

## Teaching performance in the technological high school. A statistical analysis of categorical data

Horacio Muñoz Durán<sup>§</sup>

Postgraduate in Educational Sciences-Higher Institute of Educational Sciences of the State of Mexico (ISCEEM). Tel. 595 1046681.

Corresponding author: horacio.munoz@isceem.edu.mx.

### Abstract

The practices of evaluation of the teaching performance by students in the institutions of the high school level have increased during the last decade; however, at present there are subsystems that do not have questionnaires or other instruments to assess, measure or evaluate the teaching function in the classroom; this is the case of the Centers of Technological High School of the State of Mexico. The purpose of this research is to present the results of the evaluation of teaching performance from the perspective of the student in the technological high schools, through a proposal of a questionnaire with multiple choice during the school year 2015-2016. The statistical treatment of the information was carried out through the analysis of categorical data considering the generalized linear models for nominal and ordinal scale variables using the RCommander software. The sample consists of 100 teachers from ten technological high schools of the school zone 021, who were evaluated by a total of 1 000 students.

**Keywords:** association, evaluation, perception.

Reception date: May 2021

Acceptance date: June 2021

Educational evaluation throughout its history has been characterized by belonging to one of the most complex and controversial topics for both those who research it and those who execute it. Evaluating in education covers several topics, the evaluation of learning, study plans and programs, institutions, teachers are among the most common. This research note shows the evaluation of teaching performance through the perception of students at the high school level carried out by Muñoz (2016), from the statistical treatment of categorical data. The structure of the work consists of four aspects; in the first, the problematization, objectives and hypotheses are presented. In the second, the theoretical references, the methodology and the description of the statistical tools used are written; for the third aspect, the results and discussion are shown and finally it culminates with the conclusions derived from the research.

Currently, CBTs of the State of Mexico have not formalized the practice of evaluating teaching performance through the perception of students. Adapting an evaluation model that includes students is a complex but necessary task, it is the student who is the user of the education and therefore it is the one who must evaluate it. A fundamental part of education in institutions is the faculty, the greater the preparation of teachers, the better the results obtained in students (Mateo, 2000). The evaluation of the teacher has, among several purposes, to obtain a diagnosis of the actions that they develop within the classroom.

Based on the above, it is evident that the continuous improvement in educational centers is associated with the performance of teachers; however, since there is no indicator in the evaluation that includes the student, this action will be incomplete. Therefore, there is a problem that refers to the lack of an instrument that measures teaching performance through the participation of students. Also, perform a statistical analysis of categorical data to determine the association between some variables that were measured when applying the instrument. The evaluation of teachers should emphasize the relationship between the individual and the institution; being this a complement between the needs and interests of both (Stronge, 2003).

The literature on the participation of students to assess teaching performance is generated to a greater extent in institutions of the university level; that is, it is the universities that started this educational practice since approximately 1980. In Rueda *et al.* (2010; 2011), a wide variety of instruments that were designed and implemented at the university level, mainly in public universities, is presented. In Conzuelo (2011); Muñoz (2016), the evaluation of teaching performance for a high school in the private sector and for CBTs of the eastern part of the State of Mexico is presented, respectively.

However, in the latter, both questionnaires and other ways of approaching teacher evaluation, such as narratives through written storytelling are scarce. The research carried out by Muñoz (2016) was the first approach to CBTs to evaluate the teaching performance from a descriptive perspective of the questionnaire.

Derived from the aforementioned research, new questions are raised in order to carry out a deeper analysis of the data obtained, since in the aforementioned work, only the description of the questionnaire through tables of frequencies, sum of scores and classification of teachers in any of the groups described below was emphasized. Therefore, the research questions are, is there an association between the disciplinary field to which the subject taught by the teacher belongs and

the numerical category of the teacher? Is there an association between the disciplinary field and the gender of the teacher? And is there an association between the category of the teacher and the overall classification of the teacher?

The main objective of this paper is to identify the possible associations between the variables disciplinary field, category of the teacher, gender and overall classification of the teacher to explain the student's perception of the teacher's role in the classrooms. The evidence of these possible similarities will be obtained through the statistical treatment of the data with the intention of testing the hypotheses that are written below.

### **For multinomial logistic regression models**

$H_{01}$ : there is no association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the numerical category of the teacher.  $H_{a1}$ : there is association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the numerical category of the teacher.  $H_{02}$ : there is no association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the gender of the teacher.  $H_{a2}$ : there is association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the gender of the teacher.

### **For ordinal regression model**

$H_{01}$ : there is no association, according to the treatment of the data with the ordinal regression model, between the category of the teacher and the overall grade of the teacher.  $H_{a1}$ : there is association, according to the treatment of the data with the ordinal regression model, between the category of the teacher and the overall grade of the teacher.

Some studies, such as Conzuelo and Rueda (2010); Rueda (2011, 2012); Montenegro (2014), highlight that mainly in IESs, students expressed themselves with severe, timely and forceful judgments about the teaching and learning processes, curricular pedagogical contents or infrastructure. Some experiences reveal that the students usually evaluate the teacher with little seriousness and the results usually present bias when analyzing them; however, the universities have done a very serious job in the evaluation of teaching performance, so much so that the institutions of the university level that have as a prelude some type of high school have decided to establish this evaluation process at the high school level.

The truth is that subjectivity increases, that bias is greater due to the lack of judgment and immaturity of the students; however, the student's comment, being they the user of education, must be considered by the institutions of high school level, a culture with evaluative approaches must be generated in the students. The student is the reason why a teacher prepares, the reason why we exist as teachers, they are a determining factor that drives us towards obtaining new knowledge in this beautiful work of teaching. Therefore, evaluation is not just another pedagogical process in institutions, it should not be a requirement imposed by the educational reform; on the contrary, it must be a culture that the more it is practiced and taught, the more truthful it will have at the different educational levels of our country.

The methodology with which this research was carried out is of the quantitative nature and of correlational scope by means of the multinomial and ordinal models following Agresti (2002), the data collection instrument was a questionnaire with most of the items with a multiple-choice answer. The dimensions of the original instrument are: (1) general evaluation of the course; (2) knowledge of the teacher; (3) pedagogical skills of the teacher; (4) attitudes and values of the teacher; (5) general evaluation of the teacher; and (6) open comments. The questionnaire is taken from Loredo (2000) and some adaptations to the context of CBTs are made.

For this study, the blocks two, three, four and six are analyzed, which have answers in Likert scale. Block (5) consists of a single question, where the result of the students was summarized numerically on a scale of 0 to 100. Block two consists of four questions: he or she completely masters his/her subject, is well informed and up to date, maintains a balance between theory and practice, and he or she has a broad general and humanistic culture. The answers for this block are multiple-choice numbered from 1 to 5, with 1 totally disagreeing, 2- disagreeing, 3- neither agreeing nor disagreeing and 4-agreeing and 5 totally agreeing.

With these scales, the students' answers for block three were also obtained, which contains the following items: he or she organizes and prepares his/her classes well, explains clearly and precisely, responds adequately to questions, awakens critical and creative thinking, stimulates the participation of the group, promotes a favorable work environment, uses the blackboard and audiovisual resources, evaluates according to the topics seen in class and feedbacks appropriately. Block four corresponds to the attitudes and values of the teacher, which is composed of six statements: the teacher had absences, starts and ends promptly, addresses students with order and respect, is demanding in class, in works and in exams, is honest and blameless in his/her behavior and promotes positive attitudes and values in class.

The categories that were determined in Muñoz (2016) to classify the teaching performance with the results of the instrument described above were divided into six on a scale of 0 to 100 points as follows: From 95-100, 'excellent'; from 85-94, 'very good'; from 75-84, 'good'; from 65-74, 'regular'; from 55-64, 'bad' and from 00-54, 'terrible'. The teachers who were evaluated were randomly chosen once the sample size ( $n$ ) was determined. The population ( $N$ ) corresponds to all teachers who belong to the ZE 021 of BT in 10 different CBTs (number of high schools in the period 2015-2016).

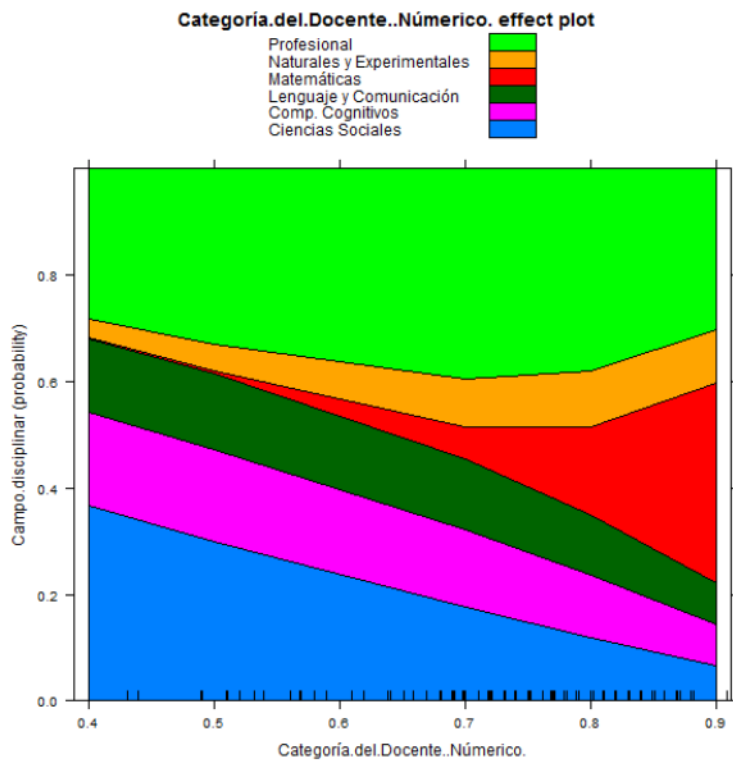
To determine the sample size, the formula for proportions when the population (450 teachers) is known was used, resulting in  $n = 100$ , with  $p$  and  $q = 0.5$ , a confidence level of 95% and an error of 5%. The type of sampling was cluster random sampling, selecting ten professors from each institution and ten students to evaluate each individual. CBTs are: San Diego, Atenco, Texcoco 2, Chiautla, Chiconcuac, Texcoco 5, Eduardo Suarez, Ecatepec, Texcoco 4 and Tepetlaotoc.

### **The results obtained are grouped as follows**

Treatment of statistical data using multinomial and ordinal regression models. For the statistical treatment through the multinomial logistic regression model and the ordinal regression model, the models corresponding to the multinomial logistic regression (MLM) model are run in the RCMDR software, to test the corresponding hypotheses. The value of the

test statistic 0.02126 is less than the p-value, which indicates that the  $H_{01}$  is rejected and the alternative hypothesis of MLM 1 is accepted: ‘there is an association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the numerical category of the teacher’.

Figure 1 illustrates the probabilities that teachers have of obtaining a better evaluation of their performance in relation to the disciplinary field to which the subject they teach belongs.

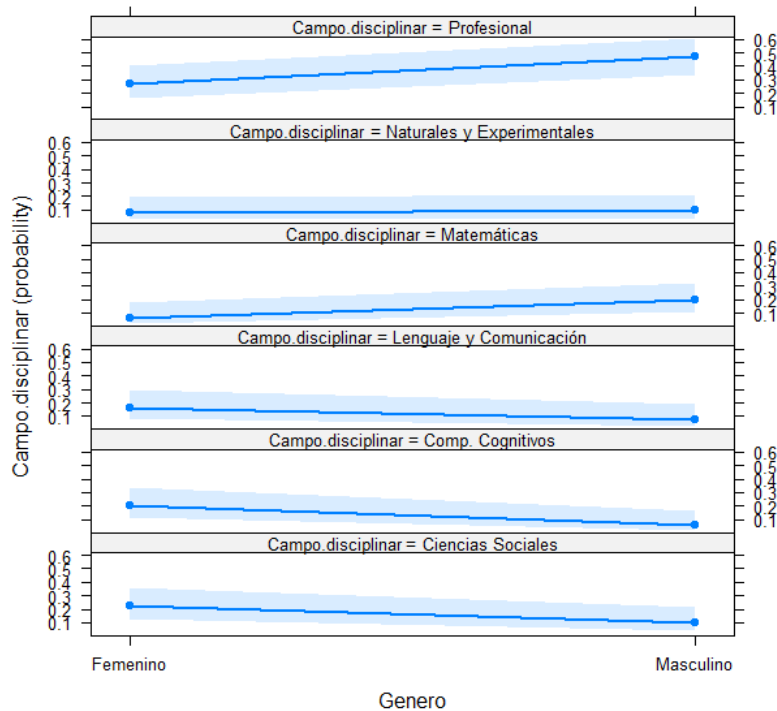


**Figure 1. Probabilities of obtaining better performance evaluation, according to the disciplinary field to which the teachers belong.**

The results of the previous graph indicate that for the disciplinary fields social sciences, cognitive components and language and communication, the probability of obtaining a high grade decreases as the perfect grade increases or is approached, to 1 or 100%. In the case of the areas of mathematics and natural and experimental sciences, teachers are more likely to obtain a higher score than that of the fields described above. Regarding the professional area, variability between probabilities and disciplines is not observed.

When running model 2. ‘Disciplinary field ~ gender: Anova for the MLM with the variables: disciplinary field and teacher’s gender’, the value of the test statistic 0.01032 is less than the p-value, which indicates that the  $H_{02}$  is rejected and the alternative hypothesis of MLM 2 is accepted: ‘there is association, according to the treatment of the data with the multinomial logistic regression model, between the disciplinary field and the gender of the teacher’. Figure 2 shows the probabilities that teachers have of obtaining a better evaluation of performance according to the gender of the teacher.

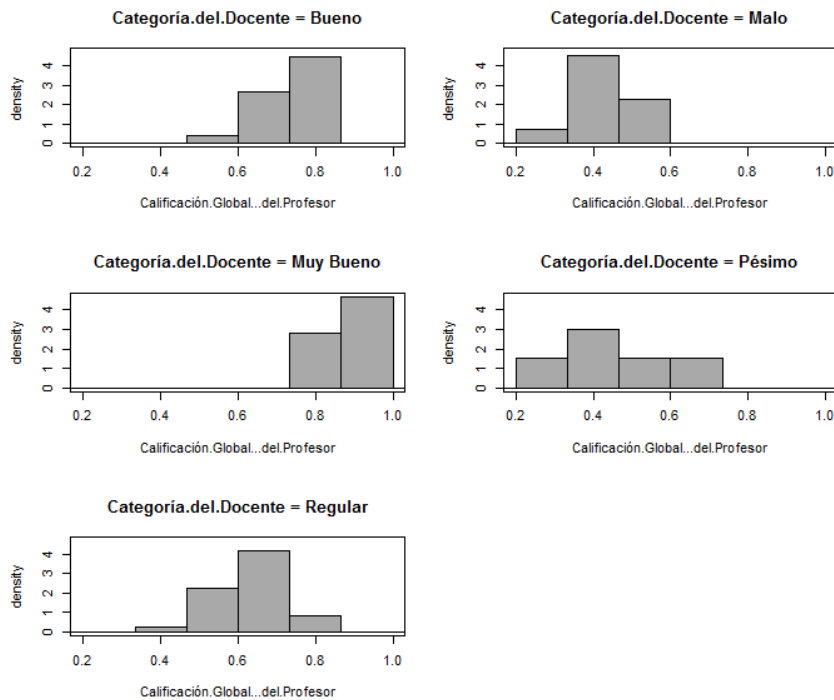
Figure 2 shows that students perceived female teachers in a better way for the disciplinary fields of social sciences, cognitive components, and language and communication. It was the opposite for the areas of mathematics and professional field, where the male gender is observed with higher grade. Finally, for the field of natural and experimental sciences, students perceive them indistinctively.



**Figure 2. Probabilities of obtaining better evaluation of the performance in relation to the gender of the teacher.**

Regarding the ordinal regression model, the statistical analysis corresponding to the linear ordinal model (OrdRegModel) is run in the Rcmdr software, to test the corresponding hypotheses ( $H_{01}$  and  $H_{02}$ ). Before that, in Figure 3, the distribution of the overall grades of the teachers by category of the teacher is presented.

In the distributions for each category, according to the density function, it can be seen that there is a minority of teachers in the ‘very good’ observed value and there is a majority in the classifications ‘regular’ and ‘good’. In other words, of the 100 teachers who were evaluated in the ZE 021 of BT, the two previous categories represent 64% of the teachers, 27% and 37%, respectively. For the OrdRegModel: Anova model for the OrdRegModel with the variables: teacher category and overall grade of the teacher, the value of the test statistic 0.01519 is less than the p-value, which indicates that  $H_{02}$  is rejected and the alternative hypothesis of OrgRegModel 1 is accepted: ‘there is an association, according to the treatment of the data with the ordinal regression model, between the teacher category and the overall grade of the teacher.’



**Figure 3. Distributions of the overall grades of teachers by category.**

The results obtained by running the linear models for categorical data show significant associations between the variables that were used for this purpose. It is evident that the performance of the teachers of the ZE 021 depends on the disciplinary field to which the subject they teach belongs, the overall grade that the student assigned to the teacher and the gender of the teacher. The results that were previously obtained with these models lead us to conclude the following: model 1, which was built to know if the disciplinary field is associated with the overall grade of the teacher, indicates that the subject taught by the teacher does present statistical significance with respect to the numerical assessment that the students gave of their teachers. In other words, we can infer that students have preference (greater probability of grading highly), when issuing a grade (numerically), with the subject taught by the teacher, but not with the comprehensive evaluation (that is, with the results of blocks 2, 3 and 4).

For model 2, the variables established have an association, which indicates that the gender of the teacher influences the disciplinary field in which they perform. In CBTs, the probability of obtaining a higher grade is that the teacher is of the male sex and that he belongs to the professional area; that is, to teach some of the subjects of the modules that have the different technical careers. In the case of social sciences, students prefer female teachers and for mathematics and natural and experimental sciences, there is no statistical preference.

Finally, the ordinal regression model was established to compare the variable 'teacher category', which, as already mentioned, presents ordinal measurement scale, with the 'overall grade'. The categories for the dependent variable have seven possible answers that were ordered from -excellent-, which no teacher reached; -very good-, where only some teachers were categorized

(16%); to -terrible-, where unfortunately 10% of them were classified with this heading. The rest is in the other categories that were already described above. In this context, the results of the ordinal regression model have the following interpretation: it is evident that the first two categories are those that are desired in CBTs; however, such an event did not occur.

## Conclusions

The evaluation of the teacher performance presents some similarity with the global evaluation of the teacher, in other words, the value judgments that the students emit in a numerical way are significantly related to the evaluation of the teacher performance and the student community of the CBT has a greater preference for male teachers for the subjects of the professional field and mathematics, express that the activities with a practical and operational approach have greater applicability than the knowledge obtained in the classroom by the subjects that belong to the field of social sciences.

## Cited literature

- Agresti, A. 2002. Categorical data analysis. 2<sup>nd</sup> (Ed.). New York: Wiley. 115-164 pp.
- Conzuelo, S. 2011. Evaluación del desempeño docente en educación media superior: un estudio de caso en el sector privado de México. Tesis de Maestría. Universidad Nacional Autónoma de México (UNAM). 27-37 pp.
- Conzuelo, S. y Rueda, M. 2010. La evaluación de la docencia en México: experiencias en educación media superior. Rev. Iberoam. Ev. Ed. 3(1):105-119. [http://www.rinace.net/riee/numeros/vol3-num1\\_e/art8.pdf](http://www.rinace.net/riee/numeros/vol3-num1_e/art8.pdf).
- Loredo, J. 2000. Evaluación de la práctica docente en educación superior: un acercamiento desde los instrumentos de evaluación. México, DF. (Ed.). Porrúa. 110-150 pp.
- Mateo, J. 2000. La evaluación educativa, su práctica y otras metáforas. Barcelona, España: (Ed.). Horsori. 95-124 pp.
- Montenegro, I. A. 2014. Evaluación del desempeño docente: fundamentos, modelos e instrumentos 1<sup>a</sup> (Ed.). México, DF. Magisterio editorial NEISA. 12-37 pp.
- Muñoz, H. 2016. Propuesta para la evaluación del desempeño docente en los CBT del Estado de México, mediante la percepción de los estudiantes aplicando los CEDA. Tesis de Maestría. Universidad Autónoma Chapingo (UACH). Texcoco, Estado de México. 8-22 pp.
- Rueda, M. 2011. Evaluar para controlar o para mejorar valoración del desempeño docente en las universidades. Bonilla Artigas (Ed.). México, DF. 5-18 pp.
- Rueda, M. 2012. La evaluación educativa: análisis de sus prácticas 1<sup>a</sup> (Ed.). México, DF. Ediciones días de santos. 13-16 pp.
- Rueda, M.; Luna, E.; García, B. y Loredo, J. 2010. La evaluación de la docencia en las universidades públicas mexicanas: un diagnóstico para su comprensión y mejora. Rev. Iberoam. Ev. Ed. 3(1e):77-92. [http://www.rinace.net/riee/numeros/vol3-num1\\_e/art6.pdf](http://www.rinace.net/riee/numeros/vol3-num1_e/art6.pdf).
- Stronge, J. 2003. Evaluating educational specialists. *In*: international handbook of educational evaluation. Kellaghan, T.; Stufflebeam Daniel L.; Wingate. L A. and Kluwer, K. A. International Handbooks of Education. 671-693 pp.