

ASSESSING SCHOOL PRINCIPAL LEADERSHIP PRACTICES. VALIDITY AND RELIABILITY OF A RUBRIC

8

(EVALUACIÓN DE LAS PRÁCTICAS DIRECTIVAS EN DIRECTORES
ESCOLARES: VALIDEZ Y CONFIABILIDAD DE UNA RÚBRICA)

Sergio Tobón

Centro Universitario CIFE, Cuernavaca, México

Luis Gibran Juárez-Hernández

Sergio Raúl Herrera-Meza

Centro Universitario CIFE, Cuernavaca, México

Cesar Núñez

University of Medellín, Colombia

DOI: 10.5944/educXX1.23894

How to reference this article/ Cómo referenciar este artículo:

Tobon, S.; Juarez-Hernandez, L. G.; Herrera-Meza, S.R. & Núñez, C. (2020). Assessing school principal leadership practices. Validity and reliability of a rubric. *Educación XX1*, 23(2), 187-210, doi: 10.5944/educXX1.23894

Tobon, S.; Juarez-Hernandez, L. G.; Herrera-Meza, S.R. y Núñez, C. (2020). Evaluación de las prácticas directivas en directores escolares: validez y confiabilidad de una rúbrica. *Educación XX1*, 23(2), 187-210, doi: 10.5944/educXX1.23894

ABSTRACT

Basic education has several priorities, such as improving students' levels of reading, calculation, and social interaction; changing pedagogical practices traditionally based on oral expositions and content learning; building a more inclusive school environment; improving security conditions; and ensuring that schools operate well. Although it is well known that rubrics are tools that allow for verifying achievements and accurately defining improvement needs, however, there are no available rubrics to assess principals' management regarding these priorities in basic education. The purpose of the present study was to design and validate an analytical rubric that could be used by principals to self-assess their practices and establish improvement actions that would benefit learning in elementary schools. The instrument was assessed by 10 judges. Aiken's V values higher than 0.75 were obtained in relevance, intelligibility of wording, and satisfaction with the instrument. The rubric was administered to a group of 645 elementary school principals, who found every item satisfactory, relevant, intelligibly worded, and suitable for use to help them to improve their managerial practices. A factor analysis was then carried out, and, in accordance with what was expected at the theoretical level, a single factor was found. Finally, it was established that the rubric has reliability of 0.877. It is concluded that the rubric for management practices self-assessment has adequate levels of content validity, construct validity, and reliability.

KEYWORDS

Principals; Primary School; leadership; measurement; rubric.

RESUMEN

La educación básica tiene diversas prioridades tales como lograr en los estudiantes mejores niveles de lectura, cálculo y convivencia; transformar las prácticas pedagógicas centradas en la exposición y el aprendizaje de contenidos; mejorar la convivencia y la inclusión; lograr instituciones seguras; y asegurar que las escuelas funcionen de manera regular, entre otras. Sin embargo, no se tienen rúbricas que permitan evaluar la gestión del director respecto a estas prioridades en la educación básica, teniendo en cuenta que las rúbricas son una herramienta que posibilita establecer con mayor rigurosidad los avances y aspectos a mejorar. El propósito del presente estudio fue diseñar y validar una rúbrica analítica para que los directores autoevalúen sus prácticas y establezcan acciones de mejoramiento que beneficien el aprendizaje en las escuelas de educación básica. El instrumento fue evaluado por 10 jueces y se obtuvieron valores de V de Aiken superiores a 0.75 en pertinencia, redacción y satisfacción con el instrumento; después, se aplicó la rúbrica a un grupo de 645 directores de escuelas de educación básica y hubo acuerdo en que el instrumento posee pertinencia, claridad y satisfacción para ser utilizado en la mejora de la gestión de los directores, a partir de un valor de V de Aiken superior a 0.75 en todas las variables evaluadas. A continuación, se hizo un análisis factorial y se encontró un único factor, acorde con lo esperado a nivel teórico. Finalmente, se estableció que la rúbrica posee una confiabilidad de 0.877. Se concluye que la rúbrica para autoevaluar las prácticas directivas posee adecuados niveles de validez de contenido, validez de constructo y confiabilidad.

PALABRAS CLAVE

Directores; educación básica; liderazgo; medida; rúbrica.

INTRODUCTION

Every nation or state recognizes that educating its people is one of the most relevant responsibilities a society must assume. For this purpose, every education national office devises complex systems, releases legal instruments, establishes institutions and authorities, issues curricular and operational guidelines, and selects, trains, and coaches stakeholders who will execute the educational policies and programs (Chamber of Deputies, 2018a; 2018b). An essential part of this social responsibility is the continuous evaluation of the pertinent execution of all these components. This represents a highly complex task, considering the wide variety of components, the multiple factors involved, and the different purposes and theoretical approaches for evaluation processes

(Organization for Economic Co-operation and Development [OECD], 2013, 2016; Department of Public Education [DPE], 2018).

It is necessary to assess teachers' and principals' performance in order to identify their specific training needs and provide them with advice, support, and courses to help them develop their activities according to what students and schools truly need. Assessing school principals' work is an issue that has been given special attention in recent years, either for recruitment and continuous improvement processes or for research purposes to identify the key characteristics and core competencies for student learning outcomes (OECD, 2013; CDHCU, 2018b). The principal's role implies several different practices in administrative, legal, labor, representation, and accountability aspects, as well as in pedagogical and human development aspects. For this reason, it has drawn special attention from various agencies committed to principals' performance improvement (Bolívar, Caballero, & García-Garnica, 2017; Elliott & Clifford, 2014; García-Garnica, 2016; OECD, 2016; Official Gazette of the Federation, 2019).

Researchers on school principals have developed several instruments. Abundant information about these instruments can be found in recent literature. Especially relevant is the VAL-ED, developed by Vanderbilt University and the University of Pennsylvania in the United States of America. This is an instrument that is administered to teachers and supervisors to assess principals' behavior and effectiveness according to the parameters of instructional leadership established by the Interstate Consortium for School Leadership of that country (ISLLC). It was designed to assess six aspects, corresponding to the basic instructional leadership components: 1) high standards for students' learning; 2) rigorous curriculum; 3) quality instruction; 4) culture of learning and teamwork; 5) connections to external communities, and 6) performance accountability. The informant responds to 72 items, which express pedagogical leadership behaviors, evaluating the principal's effectiveness in each on a 5-point Likert scale. Elliott and Clifford (2014) report VAL-ED's item construction and expert review processes, as well as the results of a pilot test conducted in 11 North American schools. After careful linguistic and contextual adaptations, VAL-ED has also begun to be used in Spain. High scores in Cronbach's alpha and the Spearman-Brown coefficient have verified the internal consistency and reliability of the Spanish version (Bolívar et al., 2017).

At the beginning of the 21st century, some instruments known as *logs* were developed to collect samples of school principals' activities throughout the day using electronic devices to record what they did and how they did it. An example of this is the experience sampling method (ESM), designed to measure principals' behaviors, beliefs, and attitudes along their normal activities and in the natural context of their daily routine (Spillane & Hunt, 2010). Another representative case of this model is the end of day (EOD) log, aimed at examining leadership practices and their relationship with student performance. The principal must take 15 minutes at the end of the day to register the activities he or she was performing at specific moments of the day, classifying each into one out of nine categories: building operations, finances and financial support for the school, community or parent relations, school district functions, student affairs, personnel issues, planning and setting goals, instructional leadership, principal professional growth. Due to the nature of these types of instruments and their content, reliability has been tested simply by collating principals' records with direct observations of their everyday activities, interviews, and student outcomes (May, Huff, & Goldring, 2012).

In Spanish-speaking environments, Moral and Amores (2014) developed a self-observation instrument or script based on the ESM to analyze principals' daily practices, with the specific purpose of detecting the obstacles that impede a full instructional leadership practice. The log-type instrument was used as a first phase of the study and complemented by semi-structured interviews with principals, aimed at categorizing their daily actions within some specific types of activity: pedagogical-curricular, institutional climate, professional growth, administrative management, and external relations. Data such as time spent, procedures used, people involved, places where these practices take place, and feelings experienced during the activities or tasks were also analyzed. Due to the descriptive and qualitative nature of this study, there is no statistical evidence of the relevant instrument's validity or reliability (Moral, Amores, & Rittaco, 2016).

The Effective Practices of Principals' Instructional Leadership questionnaire was also designed for the purpose of evaluating instructional leadership as a key characteristic of school principals (García-Garnica, 2016; García-Garnica & Caballero, 2015). This instrument was designed to be used in a 360-degree assessment process, that is, it collects information from various stakeholders around the school principal—such as heads of

study, teachers, and administrative staff—to determine how effectively the principal's practices influence instruction quality, strategic resource management, collaboration outside school, and the setting and evaluation of educational goals. The questionnaire consists of two scales: one to gather informants' appreciation of what truly happens in schools regarding principals' leadership behaviors and the other to indicate the degree to which certain leadership practices might be important to improve school operation. The questionnaire's content validity was assessed through the judgment of experts, who evaluated each item's relevance, intelligibility of wording, and adequacy and suggested amendments for the final version. The instrument, which has a high Cronbach's alpha reliability coefficient, was administered to 329 stakeholders from 207 public elementary schools in southern Spain.

With a broader perspective of the managerial role, as opposed to an exclusive focus on the leadership dimension, a model based on the parameters of the International Successful School Principalship Project (ISSPP) was designed in Mexico to analyze the characteristics that distinguish a successful principal, expressed in 18 managerial competencies grouped into 3 dimensions: school operation, social interaction, and educational program monitoring. The model was first reviewed by experts in different areas, such as ethics and values, educational administration, and competencies. It was then validated through a 122-item questionnaire, in which 114 of the items are presented as a Likert-type scale to measure how frequently a principal performs actions corresponding to the competencies defined in the model. There was a pre-pilot stage with 10 principals and a pilot stage with 58 principals, assistant principals, supervisors, and teachers working in Mexican schools. According to Cronbach's alpha coefficient, the instrument exhibits high levels of internal consistency (Villela-Treviño & Torres-Arcadia, 2015).

The studies that have been conducted on subjects related to school principals, as well as the instruments used in them, are rather descriptive and specifically focused on the leadership style principals adopt, especially its pedagogical or instructional dimensions. Little research, however, has been conducted on the subject for assessment and performance improvement purposes. In this sense, the most significant challenges lie, on the one hand, in establishing a definition of the analysis unit that enables quantitative and qualitative measurements that consider the complexity involved in school

principalship and, on the other hand, in ensuring improvement or metacognitive orientation. It is necessary to determine a unit that, despite being comprehensive, is not as wide as competency but also not as limited, narrow, or dualistic as completion or omission of certain tasks or actions.

In the present study, the concept of *practices* is chosen as an alternative to *competencies and tasks* so as to enable a precise and complete evaluation that can facilitate principals' performance improvement. Certainly, the concept of practices is found in the literature as a secondary and generic term or as a synonym for other terms, such as behaviors and actions (Bolívar et al., 2017), and has not been clearly defined, even in papers where the term *practice* appears stated as a study object (Izquierdo, 2016), for instance, when studies explore *good* or *successful practices* (Ibero-American States Organization, 2017). Some studies, however, refer to practices as important components of what is assessed or studied, for example, as elements that provide a more concrete notion for competencies provided they enclose and express principals' skills, knowledge, attitudes, values, and motivations (Barrientos-Piñeiro, Silva-García, & Antúnez-Marcos, 2016; Servín-Ramírez, & Vázquez-Sánchez, 2019). Similarly, studies that focus on standards as measurement units for performance refer to practices as indicators for such standards (Center of Study for Policies and Practices in Education [CEPPE], 2013). In the socioformative approach, practices are understood as actions mediated by principals to provide thorough education in schools by attaining active participation from every stakeholder: administrative staff, teachers, advisors, supervisors, teaching heads, parents, community people, politicians, and students (Tobon, 2018).

Recent research refers tangentially to certain characteristics of practices as elements that reflect implicit principal traits, such as values, beliefs, attitudes, and even identity (Bolívar & Ritacco, 2016; Crow, Day, & Møller, 2016; Gómez-Hurtado, 2013). At the same time, principals' practices contain, reflect, and actualize the educational policies inside the school and allow for verifying their accomplishments as well as the possible transformations they intend to operate (Contreras, 2016; Reyes, 2017). Rather than merely representing a series of operational tasks, however, practices constitute the implementation of the concrete and precise actions a principal performs every day to manage the prioritized organizational and instructional issues in a school and might be observable, verifiable, and even measurable at a certain point (Bolívar et al., 2017;

CEPPE, 2013) on the basis of criteria such as the following: 1) adhesion to educational purposes; 2) coherence regarding the specific educational approach; and 3) effectiveness regarding the expected achievements in terms of learning, thorough education, and the quality of life of all the members of the school community.

At the same time, school principals must confront in their schools a series of priorities that go beyond institutional leadership to focus on other value-added issues, such as (Official Gazette of the Federation, 2019) 1) seeking for all students to have acceptable levels of reading, writing, and calculation; 2) promoting peaceful coexistence; 3) accomplishing operation standards at the normality level, at least; 4) reducing school dropout and learning deficiency; 5) promoting a safe school environment; and 6) cooperating with teachers to improve their pedagogical practices in order to transform content-centered education into instructional processes based on problem solving through collaboration, transdisciplinarity, creativity, entrepreneurship, and continuous value strengthening, as proposed by socioformation (Arturo, 2019; Fuentes-Arismendi, 2019). The instruments designed to assess school principals' performance, however, seldom include these priorities, despite the strong need to address these priorities.

It is also important to analyze the types of instruments that have been proposed to assess principals' leadership. In most cases, they are estimation scales and checklists consisting of a series of indicators that do not allow principals to identify quality and progress in their performance, because they do not present descriptors for the different possible levels (Marin-Garcia, Ramirez Bayarri, & Atares -Huerta, 2015). Allowing principals to identify quality and progress in their performance is especially necessary when promoting self-assessment. For this reason, rubrics are preferred because they allow for identifying specific levels of attainment and help principals to determine what they need to do to perform at a higher standard. This is, no doubt, a very important feature to be considered for continuous training programs. It is therefore necessary to design rubric-type instruments in order to help principals assess their own achievements and their improvement gap.

Considering the above, the present study focused on the following objectives: 1) to design a rubric to assess principals' essential practices from the socioformative perspective in order to enhance their performance according to what have been established as management priorities in elementary schools; 2) to conduct an instrument

content validity study through experts' judgment in order to evaluate item relevance and intelligibility of wording ; 3) to determine the relevance, degree of understanding, and degree of satisfaction with the instrument for elementary school principals; and 4) to analyze the instrument's construct validity and reliability.

METHODOLOGY

Type of study

An instrumental study was conducted (Montero & León, 2007) to design a rubric-type instrument and evaluate its validity and reliability (Juárez-Hernández, 2018) according to the suggestions of Marín-García & Santandreu-Mascarell (2015). The rubric is intended to be used by school principals to assess their own management practices, taking into account certain school priorities, such as achieving the expected learning in students, developing essential skills for life, and transforming teachers' pedagogical practices. The rubric was administered through Google Forms.

Process phases

The study was developed in five phases, described below:

Phase 1. Rubric design and experts' review

An analytical rubric was devised according to the socioformative methodology (Sánchez-Contreras, 2019), through the following steps: 1) review of recently published instruments to assess principals' instructional leadership competencies; 2) analysis of several documents on the main priorities in elementary school education (Department of Public Education, 2019a, 2019b; Official Gazette of the Federation, 2019; UN, 2015); and 3) gathering opinions from three experts in school management regarding the priorities that should be assessed by the rubric. The analytical rubric was elaborated on the basis of these three sources of information; it was then reviewed by three experts with the following: 1) Ph.D. degree in education, with an emphasis on school management; 2) over 15 years of experience and expertise in instrument validation; and 3) at least five publications on school management.

Phase 2. Content validity study

Once the rubric had been prepared and its content had been enriched by the experts' collaboration, it was assessed by 10 judges with broad experience in school management and learning processes (see Table 1). This group assessed the rubric by employing a three-indicator scale to measure the relevance of the indicators, the intelligibility of the wording, and their level of satisfaction with the instrument. The first two indicators were assessed on a scale of 1 to 4 (where 1 represented a very low level and 4 represented a very high level), and the third indicator was measured on a scale of 1 to 5 (where 1 represented very low satisfaction and 5 represented very high satisfaction). The judges were also invited to provide suggestions for improving the instrument, such as adding or removing practices or modifying descriptors. To assess the degree of agreement in the experts' judgments, Aiken's V statistical procedure was used, and values higher than 0.75 were accepted (Penfield & Giacobi, 2004).

Table 1
Data of the judges who participated in the validity of content

N	10 judges
Sex (%)	Women: 20% Men: 80%
Age (mean + standard deviation)	46.5 (\pm 6.7)
Highest level of study completed (%)	Master's degree: 50% Doctorate: 50%
Areas of professional experience	School management
Years of experience as a principal (mean + standard deviation)	11 (\pm 4.6)
Number of hours of didactic training and evaluation in the past 2 years (mean + standard deviation)	265 (\pm 114.7)
Number of articles published in the area (mean + standard deviation)	1.9 (\pm 4.6)
Number of books published in the area (mean + standard deviation)	0.1 (\pm 0.31)
Number of book chapters (mean + standard deviation)	0.1 (\pm 0.31)
Number of lectures (mean + standard deviation)	2.8 (\pm 2.3)
Percentage of experts with experience in the review, design, and/or validation of a specific instrument for investigation or evaluation of educational processes	100%

Phase 3. Application in a pilot group

After content validation, the rubric was administered to a pilot group comprising 25 elementary school principals, all of whom worked in the Mexican state of Guanajuato. This was done to test the degree to which the instructions, practices, and indicators appeared comprehensible; to measure the average time needed to answer the rubric; to assess each item's relevance and correct writing; and to evaluate how satisfactory the rubric was from the perspective of users. This final point was tested using the same scale the judges had used. Likewise, the members of the pilot group were asked to provide suggestions on how the items' wording could be improved. Table 2 describes the characteristics of the pilot group.

Table 2
Data of the pilot group

N	25 principals of elementary schools
Sex (%)	Women: 71% Men: 29%
Age in years (mean + standard deviation)	40.04 (\pm 8.6)
Years of experience as a principal (mean + standard deviation)	7 (\pm 3.9)

Phase 4. Evaluation of relevance, writing, and satisfaction

After the pilot test, some amendments were made to the instructions and descriptors, primarily to address wording problems. The improved version of the rubric was then administered to a non-probabilistic sample composed of principals in charge of elementary public schools in the state of Guanajuato, Mexico. Principals were sent email invitations, and 645 principals answered, agreeing to participate. The respondents completed the instrument online and assessed the relevance, the intelligibility of the wording, and their satisfaction with the rubric, using the same scale the experts and pilot group participants had used. Sample characteristics are described in Table 3.

Table 3
Sociodemographic data of the sample

N	645 principals of public basic education schools
Sex (%)	Women: 70% Men: 30%
Age (mean + standard deviation)	38.97 (\pm 7.7)
Marital status (%)	Single: 24% Married: 69% Divorced: 5% Widower: 2%
Years of experience as a principal (mean + standard deviation)	2.0883 (\pm 2.4)

Phase 5. Construct validity and reliability

After the relevance, wording intelligibility, and satisfaction analysis, the construct validity of the rubric was analyzed using the same sample of 645 school principals. After the requirements had been verified using a correlation matrix, the Kaiser Meyer Olkin (KMO) index, and the Bartlett sphericity test (Costello & Osborne, 2005; Hefetz & Liberman, 2017; Juárez-Hernández, 2018), exploratory factor analysis (EFA) was applied. Subsequently, the main factor extraction method was selected (De Winter & Dodou, 2012; Hefetz & Liberman, 2017; Juárez-Hernández, 2018). The number of factors to be retained was set according to the Gutman-Kaiser rule (Gorsuch, 1983). Finally, the rubric's reliability was evaluated through Cronbach's alpha coefficient (Cronbach, 1951).

Ethical aspects

This study observed the requirements of Mexican personal data protection laws (Official Gazette of the Federation, 2010). All the participants had been informed about the purpose of the study and had signed online acceptance letters. All the participants had the option of ceasing to answer the rubric at any time if they so desired, without any

consequences. Once they had answered the rubric, all the participants were granted access to their results so that they could learn about some ways to improve their performance, as a result of the assessment process.

RESULTS

Rubric design

Hints from the three experts were integrated to construct the rubric, whose structure was based on the eight fundamental principalship practices (Table 4). Each item expresses a practice to be assessed using a 5-point scale containing performance descriptors ranging from the lowest to the highest or mostly desirable. In addition, a set of general demographic questions was included at the beginning of the rubric, such as age, sex, and years of experience as a principal. The rubric can be seen at the following link: <https://bit.ly/36xqrU7>

Table 4
Practices assessed by the rubric on principalship practices

Practices Description	Practices Description
Practice 1. Resource and space management	Degree of collaborative work to provide the school with the resources and physical spaces that are needed to achieve thorough education, school coexistence, recreation, and sports.
Practice 2. Expected learning	Degree to which concrete actions are managed and implemented to achieve students' expected learning according to curriculum standards, at the sufficient or basic level as a minimum.
Practice 3. Essential competencies	Degree to which all the school stakeholders implement specific and collaborative actions to assess and achieve the development of the main basic competencies: reading, writing, mathematics, and school coexistence.
Practice 4. Dropout	Degree to which every school stakeholder implements concrete and collaborative actions to reduce school dropout.
Practice 5. Pedagogical practices	Degree to which the principals implement actions to support the improvement of teachers' pedagogical practices, with an emphasis on problem solving, creativity development, and continuous-formative evaluation.

Practice 6. School safety	Degree to which stakeholders diagnose school safety issues and implement collaborative actions to prevent and manage disasters and contingencies.
Practice 7. Inclusion and coexistence	Degree to which inclusion and school coexistence are achieved by all stakeholders' participation in collaborative work.
Practice 8. Minimum normality	Degree to which the schools operate to provide educational service with normality, regularity, and punctuality, making full use of time in learning activities.

CONTENT VALIDITY

There was agreement among the judges that the principalship practices rubric is relevant, clearly formulated, and satisfactory, since every aspect obtained Aiken's V values higher than 0.75. This demonstrates that the instrument has content validity (Table 5).

Table 5
Content validity results

n=10

Variables	Mean (\pm Standard deviation)	Aiken's V
Items' pertinence to assess essential principalship practices (levels 1-4).	3.1 (\pm 0.316)	0.76
Intelligibility in the wording of questions and level descriptors (levels 1-4).	3.5 (\pm 0.527)	0.833
Satisfaction with the rubric (levels 0-5).	4.3 (\pm 0.674)	0.825

Application to the pilot group

Members of the pilot group consisting of 25 school principals also expressed agreement regarding the level of relevance of the rubric, the intelligible wording of practices and descriptors in the rubric, and satisfaction with the rubric because all the Aiken's V values were over 0.76. In addition, the rubric's average score in each of these variables was higher than 3.0, which is an acceptable level (Table 6).

Table 6

Assessment of the instrument by the pilot group

n=25

Application average time (in minutes) = 25 (\pm 6.5)

Variables	Mean (\pm Standard deviation)	Aiken's V
Items' pertinence to assess the essential principalship practices (levels 1-4).	3.33 (\pm 0.48)	0.776
Intelligibility in the wording of questions and level descriptors (levels 1-4)	3.625 (\pm 0.49)	0.875
Satisfaction with the rubric (levels 0-5)	4.1616 (\pm 0.56)	0.791

RELEVANCE, INTELLIGIBILITY OF WORDING, AND SATISFACTION

The application of the rubric to the 645-principal sample shows that there was a high level of agreement that the instrument has relevance, intelligible wording in the practices and descriptors, and acceptable satisfaction, since the Aiken's V values were, as expected, higher than 0.75 (Table 7). For the variables of relevance and intelligibility of wording, the minimum acceptable level in the mean was 3.0. In terms of satisfaction, a value of 4.0 or more indicated high satisfaction (the minimum acceptable level was 3.0).

Table 7

Assessment of the instrument by the target population

N= 645 principals

Items	Mean (\pm Standard Deviation)	Aiken's V
Items' pertinence to assess the essential principalship practices (levels 1-4).	3.30 (\pm 0.57)	0.768
Intelligibility in the wording of questions and level descriptors (levels 1-4).	3.50 (\pm 0.54)	0.834
Satisfaction with the rubric (levels 0-5)	4.17 (\pm 0.72)	0.794

CONSTRUCT VALIDITY AND RELIABILITY

Initially, the relevance of the application of the AFE was verified by means of the Bartlett's test results (X^2 : 2160.097; $p < 0.001$) and the Kaiser-Meyer-Olkin index (KMO: 0.920). It is specified that the correlations were statistically significant among all management practices (see Table 8).

Table 8
Correlation matrix
 Note: *= $p < 0.05$

	Practice 1	Practice 2	Practice 3	Practice 4	Practice 5	Practice 6	Practice 7	Practice 8
Practice 1	1							
Practice 2	.473*	1						
Practice 3	.448*	.564*	1					
Practice 4	.490*	.503*	.461*	1				
Practice 5	.486*	.571*	.485*	.530*	1			
Practice 6	.446*	.447*	.393*	.498*	.574*	1		
Practice 7	.491*	.537*	.425*	.528*	.592*	.578*	1	
Practice 8	.345*	.419*	.412*	.435*	.461*	.470*	.516*	1

Communalities showed adequate values represented in the factorial model (Table 9). A single factor was identified, explaining over 55% of the variance. This factor included the eight essential principalship practices proposed (with loads higher than 0.60) (Table 9).

Table 9
Communalities and factorial loads

	Communality	Factorial load
Practice 1	.417	.646
Practice 2	.524	.724
Practice 3	.416	.645
Practice 4	.499	.707
Practice 5	.595	.771
Practice 6	.490	.700
Practice 7	.581	.762
Practice 8	.382	.618

Finally, the reliability was obtained, which was 0.877, measured by Cronbach's alpha coefficient.

DISCUSSION

This research offers a new instrument to assess school principals' performance. It is called the Essential Principalship Practices Rubric (RPD-8). The study revealed that, based on the evaluation of experts in the area, this instrument has adequate levels of content validity, as the Aiken's V values were higher than 0.75, which is a highly acceptable score (Bulger & Housner, 2007; Penfield & Giacobbi, 2004). A content validity test is only acceptable if the judges possess the level of competence required to assess the instrument. There are several proposals in this regard. For example, Robles & Rojas (2015) suggest that a judge must have experience in the research field, investigative expertise, and knowledge regarding the process of designing instruments of the same nature as the one to be assessed. This requirement was fulfilled in the present study. All the expert judges presented over 10 years of experience in the area of school management; an average of 250 hours of professional updating in the past 2 years; at least one publication on the subject; and experience in the design, review, or improvement of instruments for research or evaluation of educational processes. Under these conditions, the process of assessment becomes reliable. On this basis, it can be said that, according to the experts, the content of the rubric is relevant, understandable, and highly satisfactory.

It is worth mentioning that not only was the content validity of the rubric assessed by expert judges, but the instrument's end users, principals currently leading elementary schools, were also included as informants to examine the degree to which they found the instrument helpful for their own professional growth. This has been a missing feature in other, similar, studies (Bolívar et al., 2017; García-Garnica, 2016). In this regard, it was found that there was a high level of agreement in the principals' opinions regarding the relevance, the intelligibility of the wording, and their satisfaction with the instrument, as they exhibited values over 0.75 in the Aiken's V index. This allows for stating that the instrument will probably be embraced by school managers as a pertinent tool for them to assess their own principalship practices, for improvement purposes.

In addition to content validity, relevance, and satisfaction, the RPD-8 has construct validity, an essential requirement in this type of study (Fernández-Cruz, Fernández-Díaz, & Rodríguez-Mantilla, 2018; Soriano, 2014). In this regard, the communalities were adequate, and, as predicted at the theoretical level, a single factor

was found to integrate the eight principalship practices. This is because all the practices refer to a level of leadership, mainly instructional, that principals must provide to ensure, at a minimum, acceptable standards of quality for students' education, that is, convergent decisions, resources, and actions to address real learning achievement for all students and the pertinence of such learning to facing the challenges encountered in society and in the relevant context (Bernal, Martínez, Parra, & Jiménez, 2015; Reyes, 2017). Finally, the rubric is reliable (Cronbach, 1951; George & Mallery, 2003), and its items exhibit internal consistency.

This study contributes a methodology to support principals, as long as, on the one hand, the sample used for validation is similar to the samples used in other important papers (Bolívar & Caballero, 2017; García-Garnica & Caballero, 2015). On the other hand, research on principalship is recent and has been focused on studying various aspects of school leadership and national policies. Little research, however, has been conducted on behalf of certain kinds of evidence-based continuous training and improvement for principals, using valid and reliable instruments (Flessa, Bramwell, Fernández, & Weinstein, 2018).

Principalship encloses huge heterogeneity in terms of goals, dimensions, and areas (Herrera & Tobón, 2017; López-Báez et al., 2018), and experts seldom agree on how to manage this in schools (Organization of Ibero-American States, 2017) because there are multiple criteria and different visions concerning this issue. Further, it is often difficult for a group of experts in the same research area to reach agreement on the content validity of instruments (Robles & Rojas, 2015). One of the most relevant outcomes of the present study was the high agreement level (measured by the Aiken's V index) in the appreciation of all 645 elementary school principals of the rubric's relevance, wording intelligibility, and adequacy. In addition to indicating high agreement levels, the scores exceeded the expected levels in these three variables (3.0). The high average score obtained in relevance (3.5 out of 4.0) might be explained by the fact that principals encounter common challenges around similar priorities at elementary schools (Reyes, 2017). This common awareness might be interpreted as progress in the field of school management, at least among the participants in this study. The high scores in wording intelligibility correspond to the fact that successive amendments were made before the rubric was presented to a large sample. Finally, the average score for satisfaction with the rubric was

over 4.0, which means a good or medium-high level. This suggests that the rubric responds to school principals' expectations for the content of an instrument to evaluate their management performance. This confirms the rubric's suitability for self-assessment purposes.

The RPD-8 shares some characteristics with other instruments in the same research area, such as the VAL-ED (Bolívar et al., 2017; Elliot & Clifford, 2014) and García-Garnica and Caballero's (2015) Effective Practices of Principals' Instructional Leadership, as it focuses on core aspects related to students' learning outcomes, such as (1) teaching practices enrichment, (2) the importance of collaborative work, (3) principals' daily actions and decisions, and (4) the availability of material resources. Notorious differences, however, can be observed, for instance, in the number of items and the specificity of some of the aspects included in the instrument. The RPD-8, with its eight items, purports to assess (1) schools' success in fostering the development of the main basic skills, such as reading, writing, calculation, and social interaction; (2) the transformation and improvement of pedagogical practices to ensure an education for the knowledge society; (3) the way school safety is managed in the face of both natural disasters and social risks, such as violence or drugs; and 4) the way inclusion and coexistence are managed and the achievements in terms of this aspect. These particular traits of the rubric respond to both the priorities in elementary school education and the socioformative vision of school principals as instructional leaders committed to a thorough human formation for the knowledge society and socially sustainable development, rather than as administration managers.

Important innovative features of the RPD-8 include its suitability to be used for predominantly formative evaluation processes (Cano, 2015) and its socioformative rationale, which represents a new educational approach that is increasingly being used in several countries (Berrelleza, 2019). Its essential characteristic is that, rather than merely focusing on certain stakeholders, it focuses on the entire education system to induce collaborative work on collective projects (Arturo, 2019). The eight essential principalship practices assessed by the instrument point to a school principal's comprehensive management to ensure an education that is authentically oriented to life and designed to progress beyond the traditional emphasis on simply learning content.

Although the analysis of the validity, relevance, adequacy, and reliability of the RPD-8 is positive, this was merely an exploratory study. To obtain conclusive results, it is necessary to conduct further studies, incorporating other regions and diverse populations both in Mexico and in other countries. In addition, other validity tests are required, such as consistency over time, convergent validity, and concurrent validity, considering their relationship with other instruments for school principals that have already been validated (Carretero-Dios & Pérez, 2005). To contrast with the findings presented here, it is also suggested that future studies employ confirmatory factor analysis.

REFERENCES

Arturo, C. (2019). The evaluation of competencies from the socioformación in transversal projects of basic education [La evaluación de competencias desde la socioformación en proyectos transversales de educación básica]. *Human Development and Socioformation (HUDS)*, 1(1), 1-21. <https://bit.ly/2uyPmJG>

Barrientos-Piñeiro, C., Silva-García, P., & Antúnez-Marcos, S. (2016). Management skills to promote participation: families in basic schools. *Educare*, 25(49), 45-62. <https://doi:10.18800/educacion.201602.003>

Bernal, D., Martínez, P., Parra, A., & Jiménez, J.L. (2015). Documentary research about quality in education in iberoamerican schools. *Revista Entramados. Educación y sociedad* (2), 107-124. <https://goo.gl/5bdfbh>

Berrelleza, N. E. (2019). The socioformative evaluation in the science subject: description of an experience in basic education. *Knowledge Society and Quality of Life (KSQ)*, 1(1), 1-21. <https://bit.ly/2RuJCJX>

Bolívar, A., Caballero, K., & García-Garnica, M. (2017). Multidimensional assessment of school leadership: keys to school improvement. *Ensaio: Avaliação e Políticas Públicas em Educação*. Publicación anticipada en línea. <https://doi.org/10.1590/S0104-40362017002500780>

Bolívar, A., & Ritacco, M. (2016). Impact of the Spanish model of school management on the professional identity of school leaders. *Archivos Analíticos de Políticas Educativas*, 24(119), 1-39. doi:10.14507/epaa.24.2512

Bulger, S. M., y Housner, L. D. 2007. Modified Delphi investigation of exercise science in physical education teacher education. *Journal of Teaching in Physical Education* 26 (1): 57-80.

Cano, E. (2015). The rubrics as an assessment tool of competency in higher education: use or abuse? *Profesorado. Revista de Currículum y Formación del Profesorado* 19(2), 265-280. Retrieved from: <https://goo.gl/52hUqy>

Carretero-Dios, H., & Pérez, C. (2005). Norms to design and review instrumental studies. *International Journal of Clinical and Health Psychology*, 5(3), 521-551.

Centre of Study for Policies and Practices in Education (CEPPE) (2013). *Learning Standards, Teaching Standards and Standards for School Principals: A Comparative Study*. Paris: OECD Publishing. doi: 10.1787/5k3tsjtp90v-en

Chamber of Deputies (Cámara de Diputados). (2018a). *General Education Law (promulgated in 1993) [Ley General de Educación]*. Mexico: Chamber of Deputies. Retrieved from: <https://goo.gl/vQPXLh>

Chamber of Deputies (Cámara de Diputados). (2018b). *General Law of the Professional Teaching Service (promulgated in 2013) [Ley General del Servicio Profesional Docente]*. Mexico: Chamber of Deputies. Retrieved from: <https://goo.gl/bvDY6r>

Contreras, T. (2016). Pedagogical leadership, teaching leadership and their role in school improvement: a theoretical approach. *Propósitos y Representaciones*, 4(2), 231-284. doi:10.20511/pyr2016.v4n2.123

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7), 1-9.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297- 334.

Crow, G., Day, C., & Møller, J. (2016). Framing research on school principals' identities. *International Journal of Leadership in Education*. doi: 10.1080/13603124.2015.1123299

Department of Public Education (Secretaría de Educación Pública). (2019a). *Framework for excellence in teaching and school management in Basic Education. Professional profiles, criteria and indicators for teachers, teaching technicians and staff with management and supervision functions* [Marco para la excelencia en la enseñanza y la gestión escolar en la Educación Básica. Perfiles profesionales, criterios e indicadores para docentes, técnicos docentes y personal con funciones de dirección y de supervisión]. Mexico: Department of Public Education. Retrieved from: <https://bit.ly/3aGGXVB>

Department of Public Education (Secretaría de Educación Pública). (2019b). *Guidelines for developing the continuous improvement school program* [Orientaciones para elaborar el programa escolar de mejora continua]. Mexico: Department of Public Education. Retrieved from: <https://bit.ly/37unPaV>

De Winter, J. C., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, 39(4), 695-710. doi: 10.1080/02664763.2011.610445

Elliott, S. N., y Clifford, M. (2014). *Principal assessment: Leadership behaviors known to influence schools and the learning of all students* (Document No. LS-5). Retrieved from: <https://goo.gl/3kvU34>

Fernández-Cruz, F. J., Fernández-Díaz, M. J., & Rodríguez-Mantilla, J. M. (2018). *Design and validation of an instrument to measure teacher training profiles in information and communication technologies*. *Revista Española de Pedagogía*, 76 (270), 247-270. doi: <https://doi.org/10.22550/REP76-2-2018-03>

Flessa, J., Bramwell, D., Fernández, M., y Weinstein, J. (2018). School Leadership in Latin America 200-2016. *Educational Management Administration & Leadership*, 46(2), 182-206. doi: 10.1177/1741143217717277

Fuentes-Arismendi, J. C., & Malagón-Micán, M. L. (2019). The evaluation of competencies from the socioformative approach. *Human Development and Socioformation (HUDS)*, 1(1), 1-13. Retrieved from: <https://bit.ly/37xunFz>

García-Garnica, M. (2016). Construction and validation of a questionnaire to measure effective practices of principals' pedagogical leadership. *Profesorado. Revista de Currículum y Formación de Profesorado*, 20(3), 493-526. Retrieved from: <https://goo.gl/YOIAcu>

García-Garnica, M., y Caballero, K. (2015). What Successful Practices of Leadership do the Andalusian Management Teams Develop in their Schools? *Revista Iberoamericana de Evaluación Educativa*, 8(2), 129-147. Retrieved from: <https://goo.gl/o2ayOh>

George, D. y Mallory, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4^a ed.)*. Boston, MA: Allyn & Bacon.

Gómez-Hurtado, I. (2013). Leadership and managing diversity in schools: towards an inclusive leadership. *Revista Fuentes*, 14, 61-84. Retrieved from: <https://goo.gl/LR1b5W>

Gorsuch, R.L. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Herrera, S. y Tobón, S. (2017). The school director from the socioformative approach. A documentary study through conceptual

mapping. *Revista de Pedagogía*, 38(102). 144-163. Retrieved from: <https://goo.gl/fG2sPn>

Hefetz, A., & Liberman, G. (2017). The factor analysis procedure for exploration: a short guide with examples. *Culture and Education*, 29(3), 526-562. doi: 10.1080/11356405.2017.1365425

Horn, A., & Murillo, F. (2016). Incidence of school principals on teachers' commitment: A multilevel study. *Psicoperspectivas*, 15(2), 64-77. doi: 10.5027/psicoperspectivas-Vol15-Issue2-fulltext-746

Ibero-American States Organization (Organización de Estados Iberoamericanos, OEI). (2017). Views on Education in Ibero-America. Professional development and leadership of school principals in Ibero-America (*Miradas sobre la Educación en Iberoamérica. Desarrollo profesional y liderazgo de directores escolares en Iberoamérica*). Madrid: OEI. Retrieved from: <https://goo.gl/ra1s9H>

Izquierdo, D. (2016). What do school principals do? Leadership practices in Spain drawn from PISA and TALIS international studies. *Revista Complutense de Educación*, 27(3), 1193-1209. doi:10.5209/rev_RCED.2016.v27.n3.4761 0

Juárez-Hernández, L. G. (2018). *Practical manual of basic statistics for research*. Florida: Kresearch.

López-Báez, I., López-Ramírez, E., Martínez-Íñiguez, J. E., & Tobón, S. (2018). Management directive: Approaches to a model for its institutional organization in higher secondary education in México [Gestión directiva: Aproximaciones a un modelo para su organización institucional en la educación media superior en México]. *Espacios*, 39(29).

Marín-García, J. A., y Santandreu-Mascarell, C. (2015). What do we know about rubrics used in higher education? *Intangible Capital*, 11(1), 118-145. Doi: <http://dx.doi.org/10.3926/ic.538>

Marin-Garcia, J., Ramirez Bayarri, L., & Atares-Huerta, L. (2015). Protocol: Comparing advantages and disadvantages of Rating Scales, Behavior Observation Scales and Paired Comparison Scales for behavior assessment of competencies in workers. A systematic literature review. *WPOM-Working Papers on Operations Management*, 6(2), 49-63. doi:<https://doi.org/10.4995/wpom.v6i2.4032>

May, H., Huff, J., y Golding, E. (2012). A longitudinal study of principal's activities and student performance. *School Effectiveness and School Improvement*, 23(4), 417-439. doi: 10.1080/09243453.2012.678866

Montero, I., y León, O. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862.

Moral, C., Amores, F.J., y Rittaco, M. (2016) Distributed Leadership and Capacity for Improvement of Secondary Schools. *Estudios sobre educación*, 30, 115-143. doi: 10.15581/004.30.115-143

Moral, C., y Amores, F.J. (2014). Resistant Architecture Determinant of Educational Leadership in Secondary Schools. *Bordón*, 66(2), 121-138. doi: [10.13042/Bordon.2014.66208](https://doi.org/10.13042/Bordon.2014.66208)

Official Gazette of the Federation (Diario Oficial de la Federación) (2019, September 30). *Decree by which the General Law of Education is issued and the General Law of the Educational Physical Infrastructure is repealed [Decreto por el que se expide la Ley General de Educación y se abroga la Ley General de la Infraestructura Física*

Educativa]. Retrieved from: <https://bit.ly/2vplpMx>

Official Gazette of the Federation (Diario Oficial de la Federación). (2010, July 5). *Federal Law on protection of personal data held by individuals [Ley Federal de protección de datos personales en posesión de los particulares]*. Retrieved from: <https://bit.ly/2U0kImU>

Organisation for Economic Co-operation and Development. (2013). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)*. PISA. OECD. doi: 10.1787/9789264201156-en

Organisation for Economic Co-operation and Development. (2016). *School Leadership for Learning: Insights from TALIS 2013*. Paris: OECD. doi: 10.1787/9789264258341-en

Penfield, R.D., y Giacobbi, Jr, P.R. (2004). Applying a score confidence interval to Aiken item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8, 213-225.

Reyes, A. (2017). School autonomy and educational change: considerations in light from the implementation of PEC-FIDE. *Revista Electrónica de Investigación Educativa*, 19(2), 12-21. doi:10.24320/redie.2017.19.2.643

Robles, P., & Rojas, M. D. C. (2015). Validation by expert judgements: two cases of qualitative research in Applied Linguistics. *Revista Nebrija de Lingüística Aplicada*, 2015(18). Retrieved from: <https://goo.gl/HeXKwF>

Sánchez-Contreras, M. L. (2019). Socioformative taxonomy: A referent for didactics and evaluation [Taxonomía socioformativa: Un referente para la didáctica y la evaluación]. *Forhum International Journal of Social Sciences and Humanities*, 1(1), 100-115. doi: dx.doi.org/10.35766/jf19119

Servín-Ramírez, E., & Vázquez-Sánchez, C. (2019). The school leader and his transition towards socioformation, a fundamental factor in education for sustainable development [El directivo escolar y su transición hacia la socioformación, un factor fundamental en la educación para el desarrollo sostenible]. *Ecocience International Journal*, 1(1), 33-36. <https://doi.org/10.35766/je19115>

Soriano, A. M. (2014). Design and validation of measurement instruments. *Dia-logos*, (14), 19-40. Retrieved from: <https://bit.ly/2sW4Arw>

Spillane, J., y Hunt, B. (2010). Days of their lives: a mixed-methods, descriptive analysis of the men and women at work in the principal's office. *Journal of Curriculum Studies*, 42(3), 293-331. doi: 10.1080/00220270903527623

Tobón, S. (2018). *School management project. Socioformation and pedagogical practices (Proyecto de gestión escolar. Socioformación y prácticas pedagógicas)*. Florida: Kresearch. Retrieved from: <https://goo.gl/D4sT8Q>

UN. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved from: <https://bit.ly/2RXb022>

Villela-Treviño, R., y Torres-Arcadia, C. (2015). Using a competency model to assess school principalship. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 8(2), 41-56. Retrieved from: <https://goo.gl/I93MMH>

Academic and professional profile of the authors:

Sergio Tobón. ORCID: <http://orcid.org/0000-0001-5543-9131>

Doctor in Educational Models and Cultural Policies. Globalization and Identity in the Knowledge Society, by the Complutense University of Madrid. He is a researcher at the CIFE University Center, Cuernavaca, Mexico (<https://www.cife.edu.mx/>) and Ekap University. Email: stobon5@gmail.com

Luis Gibrán Juárez-Hernández. ORCID: <http://orcid.org/0000-0003-0658-6818>

PhD. in Biological and Health Sciences from the Autonomous Metropolitan University. Professor-Researcher of the CIFE University Center. His lines of research are ecology, diagnosis and environmental management. Email: luisgibran@cife.edu.mx

Sergio Raúl Herrera-Meza. ORCID: <http://orcid.org/0000-0003-2605-4878>

Doctorate in Socioformation and Knowledge Society. Professor and researcher at CIFE University Center. His current research focuses on training processes and management of educational institutions. Expert in design of research and assessment. Email: sergioherrera@cife.edu.mx

Cesar Núñez. ORCID: <http://orcid.org/0000-0001-8925-993X>

Mg. Associate professor. Director of the research, development and innovation system in the Psychology Program of the University of Medellín, Colombia. Director of the research group in Psychology and Clinical-Social Processes at the same University.

Email: cnunez@udem.edu.co

Date Reception of the Article: February 15, 2019
Article Modification Date: July 08, 2019
Article Acceptance Date: July 29, 2019
Revision Date for Publication: January 27, 2020