SSRG International Journal of Economics and Management Studies
ISSN: 2393 – 9125 / https://doi.org/10.14445/23939125/IJEMS-V7I5P121

Volume 7 Issue 5, 130-142, May, 2020

Original Article

Validation of a Scale for Assessing the Quality of University Training Devices in Sciences and Techniques of Physical and Sports Activities (SAQ-UTD / STPSA)

Raja Lotfi[1], Touria Neggady Alami [2]

[1,2] *Management Techniques, Laboratory for Research in Organizational Management (LAREMO), Higher School of Technology (EST), Hassan II University, 20020 - Casablanca, Morocco.*

Received Date: 10 April 2020
Revised Date: 21 May 2020
Accepted Date: 26 May 2020

Abstract *- The aim of our study is to develop and validate a scale for assessing the quality of training devices in sciences and techniques of physical and sports activities in five Moroccan higher schools with regulated access.*
We borrowed Churchill's validation model and used measurement theory. From a multi-staged process of item generation, selection and scaling, field testing, and refinement procedures, exploratory factorial analysis (EFA) were verified with a sample of 267 students in STPSA. The construct validity of the emerging dimensions was also tested for each factor retained.

The results reveal that the scale identifies 34 items, structured into eight quality dimensions and having a good internal consistency (α =.85), an acceptable internal coherency (r = .803), and satisfactory temporal stability (r = .73). Confirmatory factor analysis allowed us to readjust and significantly improve our developed model (Δ χ2 (10) = 644.64; p< .000). We were also able to extract a general implicit factor called the "Training Quality Index" (TQI). This controls, only but significantly (p<.05), three factors of our scale.

The QAS-UTD/STPSA scale has valid and very satisfactory psychometric properties. In fact, eight distinct dimensions have been identified, explored, and confirmed. It constitutes a sufficiently reliable and appropriate tool. It explores parameters little or not taken into account by the usual instruments. Therefore, except for a few minor limitations noted, it can be used in other similar academic institutions to continuously improve their quality of university training devices.

Keywords *- Validation, Scale, Measurement, Assessment, Quality, Training device, STPSA.*

I. INTRODUCTION

The quest for quality teaching has never been more topical than in the past decade. It has become an international watchword. Today we are witnessing a generalization of strategies for evaluating the quality of education in general and higher education in particular. You just to take stock of the many colloquia and symposia on higher education and browse the abundant literature that deals with its "quality" to realize it. Likewise, the proliferation of quality evaluation agencies does not seem to meet any significant opposition. Who can say that he is against quality in higher education? Nobody obviously. But what quality is it?

Indeed, in the absence of a unanimous and consensual definition of what 'quality of higher education training device might be, and of how to measure it well and assess it better (De Ketele, 2014), any quest for quality remains a pipe dream. The expected quality risks becoming much more a means than an end in itself (Bouchard & Plante, 2002). Quality is therefore confronted with two major problems: that of stable, precise, and unanimous theoretical foundations and that of scientific rigour in the evaluation method.

A priori, three main categories of variables influence the quality of higher education (Endrizzi, 2014): the institution (its technical, structural, environmental, logistical, and budgetary dimensions), the training design (contents, programming, management, and organization) and the training actors (student, teacher, pedagogical manager, administrative manager, etc.).

There are many studies in the literature addressing quality issues that deal with one or other of these variables, such as those by Bigras, Lemire & Eryasa (2017), Mottier Lopez et al. (2018), and Simard (2018), who respectively measured the quality of TD by the level of student performance, the evaluation of teaching discerned and by training effectiveness.

However, we found very little research that addresses this issue in terms of the interaction between these three categories of variables or that measures them jointly.

These include, for example:

- Postiaux and Salcin (2009) present an approach at the crossroads between the evaluation of program content and that of teaching styles and methods;

- Gerard, Hugonnier & Varin (2018) measured the quality of the training device through four variables: teacher commitment, effectiveness, efficiency, and equity of teaching;
- Rege Colet (2009) and Younès (2015) advocate a global approach to teaching assessment by students.

Based on these epistemological and methodological considerations, and the fact that existing instruments do not jointly measure the different constructs or potential descriptors of the quality of training devices (TD) (Vinokur, 2013), we will attempt, during this study, to determine the attributes of the quality of higher education training devices. These potential descriptors will be formalized in a measurement instrument and tested for validation in the field.

The aim of our study is to develop and validate a multidimensional measurement and assessment scale for the quality of STPSA training devices in five Moroccan university institutions.

II. THEORETICAL FRAMEWORK

In the scientific and professional literature, the difficulty of measuring and assessing the quality of training devices is notable. The components of these are still under-evaluated (Meignant, 2014). On the basis of a brief review of the literature, we first define the concept of quality, then we establish a critical assessment of the studies on the quality of the university training devices and the methods and instruments of its measurement.

A. Quality is a floating and plural concept

Quality is synonymous with perfection (Aspisi, 1995; Poole, 2010), freedom from defects (zero defect) (Crosby, 1979). According to Parasuraman et al. (1985; 1991), it refers to implicit and explicit needs (Juran, 1993), customer requirements, and customer satisfaction (Tempus, 2001). In fact, the definition most often used in the contemporary literature on quality is generic. It is that of the International Organization for Standardization (ISO), which equates quality with "fitness for purpose", i.e., "all the properties and characteristics of a product or service which confer the ability to satisfy expressed or implicit needs "(AFNOR, 2015). Indeed, Hersan (1999) argues that service is qualified as good quality if it satisfies the customer's requirements in terms of technical and functional characteristics, deadlines, costs, and adaptation to the various constraints of the context. Quality is sometimes synonymous with excellence (Doherty-Delorme & Shaker, 2001), an effort towards the best or exceptional (Arcaro, 1995). Sometimes it refers to efficiency, the attainment of benchmarks, norms, standards, and indicators (Beamish, 2004). It also reflects the implementation of certain professionalism: a job well done according to the rules of the art (Figari, 2008). It is also the expression of conformity to particular specifications (Moss, 2005) or in relation to what is required (Roegiers, 2012).
Conversely, it is through "non-quality" that we could better approach what quality could really be (Stake & Schwandt, 2006). "Non-quality" reflects the negative gap noted between the quality targeted and the quality obtained. It would correspond to products or services that do not comply with customer norms, standards, and requirements. It translates critical defects and deficiencies detected during quality controls.

B. The dimensionality of the Quality concept

Given its multiple resonances and sometimes even its dissonances in those who target it, quality remains relative and polysemous. It has difficulty stabilizing. It corresponds to an ideal towards which all entities must strive (Bouchard & Plante, 2003). Quality is, therefore, a plural and extensive concept, a dynamic and inclusive construct, which further complicates its delimitation.
To this polysemy of the concept is added an additional difficulty: how can we quantify a measure of quality that is qualitative in nature? Indeed, the quality of training devices then becomes a multidimensional construct resulting from a double complexity: that of describing it well and that of better measuring and evaluating it. All this explains why, to date, there is no exhaustive, consensual, and validated measure of the quality of TD.

III. CRITICAL LITERATURE REVIEW OF TRAINING DEVICES QUALITY MEASUREMENT INSTRUMENTS

Most of the TD quality measurement instruments used to date have been developed from a purely theoretical rather than empirical basis (Hannon & Peterson, 2017; Zaslow et al., 2010). Some tools contain evaluation grids based on criteria, indices, and ratios. Others are simple thematic questionnaires, which are unanchored and developed by non-practitioners authors.

Likewise, these instruments have not undergone validation tests of their internal structures and confirmatory tests. This leads us to wonder whether these tools have the necessary psychometric and metrological properties to measure the construct of training quality in a precise, sensitive and meaningful way.

In fact, three organizations, the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME), point out that the ability of an instrument to properly quantify a measured variable is expressed by its validity and reliability (AERA, APA, and NCME, 2014).

However, the few measurement scales that are sufficiently reliable, valid, and approved [such as the Educational Quality Observation Scale (EOQS) (Bourgon & Lavallée, 2004) and the Classroom Assessment Scoring System (CLASS) (Pianta, La Paro & Hamre, 2008)], assess only one or two dimensions at most of the TD quality. We consider this to be very simplistic in relation to the multiple meanings we give to the concept of "quality".

Also, the SERVQUAL instrument (Parasuraman, Zeithaml, and Berry, 1985, 1991) proposes a conceptualization of perceived quality in five dimensions: service reliability, helpfulness, empathy, assurance, and the existence of tangible elements. Although it is by far the most widely used standardized tool for assessing service quality, it is strongly criticized in the literature. For many researchers, the proposed dimensions of service quality are not necessarily transposable and generalizable to different contexts (Babakus and Boller, 1992; Carman, 1990; Cronin and Taylor, 1992). SERVQUAL most often requires modification to be adapted to the specific context studied.

Several authors have worked on dimensions of service quality. According to Grönroos (1988), perceived quality is based on two dimensions: technical quality (which refers to the result of the service provided) and functional quality (the way in which the service was rendered; e.g., with courtesy, speed, and professionalism…). Rust and Oliver (1994) add a third dimension to Grönroos' proposal: the environment in which the service is rendered. Nevertheless, these three dimensions (Outcome, Process, and Environment) remain generic and difficult to measure.

Consequently, we believe that the quality of service is a function of the specificities of the activity sector under consideration. The number and nature of the dimensions of service quality are directly related to the types of service that the researcher analyses.

On the other hand, the measurement of quality in STPSA training devices is still in its embryonic stage (in its infancy). It highlights a certain degree of uncertainty or disagreement about the set of factors or the conceptual model that can adequately describe them. Moreover, only a very limited number of scales have been developed and deemed relevant for the sports sector in general and the STPSA university discipline in particular.

Chelladurai, Scott, and Haywood-Farmer (1987) developed a Scale of Attributes of Fitness Services (SAFS) to evaluate the service quality provided by fitness clubs to Canadian customers. It includes 30 items divided into 5 dimensions. However, the factor structure of the SAFS was not examined by the exploratory factor analysis.

In other research, Kim and Kim (1995) developed the Quality Excellence of Sports Centers (QUESC) instrument to assess the quality of service in sports centres. QUESC was based on a Korean sample and included 33 items in 11 dimensions supported by exploratory factor analysis. However, three factors (price, privilege, and stimulation) of the QUESC include only one item each. This makes the stability of a single-item factor very questionable.

Papadimitriou and Karteroliotis (2000) re-examined QUESC on a sample in Greece. They used exploratory factor analysis and concluded that the factor structure of QUESC is unsustainable. To remedy this, they suggested and tested a 4-factor model including (1) instructor quality, (2) attraction and operation of facilities, (3) availability and execution of the program, and (4) other services.

However, these four dimensions explain only 57.1% of the total variance. Similarly, CSC, of Korean origin, was taken up and tested in a Greek environment without assessing its transferability and without transcultural adaptation (translation validity and generalizability validity).

Although merits can be associated with the above-mentioned scales, their content and structural validities are questionable. They are more or less generic, designed to measure service quality in a global and superficial way. They appear to require more in-depth additional content. Therefore, they do not provide precise information to managers, allowing them to assess different aspects of the service quality rendered in order to improve it.

IV. METHODOLOGY

In order to develop and validate an instrument for measuring the quality of TD, which would attempt to fill the gaps and shortcomings of the tools identified in the literature review, this work aims to cover as many dimensions as possible and also to take into account the point of view of students, as the first clients and beneficiaries of training. To do so, through a mixed methodological approach (Creswell & Plano Clark, 2018), we borrowed Churchill's validation model (1979) and exploited measurement theory (Brennan, 2006; Guilford, 1954; Thorndike, 1971).

A. *Development of items corpus*

Our item generation process was done in three distinct steps. The first phase concerns the development of items. They are developed from the literature review. This phase made it possible to constitute the first list of 116 statements related to attributes, indicators, and criteria of the training devices' quality in higher education.

In the second phase, the set of statements is first submitted to a panel of experts consisting of four researchers to select the most relevant and coherent items capable of covering each dimension (Bardin, 2007). Then, through a phase of qualitative analysis in a "focus group" (Narang, 2012; Paillé & Muchielli, 2012), the sorting, formulation, clarity, and precision of the statements were improved.

At the end of this phase, 62 statements are eliminated. The rules for deleting items are either semantic imprecision, conceptual impertinence, difficulty in measuring the item, lack of neutrality so as not to generate bias in the responses (Mayer & Ouellet, 1991), or the item has a relationship with several dimensions. Thus, 54 items were selected to constitute the first version of the scale that will be the subject of a psychometric validation study.

B. *Description and administration of the first version of the scale*

This draft, consisting of 54 items, is called the Quality Assessment Scale of University Training Devices in Sciences and Techniques of Physical and Sports Activities (QAS-UTD / STPSA).

It is administered to a sample of 267 students belonging to five regulated access institutions of Moroccan higher education. We've targeted seven STPSA courses (Professional Licenses in Physical Education and Sport (80%), Sport and Leisure (10%), and Sport Management (10%)). The sample is made up of 188 males and 79 females (70.41% and 29.58%, respectively), aged 19 to 23. Participants were asked to rate each item using a four-level Likert scale (Likert, 1932), ranging from 1 (strongly disagree) to 4 (strongly agree).

C. Scale psychometric properties analysis

a) Reliability analysis methods

The internal consistency of the scale is tested by the split-half method (Nunnally, 1978). At the same time, internal coherency is achieved by the Cronbach alpha coefficient (Cronbach, 1951), for which a threshold of .7 is considered very acceptable. Note that we re-examined the same indices after eliminating the items whose skewness and Kurtosis coefficients are greater than ±1 and whose saturation coefficients in exploratory factor analysis (EFA) were less than .4.

Stability analysis is carried out by the test-retest method (Benzécri et al., 1973). One month after the first test, a retest of the scale was performed on a sample of 53 students, whose characteristics were almost unchanged, to explore the reproducibility of the results.

The correlation between the scores obtained on the first test and those obtained during the retest was analyzed. With a correlation coefficient greater than .70, it was possible to ensure that the instrument used is stable, that it reproduces the same results in both tests (test-retest), and that it does not depend on changes in mood or opinion of the respondents (Hendrickson et al., 2004).

b) Exploratory Factor Analysis (EFA) method

Exploratory factor analysis was used to identify latent factors from the measured variables (Costello & Osborne, 2005). The factor structure was examined, via SPSS 25 software, using the maximum likelihood extraction method with axis rotation (Varimax), assuming moderate inter-factor correlations (Kieffer, 1998). It thus makes it possible to study the factor structure of data collected without reference to predetermined dimensions. We retained the maximum of interpretable factors, whose eigenvalue is greater than 1, and the explained variance is greater than or equal to 50% (Guttman, 1954). The indices used in the factor analysis are the KMO (Kaiser-Mauer-Olin) index and the determinant of the correlation matrix. These are two indices showing the existence of correlation patterns between the items of the scale (Bourque et al., 2006). A saturation coefficient greater than .40 allows the items to be retained on the factor.

c) Confirmatory factor analysis (CFA) method

Based on the results of the exploratory factor analysis (EFA), we carried out a confirmatory factor analysis to assess whether the factor structure of the scale adequately matched the data. For this purpose, we adopted the maximum likelihood estimation method with standardized coefficients. Indeed, we considered several indices of adequacy as suggested by Hoyle (1995) and Hu & Bentler (1999):

- The Chi-square (X^2) and its ratio with the number of degrees of freedom (X^2/pdf) as an index of parsimony (Bollen, 1990),
- The Comparative Fit Index (CFI) as an incremental index (Bentler, 1990),
- The Root Mean Square Error of Approximation (RMSEA) as an absolute adjustment index (Steiger, 1990).

A non-significant Chi-square (X^2) test reveals a suitable model. However, the Chi-square test is highly sensitive to sample size, so other indicators are usually examined when estimating a model (Bollen, 1990). Thus, the authors suggest that an X^2/df ratio of less than 3 and CFI values greater than 0.90 indicate an adequate model (Bentler, 1990; Bollen, 1990).

Similarly, RMSEA values below 0.05 indicate that the proposed model is excellent, while values between 0.05 and 0.08 indicate an acceptable fit of the model to the data (Browne & Cudeck, 1993).

However, if the proposed model does not meet the suggested adequacy standards, indices of modification described by Jöreskog and Sörbom (1984) are provided. The decision to include modifications to the model is made carefully and is based on theoretical foundations (Silvia & MacCallum, 1988).

Once adjustments were made, it was subsequently verified whether the adjusted model contained an implicit second-order factor controlling all or some of the factors derived from the EFA. The relationships are then presented as standardized regression coefficients.

These confirmatory factor analyses were conducted using Amos version 25 software. The significance of the results is set at $p < .5$ (Arbuckle, 2006).

V. RESULTS

A. Exploratory Factor Analysis (EFA)

Table 1 presents the measure of the KMO (Kaiser-Mauer-Olin) index and the determinant of the correlation matrix. The measured KMO value is .82. This is well above the recommended threshold (.70). Our items, therefore, have fairly compact correlation patterns, allowing us to clearly distinguish between factors (Neuville & Frenay, 2010). As for the correlation matrix determinant, it represents a small value (.004) but not zero. What's right is the norm.

Table 1. KMO Index and The Correlation Matrix Determinant of Scale

Indices		QAS-UTD / STPSA
KMO[1] index for the measurement of sampling quality		.796
Bartlett Sphericity Test	Approximate Chi-square	8103.027
	DF	561
	Meaning of Bartlett	.000
Correlation determinant		1.368E-14

[1]: Precision measurement of Kaiser-Meyer-Olkin sampling.

The results of the exploratory factor analysis are presented in Table 2, including the factor structure obtained, the eigenvalue criteria, and the percentage of explained variances (Reckase, 2009).

Table 2. Eigenvalues and Total Explained Variance of Each Factor Identified by the EFA.

Factors	Eigen-value	Per cent of the variance	Cumulative per cent variance
F1 (COG)	6.371	18.737	18.737
F2 (CUR)	3.808	11.201	29.938
F3 (HR_QS)	3.440	10.117	40.055
F4 (MRM)	3.105	9.133	49.188
F5 (TM_LS)	3.038	8.936	58.124
F6 (ICT)	2.671	7.854	65.978
F7 (ATT_B)	2.306	6.783	72.761
F8 (ED)	1.759	5.173	**77.934**

The EFA results highlight eight major factors forming the QAS-UTD / STPSA scale, which have an eigenvalue greater than 1. These eight factors explain 77.93% of the total variance. This is a satisfactory proportion (Gorsuch, 1983), knowing that the threshold for retaining the number of factors generated is at least 60% of the explanation of the variance, with eigenvalues exceeding 1 (Guttman, 1954). Thus, the factorial matrix shown in Table 3 summarizes the saturation coefficients of each item by factor.

Factor 1 explains 18.74% of the total variance. It has an eigenvalue of 6.37 and includes five items assessing the conditions of training devices organization and governance (COG).

The second factor is also made up of five items assessing the quality of the curricula and programs adopted (CUR). It explains 11.2% of the total variance.

The third factor consists of four items reflecting the quality of the qualifications and competencies of the human resources involved in the TD (HR_QS). It explains 10.12% of the variance. Its items make up two sub-factors belonging to different semantic fields: one sub-factor refers to the technical, methodological, and professional skills of teachers and educational managers (items HR_QS1 and HR_QS3). The second sub-factor assesses the level of qualification- requalification of the institution's human resources (items HR_QS2 and HR_QS4).

The fourth factor includes four statements relating to the Quality of Material Resource Management (MRM). It relates the features of physical installations and equipment. It has an eigenvalue of 3.1 and explains 9.13% of the variance.

The fifth factor explains 8.93% of the variance. It includes four items evaluating teaching-training methods and learning styles (TM_LS).

The sixth factor (ICT) includes four items in relation to the integration of NICT in the training devices. It has an eigenvalue of 2.67 and explains almost 7.85% of the variance.

The seventh factor is formed by four items, all dealing with the nature of students' attitudes and behaviours (ATT_B). It has an eigenvalue of 2.3 and explains 6.78% of the variance.

Finally, the eighth factor (ED), with approximately 5.17% of the variance explained and 1.76 of eigenvalue, relates to professional ethics and deontology. It includes four items. For clarity, the three items of this factor have been slightly reworded.

Table 3. Factor Matrix a of the QAS-UTD / STPSA Scale and its Statistical Characteristics

Item Code	8 factors explaining 77.93 %							
	F1 (COG)	**F2** (CUR)	**F3** (HR_QS)	**F4** (MRM)	**F5** (TM_LS)	**F6** (ICT)	**F7** (ATT_B)	**F8** (ED)
COG1	.931							
COG2	.922							
COG3	.921							
COG4	.911							
COG5	.854							
CUR1		.926						
CUR2		.912						
CUR3		.812						
CUR4		.780						
CUR5		.691						
HR_QS1			.941					
HR_QS2			.927					
HR_QS3			.886					
HR_QS4			.846					
MRM1				.969				

MRM2				.963				
MRM3				.937				
MRM4				.662				
TM_LS1					.881			
TM_LS2					.881			
TM_LS3					.869			
TM_LS4					.830			
ICT1						.892		
ICT2						.862		
ICT3						.846		
ICT4						.820		
ATT_B1							.890	
ATT_B2							.873	
ATT_B3							.868	
ATT_B4							.639	
ED1								.892
ED2								.846
ED3								.794
ED4								.774
Variance explained	18.737*%*	11.201*%*	10.117*%*	9.133 *%*	8.936*%*	7.854*%*	6.783*%*	5.173*%*
M & SD	4.01±2.01	4.38±1.96	4.09±1.97	4.24±1.96	5.48±1.70	4.94±1.61	4.99±1.68	3.23±1.99

a- Attempt to extract 8 factors. Convergence of the rotation in 6 iterations. (Convergence= .017).

Extraction method: Principal component analysis with maximum likelihood. Rotation method: Varimax with Kaiser normalization. Coefficients of items less than 0.40 are eliminated and not taken into account for our scale. M & SD: Mean and Standard Deviation.

B. Reliability analysis of the scale

Based on the results of the exploratory factor analysis, the new version of the scale has eight dimensions describing the concept of "quality of the STPSA training device". We examined the internal consistency and internal coherency of the eight subscales and then re-examined the same indices after removing 20 items with an asymmetry (Kurtosis) and flattening (Skewness) coefficients greater than ±1, and whose saturation coefficient is less than .4 (Thode Jr., 2002). The results are presented in Table 4. After eliminating the 20 items, the value of internal coherency, expressed by Cronbach's Alpha, increased from .719 to .850. Likewise, the value of the internal consistency, expressed by the correlation coefficient, improved from .592 to .803.

Thus, all the values of the eight subscales have improved greatly, exceeding the threshold of .70 set by Nunally (1978) and recommended by Plano Clark & Ivankova (2015) and Creswell (2013). In fact, some researchers (Field, 2005) argue that an alpha greater than or equal to .80 indicates good
internal coherency. The results of the scale's temporal stability showed a correlation coefficient of .732, which is considered sufficiently acceptable.

Table 4. Characteristics of the Internal Consistency and Internal Coherence of The Scale.

Subscales	**Code**	**Number of items**			**Correlation between the 2 parts [b]**		**Cronbach's Alpha [c]**	
		Raw items	Items removed [a]	Retained items	Befor e	After	Before	After
Conditions of organization and governance of the TD	**COG**	7	2	5	.700	.886	.740	.893
Curricula & programs	**CUR**	8	3	5	.558	.823	.711	.881
HR's Qualifications and Skills	**HR_Q S**	6	2	4	.784	.856	.720	.877
Material Resource Management	**MRM**	9	5	4	.606	.866	.714	.878
Teaching - training methods and learning	**TM_L S**	5	1	4	.660	.877	.707	.760

styles								
Integration of NICT in the TD	**ICT**	6	2	4	.481	.506	.656	.670
Student Attitudes and Behaviours	**ATT_B**	6	2	4	.583	.779	.501	.693
Ethics and deontology	**ED**	7	3	4	.509	.706	.607	.755
Scale	**QAS-UTD/ STPSA**	**54**	**20**	**34**	**.592**	**.803**	**.719**	**.850**

a. Items with an asymmetry (Kurtosis) and flattening (Skewness) coefficients greater than ±1, and whose saturation coefficient is less than .4, are removed. b. Internal consistency coefficient. c. Internal coherency coefficient.

C. Confirmatory factor analysis (CFA)

In order to validate the factor structure of the scale, we present here the findings of the confirmatory analysis. The results of the initial model led to the observation that it did not fully represent the data produced. Indeed, not all the suitability indices sufficiently meet the standards set: ($\chi2$ (69) = 938,543, p< .000 ; $\chi2$/DFL = 1.784; CFI= .848 ; RMSEA= .064).

As a result, modifications were made in terms of introducing error covariances between peer items on the same factor. These modifications appeared theoretically founded because each dimension of our scale is made up of items that characterize the one and same semantic field. In fact, all we had to do was add two terms of covariance between the errors, so that the suggested eight-factor model provides very satisfactory indices of adjustment: ($\chi2$ (79) = 817,930, p = .13 ; $\chi2$ /df = 1.585 ; CFI = .962 ; RMSEA = .047). The first term is Between the items (CUR 3) and (CUR 4) on the Curricula & programs factor. The second term is between (ED2) and (ED3) items on the factor "Ethics and deontology". Thus, out of the 28 correlations calculated between factors (Table 5), we obtained 12 coefficients that are significant, i.e., 42.86%, but with low intensity from .1 to .3.

Our results also showed that the model, thus readjusted, has a general implicit factor that we called the "Training Quality Index" (TQI). This significantly controls (p <.05) three of the eight factors in the model (Figure 1): Conditions of organization and governance of the training device (COG); Curricula & programs (CUR), and Qualifications and skills of human resources (HR_QS). They have respectively standardized regression coefficients (Table 5), with average intensity between .5 and .6 with this Training Quality Index.

Table 5. Correlation Matrix Between the Scale Factors (Expressed by Bravais-Pearson Coefficient) and The Training Quality Index (TQI) Expressed in Terms of Standardized Regression Coefficients.

Factors	COG	CUR	HR_QS	MRM	TM_LS	ICT	ATT_B	ED
Conditions of organization and governance of the training device (**COG**)	1							
Curricula & programs (**CUR)**	.27**	1						
HR's Qualifications and Skills (**HR_QS**)	.319*	.235*	1					
Material Resource Management (**MRM**)	.108*	.053	.111**	1				
Teaching-training methods and learning styles (**TM_LS**)	.084	.106	.071	-.01	1			
Integration of NICT in the TD (**ICT**)	.046	-.013	-.024	.105	.066	1		
Student Attitudes and Behaviours (**ATT_B**)	.04	.136*	.066	.138*	.284**	.035	1	
Ethics and deontology (**ED**)	.19**	.124*	.011	-.032	-.145*	-.069	.141*	1
Training Quality Index (**TQI**)	.616*	.457*	.502*	.180	.127	.033	.164	.171

: significant p<.05 ; **: significant p<.01 ; *: significant p<.001.*

These modifications also provided a significant improvement to the initial model (Δ $\chi2$ (10) = 644.64; p <.000). All coefficients in the model were found to be significant (p < .001). Under these conditions, we obtained the best fit of the model to the empirical data. Therefore, the aggregation of all these results confirms the multidimensional aspect of the QAS-UTD/STPSA scale.

Since the correlated eight-factor model has the best fit qualities, we will try to further improve it by adding covariance links between errors. In this regard, the AMOS software suggests that we correlate certain measurement errors, notably those of the items (CUR 3) and (CUR 4) for the factor "Curricula and programs"; and those of items (ED 2) and (ED 3) on the factor "Ethics and deontology". The final model is therefore presented in Figure 1.

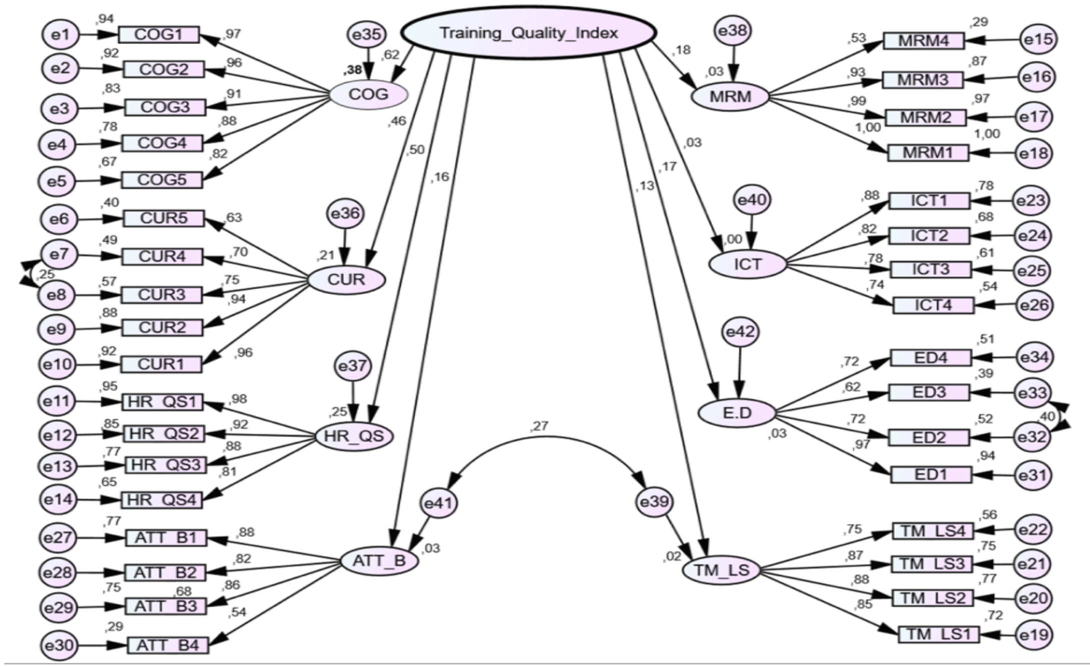

Fig. 1 Confirmatory Factor Analysis Model Generated.

Parameter estimates are standardized. All parameters and error terms are significant at p< .001.

D. Analysis of the construct validity of the scale

To verify the construct validity of the eight emerging dimensions, we identified for each factor of the scale the analysis models justifying their theoretical basis (Table 6).

Table 6. Analysis Of The Construct Validity of the Scale Through the Theories and Analysis Models.

N	Identified Factors	Theories and Models
1	Conditions of Organization and Governance of the TD (**COG**) (*Performance, Effectiveness - Efficiency)*	Quality approach and Continuous Improvement (AFNOR-ISO 9001 Standard / Version 2015; Chaize et al., 2018).
		Ergonomic approach of the TD (Bué, Coutrot & Puech, 2004).
		Student quality of life : (Felouzis, 2001; Vaez & Laflamme, 2008; Belgith et al., 2017).
2	Curricula and Programs (**CUR**)	Anglo-Saxon curicular model (Schiro, 2008 ; Jonnaert et al., 2009).
		Franco-European curricular model (Roegiers, 2012 ; Tessaro et al., 2017).
		Program theory (Chouinard, 2013; Develay, 2015).
3	HR's Qualifications and Skills (**HR_QS**)	Competency-based approach (Le Boterf, 2018; Wittorski, 2016; Zarifian, 2001).
4	Material Resource Management (**MRM**)	Resource Management Theory (Barabel, Meier & Teboul, 2013; Acedo, Barroso & Galan, 2006).
5	Teaching-training methods and learning styles (**TM_LS**)	Learning theories (Bruner, 1966 ; Kolb, 1985 ; Astolfi & Develay, 1991 ; Chartier, 2003).
6	Integration of NICT in the TD (**ICT**)	Model of ICT integration in education (Raby, 2005; Farrell and Shafika, 2007; Mastafi, 2013).
		Innovation approach (Karsenti et al., 2011; Maaroufi, 2016).
7	Student Attitudes and Behaviours (**ATT_B**)	Tripartite Model of Attitude: Cognitive Dimension, Affective Dimension, and Conative or Behavioural Dimension (Olson & Kendrick, 2008 ; Stone & Fernandez, 2008).
		Behavioural change theories (McGuire, 1985; Hejase, Al Kaakour, Halawi & Hejase, 2013).
		Commitment theories (Joule & Beauvois, 2013; Fointiat et al., 2013).

		Reasoned action theory (Hagger & Chatzisarantis, 2009).
		Planned Behaviour theory (Ajzen, 2005; Detroz,2014).
8	Ethics and deontology (**ED**)	Ethical theories: Objective-design (Deslandes, 2012); Subjective design (Jeffrey et al., 2008); Disciplinary design (Gohier, 2013); Incentive design (Naudet, 2011); Coercive design (Abbes, 2013); Deontological or normative design (Morin, 2004); Business ethic (Grunig, 2014).

VI. DISCUSSION

The objective of our study was to develop and validate a multidimensional scale for measuring and assessing the training devices quality in five Moroccan university centres that provide STPSA training courses. The relevance of our study lies in the dual fact that it covers several dimensions of the quality of university training devices on the one hand, and on the other hand that it addresses a branch that is often neglected and relegated to the background in this case, STPSA. Indeed, to our knowledge, no studies have yet been conducted to obtain proof of the validity and reliability of scales measuring the STPSA university training devices' quality.

The methodological approach adopted was based on the classical theory of scale validation, namely the Churchill paradigm. This model, initiated by Churchill in 1979 and supported by Bailey & Pearson (1983), is strongly recommended by several authors, such as

DeVellis (2003) and Roussel (2005). Our approach is also nourished by measurement theory, which has become increasingly codified (Brennan, 2006; Guilford, 1954; Thorndike, 1971; Bertrand & Blais, 2004). In the literature, few researchers have attempted to assess the quality of university training devices using this theoretical framework and to test the metrological properties of the scale used.

After completing the preliminary steps in this process, including item generation, expertise, and evaluation of the tool format, we proceeded to administer the initial version of the scale. A sample of 267 students from five university institutions with regulated access was mobilized. The size of the study sample is largely satisfactory since Hair et al. (2006) estimate that a satisfactory sample should consist of at least 200 students. This approach is used in the majority of field studies (Lagrosen et al., 2004; O'Neill and Palmer, 2004), as the main objective at this stage is to test the proposed scale on a sample rather than generalize the results to the population. From a methodological point of view, the approach taken during our study is in line with the current trend in the construction and validation processes for measuring instruments (Borsboom & Markus, 2013; Newton & Shaw, 2014; Laveault & Grégoire, 2014). However, this approach cannot claim to have highlighted all the potential determinants of the quality of STPSA training devices: some of them, which are particularly subjective and difficult to verbalize and measure, may have resisted the protocol of generation and evaluation of the item.

We verified the main metrological qualities of the QAS-UTD/STPSA through reliability and validity indices (Fortin, 2010; Newton & Shaw, 2014). In this regard, exploratory factor analysis (EFA) allowed us to identify the factor structure of our scale. It revealed eight major factors, whose eigenvalue is greater than 1, and which explain 77.93% of the total variance very satisfactorily. What constitutes a good value in psychometry (Hair et al., 2006). In terms of semantic coherence, the items having weak correlations with their respective factors, and providing little explanation, were automatically deleted. The KMO index reached a very satisfactory value of .796 for a recommended threshold of (.70). Likewise, the results of the temporal stability of the scale showed a correlation coefficient of .73, which is considered sufficiently acceptable.

The answer to the question regarding the ability of the scale to measure the quality of the STPSA university training devices is provided by the analyses of internal consistency and internal coherency. Indeed, the internal coherency of the scale, expressed by Cronbach's Alpha, is .850. Cronbach's alpha values above .70 are considered good (Nunally, 1978). Similarly, after eliminating non-homogeneous items, the correlation coefficient showing the level of internal consistency of the scale reached a fairly high value of .803.

Therefore, the overall results of these metric tests are a good guarantee of the validity and reliability of our scale. The QAS-UTD/STPSA measurement scale demonstrates good psychometric properties. The results of the confirmatory factor analysis allowed us to readjust and significantly improve our built model ($\Delta \chi 2$ (10) = 644.64; $p< .000$). We were also able to extract a general implicit factor, which we called the "Training Quality Index" (TQI). This controls, only but significantly ($p<.05$), three factors of our scale. These confirmatory analyses argue that the items can be classified into eight conceptually distinct dimensions. They reflect different aspects of the quality of STPSA training devices. This thus confirms the hypothesis of the multidimensionality of the STPSA training devices Quality.

Our study identified eight complementary dimensions. Each of them is fed by referenced theoretical models (Cronin & Taylor, 1992) and is more or less corroborated by the work of other researchers.

It allowed us to specify more dimensions of quality, until now little or not reported by the measuring instruments, such as NICT, ethics, organizational and governance conditions of the training device, attitudes and behaviours students, and management of material resources. These dimensions appear particularly important for the students requested and are strongly recommended by certain authors.

Indeed, if the NICT component has long been ignored in assessing the quality of the training system, then Mastafi (2013) and Farrell & Shafika (2007) have stressed its importance in any assessment process. They attribute to them a dual role of innovation and inclusion. Other researchers have demonstrated the positive effects of digital, multimedia, and audiovisual technologies in STPSA on student motivation and commitment on the one hand (Gaudin et al., 2013) and on improving the student learning quality on the other (Casey et al., 2016).

Deslandes (2012) and Morin (2004) consider the ethical dimension as a transversal skill required for any TD and which is acquired during training. For Rondeau (2003), ethics is a tool for school and university democratization. In the STAPS training system, a set of attributes (values, codes, beliefs, standards, representations, and attitudes) constitute the ethical guidelines or what is commonly known as sports ethics (Morgan, 2007). Moreover, the emergence of ethics committees and ethics charters within sports organizations greatly reflects the interest of this dimension in the moralization of both sports practice and student life.

While other authors integrate the ergonomic variable (Paquay et al., 2014) and that relating to the quality of student life (Felouzis, 2001; Vaez & Laflamme, 2008; Belgith et al., 2017) as two contextual sub-factors corresponding to the technical conditions of the training system. This is partially similar to the first dimension identified by our scale, namely "organizational and governance conditions of the TD".

As regards the revealed dimension "Student attitudes and behaviours", it is fed by theoretical models of commitment on the one hand (Eagly & Chaiken, 2007) and by those of behavioural change (Hejase, Al Kaakour, Halawi & Hejase, 2013) on the other hand.

For Joule & Beauvois (2013), students' commitment depends on the product of their degree of motivation and their expectations of the training course. Moreover, several social psychology studies show that the degree of commitment and perseverance in a training cycle is positively associated with the student's well-being, accommodation, catering, and transportation conditions (Brault-Labbé & Dubé, 2008; Jodoin, 2000). Thus, according to Fointiat et al. (2013) and Younès (2015), soliciting optimal levels of student commitment would allow them to consciously change their behaviours and adopt attitudes favourable to success.

Furthermore, by identifying the "Material Resource Management" dimension, our scale (QAS- UTD/STPSA) joins the first dimension of the SERVQUAL instrument (Parasuraman et al., 1988) respectively, and the third dimension of the CERM-CSQ scale (Center for environmental and leisure management - Customer Service Quality) of Howat, Absher, Crilley, and Milne (1996). The first refers to tangible elements or material assets, and the third relates to "general facilities". These two factors ("tangibility" and "General facilities") describe the physical, material, equipment, and logistics properties of the training device. While the other factors of SERVQUAL and CERM-CSQ refer to the intangible and immaterial aspects.

Anyway, our QAS-UTD/STPSA scale has at least two strengths:

- The most important dimension for any university training device, namely "Programs & Curricula", is not found in the SERVQUAL instrument;
- The "HR's Qualifications & skills" dimension can easily replace and encompass four SERVQUAL dimensions: Reliability, Responsiveness, Assurance, and Empathy.

Nevertheless, our study - being, of course, not exhaustive - would have certain limitations which are worth noting. Firstly, we cannot claim to have highlighted all the variables of the quality of the STPSA training device. Some hidden dimensions can always be identified and exploited, by other measurement instruments, to complete the description of the TD quality.

Others are particularly difficult to assess due to their sensitivity and inaccessibility, such as the financial, accounting, employability, and vocational integration dimensions. This opens the door to numerous avenues of research, which will make it possible to examine more factors predicting the quality of training devices.

As a corollary, the sampling method chosen is both simple random and voluntary (Satin & Shastry, 1993; Laveaux & Grégoire, 2014). Our study was conducted with STPSA students only for reasons of feasibility and accessibility. It would therefore be important to obtain evidence of the validity of this scale from samples belonging to other disciplines and universities.

Furthermore, our instrument measures, not an objective quality but subjective quality as perceived by students regarding their training devices (Boyer & Nefzi, 2009). Thus, the eight factors were assessed according to the culture and social representation of our respondents.

Hence the emergence of different perceptions (Zeithmal et al., 2013) and approximate judgments (Bouffard et al., 2013), which more or less influence the validity of our scale. These differences in perceptions and representations have been evoked and discussed by other researchers (Ardouin, 2006).

Consequently, the QAS-UTD/STPSA does not aim to replace the few conventional measurement instruments used to assess the quality of university training devices (UTD), but it attempts to provide further explanation of the determinants of UTD quality in STPSA. Triangulated with other tools, it could lead to useful applications in academia.

VII. CONCLUSION

At the end of our research, we confirm our initial assumption that the quality of a UTD is as difficult to define as to measure. Measuring it first requires an assessment of all its potential components. The QAS-UTD/STAPS, therefore, has a multifactorial and heterogeneous structure. It's statistically robust and has fairly good psychometrical and metrological features. It has a good internal consistency, and its temporal stability is sufficiently acceptable. It consists of 34 items divided into eight “predictor” factors of the UTD quality in STPSA. It explains 71.43% of the total variance very satisfactorily and has an eigenvalue greater than 1. We were able to extract from it a general implicit factor called the

"Training Quality Index" (TQI). This controls, only but significantly (p<.05), three factors of our scale.

Overall, the QAS-UTD/STPSA scale has at least the merit of jointly covering eight different dimensions, which can satisfactorily assess the quality of UTD in STPSA. It explores parameters little or not taken into account by the usual instruments, or they cover them separately and never together or in interaction with other dimensions.

REFERENCES

[1] Abbes, N., L'entreprise responsable : de la responsabilité sociétale à la communication environnementale, Ed. L'Harmattan, Paris., (2013).

[2] Arbuckle J., IBM SPSS Amos 4.0 User's Guide. Amos Development Corporation., (2013).

[3] Acedo F.J., Barroso C., Galan J.L., The resource-based theory: dissemination and main trends, Strategic Management Journal, 27(7) (2006) 621-636.

[4] AFNOR ., ISO 9001 Standard, V 2015 : vers un nouveau système de management de la qualité. Une norme révisée et adaptée au contexte économique actuel. 8M Management., (2015).

[5] Ajzen, I., Attitudes, personality, and behaviour (2e éd.). Maidenhead, England New York: Open University Press., (2005).

[6] American Educational Research Association, American Psychological Association & National Council of Measurement in Education, Standards for Educational and Psychological Testing. Washington, DC: AERA., (2014).

[7] Arcaro, J. S., Quality in education: An implementation handbook, Delray Beach, FL: St. Lucie Press., (1995).

[8] Ardouin, T., Ingénierie de la formation pour l'entreprise, 2ème édition, Dunod, (2006).

[9] Aspisi, A., « Mobiliser sur la qualité », in L'Expression d'Entreprise, Aout , (1995) 45.

[10] Astolfi J.-P. & Develay M., La didactique des sciences, Presses Universitaires de France, Paris, (1991).

[11] Babakus E. et Boller G. W., An Empirical Assessment of the SERVQUAL Scale, Journal of Business Research, 24 (1992) 253-268.

[12] Bailey, J. E.; Pearson, S. W., Development of a Tool for Measuring and Analyzing Computer User Satisfaction, Management Science, 29 (5) (1983) 530-545.

[13] Barabel, M., Meier, O. et Teboul, T., Les fondamentaux du Management., Collection Management sup – Management, ressources humaines Ed. Dunod, 2ème édition, (2013) 228

[14] Bardin, L, L'Analyse de contenu. Paris : Presses Universitaires de France ,(2007).

[15] Beamish, W., Consensus about program quality: An Australian study in early childhood special education, Brisbane, Australia: Griffith University, (2004).

[16] Belgith, F., Giret, J.-F., Ronzeau, M., & Tenret, E., Panorama 2016. Conditions de vie des étudiants. OVE infos, n°34, (2017).

[17] Bentler, P.M., Comparative Fit Indexes in Structural Models," Psychological Bulletin, 107 (2) (1990) 238-246.

[18] Benzécri, J.P. and al., : l'Analyse de Données. 1 (1973) Dunod, Paris.

[19] Bernier, J.J., Pietrulewicz, B., La psychométrie. Traité de mesure appliquée. Montréal. Gaëtan Morin éditeur., (1997).

[20] Bertrand, R., & Blais, J., Modèles de mesure : l'apport de la théorie des réponses aux items. Ste -Foy, Québec : Presses de l'Université du Québec., (2004).

[21] Bigras, N., Lemire, J. & Eryasa, J., Analyse comparative de trois instruments d'observation de la qualité (ECERS-R, EOQÉ et CLASS Pre-K) des services de garde à la petite enfance. Les Dossiers des sciences de l'éducation, 37 (2017) 35-54.

[22] Bollen, K. et Long, J., Testing Structural Equation Models, Sage., (1993).

[23] Borsboom, D., & Markus, K. A, Truth and evidence invalidity theory, Journal of Educational Measurement, 50(1) (2013) 110-114.

[24] Bouchard, C., Plante, J., La qualité : mieux la définir pour mieux la mesurer. Les Cahiers du Service de Pédagogie expérimentale (Université de Liège), 11-12 (2002) 219-236.

[25] Bouchard, C. et J. Plante , « La qualité : mieux la définir pour mieux la mesurer », Les Cahiers du Service de Pédagogie Expérimentale, (11/12), Service de pédagogie expérimentale de l'université de Liège, (2003) 219-236.

[26] Bouffard, T., Pansu, P. & Boissicat, N., Quand se juger meilleur ou moins bon qu'il ne l'est s'avère nuisible ou profitable à l'élève." Revue française de pédagogie, 182 (2013) 117-140.

[27] Bourgon, L. & Lavallée, C., Échelle d'observation de la qualité éducative : les services de garde en installation pour les enfants de 18 mois ou plus. Québec, QC : Gouvernement du Québec., (2004).

[28] Bourque, J., Poulin, N. et Cleaver, A.F., Évaluation de l'utilisation et de la présentation des résultats d'analyses factorielles et d'analyses en composantes principales en education, Revue des sciences de l'éducation, 32(2) (2006) 325-344.

[29] Boyer A. & Nefzi A., « La perception de la qualité dans le domaine des services : vers une clarification des concepts », La revue des Sciences de Gestion, Direction et Gestion, n° 237-238 (2009) 43-54.

[30] Brault-Labbé, A. & Dubé, L., Engagement, surengagement et sous-engagement académiques au collégial : pour mieux comprendre le bien-être des étudiants. Revue des sciences de l'éducation, 34 (3) (2008) 729–751.

[31] Brennan, R. L., Educational Measurement (4th ed.). Westport, CT: Praeger Cizek, G. J., & Bunch, M. B. (2007). A guide to establishing and evaluating performance standards on tests. Thousand Oaks, CA: sage Publications.

[32] Browne, M. W., & Cudeck, R., Alternative ways of assessing model fit, In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models, (1993) 136-162. Newbury Park, CA: Sage.

[33] Bruner, J.S.., Towards a Theory of Instruction, New-York, W.W Norton, (1966).

[34] Bué,J., Coutrot, T., Puech, I., Conditions de travail : les enseignements de vingt ans d'enquêtes, Toulouse: Octarès, (2004) 166 .

[35] Carman J. M., Consumer Perceptions of Service Quality: An Assessment of SERVQUAL, Dimensions, Journal of Retailing, (1990) 33-55.

[36] Casey, A., Goodyear, V.A., & Armour, K.M., Digital technologies and learning in physical education: Pedagogical cases. London; Routledge, (2016).

[37] Chaize, J., Fiume, O., Ballé, M. & Jones. D.., The Lean Strategy – Using Lean to Create Competitive Advantage, Unleash Innovation, and Deliver Sustainable Growth, McGraw-Hill Education, (2018).

[38] Chartier, D. ,Les styles d'apprentissage entre flou conceptuel et intérêt pratique, Savoirs, 2 (2003) 9-28.

[39] Chelladurai, P, Scott, FL & Haywood-Farmer, J., Dimensions of fitness services: development of a model, Journal of Sport Management. 1(2) (1987) 159-172.

[40] Chouinard, J.A., The case for participatory evaluation in an era of accountability, American Journal of Evaluation, 34 (2) (2013) 237 – 253.

[41] Churchill, G.A., A paradigm for developing better measures of marketing constructs, Journal of Marketing Research, 16(1) (1979) 64-73.

[42] Costello, A.B. et Osborne, J.W., Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. Practical Assessment, Research and Evaluation, 10(7) (2005).

[43] Creswell, J. W., & Plano Clark, V.L., Designing and Conducting Mixed Methods Research (3rd ed.) Thousand Oaks, CA: Sage Publications, (2018).

[44] Creswell, J. W., Research design: Qualitative, quantitative, and mixed methods approach, Thousand Oaks, CA: SAGE Publications, (2013).

[45] Cronbach, L. J., Coefficient alpha and the internal structure of tests". Psychometrika, 16(3) (1951) 297-335.

[46] Cronin J. J. et Taylor S. A., Measuring Service Quality: A Reexamination and Extension, Journal of Marketing, 56 (1992) 55-68.

[47] Crosby, P., Quality Is Free, McGraw-Hill, New York, NY., (1979).

[48] De Ketele, J.-M., Évaluation des systèmes de formation. Dans A. Jorro (dir.), Dictionnaire des concepts de la professionnalisation, (2014) 117-120. Louvain-la-Neuve : De Boeck Supérieur.

[49] Deslandes G., Le management éthique, Editions Dunod, Paris, (2012).

[50] Detroz, P. « Vers une démarche balisant la mise en place (ou la refonte) d'un dispositif d'évaluation des enseignements ». In Actes

du 28e Congrès de l'Association internationale de pédagogie universitaire (AIPU), Mons, 19-22 mai 2014. Mons : Université de Mons., (2014).
[51] Develay, M., D'un programme de connaissances à un curriculum de compétences. Louvainla-Neuve, Belgique : De Boeck Supérieur, (2015).
[52] De Vellis, R. F., Scale development: theory and applications (2nd ed. Vol. 26). Thousand Oaks, CA: Sage Publications, (2003).
[53] Doherty-Delorme, D., & Shaker, E., Missing pieces II. An alternative guide to Canadian post-secondary education 2000/2001 provincial rankings: Where do the provinces stand on education? Canadian Centre for Policy alternative, (2001).
[54] Endrizzi, L., La qualité de l'enseignement : un engagement des établissements, avec les étudiants ? Dossier de veille de l'"IFE, (93) (2014) 1-44.
[55] Farrell, G. et Shafika, I., Survey of ICT and education in Africa: A summary report, based on 53 country surveys, Enquête sur les TIC et l'Éducation en Afrique. Washington DC : infoDev / World Bank., (2007)
[56] Felouzis, G., La Condition Etudiante. Sociologie des étudiants et de l'Université Collection Sociologie d'Aujourd'hui. Presses Universitaires de France, (2001).
[57] Field, A, Discovering Statistics Using SPSS (2nd ed.). Thousand Oaks, CA: Sage Publications, (2005).
[58] Figari, G, L'évaluation des dispositifs Educatifs. Mesure et Evaluation en Education, 31(3) (2008) 77-93.
[59] Fointiat, V., Girandola, F., & Gosling, P., La Dissonance cognitive. Quand les actes changent les idées, Paris : Armand Colin., (2013).
[60] Fortin, M. F., Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives (2ème éd.)". Montréal, Canada : Chenelière Éducation., (2010).
[61] Gaudin, C., Tribet, H., & Chaliès, S., Analyse de l'activité d'enseignants novices dans un dispositif de formation professionnelle par alternance exploitant la vidéo : de l'observation vidéo à la pratique de classe. Congrès de l'Actualité de la recherche en éducation et en formation (AREF). Université de Montpellier : (2013) 27-30 août.
[62] Gerard, F.-M., Hugonnier, B. & Varin, S., Mesure de la qualité des systèmes éducatifs des pays de l'OCDE, in ADMEE-Europe, L'évaluation en éducation et en formation face aux transformations des sociétés contemporaines, Actes du colloque, Esch-sur-Alzette : Université de Luxembourg, (2018) 131-143.
[63] Gerbing D.W. and Anderson, J.C., Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach, Psychological Bulletin, 103 (1988) 411-423.
[64] Gohier, C., Éthique et formation : discussion sur les vertus ou vertus de la discussion? Formation et profession, 21(3) (2013) 70-80
[65] Gorsuch, R, Factor analysis, Hillsdale, NJ : L. Erlbaum Associates, (1983).
[66] Grönroos, C., Service Quality: The six criteria of good perceived service, Review of Business, 9(3) (1988) 10-13.
[67] Grunig, J. E., Introduction. Ethics problems as theories in public relations, Revue internationale Communication sociale et publique, (2014).
[68] Guilford, J. P., Psychometric Methods (2nd ed.). New York, NY: McGraw-Hill, (1954).
[69] Guttman, L., «A Note on Sir Cyril Burt's Factorial Analysis of Qualitative Data», British Journal of Statistical Psychology, 6 (1954) 21-4.
[70] Hagger, M. S., & Chatzisarantis, N. L., Integrating the theory of planned behaviour and self-determination theory in health behaviour: A meta-analysis, British Journal of Health Psychology, 14 (2009) 275-302.
[71] Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. Multivariate data analysis (6th ed.), Upper Saddle River, N.J.: Pearson Prentice Hall., (2006).
[72] Hannon, V. et Peterson A., Thrive: Schools Reinvented for the real challenges we face, Innovation Unit Press, London., (2017).
[73] Hejase, A. J., Al Kaakour, R. S., Halawi, L. A., & Hejase, H. J., Students' Perceptions of Student Evaluation of Teaching (SET) Process, International Journal of Social Sciences and Education, 3(3) (2013)
[74] Hendrickson, S., Martin, T., Manuel, J., Christopher, P., Thiedeman T., & Moyers, T. B., Assessing Reliability of the Motivational Interviewing Treatment Integrity Behavioral Coding System Under Limited RangeQ, Poster presented at the 27th Annual Convention of the Research Society on Alcoholism, Vancouver, BC, (2004).
[75] Hersan, C., Vade-mecum assurance qualité, Lavoisier Tec & Doc, 3e édition, , (1999) 254
[76] Howat, G., Absher, J., Crilley, G., & Milne, I., Measuring customer service quality in sports and leisure centres. Managing Leisure, 1 (1996) 77-89.
[77] Jeffrey, D., Harvengt, D., Vachon, M.-C. et Deschesnes, G. Le droit et l'éthique dans la profession enseignante. Dans F. Jutras et C. Gohier (dir.), Repères pour l'éthique professionnelle des enseignants , (2008) 75-92. Québec, QC : Presses de l'Université du Québec.
[78] Jodoin M., Engagement et surengagement au travail : composantes psychologiques, antécédents potentiels et association avec le bien-être personnel. Thèse de doctorat inedited, Université de Montréal, Montréal, Canada, (2000).
[79] Jonnaert, P., Ettayebi, M., Defise R., Curriculum et Compétences. Un cadre opérationnel. Bruxelles : Editions De Boeck Université, Coll. « Perspectives en éducation & Formation ». , (2009).
[80] Jöreskog, K.G., & Sörbom, D., LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods, Mooresville, IN: Scientific Software., (1984).
[81] Joule, R-V. & Beauvois, J-L., Petit traité de manipulation à l'usage des honnêtes gens. Grenoble: Presses Universitaires de Grenoble, (2013).
[82] Juran, J., Quality planning and analysis., Maidenhead, uK: McGraw-hill, (1993).
[83] Karsenti, T., Komis, V., Depover, C., et Collin, S., Les TIC comme outils de recherche en sciences de l'éducation. Dans T. Karsenti et L. Savoie-Zajc (dir.), La recherche en éducation : étapes et approches , (2011) 168-192. SaintLaurent : ERPI.
[84] Kieffer, K.M., Orthogonal versus oblique factor rotation: A review of the literature regarding the pros and cons. Communication présentée dans le cadre du Annual Meeting of the Mid-South Educational Research Association, Nouvelle-Orléans, LA., (1998).
[85] Kim, D. & Kim, S., QUESC: An instrument for Assessing the Service Quality of Sports Centers in Korea, Journal of Sport Management. 9(2) (1995) 208-220.
[86] Kolb D. A., The learning-style inventory: self-scoring inventory and interpretation booklet, Boston, Mass., McBer, and Company, (1985).
[87] Lagrosen, S., Sayyed- Hashemi, R. & Leitner, M., Examination of the dimension of quality in higher education Quality Assurance in Education, 12(2) (2004) 61-69.
[88] Laveault, D. & Grégoire, J. Introduction aux théories des tests en psychologie et en sciences de l'éducation, Coll. Méthodes en sciences humaines, 3ème édition, De Boeck., (2014)
[89] Le Boterf, G., Développer et Mettre en œuvre la Compétence. Comment investir dans le Professionnalisme et les Compétences. Paris Ed. Eyrolles, (2018).
[90] Likert, R., « A Technique for the Measurement of Attitudes », Archives of Psychology, 140 (1932) 1-.55
[91] Maaroufi, F., Effets des TIC sur les pratiques pédagogiques dans un établissement d'enseignement supérieur marocain., (2016).
[92] Mastafi, M., Intégration et usages des TIC dans le système éducatif marocain : Attitudes des enseignants de l'enseignement primaire et secondaire. Revue Adjectif., (2013).
[93] Mayer, R. & Ouellet, F., Méthodologie de recherche pour les intervenants sociaux, Boucherville, Gaëtan Morin Éditeur, (1991) 537 .
[94] McGuire, W. J., Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology, 2 (1985) 233Ð346. New York: Random House
[95] Meignant, A., Manager la formation. Un nouveau souffle après la réforme de ? L'ouvrage de référence des professionnels. Editions Liaisons, (2014).
[96] Morgan, d. l., Focus groups as qualitative research. Thousand oaks, CA: Sage, (1997).
[97] Morgan W.J., Ethics in sport, Champaign, Human Kinetics, (2007).
[98] Morin, E., Ethique, La méthode 6, Paris, Le Seuil, (2004) 22-25.

[99] Moss, P., Section I: Theoretical examinations of quality: Making the narrative of a quality stutter, Early Education and Development, 16(4) (2005) 405-420.
[100] Mottier Lopez, L., Lenzen, B., Leutenegger, F., Ronveaux, C., Évaluation des enseignements de formations continues diplômantes : enjeux de reconnaissance et de collaboration entre les parties prenantes. Phronesis, (2018) 7.
[101] Narang, R.., How do management students perceive the quality of education in public institutions? Quality Assurance in Education, 20(4) (2012) 357-371.
[102] Naudet J.Y.,L'éthique de l'entreprise n'existe pas », Bulletin de L'ALEPS, (2011).
[103] Neuville, S. & Frenay, M., « La persévérance des étudiants de 1er baccalauréat à la lumière du modèle expectancy-value ». In C. Michaut & M. Romainville (dir.), Réussite, échec et abandon dans l'enseignement supérieur [sous presse]. Bruxelles : De Boeck., (2010).
[104] Newton, P., & Shaw, S., Validity in educational and psychological assessment, London: Sage.,(2014).
[105] Nunnally, Jum C., Psychometric Theory, New-York, NY, McGraw-Hill, (1978).
[106] Olson, M. A., et Kendrick, R. V., Origins of Attitudes. Dans D. W. Crano et R. Prislin (Dirs.) Attitudes and Attitude Change, (2008) 111-130. New York (NY, US): Taylor and Francis Group.
[107] O'Neill, M.; Palmer, A, Importance-performance analysis: a useful tool for directing continuous quality improvement in higher education, Quality Assurance in Education. 12(1) (2004) 39-52.
[108] Paillé, P. & Mucchielli, A., L'analyse qualitative en sciences humaines et sociales, Paris, Armand Colin, 3ème éd, (2012).
[109] Papadimitriou, D. & Karteliotis, K., The Service Quality Expectations in Private Sport and Fitness Centers: A re-examination of the factor structure, Sports Marketing Quarterly. 9(3) (2000)157-164.
[110] Parasuraman, A., Zeithaml, V.A. and Berry, L.L., A Conceptual Model of Service Quality and its Implications for Future Research, Journal of Marketing, 49 (1985) 41-50.
[111] Parasuraman, A., Zeithaml, V.A. and Berry, L.L., SERVQUAL: a Multiple Item Scale for Measuring Consumer Perceptions of service quality, Journal of Retailing, 64(1) (1988) 12-40.
[112] Parasuraman, A., Zeithaml, V.A. and Berry, L.L., Refinement and reassessment of the Servqual Scale, Journal of Retailing, 67(4) (1991) 420-50.
[113] Peeters, H., & charlier, P., Contributions à une théorie du dispositif. Hermès, 25 (1999)15-23.
[114] Pianta, RC, La Paro, K., Hamre, BK , Système de notation d'évaluation en classe (CLASS). Baltimore, MD : Brookes., (2008)
[115] Plano Clark, V. L., & Ivankova, N. V., Mixed methods research: A guide to the field. Thousand Oaks, CA: SAGE Publications., (2015).
[116] Poole, B., Quality, semantics, and the two cultures. Quality Assurance in education, 18(1) (2010) 6-18.
[117] Postiaux, N., & Salcin, A. Au croisement de l'évaluation de l'enseignement et de l'évaluation de programme : les étudiants finissant évaluent leur apprentissage à partir d'un référentiel de compétences. In M. Romainville & C. Coggi (Eds.), L'évaluation de l'enseignement par les étudiants - Approches critiques et pratiques innovantes. Bruxelles : De Boeck Universités., (2009).
[118] Raby, C., Le processus d'intégration des technologies de l'information et de la communication. In T. Karsenti et F. Larose (dir.), L'intégration pédagogique des TIC dans le travail enseignant : recherches et pratiques , (2005) 79-95. Sainte-Foy, Québec : Presses de l'Université du Québec.
[119] Reckase, M. D., Multidimensional item response theory. Statistics for Social and Behavioral Sciences, Michigan, United States: Springer., (2009).
[120] Rege Colet, N., « L'évaluation de l'enseignement au cœur des processus d'assurance qualité : l'arbre qui cache la forêt », in M. Romainville & C. Coggi (dir.), L'évaluation de l'enseignement par les étudiants, Bruxelles : De Boeck, (2009) 235-254.Roegiers, X. & coll. (2010). Des curricula pour la formation professionnelle initiale. Bruxelles : De Boeck.
[121] Roegiers, Xavier., Quelles réformes pédagogiques pour l'enseignement supérieur ? Préface par Nicole Rege Colet. Collaborateurs : Mohamed Miled, Ioan Ratziu, Caroline Letor, Richard Etienne, Gaëlle Hubert, Mohamed Dali. Bruxelles, Ed. De Boeck, (2012)313 .
[122] Rondeau, D., « La délibération éthique, un outil pour la démocratie scolaire », Le point en administration scolaire, (2003) 18-19.
[123] Roussel, P., Méthodes de développement d'échelles pour questionnaires d'enquête, In P. Roussel et F. Wacheux, Management des Ressources Humaines : Méthodes de Recherche en Sciences Humaines et Sociales, Bruxelles : De Boeck, Chapitre, 9 (2005) 245-276.
[124] Rust, R. T., & Oliver, R. L., Service Quality: Insights and managerial implications from the frontier In R. T. Rust & R. L. Oliver (Eds.), Service Quality: New directions in theory and practice, (1994) 1-19. Thousand Oaks, CA: Sage Publication.
[125] Satin, A., et Shastry, W., L'échantillonnage : un guide non mathématique. Ottawa : Statistique Canada, Ministère de l'Industrie, des Sciences et de la Technologie, (1993).
[126] Schiro, M.S., Curriculum Theory. Conflicting Visions and Enduring. London: Sage Publications, (2008).
[127] Silvia, E.S.M., & MacCallum, R.C., Some factors affect the success of specification searches in covariance structure modelling. Multivariate Behavioral Research, 23 (1988) 297-326.
[128] Simard, Y., L'efficacité de la formation à distance au niveau postsecondaire : une méga-analyse (Thèse de doctorat non publiée). Université Laval, Québec, (2018).
[129] Stake, R., & Schwandt, T., On discerning Quality in Evaluation. In I. F. Shaw, J. C. Greene, & M. M. Mark (eds), Handbook of Evaluation. London: sage, (2006)
[130] Steiger, J.H., Structural model evaluation and modification, Multivariate Behavioral Research, 25 (1990) 214-12.
[131] Stone, J., et Fernandez, N. C., How Behavior Shapes Attitudes. Cognitive Dissonance Processes. Dans D. W. Crano et R. Prislin (Dirs.), Attitudes and Attitude Chang, (2008) 313-334. New York (NY, US): Taylor and Francis Group.
[132] Tempus., Glossary of the terms related to quality assurance. Development of Quality Assurance System in Higher Education (QUASYS) Tempus Joint European Project, UM JEP16015-2001, (2001).
[133] Tessaro, W., Gerard. F.-M. & Giglio, M., Changements curriculaires : un levier pour les pratiques évaluatives des enseignants ? Evaluer, Journal international de Recherche en Education et Formation, 3(1-2) (2017) 51-60.
[134] Thode Jr., H.C., Testing for Normality, Marcel Dekker, New York., (2002).
[135] Thorndike, R. L., Educational Measurement (2nd ed.). Washington, DC: American Council on Education, (1971).
[136] Vaez, M. & Laflamme, L. Experienced stress, psychological symptoms, self-rated health, and academic achievement: a longitudinal study of a Swedish university student, Social behaviour and personality, 36(2) (2008) 183-195.
[137] Vinokur, A., La normalisation de l'université, in Gouverner par les standards et les indicateurs : De Hume aux rankings, (2013) 235-261, Bruxelles.
[138] Weisser, M., « Dispositif didactique ? Dispositif pédagogique ? Situations d'apprentissage ! », Questions Vives, 4(13) (2010).
[139] Wittorski, R., La professionnalisation en formation : textes fondamentaux. Mont-Saint-Aignan, France : Presses universitaires de Rouen et de Le Havre, (2016).
[140] Younes, N., L'évaluation de l'enseignement par les étudiants : le tournant du partage. Evaluer- Journal international de recherche en éducation et formation, 1 (2015) 79-90.
[141] Zarifian P., Le modèle de compétence. Trajectoire historique, enjeux actuels et propositions. Malmaison: Éditions Liaisons, (2001).
[142] Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M., Quality, dosage, thresholds, and features in early childhood settings: A review of the literature, (OPRE 2011-5), (2010).
[143] Zeithaml, Valarie A., Bitner, Mary Jo and Dwayne D. Gremler. Services Marketing. Integrating Customer Focus Across the Firm, 6th Edition, New York: McGraw-Hill Irwin, (2013).