# Approaches and Tools for Quality Measurement Scales Validation in Education: An Initial Systematic Literature Review Collection

https://doi.org/10.3991/ijet.v17i16.31685

Hicham Berbar[1](✉), Said Lotfi[2], Mohammed Talbi[1]
[1] Faculty of sciences Ben M'sik, Hassan II University of Casablanca, Casablanca, Morocco
[2] Normal Superior School, Hassan II University of Casablanca, Casablanca, Morocco
hichamberbar00@gmail.com

**Abstract—**Studies in psychometrics related to the construction and validation of measurement scales have proved their reliability and construct validity in different domains. This systematic literature review study on instruments for measuring quality and indicators has two main objectives: Identifying research construction of validated measurement instruments and building an initial and predictive database for researchers and practitioners. The study gave satisfying results on different domains that reported important methodological features. This analysis is not a macro-analysis but an overview of measurement scales in psychometrics that have strengths and weaknesses. Further research is needed to characterize and explore the advantages and disadvantages of each procedure.

**Keywords—**psychometrics studies, tools, measurement, quality

## 1 Introduction

Quality improvement is a significant area of concern in different sectors in every significant problem for performance. The teaching and education sector is constantly committed to improving quality in small entities such as schools. Performance has become an inescapable imperative in the reform and management of the school. According to a national referential ambition is to raise the quality to lead to an attractive school. Thus, it is necessary to think first about the quality of school as the core of the educational system. Such a strategic choice is reflected in the adoption by the Ministries of National Education of countries. Especially**,** those in the development process, reform strategies, renovation of teaching, training, and management professions, which remains a first prerequisite for improving quality in schools[1]. On the other hand, constructing an evaluation system of indicators to test the quality remains a continuous challenge[2]–[5].

It is unconceivable to imagine adapting a quality model from one country to another country or academic system. In order to achieve this, it is essential to select the key parameters that can fit into the national context and culture and the modalities characterizing the federal system [6]–[8]. The primary concern of all developing countries is

to build on the models and quality systems already in place and tested, which offers excellent opportunities for them. However, the question that arises concerning this transfer of quality approaches to such a complex sector as schools, and regardless of the level of development of the country, comes up against several challenges and constraints, relating to national identity and cultural and religious sensitivity, as shown by [9].

The construction of a quality evaluation system based on indicators that would allow quality to be witnessed, it is highly desirable by many educational systems to optimize and validate a quality system consistent with hands that give a meaningful representation, based on rigorous scientific research, evaluating the quality of school[10], [11].

The lack of consensus on the definition of quality in schools and the relevance of a measurement scale capable of extracting its different facets. The literature review we consulted highlighted the importance and the growing interest and proliferation of research on the subject [12], [13]. The authors who were interested in this kind of topic**,** **t**hey proposed an extensive literature review but highlighted a useful starting point of knowledge. The researcher can build to readapt them to new contexts specific to the research objects conducted. The dimensions of the SERVQUAL approach are used to compare them with those taken from the qualitative exploratory study conducted[13]. However, the lack of a comprehensive and systematic study on the psychometric properties of high school quality measurement scales**,** as a carrier medium**,** will prevent practitioners from effectively basing and using an appropriate instrument for school quality assessment. Despite the increase of studies on school quality and its dimensions and related indicators, they remain numerous, diverse, and typical in the context of the research [14]–[16].

Many have developed particular measurement instruments in higher education and not on the entire entity of the institution [17], [18]. Many studies have relied on existing literature models despite the importance of empirical approaches adopted [19]–[22].The literature review analysis quality and validation of instruments measuring quality in secondary schools remain under-explored compared to school education.

Establishing a quality assessment system in schools is fundamental to this educational training entity's performance development and self-management. However, it is crucial to identify, on the one hand, the indicators reflecting the quality in its schools, and on the other hand, to focus on all the instruments and tools of psychometric measurement. The study suggests a systematic review of the literature on available measurement instruments in this perspective.

The systematic literature review on tools and instruments for measuring quality and its indicators has two main objectives. First, is to identify the psychometric properties of tools available in the literature on validation and construction of measurement instruments, considering steps validated and adopted by researchers. Second, is to build an initial and predictive database for researchers and practitioners.

## 2 Dimensions of quality in education

Approaching the literature regarding the purpose of the study, and the concepts investigated in this research, it is emerging that there is a debate about whether speaking of expectations or perceptions or between the two constitutes a better measure of quality (service, pedagogical, administrative school).

Generally, in this type of research: it is necessary to recommend a conceptual measure that connects the theoretical level of the concept studied (definition of the phenomenon learned) to a practical level (description of indicators representing this phenomenon and on which the concrete measurement operations are based)[23]. Other researchers and authors refer to a set of parameters that can influence the whole process of validation of indicators, asking the question about the objective of the measurement and the object of size to deduce the relevance of a constructor or not [24]. The latter is considered as a phenomenon of theoretical interest whose conceptual definition must include (1) the object and its components, (2) the attributes and its components, and (3) the respondents to indicate how the construct will be measured operationally. Others focus on the limitations of psychometric assessment [25], [26].

Another approach has been the subject of many studies on higher education, which have made the pairing between the dimensions of process, engagement, and content such as: my curricula, innovative practices, student motivation, and subject diversity. [27] They are determinants of quality that were based on the perceptions of actors. However, given the complexity of measuring quality in higher education institutions, the results obtained are minimal [28].

Another dimension that has been interest of authors is the quality of service that is in correlation with its evaluation of how they emphasize service [29], [30]. Others talk about quality management through three dimensions: inputs, processes, and outputs [31]. Some qualify the measurement of quality between expected and perceived service [32] whose model most known in the literature review is that of the SERVQUAL model of Zeithaml and Parasuraman. They proposed this model with a scale with multiple elements to measure the quality of service; it is called a model of the gap identified between expectations and perceptions that must be closed to satisfy the quality. It has been used and applied in several research on education [33]–[35].

Yet, the models remain highly critical, and several authors do not support the five-factor SERVQUAL [36], and the administration of expectancy items is another essential element [37]–[40] have been particularly vociferous in their criticism. The development of their performance-based measure, called SERVPERF. The SERVPERF scale is the unweighted perceptual component of SERVQUAL, consisting of 22 perceptual items, thus excluding any consideration of expectations.

In their empirical work, [39] found that the unweighted SERVPERF measure (performance only) performs better than other unweighted measure of service quality and has greater predictive power (ability to provide quality services.) An accurate service quality score) than SERVQUAL. They argue that the best reflection of customers' perceptions of service quality and performance is not part of this concept.

Similarly, [41] rejects the value of an expectations-based approach. SERVQUAL, and agree that service quality is influenced only by perceptions. Across the studies we

consulted, consent is a common feature of a range of measurement scale validation approaches that emphasize content and purpose validation in a qualitative manner-usually in the exploratory phase, especially in the absence of consensus on the concepts being studied [42], [43].

## 3 Quality indicators according to national and international organizations

Quality indicators are essential in education, especially in schools [3]. Developing an evaluation scale requires the availability of a minimum number of statements [43], [44] for each of the dimensions related to the studied concept. Therefore, the measurement instruments currently available in the literature review do not allow to respond to the aspects studied, which are either related to different fields other than the field of education and teaching or affect other aspects and sectors of activity (psychometric measurements). However, non-governmental organizations have been able to disclose indicators that reflect the quality and good governance of certain countries, the example of African countries. Nevertheless, it remained an approach to determine its hands based on institutional documentation and government figures far from any scientific validation by researchers. UNESCO specialists (2005), within the framework of the Ecole Pour Tous (EFA) in Côte d'Ivoire, have taken stock of the indications (eight indicators) on the effectiveness or otherwise of elementary school (qualitative and quantitative).

The multidimensional nature of the concept of quality and especially in a much more complex entity (several stakeholders and actors) complicates the evaluation of the latter's quality. In this sense, the last report of [45] on the quality of schools had recommended indicators and fields relevant to the improvement of the performance of schools with a balanced involvement of all actors (school leaders, administrators, inspectors, teachers, students and parents of students). The issue of school performance has been the subject of a growing number of studies in various national and international institutions [46]. In this perspective, the OECD countries are launching scientific debates on the effectiveness of the quality of global education. In addition to the declaration that quality education is fundamental for social and economic development education ministries of 153 member countries of UNESCO at the 48th International Conference on Education in November 2008 in Geneva. Indeed, the quality remains evident for the educational system and the school. However, this concept generates ambiguity and changes over time and affects several aspects [22], [47]. Although the school is a local entity of the educational management system, it is complex to assess its quality [48]. The diversity of dimensions attributes is difficult to conduct systematic investigations and frame the concept [49], [50].

Clearly, and in the face of the continuing demand for a collection of specialized studies on indicators measuring quality in education and schools. This study is an encouraging step in a literature review to other studies focusing on approaches to validat-

ing quality measurement tools. It can serve as a primary, transparent tool for researchers, practitioners, and educational leaders to reflect on quality in schools. The following sections contain the methods and results obtained.

# 4 Methodology

## 4.1 Selection criteria and data sources

The systematic review was conducted independently and qualitatively by Two experts and evaluation researchers (SL, HB). Used it to identify relevant studies related to our topic based on the search questions. Full texts were obtained, referenced, and reviewed for relevant studies using the bibliographic reference management software Zotero and the Excel spreadsheet. The following search strategy was combined with the following keyword items: (measurement scale); (psychometrics); (psychometric development of measurement scales); Or (validation of measurement scales); Or (quality indicators; And service quality); (systematic review of psychometric properties of quality measurement instruments); Or (quality of measurement scales and tools in psychometrics). This search strategy was adapted to the other databases of the keywords (concepts: quality, evaluation, and audit). All research references were imported into the bibliographic data of ZOTERO and MENDELEY. Duplicates were eliminated before the selection process.

## 4.2 Literature search strategy

An in-depth search was performed using the database of major journals and newspapers known for their scientific rigor and requirements namely, Scopus, Science Direct, Web of Science, Cairn, and Google scholar. Other essays were found from article reference lists and author bibliographic databases. We have included articles published up to December 2019 and limited electronic database searches to English-language publications in this review. Still, we applied the following exclusion criteria to the title/abstract and full text to identify relevant studies: letters, commentaries, and consensus reports, descriptive notes, tools that are not based on reliable and valid methods approved by researchers in the literature. This literature review search was based on two conceptual blocks that constitute the keywords of our search (See Figure 1): quality in education (and service quality within the school) and studies of the construction and validation of measurement scales. In addition, open studies that address other themes out of context or general are not mentioned unless they are relevant or have added value in terms of methodological and conceptual aspects of measurement tools and scales and their validation. The choice of treatments is based mainly on precise results and findings from the scientific literature. In this regard, we focused on recent studies and previous studies to emerge with explicit and comprehensive recommendations. We analysed mainly articles written in English after their translation into French, including those in French but with a low rate.

We selected the articles by abstracts and complete texts carried out in two steps: the first transversal reading with a translation of the articles in English and Spanish into French. And a second reading by two reviewers: post-doctoral and academic researchers' experts in evaluation. All studies were extracted and exported to a Microsoft Excel spreadsheet, version 2016. Any disagreement was resolved by discussion between the two authors until a consensus was reached. An additional English-language academic reviewer was consulted.

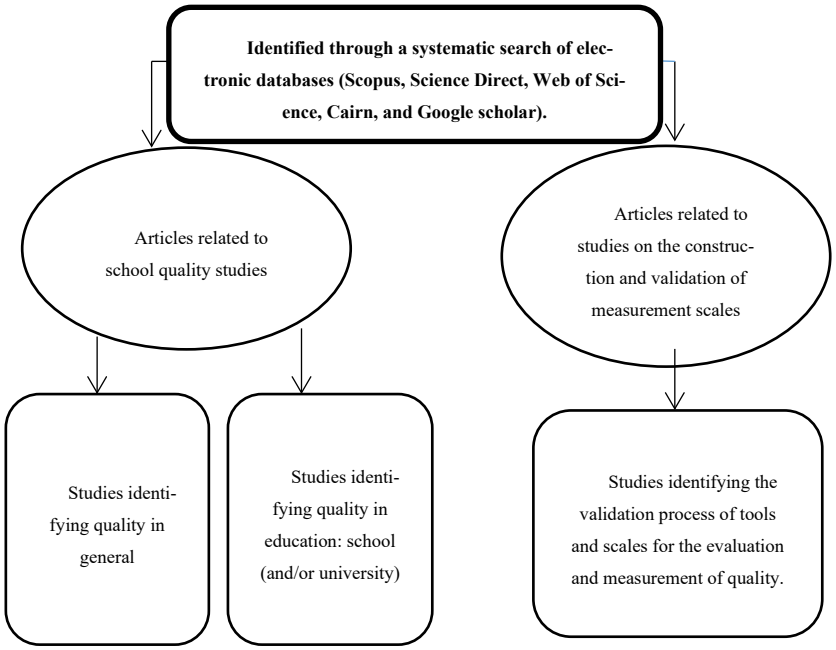


**Fig. 1.** The approach to the collection of studies

### 4.3 Data extraction and limitations

The approach to article analysis was purely narrative and based on descriptive content analysis[51]. This decision was informed by recent methodical exploration of literature review and argued that not conducting a meta-analysis when studies are too diverse in terms of study design, study populations, questions measured, etc… It is sufficient to report the results descriptively using a textbook and tables.

The study selection process consisted of the following two phases:

Level 1 screening: The two expert reviewers (SL and HB) independently reviewed the titles and abstracts of studies identified from electronic databases for eligibility according to the inclusion and exclusion criteria.

Level 2 screening: The same researchers (SL and HB) independently reviewed the full texts of the studies selected in the first screening for eligibility, using the same inclusion and exclusion criteria.

We extracted data from the literature through a Microsoft Excel spreadsheet, containing the following parameters: Original article title, author, journal or newspaper, year of publication, country, study category, study purpose(s), conceptual or theoretical framework, study focus areas, samples (sample size, sample types, age, and gender), measurement instruments or methods, variables assessed, data analysis methods, keywords, results, conclusions, and recommendations. We identified and entered the articles into the Excel spreadsheet (version 2016). The articles included in the study contained approach analysis studies of the validation of measurement and evaluation tools and instruments. They represented different fields of intervention and sectors of activity.

# 5 Results

## 5.1 Studies identified

A total of 103 articles were retrieved from the raw database, of which fifty (50) papers were obtained and retained in the electronic database related to the dimensions of quality. Eleven articles were included and identified in the literature review, presenting the psychometric validation processes and results of 4 instruments related to validating tools and measurement instruments (See Table 1,2 and 3). Thirty articles were excluded from selecting papers according to the criteria administered and cited above, and nine were excluded after consensus among the reviewing authors.

## 5.2 Study characteristics

The studies were diverse and were conducted in different continents of the world (America, Europe, Asia, Australia). Some studies were conducted on quality indicators in education in general, and others identified service quality indicators in the entire education sector. Moreover, we found studies that did not focus on educational organizations such as schools or academic institutions but rather on quality indicators related to the programs and contents taught and teachers' skills. A small number of studies examined quality in schools and how they operate.

**Table 1.** Summary of the characteristics approach studies (Continued)

| Study or research categories | Authors and publication years | Areas and fields studied | Country of data collection | Methods/Scales | Results |
|---|---|---|---|---|---|
| Development and construction of scales | C. Dayan et al. 2016 | Medical:neurology | France | -DELPHI method / Focus group/<br>-the QOL-PMD scale developed by [52] | - Quality of Life Scale for persons with multiple disabilities Construct a quality-of-life scale for children with various disabilities aged 6 to 14 years.<br>- Construction of a personal and environmental characteristics grid. |

| | | | | | |
|---|---|---|---|---|---|
| Development and construction of scales | Cheng et al. 2010 | Educational and PsychologicalMeasurement | Taiwan | -The Rasch Model (Rasch,1960)<br>-The Olweus Bully/Victim Questionnaire (OBVQ): an instrument for investigating school bullying. | The purpose of the study is to develop three school bullying scales: the bullying scale, the victim scale, and the bystander scale, to assess high school students' bullying behaviors, including physical bullying, verbal bullying, relational bullying, and cyberbullying. |
| Development and construction of scales | William Gunawan, Peter A. Creed, and A. Ian Glendon.2018 | Applied-psychology | Australia | -Exploratory and confirmatory factor analysis: EFA and CFA.<br>-Likert scale with 5 points | Development and initial validation of a scale to measure young adults perceptions of their future employability. |
| Development and construction of scales | N.Roussiau, N.Bailly, E.Renarda.2017 | Psychology | France | The DSES, the Spiritual Well-Being Questionnaire, and the Spiritual Transcendence Scale;<br>A Likert scale with levels;<br>Parallel analysis following the procedure described by O'Connor (2001);<br>The Diener scale [53]. | Construct and validate an areligious spirituality scale applicable to the student population. |

**Table 2.** Summary of the characteristics approach studies (Continued)

| **Study or research categories** | **Authors and publication years** | **Areas and fields studied** | **Country of data collection** | **Methods/ Scales** | **Results** |
|---|---|---|---|---|---|
| Development and construction of scales | Nancy Gaudreau, Éric Frenette et Stéphane Thibodeau.2015 | Measurement and evaluation in education | Canada | - Scale was developed according to the recommendations of Bandura (2006) and the procedure proposed by Dussault, Valois, and Frenette (2007).<br>- Scale covers the five dimensions of classroom management suggested by O'Neill and Stephenson (2011).<br>- Rash's model (1960).<br>- 6-level Likert scale | The development of a scale dealing with teachers' sense of self-efficacy in classroom management (with 28 statements). Developed according to the recommendations of Bandura (2006) and the procedure proposed by Dussault, Valois, and Frenette (2007), this scale addresses the five dimensions of classroom management suggested by O'Neill and Stephenson (2011). |
| Development and construction of scales | Osman M. Karatepe, , Ugur Yavas, Emin Babakus.2005 | Banking sector Service quality management | Nord de Chypre | -Churchill's (1979) Paradigm Steps.<br>- The five-dimensional SERVQUAL model of Parasuraman et al. (1988).<br>LISREL instrument (Jo reskog and So rbom, 1993),<br>- Likert scale. with 5 points | To develop and test a service quality instrument using retail banking services in Northern Cyprus. With the following dimensions: service environment (four items), interaction quality (seven things), empathy (five items), and reliability (four items). |

| | | | | | |
|---|---|---|---|---|---|
| Development and construction of scales | Igalens, J., & Tahri, N. (2012) | Management: Human Resources Management | France | 1) Academic measures of ESP [54].<br>2) Churchill's (1979) paradigm with a synthesis of 3 steps: Step 1: Definition of the conceptual domain (Theoretical reflection on the subject and precise definition of what we are looking for); Step2: Exploratory phase (Generalization of items and purification of the measure: Alpha, AFE); Step 3: Validation phase: Validation study (Validity, reliability).<br>3) Fit indices: the norms recommended by [55], [56]. | It created a 9-item, three-dimensional scale that captures 69% of the variance in employees' perception of CSR. And to measure the effects of perceived socially responsible practices on employees' attitudes and behaviors at work. |

**Table 3.** Summary of the characteristics approach studies

| **Study or research categories** | **Authors and publication years** | **Areas and fields studied** | **Country of data collection** | **Methods/ Scales** | **Results** |
|---|---|---|---|---|---|
| Development and construction of scales | Igalens, J., &Tahri, N. (2012) | Management: Human Resources Management | France | 1) Academic measures of ESP (Igalens and Gond, 2008): Carroll (1979); Wood (1991); Clarkson (1995);<br>2) Churchill's (1979) paradigm with a synthesis of 3 steps: Step1: Definition of the conceptual domain (Theoretical reflection on the subject and precise definition of what is sought). Step 2: Exploratory phase (a generalization of items and purification of the measure: Alpha, AFE); Step 3: Validation phase: Validation study (Validity, reliability).<br>3) Fit indices: the norms recommended by [55], [56]. | It has created a 9-item, three-dimensional scale that captures 69% of the variance in employee perception at work. |
| Critical analysis on the use of SERVQUA | Miguel Morales Riadh Ladhari Simon Perreault Simon Nyeck1.1 998 | Critical evaluation of the use of SERVQUAL | Canada | The criteria of the evaluation grid were divided into four headings to identify the articles: general characteristics, problem formulation, data collection method, and data analysis. | The SERVQUAL instrument is reliable and demonstrates face and predictive validity, while results are inconsistent construct validity. Other psychometric properties of the SERVQUAL instrument are unstable and unestablished, far from consensus across the work reviewed. |

## 6 Discussion and conclusion

Psychometrics proposes several rigorous methods and approaches for developing measurement instruments [57] to create innovative devices and tools for good governance and management practices in different areas. We conducted a literature review on developing and validating a measurement scale, which is not a contextual element in itself but rather one of the inseparable parameters of evaluating the quality and good governance of an organization, institution, or establishment. However, some of the research conducted referred to the literature review to adopt in two stages the inductive and deductive approach following the paradigm steps[4], [58], [59] (See Table 1,2 and 3). The first stage validates the instrument and the measurement tool, and the second stage identifies the research object's dimensions. Furthermore, the first approach allowed for the formulation of the sizes raised as research hypotheses and the second (deductive) in the sense of testing them [60], [61].

Some authors had researched the notion of quality in other dimensions related to the medical, banking, and psychological fields[4], [12], [38], [59], [60], [62] but it falls into the convergent category of measurement instrument validation studies. Studies were conducted to construct a scale to measure well-being and service quality. One article was based on an approach adapted and corrected from the Churchill paradigm (1979), called the C-OAR-SE approach [24], which respects the different analysis of the validation of scales (AFE, AFC, reliability, validity, and Principal Component Analysis PCA) [63]. Therefore, authors focused on the quality of service in entities to determine a measurement scale determining the perception of people (or customers) towards quality had a tool called SERVQUAL model with five dimensions of [36] (Table2): service environment (four items), interaction quality (eight items), reliability (five items), empathy (10 items) and technology (four items). Likewise, the SERVPERF model, which assesses customer perceptions on five criteria: reliability, responsiveness, safety, empathy, and tangibles. The latter was used in research to evaluate the service quality of the banking system in Mexico [12]. These models or dimensions jointly are widely used by researchers in management and service quality.

Other authors were interested in fields similar to the present study, which is the field of education. An article that considers the development of a scale and validation study on teachers' sense of self-efficacy in classroom management. This study was based on the procedure proposed by[64] (See Table 4 and 5) and focused on pre-established classroom management dimensions suggested by [65]. In contrast, a recent Australian study focused on developing and validating a scale for measuring young adults' perceptions of their future employability. The latter adhered to multi-phase validation analysis (construct validation through experts, AFE, AFC, construct validity was examined by correlating scale scores with measures of career ambition) [62].

We excluded some studies [66] with approaches and validation methods from the field of psychometrics (Rash's process) for their complexity, and because their procedures are not suitable for our research objective: to establish a school quality system (E.S).

An initial observation that emerged from the literature review showed a multitude of research on the construction and validation of scales and measurement tools in recent

years, but it remains challenging in gaining a universal agreement of researchers; given the specificity of each field on which the study was carried out and the limitations of the psychometric characteristics of each approach to scale validation. In this perspective, we focused our research review on synthesis articles that examined the strengths and weaknesses of each method and procedure and the measurement instruments used by the researchers, which would be potentially compatible with our study object (quality in HE). However, it should be emphasized that none of them allows us to measure it by considering the various aspects of schools' pedagogical and administrative quality and management. The multidimensional nature of the concept of quality, in a more complex entity (several stakeholders and actors), makes it more challenging to assess the latter's quality. Nevertheless, the SERVQUAL scale (See Table 4 and 5) was developed by [36]. It comprises 22 statements grouped into five dimensions (tangible elements, reliability, helpfulness, assurance, and empathy) that had known a vital hearing and important use of the researchers. However, it has limitations concerning the instability of its factorial structure and its invariance across various sectors. This observation is underlined by research published in 1998 in the faculty of sciences of administration-Université Laval-[67]. These authors conducted a comparative evaluation of works (61 articles) that used SERVQUAL or a modified SERVQUAL approach. They were able to draw some conclusions emphasizing the limits of using the SERVQUAL approach, despite good reliability indicators. The other psychometric properties of the instrument are not established, although the original study of Parasuraman and collaborators, which proposed five "universal" dimensions supposed to measure the quality of services whatever the sector. According to each field of application, the authors intensely observed a variable and unstable dimensional structure (different from 5 dimensions). This research was able to enlighten us by comparing the psychometric characteristics of the SERVQUAL approach and that of Churchill. The conclusion shows that the first (SERVQUAL) had research design characteristics (sample size, number of statements, and method of administering the instrument) had no relationship with Cronbach's Alpha coefficient. As for, the second had a strong relationship (positive or negative influence) on the Alpha and the impact of the area of study on dimensionality.

In summary, the psychometric evaluation is both diverse and limited, given the numerous measurement scales developed by the authors in the research.

Studies have conceptualized scales in the literature with an existing set of outcome measures or the construct of interest on which the items that measure them will be based [68]. The approaches are proposed by Timothy [69] in his article, "A 37 Brief Tutorial on the Development of Measures for Use in Survey Questionnaires." Hinkin presented six steps to the development process of a measurement scale (See Table 4 and 5), Rossiter who also proposed an approach adapted from Churchill's (1979) six-step paradigm (The C-OAR-SE Procedure). The following table summarizes the approaches to validating the scales identified in the literature.

Developing a scale for evaluating educational and administrative quality requires the availability of a minimum number of statements [43] for each of the dimensions related to the concept under study. Therefore, the measurement instruments currently available in the literature review do not respond to the aspects studied (associated with fields

other than education and teaching) or affect other elements and sectors of activity (measures in psychometry).

Moreover, the measurement instruments differ according to the period of reference of the research, the length of the scale, the categories and dimensions raised, and the number of responses counted. Also, there is a lack of consensus among researchers in developing the items. We find a version that aims to develop indicators from previous theoretical sources already established. However, some authors do not opt for this version, as [24], who does not consider the concept studied as a construct but rather an attribute that needs to be embodied in a specific field to indicate how the construct will be measured [63]. He invites us to pay more attention and write predictive items in the exploratory phase. It is not up to the statistical software to choose the right things, but to the researcher; otherwise, he recommends adopting a more theoretical, qualitative, and adapted approach to constructing the measurement scale. This position suspects the systematic application of the Churchill (1979) paradigm with suspect measures. It was not the only limitation issued on such an approach. The SERVQUAL approach also has limitations; recently, it was stated that managers should avoid using the SERVQUAL scale on a global scale and instead develop "a new, culturally limited measure of service quality" [70], [71].

Consent can be a common feature of a set of measurement scale validation approaches that emphasizes the content and purpose of the research in a qualitative manner, especially in the exploratory phase and in the absence of consensus on the concepts being studied. Measuring an ideal means establishing a "correspondence between a theoretical level (definition of the phenomenon reviewed) and a practical level (description of the indicators representing this phenomenon and on which the concrete measurement operations are based) [72].

Our literature review on the construction and validation of measurement scales in psychometrics constitutes an essential database for evaluating quality in schools and the first at the Moroccan national level.

**Table 4.** Characteristics and limitation measurement scales (continued)

| Authors | Concept measured | Dimensions | Items | Logical reasoning | Steps/procedure | Limitations (and advantages?) |
|---|---|---|---|---|---|---|
| Procedure C-OAR-SE de Rossiter (2002) | Concept and validation dimensions of the scales: e.g., Self-esteem, quality of work-life, etc. | Definition of the construct Classification of the object Classification of the attribute Refinement of the purpose of the construct Identification of the respondents' Creation of | 6 to 8 items perdimension | Hypothetical-inductive and deductive | Step 1: Definition of the construct<br>Step 2: Classification of the object<br>Step 3: Attribute classification Refinement of the construct definition<br>Step 4: Identification of respondents<br>Step 5: Creation of the scale | - Pay much more attention to the writing of the items. The researchers' role is to choose the right items and not the statistical software.<br>- The systematic application of Churchill's (1979) paradigm has resulted in the development of suspect measurement scales. On the other hand, a more theoretical approach and procedure for constructing measurement scales are recommended. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | the scale calculation of the score. | | | Step6: Calculating the score | |
| Timothy Hinkin (1998) | Concept and validation dimensions of the scales: e.g. (mental health climate, etc.) | Not assigned | Indeterminate | Deductive | Step 1: Item generation<br>Step 2: Questionnaire Administration<br>Step 3: Initial Item Reduction<br>Step 4: Confirmatory Factor Analysis<br>Step 5: Convergent/Discriminant Validity<br>Step6 : Replication | The stability of the factor structure is not demonstrated, nor is its invariance across various sectors. |

**Table 5.** Characteristics and limitation measurement scales

| Authors | Concept measured | Dimensions | Items | Logical reasoning | Steps/procedure | Limitations (and advantages?) |
|---|---|---|---|---|---|---|
| De Vellis (2012) | Multiple | Definition of the construct<br>Classification of the object<br>Creation of the scale<br>Calculation of the score | Undetermined | Deductive | Step 1: Identify the purpose of measurement: definition of the concept<br>Step 2: Generate an initial pool of indicators from the interviews<br>Step 3: Determine the format of the response device<br>Step 4: Submit the collection of indicators generated in Step 2 for expert review Analysis)<br>Step 5: Administer the indicators to a sample<br>Step 6: Evaluate the indicators<br>Step 7: Optimize the length of the measurement scale | Optimization of the scale needs more confirmation through research;<br>Minimal statistical analysis of scale validation; |
| Modèle SERVQUAL de Parasuraman et coll. (1988) | Quality of service | Five dimensions (05): (tangibles, reliability, helpfulness, assertiveness, and empathy) | 22 statements: Tangible elements (4 statements)<br>Reliability (5 statements)<br>Helpfulness (4 statements)<br>Assurance (4 statements) | Deductive and indicative | Qualitative and conceptual analysis;<br>Exploratory Factorial Analysis;<br>Confirmatory statistical factor analysis;<br>(Unidimensionality, reliability, validity, and variance explained) | The stability of the factor structure is not demonstrated, nor is its invariance across different sectors.<br>The SERVQUAL instrument is reliable and demonstrates face and predictive validity, while results are |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | Empathy (5 statements) | | | divergent in terms of construct validity. [67] |
| HEdPERF (Higher Education PERFormance Modèle proposed by Firdaus (2004) | Quality of service | Four Dimensions (04): Non-academic aspects; Academic Aspects; Reliability Empathy | 41 items, some adapted from SERVPERF, and 28 items generated by the qualitative literature review | Deductive and indicative | Qualitative analysis; Literature review; Exploratory Factorial Analysis; Confirmatory statistical factor analysis; Regression analyses (Unidimensionality, Reliability, Validity and Explained Variance) | Limited to a single service industry, this claim of generalizability should be validated by further research and other high school settings. |

## 7 Limitations

Our research remains limited, as we did not conduct an exhaustive list of articles related to the construction of measurement scales. Furthermore, the research relevance measure analysis was based on descriptive and expert consensus. We could not carry out an in-depth investigation concerning our theme with known statistical indicators such as the methodological protocol for analyzing the psychometric properties of COSMIN for research relevance, which used the different measurement scales. The fields of research were different. The caution is that implementing a quality measurement approach in education was found in the results of this research. Future research could explore and further emphasize the strengths and drawbacks of each method and procedure for validating measurement scales and analyzing the influence of the field of study on the dimensionality of the tool constructed.

However, this study can support practitioners and researchers on the scales of measurement of the quality of entities and organizations of teaching and education.

## 8 References

[1] S. Jia et Y. Pang, « Teaching Quality Evaluation and Scheme Prediction Model Based on Improved Decision Tree Algorithm », *Int. J. Emerg. Technol. Learn.*, vol. 13, nº 10, p. 146, oct. 2018. https://doi.org/10.3991/ijet.v13i10.9460

[2] H. Berbar, S. Lotfi, et M. Talbi, « Validation and Development of a Moroccan School Quality Evaluation System », *Education Research International*, vol. 2021, p. 1829259, déc. 2021. https://doi.org/10.1155/2021/1829259

[3] E. Talbott, D. M. Maggin, E. Y. Van Acker, et S. Kumm, « Quality Indicators for Reviews of Research in Special Education », *Exceptionality*, vol. 26, nº 4, p. 245-265, oct. 2018. https://doi.org/10.1080/09362835.2017.1283625

[4] C. Dayan, R. Scelles, A.-M. Boutin, G. Ponsot, C. Arnaud, et M. Storme, « Création et validation d'une échelle de qualité de vie et d'une grille d'observation des caractéristiques personnelles et environnementales pour les enfants polyhandicapés : enseignements d'une étude », *Motricité Cérébrale : Réadaptation, Neurologie du Développement*, vol. 37, nº 4, p. 127-138, déc. 2016. https://doi.org/10.1016/j.motcer.2016.10.004

[5] T. Machado, P. Desrumaux, et A. Van Droogenbroeck, « Indicateurs organisationnels et individuels du bien-être. Étude exploratoire auprès d'aides-soignants et d'infirmiers », *Bulletin de psychologie*, vol. Numéro541, nº 1, p. 19, 2016. https://doi.org/10.3917/bupsy.541.0019

[6] P. Dahler-Larsen, « Publication des données sur la qualité des établissements scolaires au Danemark : contingence et non-linéarité », *Education et sociétés*, vol. 28, nº 2, p. 21, 2011. https://doi.org/10.3917/es.028.0021

[7] D. van Damme, « Tendances et modèles de l'assurance internationale de la qualité de l'enseignement supérieur en relation avec le commerce des services d'éducation », *Gestion de l'enseignement supérieur*, vol. 14, nº 3, p. 107-158, déc. 2002. https://doi.org/10.1787/hemp-v14-art21-fr

[8] M. Morales, R. Ladhari, S. Perreault, et S. Nyeck, « La validité de la mesure de la qualité des services: Une évaluation de 10 ans d'utilisation de SERVQUAL », *Document de travail*, vol. 14, 1998.

[9] D. Billing et H. Thomas, « The International Transferability of Quality Assessment Systems for Higher Education: The Turkish experience », *Quality in Higher Education*, vol. 6, nº 1, p. 31-40, avr. 2000. https://doi.org/10.1080/13538320050001054

[10] Y. Yu, « Student Assessment System for Distance Education on Financial Management », *Int. J. Emerg. Technol. Learn.*, vol. 11, nº 10, p. 11, oct. 2016. https://doi.org/10.3991/ijet.v11i10.6265

[11] Z. Wu, H. Li, X. Zhang, Z. Wu, et S. Cao, « Teaching Quality Assessment of College English Department Based on Factor Analysis », *Int. J. Emerg. Technol. Learn.*, vol. 16, nº 23, p. 158-170, déc. 2021. https://doi.org/10.3991/ijet.v16i23.27827

[12] J. Torres Fragoso et I. Luna Espinoza, « Assessment of banking service quality perception using the SERVPERF model », *Contaduría y Administración*, vol. 62, nº 4, p. 1294-1316, oct. 2017. https://doi.org/10.1016/j.cya.2017.06.011

[13] O. M. Karatepe, U. Yavas, et E. Babakus, « Measuring service quality of banks: Scale development and validation », *Journal of Retailing and Consumer Services*, vol. 12, nº 5, p. 373-383, sept. 2005. https://doi.org/10.1016/j.jretconser.2005.01.001

[14] C. Hildesheim et K. Sonntag, « The Quality Culture Inventory: a comprehensive approach towards measuring quality culture in higher education », *Studies in Higher Education*, vol. 45, nº 4, p. 892-908, avr. 2020. https://doi.org/10.1080/03075079.2019.1672639

[15] A. Badran, E. Baydoun, et J. R. Hillman, Éd., *Major Challenges Facing Higher Education in the Arab World: Quality Assurance and Relevance*. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-03774-1

[16] C. Sattler, K. Sonntag, et K. Götzen, « The Quality Culture Inventory (QCI): An Instrument Assessing Quality-Related Aspects of Work », in *Advances in Ergonomic Design of Systems, Products and Processes*, B. Deml, P. Stock, R. Bruder, et C. M. Schlick, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, p. 43-56. https://doi.org/10.1007/978-3-662-48661-0_3

[17] M. Karaca et S. Inan, « A measure of possible sources of demotivation in L2 writing: A scale development and validation study », *Assessing Writing*, vol. 43, p. 100438, janv. 2020. https://doi.org/10.1016/j.asw.2019.100438

[18] Y. Ozdemir, S. K. Kaya, et E. Turhan, « A scale to measure sustainable campus services in higher education: "Sustainable Service Quality" », *Journal of Cleaner Production*, vol. 245, p. 118839, févr. 2020. https://doi.org/10.1016/j.jclepro.2019.118839

[19] D. S. Silva, G. H. S. M. d Moraes, I. K. Makiya, et F. I. G. Cesar, « Measurement of perceived service quality in higher education institutions: A review of HEdPERF scale use »,

*Quality Assurance in Education*, vol. 25, nº 4, p. 415-439, sept. 2017. https://doi.org/10.1108/QAE-10-2016-0058

[20] A. A. Shah, M. A. Uqaili, et A. S. Qureshi, « Adoption of quality culture — A case study of Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan », in *2017 IEEE Global Humanitarian Technology Conference (GHTC)*, San Jose, CA, oct. 2017, p. 1-5. https://doi.org/10.1109/GHTC.2017.8239254

[21] N. Huson, « Oman. Quality Culture in Higher Education A Good-Practice Example », *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, vol. 20, nº 2, 2015.

[22] F. Abdullah, « HEdPERF versus SERVPERF: The quest for ideal measuring instrument of service quality in higher education sector », *Quality Assurance in Education*, vol. 13, nº 4, p. 305-328, janv. 2005. https://doi.org/10.1108/09684880510626584

[23] Y. Evrard, B. Pras, E. Roux, P. Desmet, A.-M. Dussaix, et G. L. Lilien, « Market - Fondements et méthodes des recherches en marketing », HAL, hal-00490724, juill. 2009. Consulté le: 12 janvier 2022. [En ligne]. Disponible sur: https://ideas.repec.org/p/hal/journl/hal-00490724.html

[24] J. R. Rossiter, « The C-OAR-SE procedure for scale development in marketing », *International Journal of Research in Marketing*, vol. 19, nº 4, p. 305-335, déc. 2002. https://doi.org/10.1016/S0167-8116(02)00097-6

[25] J. Vessey, T. D. Strout, R. L. DiFazio, et A. Walker, « Measuring the youth bullying experience: A systematic review of the psychometric properties of available instruments », *Journal of School Health*, vol. 84, nº 12, p. 819-843, 2014. https://doi.org/10.1111/josh.12210

[26] A. M. Vivolo-Kantor, B. N. Martell, K. M. Holland, et R. Westby, « A systematic review and content analysis of bullying and cyber-bullying measurement strategies », *Aggression and violent behavior*, vol. 19, nº 4, p. 423-434, 2014. https://doi.org/10.1016/j.avb.2014.06.008

[27] E. Varouchas, M.-A. Sicilia, et S. Sánchez-Alonso, « Towards an integrated learning analytics framework for quality perceptions in higher education: a 3-tier content, process, engagement model for key performance indicators », *Behaviour & Information Technology*, vol. 37, nº 10-11, Art. nº 10-11, nov. 2018. https://doi.org/10.1080/0144929X.2018.1495765

[28] D. Chalmers, « A review of Australian and international quality systems and indicators of learning and teaching », *Carrick Institute for Learning and Teaching in Higher Education*, vol. 1, nº 2, p. 1-122, 2007.

[29] M. A. Khalaf et N. Khourshed, « Performance-based service quality model in postgraduate education », *Int J Qual & Reliability Mgmt*, vol. 34, nº 5, Art. nº 5, mai 2017. https://doi.org/10.1108/IJQRM-04-2015-0059

[30] L. L. Rodrigues, G. Barkur, K. Varambally, et F. G. Motlagh, « Comparison of SERVQUAL and SERVPERF metrics: an empirical study », *The TQM Journal*, 2011. https://doi.org/10.1108/17542731111175248

[31] E. Bayraktar, E. Tatoglu, et S. Zaim, « An instrument for measuring the critical factors of TQM in Turkish higher education », *Total Quality Management & Business Excellence*, vol. 19, nº 6, p. 551-574, juin 2008. https://doi.org/10.1080/14783360802023921

[32] C. Gronroos, « Service quality: The six criteria of good perceived service », *Review of business*, vol. 9, nº 3, p. 10, 1988.

[33] L. F. Dourado et J. F. de Oliveira, « A qualidade da educação: perspectivas e desafios », *Cadernos Cedes*, vol. 29, p. 201-215, 2009. https://doi.org/10.1590/S0101-32622009000200004

[34] C. Zafiropoulos et V. Vrana, « SERVICE QUALITY ASSESSMENT IN A GREEK HIGHER EDUCATION INSTITUTE », *Journal of Business Economics and Management*, vol. 9, nº 1, p. 33-45, mars 2008. https://doi.org/10.3846/1611-1699.2008.9.33-45
[35] K. C. Tan et S. W. Kek, « Service quality in Higher Education using an enhanced SERVQUAL approach », *Quality in Higher Education*, vol. 10, nº 1, p. 17-24, avr. 2004. https://doi.org/10.1080/1353832242000195032
[36] A. PARASURAMAN, « SERVQUAL: A Multiple-item scale for measuring consumer perceptions of service quality », *Journal of Retailing*, vol. 64, nº 1, p. 12-40, 1988.
[37] A. Parasuraman, « Alternative scales for measuring service quality: A comparative assessment based on psychometric and diagnostic criteria », *Journal of Retailing*, vol. 70, nº 3, p. 201-230, 1994. https://doi.org/10.1016/0022-4359(94)90033-7
[38] E. Babakus et G. W. Boller, « An empirical assessment of the SERVQUAL scale », *Journal of Business Research*, vol. 24, nº 3, p. 253-268, mai 1992. https://doi.org/10.1016/0148-2963(92)90022-4
[39] J. J. Cronin et S. A. Taylor, « Measuring Service Quality: A Reexamination and Extension », *Journal of Marketing*, vol. 56, nº 3, p. 55-68, juill. 1992. https://doi.org/10.1177/002224299205600304
[40] J. M. Carman, « Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions », *Journal of Retailing*, vol. 66, nº 1, p. 33, Spring 1990.
[41] W. Boulding, A. Kalra, R. Staelin, et V. A. Zeithaml, « A dynamic process model of service quality: from expectations to behavioral intentions », *Journal of marketing research*, vol. 30, nº 1, p. 7-27, 1993. https://doi.org/10.1177/002224379303000102
[42] R. Cohen, D. Kincaid, et K. E. Childs, « Measuring school-wide positive behavior support implementation: Development and validation of the benchmarks of quality », *Journal of Positive Behavior Interventions*, vol. 9, nº 4, p. 203-213, 2007. https://doi.org/10.1177/10983007070090040301
[43] C. Rodary, V. Pezet-Langevin, et C. Kalifa, « Qualité de vie chez l'enfant : Qu'est ce qu'un bon outil d'évaluation ? », *Archives de Pédiatrie*, vol. 8, nº 7, Art. nº 7, juill. 2001. https://doi.org/10.1016/S0929-693X(00)90309-9
[44] A. Bandura, « Guide for constructing self-efficacy scales », *Self-efficacy beliefs of adolescents*, vol. 5, nº 1, p. 307-337, 2006.
[45] S. Maghnouj, J. Belanger, M. Clarke, E. Fordham, H. Kitchen, et I. McGregor, « Examens de l'OCDE du cadre d'évaluation de l'éducation: Maroc: OECD Reviews of Evaluation Framework Education: Morocco », 2018. https://doi.org/10.1787/9789264301832-fr
[46] P. Dalin, *How schools improve: An international report*. Burns & Oates, 1994.
[47] J. Soria-García et Á. R. Martínez-Lorente, « Development and validation of a measure of the quality management practices in education », *Total Quality Management & Business Excellence*, vol. 25, nº 1-2, p. 57-79, févr. 2014. https://doi.org/10.1080/14783363.2011.637790
[48] C. Winch, *Quality and Education*. Wiley-Blackwell, 1996.
[49] E. Sallis, « Total Quality Management in Education. Falmer », 2002.
[50] B. A. Winn et K. S. Cameron, « Organizational quality: An examination of the Malcolm Baldrige national quality framework », *Research in higher education*, vol. 39, nº 5, p. 491-512, 1998. https://doi.org/10.1023/A:1018745505108
[51] D. Gijbels et F. Dochy, « Students' assessment preferences and approaches to learning: can formative assessment make a difference? », *Educational studies*, vol. 32, nº 4, p. 399-409, 2006. https://doi.org/10.1080/03055690600850354

[52] K. Petry, B. Maes, et C. Vlaskamp, « Measuring the quality of life of people with profound multiple disabilities using the QOL-PMD: First results », *Research in Developmental Disabilities*, vol. 30, nº 6, p. 1394-1405, 2009. https://doi.org/10.1016/j.ridd.2009.06.007
[53] E. Diener, R. A. Emmons, R. J. Larsen, et S. Griffin, « The Satisfaction With Life Scale », *Journal of Personality Assessment*, vol. 49, nº 1, p. 71-75, févr. 1985. https://doi.org/10.1207/s15327752jpa4901_13
[54] J.-P. Gond et J. Igalens, *La responsabilité sociale de l'entreprise*. Presses universitaires de France, 2008. https://doi.org/10.3917/puf.gond.2008.01
[55] P. Roussel, « Chapitre 9. Méthodes de développement d'échelles pour questionnaires d'enquête », in *Management des ressources humaines*, Louvain-la-Neuve: De Boeck Supérieur, 2005, p. 245-276. https://doi.org/10.3917/dbu.rouss.2005.01.0245
[56] K. J. Spencer et L. P. Schmelkin, « Student perspectives on teaching and its evaluation », *Assessment & Evaluation in Higher Education*, vol. 27, nº 5, p. 397-409, 2002. https://doi.org/10.1080/0260293022000009285
[57] P. Dickes, J.-L. Kop, A. Flieller, et J. Tournois, *La psychométrie: théories et méthodes de la mesure en psychologie*. Presses universitaires de France, 1994.
[58] G. A. Churchill, « A Paradigm for Developing Better Measures of Marketing Constructs », *Journal of Marketing Research*, vol. 16, nº 1, p. 64-73, févr. 1979. https://doi.org/10.1177/002224377901600110
[59] J. Igalens et N. Tahri, « Perception de la RSE par les salariés : construction et validation d'une échelle de mesure », *Revue de gestion des ressources humaines*, vol. 83, nº 1, Art. nº 1, 2012. https://doi.org/10.3917/grhu.083.0003
[60] N. Roussiau, N. Bailly, et E. Renard, « Premières étapes de construction et de validation d'une échelle de spiritualité explicite areligieuse », *Pratiques Psychologiques*, vol. 24, nº 3, p. 277-291, sept. 2018. https://doi.org/10.1016/j.prps.2017.11.001
[61] R. Abou Nasr, « A Study of Job Motivation, Satisfaction, and Performance Among Bank Employees. », 2015.
[62] W. Gunawan, P. A. Creed, et A. I. Glendon, « Development and Initial Validation of a Perceived Future Employability Scale for Young Adults », *Journal of Career Assessment*, vol. 27, nº 4, p. 610-627, nov. 2019. https://doi.org/10.1177/1069072718788645
[63] L. Nasr, J. Burton, et T. Gruber, « When good news is bad news: the negative impact of positive customer feedback on front-line employee well-being », *Journal of Services Marketing*, 2015. https://doi.org/10.1108/JSM-01-2015-0052
[64] M. DUSSAULT, P. VALOIS, et E. FRENETTE, « Validation de l'échelle de Leadership Transformatif du directeur d'école », *Psychologie du Travail et des Organisations*, vol. 13, nº 2, p. 37-52, 2007.
[65] N. Gaudreau, É. Frenette, et S. Thibodeau, « Élaboration de l'Échelle du sentiment d'efficacité personnelle des enseignants en gestion de classe (ÉSEPGC) », *Mesure et évaluation en éducation*, vol. 38, nº 2, p. 31, 2015. https://doi.org/10.7202/1036762ar
[66] Y.-Y. Cheng, L.-M. Chen, K.-S. Liu, et Y.-L. Chen, « Development and Psychometric Evaluation of the School Bullying Scales: A Rasch Measurement Approach », *Educational and Psychological Measurement*, vol. 71, nº 1, p. 200-216, févr. 2011. https://doi.org/10.1177/0013164410387387
[67] M. Morales, R. Ladhari, S. Perreault, et S. Nyeck, « DOCUMENT DE TRAVAIL 1998-014 », p. 22.
[68] A. Hirschi, P. A. Freund, et A. Herrmann, « The Career Engagement Scale: Development and Validation of a Measure of Proactive Career Behaviors », *Journal of Career Assessment*, vol. 22, nº 4, Art. nº 4, nov. 2014. https://doi.org/10.1177/1069072713514813

[69] T. R. Hinkin, « A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires », *Organizational Research Methods*, vol. 1, nº 1, p. 104-121, janv. 1998. https://doi.org/10.1177/109442819800100106

[70] U. Yavas et M. Benkenstein, « Service quality assessment: a comparison of Turkish and German bank customers », *Cross Cultural Management*, vol. 14, nº 2, p. 161-168, mai 2007. https://doi.org/10.1108/13527600710745769

[71] B. C. Imrie, J. W. Cadogan, et R. McNaughton, « The service quality construct on a global stage », *Managing Service Quality*, vol. 12, nº 1, p. 10-18, févr. 2002. https://doi.org/10.1108/09604520210415353

[72] B. Pras, Y. Evrard, et E. Roux, *Market: fondements et méthodes des recherches en marketing*. Dunod, 2009.

## 9 Authors

**Hicham Berbar** is a pedagogical inspector of secondary education at the regional academy of education and training of the Oriental. He is a PhD student in training engineering and science didactics at the Faculty of Sciences of BenM'sik in Casablanca. Morocco. Within the framework of pedagogical and didactic research projects, he coordinates work on evaluating quality in secondary schools in the provincial delegation of education of Berkane. Morocco.

**Saïd Lotfi** is a PhD in Education assessment measure. He is a member of Laboratory of Education and Training Sciences. He is a professor at the Normal Superior School at Hassan II University, Casablanca. Morocco. He pursues a career in training engineering and research methodology by combining consulting and training. Director of a research laboratory at the ENS. Throughout his professional career, He has led projects to implement management evaluation systems within institutions. He is a member of expert committees solicited by the Ministry for several structuring projects.

**Mohammed Talbi** is Dean of the Faculty of Sciences of BenM'sik of Casablanca. Doctor of State in sciences and analysis processes at the Pierre and the Marie Curie University of Paris, Professor at the Hassan II University of Casablanca in Morocco. The fields of teaching and research include educational technologies, distance learning engineering, engineering, and techniques of education and training. Within the Observatory of Research in Didactics and University Pedagogy framework, he coordinates work on the teaching-learning process, didactics of science, the project, and training. He is also interested in evaluating learning, evaluation in project situations, and analysing policies and educational systems.

Article submitted 2022-04-14. Resubmitted 2022-05-28. Final acceptance 2022-05-29. Final version published as submitted by the authors.