Instrument to Measure Teacher Learning as a Product of the Performance Evaluation Process

Oscar Luis Ochoa[1], Belia Cháidez[2], Gonzalo Arreola[3]

[1]Universidad Pedagógica de Durango, Carretera al Mezquital Km 4.2 s/n Fracc, Predio Rústico Calleros, 34162 Durango, México
[1]*chokar128[at]hotmail.com*
[2]*beliachaidez05[at]hotmail.com*
[3]*garreolamupd[at]hotmail.com*

Abstract: *The objective of this work was to build an instrument to measure teacher learning as a product of the performance evaluation process (IEAD). The study was considered instrumental and its structure was made up of 43 items measurable through a Likert scale, distributed as follows: 17 items in dimension 1, "class organization"; 16 items in dimension 2, "class management" and; 10 items in dimension 3, "evaluation". For data collection, an intentional sample of 108 academic figures who participated in the application of the IEAD was used. In the measurement of its psychometric properties, the reliability statistic of Cronbach's alpha yielded a value of 0.98 and, in the construct validity carried out through the "total domain" correlation test, Spearman's Rho coefficient values were obtained that ranged between 0.91 and 0.98; the validity of these results were confirmed by the multiple linear regression test and among them, the confirmation of the absence of collinearity between the dimensions of the IEAD through the result of the Variance Inflation Factor (VIF) and Durbin Watson (D) statistics stands out. Finally, the numerical description of the level of learning shown by the 108 teachers evaluated according to the Likert scale, reflected that the IEAD objectively measures the construct for which it was designed.*

Keywords: Instrument for measurement, learning, teacher evaluation

1. Introduction

Teacher evaluation is one of the areas within educational research that has stood out for being an object of interest; in its origins, at the beginning of the 20th century, evaluating teachers consisted solely of supervising their work, where the sole purpose was to inspect the work they carried out within the educational institution; actually, in most cases efforts are made to elaborate and develop new evaluation models that contribute to the improvement of educational quality. Through time, both European and American countries have proposed teacher evaluation models whose objectives are aimed at strengthening the professional development of teachers, as it is considered to be one of the main ways to improve educational quality.

1.1 Statement of the problem

On September 4, 2013, in the plenary session of the Senate of the Republic, the General Law of Professional Teaching Service (LGSPD) [1] was approved and it was in 2015 when the Evaluation of Teaching Performance was implemented, with a first group of personnel in service at the national level, based on what is established in the Profiles, Parameters and Indicators (PPI) instrument that identifies the basic performance characteristics of the Professional Teaching Service Personnel. In this regard, the teachers stated that they agreed with the evaluation of their performance but not with the fact that their permanence in the service was threatened, a situation that ultimately caused great social problems in the country.

With the awareness of the problems that originate the evaluation processes faced by teachers, specifically related to the performance of their educational practice; in this study, the challenge was taken to build an instrument that would offer reliable information on the degree of learning obtained by teachers as a result of their performance evaluation proposed by the Secretary of Public Education (SEP).

1.2 Objective

Build an Instrument to measure teacher learning as a product of the performance evaluation process (IEAD).

2. Theoretical Framework

The learning that is generated in class, has as a preponderant factor the ability of the teacher to conduct his educational process, the quality of the teaching practices of the teachers is one of the school factors that has the greatest impact on student learning; improving the educational service offered in Basic Education schools means strengthening the work of teachers in the classroom.

To strengthen the abilities of teachers, it is necessary to identify those aspects that can be improved in their teaching practice, and this can be achieved through performance evaluation; in this way, the evaluation of teacher performance will contribute to the strengthening of teaching practices, because its results will provide information that can be used to improve their learning, in addition, this information will serve as a reference to develop their continuous training process.

In congruence with the observations related to learning and teaching performance, this research work was developed based on the specific description of Carlos Marcelo (2009)[2], who states that teacher professional development is conceived as a learning process, not linear and evolutionary, whose result is not only perceived in the

change of teaching practices, but also in the thought about the how and why of this practice. This thinking is what is called a "personal interpretive scheme": a set of mental reflections and representations that operate as lenses through which teachers view their own work and give it meaning.

3. Methodology

The objective of this research was the construction of an Instrument for the previously established purposes, for this reason, the study was considered instrumental in nature since "(…) All studies aimed at developing tests and devices are considered to belong to this category, including both their design (and adaptation) and the study of their psychometric properties". Montero y León [3].

3.1 Instrument Structure

The construction of the IEAD was supported by the operationalization of the dimensions of two instruments: a) the Class Observation Instrument (IOC), which was an adaptation of the Tsang Hester Observation Rubric (THOR) instrument, designed to measure the performance of teachers in the dimensions "class management", "teaching conduction" and "learning evaluation", "(…) Is based on the study by Tsang (2004) whose original version has already been evaluated and validated by Good (2006)". Contreras, et al. [4]. The IOC was adopted to the Chilean educational context and was built with the objective of being able to determine the level of performance of teachers that is practiced in that context, andb) the instrument Profiles, Parameters and Indicators (PPI), proposed by the Secretary of Public Education, SEP [5]. The structure of the IEAD was made up of 43 items distributed as follows: 17 items in dimension 1, "class organization" (the name of this dimension was instead of "class management" for reasons favorable to the Mexican context); 16 items in dimension 2, "development of class" and; 10 items in dimension 3, "evaluation"; the IEAD is in the Appendix section (see appendix 1). The response to the items of the IEAD dimensions were rated with a Likert scale of four levels suitable for measuring teacher performance; that is, to their intervention in the classroom in promoting their students' learning.

3.2 Sample selection

The IEAD was built to measure teacher learning as a product of the performance evaluation process in the three levels of basic education and, to obtain the appropriate information, the instrument was applied to 108 academic figures to evaluate their teachers, including: Heads of Sector, Supervisors, Directors, Deputy Directors and Technical Pedagogical Advisors, all knowledgeable about the work that their teachers should and did at the time of the evaluation. The selection of the sample was of the "intentional" type; this procedure is conceptualized by Otzen and Manterola [6], as follows, "It allows selecting characteristic cases from a population, limiting the sample only to these cases. It is used in scenarios in which the population is highly variable and consequently the sample is very small".

3.3 Information analysis tools

The classification and organization of the data and the psychometric tests of the IEAD were carried out with the use of the Excel spreadsheet and the Statistical Package for the Social Sciences SPSS program in version 22, respectively.

3.4 IEAD scores

The scores obtained in the IEAD application were the responses that the 108 participants gave to each of the items, being coded for the SPSS program as follows: a) Teaching performance (DESDOCEN), contains the score of all the items; b) Dimension 1, Organization of the class (D1ORCLAS), contains the score of its items; c) Dimension 2, Conduct of the class (D2COCLAS), contains the score of its items and; d) Dimension 3, Evaluation (D3EVALUA), contains the score of its items. Based on these variables, the scores equivalent to the level of the Likert scale obtained by each participant were determined, this in order to be able to know the level of learning obtained by the teachers who participated in the teaching performance evaluation process; these were coded as follows: a) Global learning of teachers (APRENDOC); b) Teacher learning in dimension 1, "class organization" (APORGCLA); c) Teacher learning in dimension 2, "class management" (APCONCLA) and; d) Teacher learning in dimension 3, "evaluation" (APEVALUA).

3.5 Normality test

According to the study sample, the test was performed using the Kolmorogov Smirnov statistic; about its application the authors Herrera, et al. [7] state that the Kolmogorov Smirnov and Lillierfors test is more convenient when the variable being analyzed is continuous or ordinal and is more effective for large samples. The normality tests were carried out according to the following hypothesis statement:

$H_0 = The variable has a normal distribution$
$H_1 = The variable does not have a normal distribution$
   The decision rule:
$if value "p" > \propto$, accept H_0

3.6 IEAD reliability test

The instrument's "reliability" test was carried out by means of the Cronbach's alpha coefficient, the concept of reliability implies the amount of error that is committed when making any measurement, in educational practice it is common to doubt about the repeatability of a test; if a result is not reproducible, then the value and utility of the test is considered poor. About this concept Dueñas [8], believes that "Reliability indicates the precision (consistency and stability) of the measurement of a test. It gives us the accuracy or precision with which the scale of a test gives us the true scores, and deals with obtaining the same results on different occasions, by the same group of subjects".

3.7 Construct validity test

The construct validity test refers to the "full domain" correlation; In other words, the correlation test of each of the IEAD domains is carried out with respect to its total domain,

in this way the correct selection of the groups of items is confirmed.

This test is validated by Morey [9] who states that "The evaluation of construct validity is multiple, and includes:"(…) correlational studies with a wide variety of relevant measures".To determine the existence of a correlation between the interval or ratio variables with a non-parametric distribution, Martinez, et al. [10] confirms that "Spearman's correlation coefficient is advisable to use it when the data present extreme values, since these values greatly affect the Pearson correlation coefficient, or when faced with non-parametric distributions.It is not affected by changes in the units of measure". The bivariate correlation test between the domain variables and with respect to the total domain variable was determined based on the following approach:

$$H_0 = there is no linear relation between variables$$
$$H_1 = there is a linear relation between variables$$

The decision rule:
$$if\ value\ "p" \leq \propto, accept\ H_0$$

3.7.1 Confirmatory validity test

In order to complement the construct validity of the IEAD, the multiple regression test was performed; for this purpose, the DESDOCEN variable was considered as the dependent variable, and the D1ORCLAS, D2COCLAS and D3EVALUA variables as independent variables, according to the following model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K + E$$

3.7.1.1 ANOVA test

The ANOVA test was used to determine the existence of an association between the domain variables, according to the following hypothesis statement:

$$H_0 = there\ is\ no\ association\ between\ variables$$
$$H_1 = there\ is\ an\ asociation\ between\ variables$$

The decision rule:
$$if\ value\ "p" \leq \propto, reject\ H_0$$

3.7.1.2 "t" test

A second inferential test that considers the values of the coefficients of the model and that determines whether the information provided by the independent variables to the explanation of the dependent variable is significant, is the hypothesis test by means of the linear regression coefficient through the test "t", the approach is as follows:

$$H_0 = the\ VI\ does\ not\ provide\ significant\ information$$
$$H_1 = the\ VI\ provides\ significant\ information$$

The decision rule:
$$si\ valor\ "p" \leq \propto, reject\ H_0$$

3.7.1.3 R2 determination test

The third test consists of determining the coefficient of determination R2, a statistic that reflects the variation of the dependent variable caused by the independent variable. The Durbin Watson (D) Botero and Vertel [11] statistic is found in this same table; the value that indicates the non-existence of collinearity between the independent variables must be close to "2"; the proposed range was as follows:

$$0.5 < D < 4$$

3.8 Teacher global learning APRENDOC

This section shows the result corresponding to the level of learning obtained by the teachers evaluated in the application of the IEAD; that is, the global score achieved in the response to the items of the three dimensions of the Instrument. To obtain the result, the range of the DESDOCEN score was divided equally and the appropriate range was obtained to locate each level of learning on the Likert scale.

4. Results

4.1 Normality of scores

In table number 1, you can find the concentrate of the normality test results for each score, it can be seen that only the score of dimension three obeys a normal distribution, for this reason, in the correlational tests the Spearman Rho statistic was used.

Table 1: Normality test concentrate

Punctuation	KS	"p"	rel	α	Dist
DESDOCEN	0.093	0.022	<	0.05	not
D1ORCLAS	0.094	0.02	<	0.05	not
D2COCLAS	0.09	0.03	<	0.05	not
D3EVALUA	0.079	0.09	>	0.05	yes

4.2 IEAD reliability

Table number 2 indicates the result of the reliability test and the number of items processed, the result of the psychometric property of reliability of the IEAD yielded a value of the Cronbach's alpha statistic of 0.984; this value indicates an excellent reliability of the Instrument.

Table 2: IEAD reliability test

Reliability test	
Cronbach's alpha	N of elements
.984	43

4.3. Construct validity results

4.3.1 Total domain correlation

Table number 3 shows the concentration of the results of the Spearman Rho correlation test, between dimensions and with respect to the total domain. The value of the correlation statistic obtained between the different combinations of variables, oscillates between 0.84 and 0.97, quite acceptable values and indicative of the relation between their scores and the validity of the instrument.

Table 3: Total Domain correlation

		DD	D1	D2	D3
DD	Rho	1	.973**	.978**	.918**
	Sig.	.	0	0	0
	N	108	108	108	108
D1	Rho	.973**	1	.932**	.847**
	Sig.	0	.	0	0
	N	108	108	108	108
D2	Rho	.978**	.932**	1	.872**
	Sig.	0	0	.	0
	N	108	108	108	108
D3	Rho	.918**	.847**	.872**	1
	Sig.	0	0	0	
	N	108	108	108	108

3.3.2 Validity confirmation

3.3.2.1 ANOVA statistic
Table number 4 shows the "p" value (sig.) and according to the hypothesis statement, the existence of an association between the variables involved was confirmed.

Table 4: ANOVA test result

	Sum of squares	gl	Square mean	F	Sig.
Regression	106621.9	3	35540.6	108060	.000b
Residue	34.205	104	0.329		
Total	106656.1	107			

a. Dependent variable: EVALDDOC
b. Predictors: (constant), D3EVALUA, D1ORCLAS, D2COCLAS

3.3.2.2 "t" Statistic
Table number 5 shows the value of the coefficients of the regression model (B) and, according to the hypothesis contrast, it was determined that the independent variables significantly influence the regression analysis; It was also observed that in none of the three dimensions the "p" value (sig.) exceeds the value of "10" in the Variance Inflation Factor (VIF), indicating that there is no autocorrelation between dimensions of the IEAD, Guerra [12].

Table 5: "t" test result

coefficients					
Model	β	t	Sig.	Tolerance	VIF
1 (constant)	.087	.423	.673		
D1 ORCLAS	1.02	82.5	.000	0.12	8.3
D2 COCLAS	1.04	73.6	.000	0.1	9.2
D3 EVALUA	.98	57.3	.000	0.2	4.8

Dependent variable: EVALDDOC

3.3.2.3. R2 statistics
In table number 6 you can see the value of R^2, indicative of the measure of variation of the response variable that is explained, by the behavior of the regression variable Carrasquilla et al. [13]. In this case, its adjusted value is equal to "1", which means that the independent variables explain the behavior of the dependent variable by 100%.

Table 6: R^2 statistics

Model Summary				
Model	R	R square	Standard error	Durbin atson
1	1.000a	1	0.573	2.163

a. Predictors: D3 EVALUA, D2COCLAS, D1 ORCLAS
b. Dependent variable: EVALDDOC

5. Results of the application of IEAD

In table number 7, there is the information that indicates the level of learning obtained by the teachers evaluated through the IEAD (APRENDOC); in the percentage values it is observed that 41.7% of the teachers obtained an evaluation of null or low learning, while 58.3% of them improved significantly; especially in level three "improved a lot", with 41 teachers of the 108 participants.

Table 7: Teacher learning

APRENDOC	Frec	%	% valid	% accumulated
Did not improve	20	18.5	18.5	18.5
Improved something	25	23.1	23.1	41.7
Improved a lot	41	38	38	79.6
Improved too much	22	20.4	20.4	100
Total	108	100	100	

The graph of figure number 1 shows the data described in table number 7; it is observed that only 20% of the teachers, equivalent to 22 of them, obtained outstanding learning in their performance evaluation process, pointing out that this measurement is also the product of their performance in educational practice.
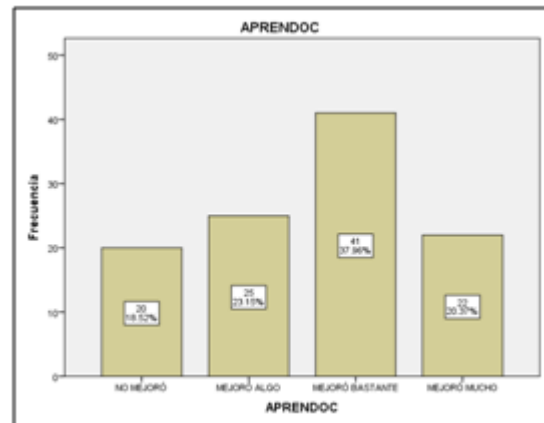

Figure 1: Teacher learning graph

6. Conclusions

The IEAD was built in the Mexican context and had as a reference the IOC instrument adapted in Chile and the PPI instrument developed in Mexico, this situation available an adequate operationalization of the variables of both instruments, a process that had a positive impact on the psychometric properties of reliability and construct validity of the IEAD. The first was understood as a statistical reliability required to guarantee the validity and precision of the statistical analysis focused on the consistency of the measures of its application, the value of the Cronbach's alpha coefficient obtained was 0.94, this value is indicative that the learning result achieved by teachers in the performance evaluation process is attached to reality and that as many times as it is applied in similar contexts, it will reproduce the results with a minimum error. The result of the various tests of construct validity were affirmative and through the "VIF" and "D" tests, the non-existence of collinearity between the independent variables was confirmed. Based on the results of the psychometric properties of reliability and validity exhibited by the IEAD, it can be affirmed that the Instrument is capable of measuring the learning that the teacher exhibits in the performance of their educational practice in each of its dimensions.

7. Recommendations

The indicators of each one of the dimensions of the IEAD are the reflection of the daily activities that the teacher carries out in his educational practice, for this reason, the

instrument not only evaluates teacher learning in its performance evaluation process that originated in the framework of the General Law for Professional Teaching Service (LGSPD) Teaching that was implemented in Mexico in 2015, but also evaluates the performance that he performs with his students on a daily basis.

On the other hand, even when the Instrument was piloted with primary school teachers, precisely because of these characteristics it is considered appropriate to evaluate the performance of upper and middle school teachers.

References

[1] Ley General del Servicio Profesional Docente (2017). Leyes. mx.com. LGSPD. Recuperado de http://leyes-mx.com/ley_general_del_servicio_profesional_docente/52.htm

[2] Marcelo, C. (2009). La evaluación del desarrollo professional docente: de la cantidad a la calidad. *Revista Brasileira de Farmacao de Professores-RBFP*, 1(1), 43-70

[3] Montero I. &León, O. (2019). A guide for naming research studies in Psychology. (E. Universidad Autonóma de Madrid, Ed.) International Journal of Clinical and Health Psycology, 7 (3), pp. 847-862, 2019. Recuperado de https://www.redalyc.org/pdf/337/33770318.pdf

[4] Contreras, I., Rittershaussen, S., Correa, E., Solís, C., Correa, E., Solís, C., Nuñez, C. & Vásquez, N. (2013) IOC un instrumento para cualificar desempeño docente en aula: Su generación y validación. Revista Estudios Pedagógicos. Vol, XXXIX (2), pp. 85-96, 2013. Recuperado de: https://www.redalyc.org/pdf/1735/173529673006.pdf

[5] Secretaría de Educación Pública (2018). Perfiles Parámetros e Indicadores para Docentes y Técnicos Docentes en Educación Básica. Recuperado de: http://www.sec.gob.mx/portal/docs/destacados/2018/01/12/PPI_DOC_TECNICO_DOCENTES_080118.pdf

[6] Otzen, T. & Manterola, C. (2017). Técnicas de Muestreo sobre una Población de Estudio. Int. J, 35(1), pp. 227-232. Recuperado de: https://scielo.conicyt.cl/scielo.php?pid=S0717-95022017000100037&script=sci_abstract

[7] Herrera, M., Caridad, C. W., Sarduy, L., García, Y.& Martínez, C. E (2011). Diferentes métodos estadísticos para el análisis de variables discretas. Una aplicación en las ciencias agricolas y técnicas. *Revista Ciencias Técnicas Agropecuarias, 21(1)*. Recuperado de: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2071-00542012000100011&lng=es&nrm=iso>. ISSN 2071-0054.

[8] Dueñas, M. (1998). El libro de los tests. Planeta Mexicana, S. A. de C. V. Distrito Federal, México

[9] Morey, L. C. (2011). Evaluación del inventario PAI. Inventario de Evaluación de la Personalidad (PAI). C. G. Psicólogos, Ed. Madrid, España. Recuperado de: cop.es/uploads/PDF/2013/PAI.pdf

[10] Martínez, R. M., Tuya, L. C., Martínez, M., Pérez, A. & Cánovas, A. M. (2009). El coeficiente de correlación de los rangos de spearman. *Revista Ciencias Médicas, la Habana, VIII, (2), pp. 1-19*. Recuperado de: https://www.redalyc.org/pdf/1804/180414044017.pdf

[11] Botero, L., Vertel, M. (2006). Modelo matemático aplicado a la curva de lactancia en Ganado vacuno doble propósito. *Revista MVZ Córdoba, vol 11, (1), pp. 765-795*. Recuperado de: http://www.scielo.org.co/pdf/mvz/v11n1/v11n1a10.pdf

[12] Guerra, C. W., Herrera, M., Vázquez, Y.& Quintero, A. B. (2014). Contribución de la Estadística al análisis de variables: categóricas: Aplicación del Análisis de Regresión. Revista Ciencias Técnicas Agropecuarias, 23, (1), pp. 68-73. doi:1010-2760

[13] Carrasquilla, A., Chacón, A., Nuñez, K., Gómez, O., Valverde, J.,& Guerrero, M.(2016). Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal. *Tecnología en Marcha, pp. 33-45*. doi:10.18845/tm.v29i8.29

Appendix

Appendix 1

Instrument to Evaluate Teaching Learning	
IEAD	
Likert scale (teacher learning)	
Level 1: did not improve	
Level 2: improved something	
Level 3: improved a lot	
	Level 4: improved too much
I CLASS ORGANIZATION	
no	ítem
1	Does the teacher know how to listen to his students?
2	The teacher rescues the positive from the observations made to your students?
3	Does the teacher promote social acceptance in the group?
4	Does the teacher develop standards to promote good coexistence?
5	Does the teacher attend to the spatial organization to develop his educational practice?
6	Does the teacher organize work groups in class?
7	Does the teacher determine the learning style of your students?
8	Does the teacher have an adequate management of times in the development of the class?
9	Does the teacher establish rules of coexistence according to the characteristics of his students?
10	Does the teacher take actions to avoid gender discrimination?
11	Does the teacher implement strategies to promote attitudes of commitment for healthy coexistence?
12	Does the teacher implement strategies to encourage collaboration for healthy coexistence?
13	Does the teacher promote equitable treatment with each one of your students?
14	Does the teacher promote permanence and success in the study of all his students?
15	Is the teacher prepared to address barriers for the learning that your students could present?
16	Does the teacher observe the behavior of his students to avoid distractions in class?
II DEVELOPMENT OF CLASS	
17	Does the teacher identify the characteristics of the didactic approaches of the subjects in learning activities?
18	Does the teacher relate the learning contents of the subjects to the achievement of educational purposes?
19	Does the teacher relate his teaching intervention to the

	corresponding expected learnings?
20	Does the teacher relate his teaching intervention to the educational needs of your students?
21	Does the teacher relate his teaching intervention to the didactic approaches of the subjects?
22	Does the teacher use didactic strategies so that your student develop cognitive skills?
23	The teacher diversifies the use of materials didactic for the achievement of its educational purposes?
24	Does the teacher prepare specific teaching material for the development of its contents?
25	The teacher uses teaching materials with support of information technologies?
26	Does the teacher show a good attitude and generate guidelines to lead the pace of the class?
27	Does the teacher use cooperative learning methods in working with his students?
28	Does the teacher use on-the-job learning workshops with his students?
29	Does the teacher use didactic strategies so that do your students learn based on their prior knowledge?
30	The teacher uses didactic strategies to that your students learn among peers?
31	Does the teacher use didactic strategies so that your students learn with the participation of all?
32	Does the teacher master the contents of the Primary Education subjects?
33	Does the teacher propose a sequence of educational content to facilitate the learning of his students?
	III EVALUATION
34	Does the teacher use the question and answer technique to assess the learning of this students?
35	To evaluate the learning of his students, does the teacher use the self-evaluation technique?
36	To evaluate the learning of their students, does the teacher use the co-evaluation technique?
37	To evaluate the learning of their students, does the teacher use the hetero-evaluation technique?
38	To evaluate the learning of his students, does the teacher use the test or questionnaire?
39	Does the teacher use the checklist to assess his students' learning?
40	Does the teacher systematize the process to evaluate the learning of his students?
41	Does the teacher use new procedures to assess the learning of his students?
42	Does the teacher participate with his peers in the analysis of your teaching practice?
43	The teacher uses theoretical references in the analysis of your teaching practice?

Author Profile

My name is **Oscar Luis Ochoa Martínez** and I am attached to the Pedagogical University of Durango, Mexico; the last academic study was carried out in 2016 obtaining the degree of Doctor in Educational Management from a Humanist Perspective, date from which I dedicate myself to educational research; at present I am a member of the Network of Researchers of the State of Chihuahua and I have the distinction of candidate in the National Service of Investigators.

My name is **Belia Cháidez Nevárez,** a graduate in Special Education from the Autonomous University of Tlaxcala, I did my doctorate studies at the Pedagogical University of Durango; currently I work as a Basic Education Supervisor and I share it with educational research.

My name is **Gonzalo Arreola Medina,** a research professor with a doctorate in Educational Evaluation and a full-time researcher at the Pedagogical University of the State of Durango, Mexico