an instrument and determines whether the interpretations made from it are consistent with the intent[6] -[36]. The concept of construct validity [67] refers to the degree to which an operationalization (instrument) to measure the concept that it is supposed to represent. The literature (see [36]) also uses the concept of validity term as a synonym and considers the associated procedures give meaning to test scores in connection with the theoretical model that is to be evaluated. Reference [39] specify that it is to determine the degree to which the test performance can be interpreted in terms of one or more constructed. This degree is established from the balance between the theoretical and empirical structure of the measured object [21]-[38].

In recent years studies have examined the validity of teaching evaluations by students. According to [57], evaluation of teaching by students is a controversial topic, few evaluations of teaching by students have been empirically validated and few of them are supported by theoretical foundations. Many of the faithful and valid questionnaires have been developed since the 70s but without academic consensus [35]. According to[45], teaching evaluations by students are multidimensional, reliable, stable and relatively valid. Factor analyzes have identified nine dimensions of Student Evaluation of Quality Educational (SEEQ): Learning, Enthusiasm, Organization, Interacts in the group, relationship with the teacher, Wealth teaching, evaluation of students, staff labor and overall judgment.

In this research, the authors propose to evaluate the dimensionality of the questionnaire of eight items used by Laurentian University through principal components factor analysis (PCA) with varimax rotation. This type of analysis helps to identify the independent components extracted and ensure the validity of the questionnaire. Exploratory factor analyzes were also conducted to determine the latent these analyzes were performed on the first 7 of the questionnaire items only as the last item related to a general dimension of the quality of education: I would recommend a course taught by this teacher; this item is not considered in the analysis.

3.5 Ethical Considerations:
Students received in the same time a consent letter in which they were informed of the nature of the research. This is a voluntary participation, students can complete the questionnaire without being forced and may withdraw from the study at any time without risk of being penalized.

The handover protocol teaching evaluation questionnaire by the students is presented in the consent form: students are in a classroom and are asked to complete the anonymous questionnaire on site. Courses will not be evaluated those taught by one or other of the researchers. Researchers or assistant ask students to evaluate courses that were not taught by themselves.

This study has been accredited by the Ethics Committee of Research at Laurentian University. The file number is 2013-03-08. This research was funded in part by the National Centre for Scientific and Technical Research of Morocco. We informed the participants that the results of the data analysis will be available for inspection at the principal investigator of the web page that posted them at the end of September 2013:
http://142.51.14.1/Laurentian/Home/Departments/Human+Kinetics/Faculty/Faculty+Members/Dr.+Jaouad+Alem/ General+info.htm?Laurentian_Lang=fr-CA

Students were also informed of the results of the research, by a display in the space provided for the information of students in their institutions.

IV. RESULTS
The main results are reported in the tables below. They have three psychometric qualities of the Laurentian questionnaire. The internal consistencies, degrees of discrimination and factorial solutions in both contexts (Algeria, Tunisia), where the eight items were tested.
Table 5: The Degrees of Discrimination, Internal Consistency of Items and the Factor Solutions (Exploratory) 8 Items Experienced in the Algerian and Tunisian Contexts

<table>
<thead>
<tr>
<th>Données de l’Algérie (83 sujets), Alpha = 0.93, 2 components extracted: 75% of total variance explained</th>
<th>Données de la Tunisie (56 sujets), Alpha = 0.83, 2 components extracted: 63% of total variance explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotated Component Matrix*</td>
<td>Rotated Component Matrix*</td>
</tr>
<tr>
<td>emerging component</td>
<td>the teaching-learning objectives</td>
</tr>
<tr>
<td>it1</td>
<td>.844</td>
</tr>
<tr>
<td>it5</td>
<td>.794</td>
</tr>
<tr>
<td>it6</td>
<td>.774</td>
</tr>
<tr>
<td>it8</td>
<td>.753</td>
</tr>
<tr>
<td>it2</td>
<td>.716</td>
</tr>
<tr>
<td>it7</td>
<td>.691</td>
</tr>
<tr>
<td>it4</td>
<td>.813</td>
</tr>
<tr>
<td>it3</td>
<td>.777</td>
</tr>
</tbody>
</table>

The Cronbach alpha coefficients, measuring the internal consistency, Algeria and Tunisia, are respectively in the range of 0.93 and 0.83. They are satisfactory (above 0.83). A principal components factor analysis by the method that maximizes the variance (varimax) was performed on the responses. The results of this exploratory factor analysis explained respectively in Algeria and Tunisia, 75% and 63% of the total variance. The factor solution for each country reveals two main components strong likeness: the learning climate that can be defined by the item 4 and item 3 and the teaching-learning objectives that can be defined by item 1, item 5 and item 6. These two components alone explain more than half of the total variance.

V. DISCUSSION OF RESULTS

Aspects of the most controversial interests us in this study was that which involved the analysis of the dimensionality of the construct that is actually measured by these evaluations. It appears that the teaching evaluations by students are as reliable calculation of Cronbach's alpha. The questionnaire shows indeed very acceptable reliability measures (0.93 and 0.83): they are usually beyond the desired threshold is 0.80. Our results agree with those of [33], where the reliabilities of the two terms, that is to say, online and paper was above the threshold of 0.80. The study of [15] reported for its part, still higher reliability: the order of 0.922.

We also analyzed the degree of discrimination of the items; these are all discriminants according to Educational Testing Service method (ETS).

Exploratory factor analysis with varimax rotation produced in the first seven items of the questionnaire identified the number of orthogonal factors (uncorrelated factors) present in the data. The factor solution (ACP) for each country reveals two main components strong likeness: the learning climate that can be identified by the item 4 and item 3 and the teaching-learning objectives that can be defined by the item 1, item 5 and item 6. These two components alone explain more than half of the total variance.

The study of [33] highlighted the presence of a single hierarchical dimension: General Skill in Teaching with an impact on the reliability of the primary dimensions; otherwise, factor structure remained the same for both versions (online and paper). These authors also suggest the evaluation of teaching by students in a more formative context.

The factor solution of study [15], the five empirical factors obtained correspond perfectly to the theoretical dimensions: Organisation of teaching, relationship with teachers, evaluation methods, Pedagogy first evaluation results. Most of the factors in this study were unambiguous, without double saturation.

In the case of this study, the factor solution does not indicate seven clearly distinct dimensions, but rather two distinct dimensions. So, it appears that the question of the dimensionality of teaching evaluations by students is even asked.

VI. CONCLUSION

The objective of this study was to analyze three psychometric properties of the instrument used to Laurentian for evaluating the quality of university teaching. These three qualities are reliability, degree of discrimination and construct validity. It appears that the reliability and the questionnaire's ability to differentiate between the levels of the quality of teachings are quite acceptable: the 8 items of teaching evaluations by Laurentian students are reliable and discriminating.
This study suggests that it is possible to eliminate redundant items and offer academic training programs in the Maghreb a grid of evaluation of teaching by students simpler, consisting of only 4 items with the following metric qualities:

- Reliability of items in terms of their internal consistency,
- The validity of the items but only in terms of their ability to assess two distinct dimensions of quality of education: the climate and the learning objectives,
- The availability of items or the facility to administrate the questionnaire in terms of number of items. Indeed, by themselves, the 4 items retained 'capture' more two-thirds of the total variance.

Finally, we will take the conclusion of the study [33] who agree that teaching evaluations by students can be considered as a valid and reliable process when the interpretation made from the measuring instrument fits into a perspective where only the overall result is considered. As future research prospects, several questions remain. There are questions on other dimensions of quality of education as the enthusiasm and other criteria of teaching evaluations by students (March, 1982): should we try to make the evaluation also by students? If yes, how?

What about the other factors possibly related to the evaluation of teaching by students such as gender of students and teachers, the age of the students and teachers, the personality of the teacher, the academic level of students, the academic performance of students, the type of evaluated courses (science and engineering versus humanities)? Should we also take into account all these factors to deliver a reliable questionnaire discriminant and valid?

REFERENCES


---

**AUTHOR’S PROFILE**

**Dr. Jaouad Alem**, Associate Professor, School of Human Kinetic, Laurentian University, Ontario-Canada. Main Research: Psychometry, quantitative research, assessment, learning, pedagogy. Phone: 1 705 675 1151 poste 1013.

email: jalem@laurentian.ca

**Ines Bouguerra**, Student- PhD in Human Studies Laurentian University, Ontario-Canada. Main Research Axis: Mental coaching, consciousness, Students valuations of teaching (SETs) Cell phone: (+216) 26 28 03 66

---

Copyright © 2015 IJIRES, All right reserved
Dr. Sabeur Hamrouni, Assistant-Professor
The High Institute of Sport and Physical Education of
Tunisia, University of Manouba, Tunisia
Main Teaching Activity: Psychophysiology, Sport
Psychology and Developmental Psychology
Main Research Axis: Mental Coaching In Sport, Cardiac
Coherence and Biofeedback in sporting activities
Cell phone: (+216) 52688647
email: sabeurhamrouni@yahoo.fr

Dr. Hakim Hariti, Professeur des universités
Directeur du laboratoire SPAPSA, université d’Alger 3
Directeur de la formation en doctorat STAPS, filière:
Education et Motricité Membre du conseil scientifique
de l’ATRSSH de la DGRSDT, Algérie
Localisation: Institut d’éducation physique et sportive de l’Université
d’Alger a, Ville nouvelle de Mehelma, Zéralda Alger, Algérie, 16832
email: hhariti@yahoo.fr
Tél. ligne 1 : +213770348940 ligne 2 : +2130665192466
Compétences: Evaluation de la performance motrice, apprentissages
moteurs, contrôle moteur