Research Article

Validation and Development of a Moroccan School Quality Evaluation System

Hicham Berbar ¹, Said Lotfi ², and Mohammed Talbi ¹

¹Laboratory of Analytical Chemistry and Physical-Chemistry of Materials, Department of Chemistry, Faculty of Sciences Ben M'sik, Hassan II University of Casablanca, Casablanca, Morocco

²Multidisciplinary Laboratory in Education Sciences and Training Engineering (LMSEIF), Sport Science Assessment and Physical Activity Didactic, Normal Superior School (ENS), Hassan II University of Casablanca, Casablanca, Morocco

Correspondence should be addressed to Hicham Berbar; hichamberbar00@gmail.com

Received 8 November 2021; Accepted 16 December 2021; Published 31 December 2021

Academic Editor: Syed Sameer Aga

Copyright © 2021 Hicham Berbar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evaluation is currently at the heart of the priorities of education systems. It is not limited to learning but affects several aspects: teachers, schools, training, management, education policies, and the system as a whole. There is a need in this area where research is extremely scarce in Morocco and especially in the teaching and education sector. The notion of transposing quality evaluation to the pedagogical side is very difficult and ambiguous. The evaluation of a school is a complex process, with varied practices and multiple actors. The first objective of this work is to present, within a rigorous methodological framework, the validation of pedagogical and administrative quality indicators in schools. This tool is a dashboard with precise indicators for the pedagogical audit of schools and educational institutions adapted to the Moroccan context. To select the best indicators, we used several techniques (structured interviews, focus groups, factor analysis, etc.) with the actors who carry out their activities. We identified three (03) fields and ten (10) criteria with indicators that form the basis of a quality assessment. The fields are management and strategic planning, administrative and sector management, and pedagogical organization.

1. Introduction

In Morocco, the Ministry of National Education, Vocational Training, Higher Education and Scientific Research plans to establish quality standards, particularly at all levels of education and training, and to encourage all educational and administrative actors to adhere to such a reform process of the education and training system.

The Ministry of National Education has been working on the development of a national plan to establish a quality system in the education and training system in all its components and programs. The Directorate of Quality of the Education and Training System was created with the aim of disseminating a culture of quality.

This strategic choice is reflected in the adoption by the Ministry of National Education of the Strategic Vision of the 2015/2030 reform [1], elaborated by the Higher Council of Education, Training and Scientific Research (CSEFRS), in Chapter 2, "For a quality school for all," Level 9, "Renovation of the teaching, training and management professions: the first prerequisite for quality improvement especially with the new development model" [2, 3].

It is in this context that our reflection began with the organization of preparatory workshops with a research team composed of trainee inspectors within the framework of training module of measures and evaluations at the Inspectors Training Center for Teaching (CFIE), in order to validate "a system of evaluation of the quality of the school."

This approach consists in elaborating and building a system composed of indicators tested with several resource persons (school directors, staff of the provincial directorate and regional education and training academies (AREFs), inspectors) of the educational sector. It is based on a refinement and definition of a set of key concepts related to the field of evaluation and quality measurement.

As an exploratory measure, an optimization procedure is also proposed to select the best indicators in the context of improvement and accreditation, through focus groups with the actors who carry out their activities followed by a debriefing by designers and specialists to eliminate and/or reformulate the selected indicators statements.

The methodology followed is based on Churchill's (1979) paradigm of specifying the domain of the construct, generating the statements, verifying that the items are well related to the notion of quality, purifying the measures, and evaluating the convergent, discriminant, and predictive validities [4].

We used different methods of structural equation [5, 6] modeling and factor analysis [7] to explore the relationships between unperceived constructs (latent variables) and observable ones.

It is also easy to note that the notion of transposing quality evaluation to the pedagogical side is very difficult and ambiguous. Nevertheless, some research is relevant to improve our understanding of the evaluation system of quality indicators and especially of accreditation methodologies [8–10]. Hence, the choice of the theme is based on the following:

- (i) Lack of a reliable and consensual system of school evaluations
- (ii) Lack of a reference system of inter- and intraschool comparison indicators
- (iii) Lack of knowledge of performance indicators on the part of new principals, head teachers, and inspectors
- (iv) Lack of validated tools based on methodological approaches approved by the scientific community

The objective of the study is to develop and validate a tool for evaluating schools in order to propose it an adequate measuring instrument for the quality of teaching for the future missions of the inspectorate. We used the following methodological approaches:

- (i) Establish as comprehensive a list as possible of indicators that could be used to measure the quality of the organizations involved
- (ii) Categorize these indicators according to the most probable qualities
- (iii) Develop the quality indicator framework using clearly defined criteria
- (iv) Optimize the quality indicator system
- (v) Filter the indicators that will serve as a system for valuing the quality being measured
- (vi) Monitor the institution's performance

From this perspective, our study seeks to answer the following questions:

What are the methodological approaches for building a system for evaluating the quality of schools? What are the reliable indicators of school quality? How can the validity, relevance, and reliability of the instrument for evaluating and monitoring schools be developed?

These questions are among an indefinite list of concerns of instructional designers, listeners, and researchers. We focus our study on these issues, which we consider important.

2. Literature Review

The improvement of quality is one of the axes of the reform adopted by the strategic vision 2015–2030 to ensure the performance of education and the quality of the school. It has become an unavoidable obligation of the reform. Hence, there is the need to build a national repository for the evaluation of the quality of schools. It is important to emphasize that this problem is not limited to the Moroccan context but has also spread to other foreign countries. In 2007, the French Ministry of Education conducted an in-depth evaluation of its schools because it considered that evaluation is necessary to lead to improvement actions, to communicate with the actors and users, and to report for public action [11]. There are many different evaluation methods: results-based management, external evaluation, self-evaluation, audit, etc.

Thierry Bossard, head of the department at the Ministry, reminded 2009 during the “Governance and Performance” conference that they are, along with Greece and Bulgaria, the last country in Europe where there is no organized, systematic, framed evaluation of our educational establishments, whereas in all the other countries, a strong point on which we are banking to improve the results of the establishments [11]. However, a number of research studies provide a more nuanced assessment. They emphasize that the construction of a system of quality indicators for the purposes of accreditation of the quality of public or private organizations, as they are considered a very complicated set of diverse and interfering processes where we find multiple and sometimes even unknown interests. Therefore, trying to account for the quality of this set of actors is difficult and a major problem. Confronted with this challenge, the construction of an evaluation system of indicators that would allow to demonstrating quality remains a challenge and a crucial problem.

The medical sector is not immune to this problem of validation and accreditation of indicators or evaluation systems. In this respect, initiatives to develop measures of the quality of care provided by health establishments are multiplying, and the justification for this is a major movement that has a perspective of aiding decision-making and developing the transparency of the hospital system vis-à-vis the general public [12].

In France, in the United States, and also at the international level, a pioneer country in this field, various initiatives have been developed over the last twenty years. This movement having accelerated in recent years, as

demonstrated by the report of the Institute of Medicine [13], which affirms this dual objective of aiding decision-making and public broadcasting or “accountability.” Similarly, the experiments were conducted in Great Britain (Commission for Health Improvement and the National Institute of Clinical Excellence), Germany (The Federal Consortium for Quality Assurance and the Cooperation for Transparency and Quality in Hospitals), and Denmark (the National Council for Quality Development).

Quality assessment tools remain without scientific basis for validation and relevance. So, this movement to accredit the quality of educational organizations, and more specifically schools, is our problem because we find that there is a set of criticisms that continue to be directed at accreditation procedures. Most of these comments stem from the fact that there is still a lack of clarity as to how to properly establish systems of indicators for measuring quality and about their validity in the education sector. Several researchers have pointed out that this approach to evaluating an organizational quality system has conceptual and empirical limitations that reduce its credibility. The former are related to the risks of adverse selection and complacency with managers [14]. The second connected more to the characteristics of the indicators identified (simplistic and reductive of the complexity of quality) and their inability to determine methodologically what needs to be done to improve quality given the particularity of the sector selected [15].

In this sense, a prospective study was carried out on 36 public and private establishments in France, within the framework of a “COMPACH” project, which sought answers to problems such as the methodology for validating an evaluation system, enabling a selection of chosen indicators to be established, and finally to discuss questions of feasibility and objectives for the use of these items by underlining their implications at the level of the design of the indicators themselves [16]. Other research speaks of compliance with the needs, expectations, and requirements of the responsible institution’s guidelines and the broad outlines of the larger societal project to which the entire institution is engaged. Indicator systems must take into account the totality of an organization’s qualities, regardless of its scope [17].

2.1. Theories of Validation of Measuring Instruments. For several years, scientific research has constituted a very broad field of reflection in different sectors and fields of intervention: political, economic and management sciences, etc, which are characterized by the diversity of their themes and thus of their problems which can affect quite a few specialties.

For this effect, the searcher always relies on previous studies similar to his or her own field of research, and if not, he or she tries to approve their statements, validate their instruments used, and verify the degree of reliability of the conclusions reached. This means that the main problem of any research is to know how to develop better tools and measures with a rigorous and scientific methodology. In the field of psychometrics, the use of such scales has become very demanded in academic research (especially in management

sciences), where many researchers are more interested in tools that prove the properties of measurement scales.

From this point of view, these scales are measurement instruments that are generally composed of several items. They are accompanied by attitude scales that allow the calculation of scores for each respondent: Likert scale, Osgood scale, etc; however, another problem that confronts the researcher is that of the validation of measurement scales, which is a necessary condition for the quality of the results, and moreover we are talking here about two primordial notions: reliability (or fidelity) and validity.

Furthermore, there are several theories or approaches that represent the references of the majority of scientists interested in the quality and validation of items; we will focus on two of them with brief descriptions of the principles of each.

We speak of two types of approaches that stand out: the Items Response Theory (Rasch model) and the classical theory of validation of measurement instruments (Churchill’s paradigm). The first, a more analytical approach (Items Response Theory (IRT)), is based on a relationship between responses to items measuring a latent trait. Contrary to the classical analysis, the relationship between the observed score and the latent trait is not necessarily linear. The IRT is based on three fundamental assumptions: one-dimensionality, monotonicity, and global independence [18]. It is considered to be a relatively recent model (second half of the 20th century) that has provided solutions and satisfactory answers to the problems that classical psychometricity still confronts.

In summary, we can say that the IRT aims at estimating the metric characteristics of the items (difficulty and discrimination parameters) and, on the other hand, at measurement the latent score of each individual (example of parameter related to his skill level). All these estimates are independent of the relative samples (group of individuals and items). We talk about the relation between the latent trait (level of competence for example) and the probability of passing an item. This relationship is formalized by a function called the characteristic function of the item and can be represented graphically by a curve (the characteristic curve of the item).

This method is considered to be conceptually very complicated and difficult for most academic researchers to accomplish. In the second part that follows, we will focus on the classical method of instrument (Churchill’s paradigm), which has been adopted and used by a large number of researchers.

The approach adopted (Figure 1) is based on the steps recommended by Churchill’s paradigm. It aims to integrate the knowledge of measurement theory and the appropriate techniques to improve it into a systematic procedure. This approach allows for the rigorous construction of assessment instruments such as multiscale questionnaires.

Churchill’s paradigm is a methodological approach recommended by the American Gilbert Churchill for the organization of research. The investigation techniques vary according to the stage of the research. In 1979, Churchill developed a methodological approach for the construction

FIGURE 1: Paradigm approach of Churchill [4].

of multiscale or multi-item questionnaires, which have since been much improved. It serves as a reference for scale development and is part of the theory of measurement to test their quality.

The objective assigned to an instrument is to strive for a perfect measurement of the phenomenon under study (true value). This quest proves difficult when the domain under study involves subjective attitudes and perceptions. Therefore, the different steps proposed in Churchill's paradigm aim at reducing two types of measurement error. First, the exploratory phase attempts to reduce random error, i.e., the exposure of the instrument to its elements such as circumstances and the mood of the respondents. This is achieved by testing the reliability of the scales. Then, the validation phase tries to reduce not only the random error but also the systematic error related to the design of the instrument. The exploratory phase includes the first four steps shown in Figure 1, while the validation phase includes the last four steps.

This research fits into this perspective through the construction and validation of a measurement scale, on the one hand, and verifying the degree of validity and reliability of the items and their consistency, which requires following the paradigm of Churchill [4] with a determined number of steps and rigorous techniques to be used and applied.

Churchill's paradigm comprises eight steps grouped into two phases (Figure 1). The scale construction phase includes 4 steps: construct specification, production of statements, initial data collection, and the editing of the statements.

The instrument validation phase is based on data collection using the purified scale. It includes 4 steps: final data collection, reliability measurement, validity, and standards production.

This paradigm is formalized as shown in Figure 1.

3. Methodology

This chapter presents the methodological framework of the research. It specifies the nature of the research and the means used to meet the expected objectives of this research:

- (i) Optimize the quality indicator system; to build and validate a system of pedagogical auditing of schools, aiming at the experimental development of a system of quality indicators in the particular context: education.
- (ii) To build a toolkit for the inspector for the pedagogical audit of schools and institutions
- (iii) To underline a methodological feature for the choice of quality indicators—which remains very rare in the Moroccan educational sector—based on an analytical framework and a prospective study that takes into consideration several parameters and objectives and can build possible standards of competition between different schools.

3.1. Overall Methodological Approach. This approach consists of developing and constructing a system composed of indicators (to be determined later) that was tested with several resource persons (school principals, staff of the provincial directorate and/or AREFs, inspectors) in the education sector. These can only be approached through a quantitative approach, based on a refinement and definition of a set of key concepts related to the field of evaluation and the instrument for measuring quality. The second approach

is qualitative, where we adopt a methodology of tool development, based on deductive approaches, triangulation of data, and statistical analysis of significant parameters from the field. The steps we present include three phases, which are described in the following chapter.

3.2. Steps of the Study. As an exploratory step, an optimization procedure is also proposed to select the best indicators in the context of improvement and accreditation, through focus groups with the actors who carry out their activities: school directors, educational inspectors, and teachers, which was followed by a debriefing carried out by designers and specialists to eliminate and/or reformulate the selected indicator statements. It is important to note that the development of a quality system for an organization involves steps that can be very delicate and a rigorous approach.

It is a methodological study composed of two phases, one qualitative and the other quantitative. The qualitative phase identifies the attributes and their valid measurement indicators of educational quality in schools. The quantitative phase estimates the strength or importance of the internal and external relationships between the different levels of indicators and criteria.

In the following lines, we take a step-by-step approach to build a system of indicators of the pedagogical and administrative quality of schools. To make it easier to read, we have broken this approach down into seven distinct stages:

Qualitative Study and Construction of the Corpus. The scheme adopted in this research is generally based on research that was tried and tested and that is part of the validation and accreditation of quality indicators in an organization similar to ours (teaching and education) [19–21]. Although there are several models that can be drawn upon, there are very few steps that can be taken to complete the construction of a quality indicator system.

Qualitative Data Collection. In order to achieve the objectives of our study, the actors first had to understand the definitions of the dimensions and their facets of this reflection, in which they could find lists of indicators through a focus group to explain the issue and the general context of the study. Each member of the group was then invited to propose up to five indicators for each of the facets identified in the previous step and to forward their proposals to the people in charge of the TRIAGE application. The proposals of each member were grouped by theme by the managers, i.e., according to each dimension and their facets, in view of a meeting during which each theme was individually analyzed by the “expert group.”

The objective of this step was to identify the raw quality-determining items. To this end, two main strategies were used to proceed with the identification of credible quality indicators of potential interest: meetings with field specialists, i.e., the various actors who are in direct contact with the schools, and a review of the relevant literature in order to establish the state of the art of the indicators available in the

field of education and teaching and, on another side, to prepare the questionnaire for the actors whose seniority varies between 10 and 30 years, which is based on a simple form with a single nominal question in two languages (Arabic and French). We discarded the financial aspect and we focused our research on the administrative and pedagogical aspects.

The survey was carried out by means of nondirective interviews, the form of which consisted of a single key question, from which the actors were asked to list statements they considered important on indicators related to the theme: “What are (in your opinion) the indicators that reflect the pedagogical and administrative quality at the level of schools (secondary cycle)?” Figure 2 summarizes this step.

4. Results

4.1. Presentation of the Field. Given the scope of this educational research and its stakes the rigor required by the chosen model the means at our disposal and the constraints of the field, we contacted all the actors (school principals: 62, inspectors: 40, administrators of the provincial directorates and AREFs: 28, teachers: 80) by telephone and by direct meeting. We scheduled appointments with others to carry out an interview with them (Table 1). In order to obtain a favorable response rate, we confirmed the anonymity of their comments. In the end, we succeeded in identifying with them the list of raw items in the same period from October 2017 to December of the same year.

This research was conducted in three (03) regional academies of education and training (AREFs): Oriental, Fez-Meknes, and Rabat-Salé-Kenitra. Each of these is run by a group in charge of the protocol and process of the research composed of inspectors-trainees of the Training Center of Education Inspectors (Table 2).

4.2. Initial Sifting of Statements. Once the data were collected, we submitted them to focus groups in different locations (the three pilot academies) and in two stages:

Internal Workshop. We submitted the statements of the different actors to another group of actors composed of four to five people for cross-validation. This method avoided misinterpretation of the statements, while allowing participants to provide additional information or to group items that had the same meaning, if they thought it would be useful. This procedure also allowed for verification of the content of the proposed items.

It should be noted that in addition to the requirements of this phase, the ethical rules specific to the formulation and respect of the meaning of each item were respected.

External Workshop. A content analysis was carried out based on the data collected in the three AREFs, at the CFIE in Rabat, where we confronted all the items and statements proposed by the participating actors in this empirical and exploratory study.

FIGURE 2: Data collection process.

TABLE 1: Resource actors during the exploratory stage of the research.

School principals	Administrators of the provincial directorates and AREFs ^a	Inspectors	Teachers
62	28	40	80
Total: 210			

^aAREFs: regional academies education and training.

TABLE 2: Pilot academies during the exploratory phase of the research.

Regional Education and Training Academies	Oriental Fès-Meknès Rabat-Salé-Kenitra
---	--

The analysis was conducted in two stages. A thematic analysis of each interview was conducted. Then, a vertical and a horizontal thematic analysis of all the interviews were carried out. This consists of giving each interview a more global thematic structure of its own (i.e., vertical analysis) and comparing all the interviews on their global thematic structures (i.e., horizontal analysis). This finally leads us not to consider the singular coherence of each interview but rather to look for a global coherence at the level of the corpus of data produced by all the interlocutors. We then look for the occurrence, the meaning, and the relevance of the themes from one interviewee to another.

The raw items proposed are in both French and Arabic, which leaves us with another dilemma, a constraining one. However, we proceeded by translating the items in Arabic into French, and to guarantee the same meaning of the items, we resorted to trainee inspectors and associate professors of the French language to establish reliable and valid lists of translated items.

The results of these analyses allowed us to identify the items with a filtering and grouping of those that have the same meaning, also a crossing and a confrontation of all the remarks and results raised.

4.3. Analysis and Emergent Categorization of Items. The steps of the qualitative analysis are shown in Figure 3.

We are aware that the objective is not to extrapolate the results to a wide range of the target population but to establish as exhaustive a list as possible of indicators that could be used to measure the quality of the organizations involved, thus, to highlight a methodological trait for the choice of quality indicators.

We use consensual techniques such as nominal group technique (TGN), focus group, and TRIAGE. The facilitation technique used to ensure optimal production of indicators is the Technique of Information Retrieval by Facilitating an Expert Group (TRIAGE). This technique was developed by Plante and Côté in 1993 [22]. We chose this technique because of the diversity of the themes to be treated at the same time.

TRIAGE is an amalgam of two techniques. The nominal group technique (TGN) developed in 1968 by two American researchers, Delbecq and Van de Ven [23]. The TGN was not chosen mainly because only one theme can be addressed at a time with the same group, which is a major constraint in current research. The TRIAGE technique is characterized by three distinct phases: individual production, group production, and prioritization within a decision framework. The process is simple. It requires the participation of individuals, who, in combination with each other, form the “expert group.” In the same way, we used separate groups of professional key informants who contributed, firstly, to the qualitative phase that allowed the identification and validation of indicators of the pedagogical quality of schools

FIGURE 3: The phases of the analysis and categorization of the items.

and, secondly, to the quantitative phase that determined the weighting (i.e., the relative contribution) of the indicators identified and validated in the qualitative phase.

Following a meeting with these “expert” professionals (trainee inspectors from the CFIE in Rabat, two audit module trainers, two people in charge of regional audit units at the level of the AREFs Oriental and Fez-Meknes, and two doctoral researchers specializing respectively in auditing and in teaching and education). We exposed, in a first step, the state of progress and the collection of its items, and through that, the remarks and observations concerning the latter. More than 289 raw items were proposed during the individual production of the actors, and we eliminated and reformulated 32 items of items that have a general aspect (not measurable). We also grouped some items that have the same linguistic concentration. In the second step, another analysis consists in establishing categories from the items grouped according to the semantical field. This step was adopted by the research team in order to eliminate vague, general statements and to reformulate those that required more precision.

The main task of each of the selected experts was therefore to judge a limited number of indicator statements classified under only one of the selected dimensions. The work performed by each of the field experts was threefold. First, they were asked to rate each indicator on a scale ranging from 0 (not at all related to the dimension under which it is associated) to 3 (strongly related). If an indicator was deemed relevant to a dimension other than the classified one, we proceeded with a consensus among the expert members in order to validate its membership in the selected criterion. And in the case where an indicator statement deemed ambiguous, the expert was invited to suggest a new formulation of this statement. At the end of its work, which took place over two successive days with half-day meetings per day, the team was able to add to the list of new indicators that it deemed relevant and usable for each criterion and each field related to it.

Nevertheless, this step was done taking into account the existing national reference systems, where we took over with reformulation of four (04) indicators of the PRQES project of the Ministry of National Education in Morocco (reference project of the quality of the secondary cycle) and five (05) indicators of the grid of quality of the schools of the primary cycle (Table 3).

4.4. Finalization of the Items and Determination of the Measurement Scale. This last phase of the qualitative stage, contrary to the requirement of the phases prescribed in the paradigm of Churchill, requires the development of a questionnaire taking into account the classification retained of the elements and items perceived and classified in order to ensure the reliability and validity of the measurement tool used, which passes essentially by the evaluation of its dimensionality [24].

A study day on the pedagogical audit was organized at the training center of education inspectors in Rabat (CFIE), March 07, 2018, which brought together 200 trainee inspectors, 12 inspector regional coordinators of regional academies of education and training, and a representative of the Ministry of National Education (MNE, the Inspector General of Educational Affairs).

The previous step allowed for the identification of relevant indicators. The results of the qualitative analysis made it possible to identify a total of more than 157 raw indicators that were proposed by the participants during the individual production. Subsequently, they stopped at 154 indicators listed, in ten (10) criteria of which three (03) indicators were eliminated. The work of the three groups of trainee inspectors from the CFIE in Rabat and the audit experts used the DELPHI technique. This choice is justified by the fact that it allows probing the priorities perceived by the participants, avoiding the confrontation of their suggestions and their comments within the group. The process was based on the use of several proposals put forward by the same expert group [25].

First, a number of open-ended, general questions were asked about the different categories and potential themes related to the indicators identified. The answers are formulated in short sentences. Then, based on the statements made and provided by the experts in this first document, which constitutes the first task, and which will be followed by a second task, each expert is informed of the degree of convergence or divergence between all the statements through a facilitator. This scenario is carried out under the same conditions but in different places and separated in terms of time: first, at the Academy of Oriental, then in Fez-Meknes, and finally in Rabat-Salé-Kenitra. These meetings of experts are always led by the same person (responsible for this study).

At the end of each task and in the three places where the experts’ meetings took place, the criteria and the proposed themes on which a consensus was not reached were submitted to the experts in the second place (Academy of Fez-Meknes) and in the third place (Academy of Rabat-Salé-Kenitra). The criteria and themes, which were not obtained

TABLE 3: Draft list of quality indicators for the secondary school (Ministry) and primary school (UNICEF).

Draft reference on the quality of secondary schools, which includes nine criteria	The quality grid for primary schools, produced in partnership with UNICEF (in July 2011), which includes seven criteria
Building commitments to vision and values	The school's environment and surroundings
Leadership and strategic planning	Inclusion of the school in its environment
Human resource management	Learning conditions in the classroom
Resource allocation and distribution	Attitudes and behaviors
Development and renewal	Instructional content
Developing performance effectiveness	Individual follow-up of children
Stakeholder satisfaction with expectations	
Results and achievements	School project
Improvement and preparation for the future	

after the second round, were rejected. At the end, out of its tasks, a consensus was reached on ten (10) criteria and three (03) major axes (Table 4).

During this study day and through consultation workshops, the audit committee could intervene. Regarding the different stages of the process to be audited, the members of the committees of researchers and experts were almost all in agreement.

This study day is linked to the objective and purpose of each stage (qualitative), consisting of producing a report on the evaluation of procedures, a summary note, and recommendations for the indicators raised.

Finally, this day, put a confrontation of the results found with the quality indicators determined by this qualitative stage. This comparison allowed the integration of certain items not identified by the first qualitative approach. We submitted these items to two other audit trainers and fifty (50) trainee inspectors in order to ask them to verify the absence of redundancy between the items as well as their ease of understanding. At the end of this day, and after all its stages, we managed—taking into account the items retained by the qualitative study—to obtain a significant number of indicators. We have replicated and listed them in fields and categories that correspond to them exclusively and exhaustively through a consensus of actors, researchers, leaders etc.

4.5. Study and Exploratory Analysis of the Scale. Initially, and during the qualitative study, a significant number of indicators were detected, or we set 154. To recapitulate these items are the results of the different steps that appear in Churchill's paradigm, this number of items has been generated. To ensure the reliability and validity of this measurement tool, three essential points must be verified: the evaluation of its dimensionality, the study of its reliability, and the assessment of its validity (exploratory analysis). To do this, we constructed questionnaires that were subjected to exploratory analyses according to the following approach (Figure 4).

To explore the structure of the school quality scale, we conducted an exploratory factor analysis, which is recognized as appropriate when testing scales under construction [26]. Exploratory factor analysis was used to identify latent factors from the measured variables [27]. The results of the

TABLE 4: List of agreed indicators and their criteria, as well as the corresponding fields, following the consensus of the qualitative stage.

Fields and axes	Total items
<i>Management forecasting and strategic planning</i>	
Institutional foresight	21
Leadership, planning, and management	22
Human resources management	17
<i>Administrative and sector management</i>	
Space and environment of the institution	20
Evaluation and monitoring	13
Conditions of achievement and learning	17
Supervision, training, and monitoring	04
<i>Pedagogical organization</i>	
Teaching contents	11
Practices and behaviors of the actors within the institution	12
School life activities	05
	154

FIGURE 4: The steps of the exploratory analysis.

exploratory factor analysis show three factors with 154 indicators forming the school quality scale, and the eigenvalue was greater than 1 [26]. This is a satisfactory proportion, with eigenvalues exceeding 1 [28].

The results showed, after a factor analysis (154 items), with variance values that exceed 1 and eliminated indicators (items) that have factor loadings less than 0.4 (<0.4 in the component matrix).

5. Discussion

5.1. Questionnaire Design. The main objective of this study was to develop and validate a multidimensional measurement device to evaluate the quality of secondary schools,

optimized and adapted to the context of education in Morocco. This study presents specific characteristics from which the relevance lies; on the one hand, it approaches the structure of the qualifying secondary school with the different facets that constitute its performance and quality, and on the other hand, it covers multiple dimensions of measurement of the quality of an educational organization, such as schools, which remains very little touched by the studies of validation of the scales of measurement.

The design and validation of this measurement and evaluation scale is based on a methodological approach related to the classical theory of scale validation, through the paradigm of Churchill [4]. This model was widely supported and recommended by several authors [29, 30]. In addition, our study is based on the theory of measurement that is increasingly explicit [31]. The exploratory study is an integrated phase in the Churchill paradigm process and consists of purifying the measurement instrument. Among the main methods of possible data collection on which the persons in charge of this study with five trainee inspectors acting as experts have independently made the choice on the method of questionnaire, of course, this last one as any other choice of methods, has its advantages and disadvantages. However, this choice was guided by the fact that this phase aims first of all to establish the dimensionality of the scales and also seeks to verify the internal consistency and their reliability and validity. The second objective is to purify the measurement instrument and extract latent indicators from the items identified; and the third is the large number of resource actors involved in this study, as well as the territory to be covered (three regional education and training academies).

This exploratory factorial analysis consists of administering all the selected items with a five-position Likert scale, where each item can only take one value.

5.2. Measuring Instrument. The exploratory factor analysis AFE was used to extract and highlight the factors that form the school quality scale. And to answer the question about the reliability and validity of our measurement scale, the results of the tests we conducted demonstrate good psychometric properties. Our sample size is acceptable since it is estimated to be sufficient in the literature [32, 33]. Ultimately, our goal is to test the scale with the selected sample and not to generalize the results to the population.

Finally, a scale is assigned to each indicator to allow the measurement of items or specific and observable characteristics of the key concept representing the general objective, that is, for this study, identification of indicators measuring the pedagogical and administrative quality of schools (secondary cycle). As a result, a relative weighting is assigned to each dimension and indicator.

It should be remembered that following the first questionnaire, which was used in the qualitative phase of the process, the different items were identified and validated. A second questionnaire was drawn up, intended for the following actors: the school's administrative body (principals, general supervisors, censors, etc.), teachers, pedagogical inspectors, and planning and guidance inspectors; this

choice of carrying out the same questionnaire for different resource actors (difference in the field of intervention and roles) is explained by the large number of indicators to be "tested" with these resource persons.

Both questionnaires were pretested with five preservice inspectors, 10 secondary school teachers, and one principal to ensure clarity and understanding.

The second questionnaire was designed to validate the indicators that did not reach consensus in the previous version. This validation was measured by the consensus rate, a rate representing the percentage of approval of the indicators proposed by the professional resource actors. For an indicator to be considered valid, the criteria of clarity and relevance had to be evaluated individually. Thus, upon return of each questionnaire, the level of consensus was calculated for the characteristics (clarity and relevance) of each of the indicators listed. The approval rule for the key indicators was a progressive scale, graduated from 1 to 5: (1) none (or nonexistent), (2) little (or under construction), (3) quite (or existing), (4) very (or existing and operational), and (5) do not know. For each quality indicator, each of the stakeholders had to give an evaluation rating for each of the two criteria (clarity and relevance). For an indicator to be retained and considered valid, it had to be judged and recognized as very (4) or fairly (3) clear and very (4) or fairly (3) relevant by at least 80% of the key informants consulted. It should also be noted that, for each criterion, the resource people who answer the questionnaire have the possibility of suggesting other items that are deemed important and that do not appear on the list of items that constitutes the questionnaire. These suggestions were also collected and analyzed using a qualitative approach.

To succeed in this important phase, we have elaborated the second questionnaire in two versions: a version with paper format of six pages, distributed to the actors, and an electronic format "Google Forms." This choice was justified, on the one hand, considering the diversity of the target actors of our research, who have concerns and daily commitments, and on the other hand, this choice was taken, considering the constraint of time, which explains the choice of the electronic version (Google Forms).

We administered the questionnaire by post and mailing and several telephone follow-up procedures in order to obtain the maximum number of responses, given the relatively large number of our target population.

Certainly, the methodological approach adopted in this research is in line with the processes and principles of construction and validation of measurement instruments; however, it did not put all the factors of the quality of schools because some determinants are difficult to extract given the difficulty to measure or verbalize items or items were subjective. This study has enriched the existing literature for the validation of measurement scales assessing the quality of the school, which remains a very complex organization in deducing the quality.

5.3. Limitations. The methodological aspects used allowed the development of a measurement scale evaluating the quality of the school based on a pragmatic approach in the

development of quality indicators that allowed to having a basic device determining the quality for schools. Nevertheless, our study is not exhaustive and has limitations. It is true that the study involved the various actors within the school in order to extract dimensions and indicators evaluating quality, but we did not take into consideration the sampling of students on quality because of the organizational order of the schools, the availability of students, and the impossibility of organizing focus groups. There are differences in perceptions of the characteristics of quality among the actors [34, 35], which could be the subject of other research that could reinforce the dimensions raised by this study. These differences in perception could be the subject of other studies including, for example, the representations and expectations of students, which could constitute another pillar for making these perceptions more exhaustive in determining quality in the school. In addition, it is difficult to evaluate some dimensions such as financial ones.

6. Conclusion

The results of the confirmatory analysis applied to the structure of each stage of the evaluation process confirm the validity of the factorial structure resulting from the exploratory phase; based on a rigorous approach to the development of measurement scales, we conducted qualitative and quantitative studies via several samples of members of the national education system (school administrative bodies, administrators, and teachers).

We identified three (03) fields and ten (10) criteria with indicators that constitute the basis for an evaluation of quality. However, this study is not yet completed and its results represent the state of progress.

Regardless of the results obtained, our study has shown the possibility of constructing a process-based quality measurement scale.

Our methodology, which is based on Churchill's [4] approach, allows us to capture the complexity of the concept of quality by expressing it through multi-item scales. We have thus demonstrated that this methodology, which has proven itself in marketing and social science research, can be used to develop scales for measuring processes as complex as the evaluation process in the educational sector.

However, this study does have limitations: in addition to the size of our sample, with which we conducted a qualitative study of 74 actors, we used a consensus between groups of trainee inspectors and an expert group during validation.

Data Availability

The data used to support this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] V. T. Mohjbilou, "The strategic vision for the reform of the education, training and scientific research system," Higher Council for Education, Training and Scientific Research, 2021, <https://vision.csefrs.ma/integral/> (consulté le mai 2021).
- [2] 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, pp. 107–158, Article ID 3, 2002.
- [3] D. Billing and H. Thomas, "The international transferability of quality assessment systems for higher education: the Turkish experience," *Quality in Higher Education*, vol. 6, no. 1, pp. 31–40, 2000.
- [4] G. A. Churchill, "A paradigm for developing better measures of marketing constructs," *Journal of Marketing Research*, vol. 16, no. 1, pp. 64–73, 1979.
- [5] R. H. Hoyle, "The structural equation modeling approach: basic concepts and fundamental issues," in *Structural Equation Modeling: Concepts, Issues, and Applications*, R. H. Hoyle, Ed., pp. 1–15, Sage Publications, Inc, Thousand Oaks, CA, USA, 1995.
- [6] M. Tenenhaus, V. E. Vinzi, Y.-M. Chatelin, and C. Lauro, "PLS path modeling," *Computational Statistics & Data Analysis*, vol. 48, no. 1, pp. 159–205, 2005.
- [7] M. Crucianu, J.-P. A. de Beauville, and R. Boné, *Méthodes Factorielles Pour l'Analyse des Données: Méthodes Linéaires et Extensions Non-Linéaires*, Hermes Science Publications, Paris, France, 2004.
- [8] W. Gunawan, P. A. Creed, and A. I. Glendon, "Development and initial validation of a perceived future employability scale for young adults," *Journal of Career Assessment*, vol. 27, no. 4, pp. 610–627, 2019.
- [9] N. Gaudreau, É. Frenette, and 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, pp. 31–60, 2015.
- [10] O. M. Karatepe, U. Yavas, and E. Babakus, "Measuring service quality of banks: scale development and validation," *Journal of Retailing and Consumer Services*, vol. 12, no. 5, pp. 373–383, 2005.
- [11] D. Odry, *L'évaluation Dans Le Système Éducatif: Ce Que Vaut Notre Enseignement*, Mardaga, Belgium, 2020.
- [12] E. Minvielle and G. Pouvoirville, "Assessment of health care quality: a public health challenge," *Rev. Épidémiologie Santé Publique*, vol. 49, pp. 113–115, 2001.
- [13] Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century*, National Academy Press, Washington, DC, USA, 2001.
- [14] E. F. Fama and M. C. Jensen, "Separation of ownership and control," *The Journal of Law and Economics*, vol. 26, no. 2, pp. 301–325, 1983.
- [15] S. G. Sutton, "Toward an understanding of the factors affecting the quality of the audit process," *Decision Sciences*, vol. 24, no. 1, pp. 88–105, 1993.
- [16] C. Grenier-Sennelier, C. Corriol, V. Daucourt, P. Michel, and E. Minvielle, "Développement d'indicateurs de qualité au sein des établissements de santé: le projet COMPAQH," *Revue d'Epidémiologie et de Santé Publique*, vol. 53, pp. 22–30, 2005.
- [17] C. Bouchard, *La Qualité d'un Organisme: Construction et Expérimentation d'un Dispositif d'Indicateurs de la Qualité*, Corpus, Canada, 2002.

- [18] J.-B. Hardouin, *Construction d'échelles d'items unidimensionnelles en qualité de vie*, Ph.d thesis, Université René Descartes, Paris, France, 2005.
- [19] S. Maghnouj, J. Belanger, M. Clarke, E. Fordham, H. Kitchen, and I. McGregor, *Examens de l'OCDE du Cadre D'évaluation de L'éducation*, Maroc: OECD Reviews of Evaluation Framework Education, Morocco, 2018.
- [20] Y. Ozdemir, S. K. Kaya, and E. Turhan, "A scale to measure sustainable campus services in higher education: "sustainable service quality"," *Journal of Cleaner Production*, vol. 245, Article ID 118839, 2020.
- [21] P. Ramseook-Munhurrin and P. Nundlall, "Service quality measurement for secondary school setting," *Quality Assurance in Education*, vol. 21, no. 4, pp. 387–401, Article ID 4, 2013.
- [22] J. Plante and M. Côté, "TRIAGE: Technique for facilitating an expert group," 1993.
- [23] A. L. Delbecq and A. H. Van de Ven, "A group process model for problem identification and program planning," *The Journal of Applied Behavioral Science*, vol. 7, no. 4, pp. 466–492, 1971.
- [24] N. Roussiau, N. Bailly, and E. Renard, "Premières étapes de construction et de validation d'une échelle de spiritualité explicite areligieuse," *Pratiques Psychologiques*, vol. 24, no. 3, pp. 277–291, Article ID 3, 2018.
- [25] C. Dayan, R. Scelles, A.-M. Boutin, G. Ponsot, C. Arnaud, and 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, pp. 127–138, 2016.
- [26] A. E. Hurley, T. A. Scandura, C. A. Schriesheim et al., "Exploratory and confirmatory factor analysis: guidelines, issues, and alternatives," *Journal of Organizational Behavior*, vol. 18, no. 6, pp. 667–683, 1997.
- [27] A. B. Costello and J. Osborne, "Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis," *Practical Assessment, Research and Evaluation*, vol. 10, no. 1, p. 7, 2005.
- [28] R. L. Gorsuch, *Factor Analysis*, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1983.
- [29] R. F. DeVellis, *Scale development: Theory and applications*, Vol. 26, Sage Publications, , Thousand Oaks, CA, USA, 2016.
- [30] P. Roussel, "Chapitre 9. Méthodes de développement d'échelles pour questionnaires d'enquête," in *Méthodes & Recherches*, pp. 245–276, Louvain-la-Neuve: De Boeck Supérieur, Belgium, 2005.
- [31] R. Bertrand and J.-G. Blais, *Modèles de Mesure: l'apport de la Théorie des Réponses Aux Items*, PUQ, 2004.
- [32] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. Tatham, *Multivariate Data Analysis*, Pearson Education, London, UK, 2006.
- [33] S. Lagrosen, R. Seyyed-Hashemi, and M. Leitner, "Examination of the dimensions of quality in higher education," *Quality Assurance in Education*, vol. 12, no. 2, pp. 61–69, 2004.
- [34] A. Wilson, V. A. Zeithaml, M. J. Bitner, and D. D. Gremler, *Services Marketing: Integrating Customer Focus across the Firm*, McGraw Hill Education (India) Private Limited, UP, India, 2012.
- [35] A. Lai Fong Cheng and H. Keung Yau, "Principals' and teachers' perceptions of quality management in Hong Kong primary schools," *Quality Assurance in Education*, vol. 19, no. 2, pp. 170–186, 2011.