Oviedo Mandujano, A. F., González Barbera, C., and Caso Niebla, J. (2025). Teaching profiles associated with assessment practices in mathematics. *Revista de Investigación Educativa, 43.*
DOI: https://doi.org/ 10.6018/rie.603421



# Teaching Profiles Associated with Assessment Practices in Mathematics

## Perfiles docentes asociados a las prácticas de evaluación en matemáticas

Angélica Fabiana Oviedo Mandujano*[1] , Coral González Barbera** and Joaquín Caso Niebla***.

*Faculty of Human Sciences. Autonomous Universidad Autónoma de Baja California (Mexico)
**Faculty of Education. Complutense University of Madrid (Spain)
***Institute for Educational Research and Development. Universidad Autónoma de Baja California (Mexico)

**Abstract**

*This research aims to identify teaching profiles associated with assessment practices in Mathematics and their differentiation in terms of sociodemographic variables. A sample of 4,664 teachers ($M_{age}$=38, SD=9.70, 77.36% women) at the primary educational level in the state of Baja California, Mexico was used. For data analysis, a Latent Class Analysis (LCA) was performed, and t-student and chi-square tests were conducted. Two teaching profiles were identified, the intensive and the moderate, whose distinction lies in the frequency with which they carry out the different documented evaluation practices. Although both profiles are clearly differentiated in most dimensions of the evaluation process, this was not the case for the variables on the use of evaluation techniques and instruments. In addition, differences were found in the sociodemographic variables of sex, age, years of teaching experience, institution of initial training, participation in continuous training courses, school funding and rural or urban location of the school, although with low effect sizes. However, the education level, school grade, type of working day and type of school did not represent differences between profiles. The*

[1]**Correspondence:** Angélica Fabiana Oviedo Mandujano,angelica.oviedo@uabc.edu.mx , Facultad de Ciencias Humanas, Universidad Autónoma de Baja California, 21100, Mexicali, Baja California, Mexico.

*hypothesis is partially accepted, which states that sociodemographic variables allow characterizing teaching profiles associated with assessment practices in Mathematics where, although their effect is low, they can guide the implementation of differentiated strategies to support the teaching work.*

Keywords: *assessment; teaching practices; profiles; Mathematics education.*

## Resumen

*Esta investigación tiene como objetivo identificar los perfiles docentes asociados a las prácticas de evaluación en matemáticas y su diferenciación en términos de variables sociodemográficas. Se contó con una muestra de 4,664 docentes ($M_{edad}$=38, DE=9.70, 77.36% mujeres) de educación primaria en la entidad de Baja California, México. Para el análisis de datos se realizó un Análisis de Clases Latentes (ACL) y se utilizaron pruebas t-student y chi-cuadrada. Se identificaron dos perfiles docentes, el intensivo y el moderado, cuya distinción radica en la frecuencia en la que se realizan las distintas prácticas de evaluación documentadas. Si bien, ambos perfiles se encuentran claramente diferenciados en la mayoría de las dimensiones del proceso evaluativo, no fue así para las variables sobre el uso de técnicas e instrumentos de evaluación. Además, se ubicaron diferencias en las variables sociodemográficas de sexo, edad, años de experiencia docente, institución de formación inicial, participación en espacios de formación continua, sostenimiento de la escuela y asentamiento rural o urbano del centro escolar, aunque con tamaño del efecto bajos. No obstante, el grado de estudios, grado escolar, tipo de jornada y la modalidad de la escuela no representaron diferencias entre perfiles. Se acepta parcialmente la hipótesis en lo relativo a que las variables sociodemográficas permiten caracterizar perfiles docentes asociados a las prácticas de evaluación en matemáticas donde, aunque su efecto es bajo, pueden orientar la implementación de estrategias diferenciadas para apoyar la labor docente.*

Palabras clave: *evaluación; práctica docente; perfiles; educación matemática.*

## Introduction

### Teaching practices in mathematics assessment

Teaching assessment practices have gained relevance in educational research, given the recognition of their ability to condition teaching and influence learning (Navarro-Mosquera et al., 2017, Tariq et al., 2023). Moreover, in recent decades they have become more strongly positioned within educational policies that advocate a constructivist approach (Acar-Erdol and Yildizli, 2018; Schildkamp et al., 2020)

When assessment is carried out within the framework of mathematics learning, assessment practices are understood as the actions that teachers deploy to obtain information about their students' learning with respect to their level of mathematical knowledge and their competence to solve problems. This requires a systematic process guided by objectives and criteria defined during planning, in order to make interpretations that allow them to make decisions and communication with other actors

Within assessment it is possible to identify at least five dimensions on the basis of which to structure this process: purpose of assessment, object of assessment, assessment strategies, interpretation of learning evidence, and communication and use of results (Chappuis et al., 2012; Demosthenous et al., 2021; Leenknecht et al., 2021; Sheikh and Manap, 2024).

In relation to the purpose of assessment, it is usually divided into two main aspects commonly referred to with the terms coined by Scriven in 1967 regarding summative and formative assessment. The first refers to the social function of assessment aimed at the certification of learning and the second to a pedagogical function (Ravela et al., 2017; Schellekens et al., 2021), associated with the improvement of teaching and learning, which promotes the development of metacognition and the capacity for self-regulation (Braund and DeLuca, 2018; Zulliger et al., 2022).

In the case of formative assessment, Black and Wiliam (2009) proposed five key strategies for developing formative assessment: sharing and clarifying expected learning and assessment criteria; designing and implementing activities that provide evidence of learning; providing feedback; activating students as a source of learning for their peers; and activating each student as responsible for his or her own learning. These strategies continue to be a reference in more recent studies such as Chan (2021) and Wafubwa and Csíkos (2021)

On the other hand, assessment always refers to a set of specific contents or skills that constitute its object. In the case of mathematics, the problem-solving approach has positioned itself as the ideal framework for its teaching, from which the development of mental operations such as analysis, synthesis, generalisation, abstraction, and the development of heuristic, flexible and creative thinking is promoted (Herbert, 2021). The approaches of this approach find their beginnings in the contributions of Polya (1965) who structured the process of problem solving in the stages of understanding the problem, planning the solution, executing the plan and evaluation; which are likely to form part of the object to be evaluated in correspondence with the didactic approach.

As for summative assessment, which is also part of teachers' core tasks, there are various mechanisms for interpreting evidence of learning in order to generate a mark. One aspect of interest concerns the combination of cognitive and non-cognitive aspects . In this regard, the specialised literature points out that considering non-academic elements within the grade reduces its validity (Dagdag and Dagdag, 2020), so this practice is not recommended to strengthen learning from assessment (Chappuis et al., 2012; Paderes and Ramos, 2024).

Finally, the dimension of communication of assessment results refers to the ways in which teachers inform different actors of the learning obtained (Dagdag and Dagdag, 2020), including the student body, the school community and, in the case of primary education, families, given their key role in student performance (Lara and Saracostti, 2019). Corresponding to the purpose of the assessment, the results may be reported for

purely informative purposes or accompanied by improvement strategies for teaching and learning.

### Problem statement

In the case of Mexico, problem solving in mathematics and the formative assessment approach have been incorporated into the pedagogical discourse of primary school curricula since the 1993 reform. However, historically, the results of large-scale assessments show low learning achievement in this area of knowledge.

An example of this is the educational results in Mathematics recorded in the PLANEA test in 2018, in which nearly half of sixth grade students (49.1%) were placed at an insufficient level of knowledge, which represents serious difficulties to continue with their school career (Instituto Nacional para la Evaluación de la Educación [INEE], 2018). These results are consistent with those recently reported by the Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA), where 65.8% of 15-year-old students in Mexico scored below the minimum level in Mathematics, as well as registering a regression with respect to the 2018 assessment (OECD, 2023).

Given this situation, and given the relevance of assessment in the acquisition of learning, it is of interest to investigate the teaching practices associated with this process. One way of approaching this is the analysis of profiles that allows us to identify patterns in teachers (Sayac and Veldhuis, 2021; Veldhuis and Van Den Heuvel-Panhuizen, 2014; Zhao et al., 2018). Within this framework, the following research questions are posed: what teacher profiles are identified in mathematics assessment practices, and what variables make it possible to differentiate these profiles? This leads to the objective of identifying profiles that characterise mathematics assessment practices in primary education, and their differentiation in terms of socio-demographic variables, hypothesising that the latter represent distinctions between groups of teachers.

## Method

### Participants

The study was conducted in Baja California, Mexico, with 14,276 primary school teachers. By means of convenience sampling, a participation of 5,665 teachers was obtained, representing 39.68% of the population. Cases with missing values were eliminated using the *listwise* method, leaving 4,674 participants, 77.36% of whom were female, with a mean age of 38 years (SD=9.70), and mainly undergraduate (81.64%) and more than 6 years of experience at this educational level (65.25%).

## Instrumen

A self-report questionnaire was used, consisting of 11 items with socio-demographic information and a scale on assessment practices in mathematics, of which 47 items were considered with a Likert-type response option that explores the frequency with which teachers make each of the statements made.

For the scale, evidence of content validity was obtained through expert judgement, which allowed adjustments to be made to the instrument in its initial stage. Subsequently, evidence was obtained regarding the internal structure, for which the ordinal alpha coefficient was calculated, presenting a value of .97, considered an excellent level of internal consistency (George and Mallery, 2003), in addition to the corrected biserial point correlation, with values greater than .30, indicating a good ability to discriminate the items (Ebel and Frisbie, 1986). Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used for both validation and data reduction purposes

For the CFA, items belonging to a theoretical dimension with a *per se* classification were entered into the model. In this procedure, the unweighted least squares estimation method was used on the polychoric correlation matrix and the values of the goodness-of-fit index (GFI [≥.95]), root mean square error of approximation (RMSEA [≤.06]), standardised root mean square root (SRMR [≤.08]), comparative fit index (CFI [≥.96]) and Tucker Lewis index (TLI [≥.95]), following Lloret-Segura et al. (2014).

For its part, the AFE was used with those items that did not present a previous classification, for which the method of extraction of unweighted least squares and varimax rotation was used, after obtaining acceptable values in Bartlett's test of sphericity $p < 0.001$ (Montoya, 2007) and in the Kaiser-Meyer-Olkin KMO statistic > 0.80 (Lloret-Segura et al., 2014). We checked that the communalities were greater than .30, and maintained factor loadings above .30, in addition to corroborating that the factors obtained had a coherent theoretical explanation.

As a result of the above, 11 factors were obtained from this procedure, the resulting values of which corroborate that the theoretical structure presents a high correspondence with the data (see Table 1).

Table 1

*Variables resulting from the factor analysis*

| Theoretical dimension | Number of items | Factor | Factor analysis |
|---|---|---|---|
| Purpose of the evaluation | 4 | Educational purpose | AFC |
| | 4 | Summative purpose | |
| Object of evaluation | 4 | Understanding of the problem and communication skills | AFE |

| Theoretical dimension | Number of items | Factor | Factor analysis |
|---|---|---|---|
| | 3 | Adjustment and evaluation of the problem solution | |
| Formative evaluation strategies | 6 | Formative evaluation strategies | AFC |
| | 8 | Alternative assessment techniques and tools | AFE |
| | 3 | Examinations as assessment tools | |
| Interpretation of evidence | 3 | Cognitive aspects that make up the qualification | AFC |
| | 5 | Non-cognitive aspects making up the rating | |
| Communication and use of results | 4 | Communication with the school community | AFE |
| | 3 | Communication with students and families | |
| | 47 | | |

*Note.* In the CFA, items were entered according to the theoretical dimensions proposed in a correlated trait model obtaining optimal goodness-of-fit indices ($x^2$= 3,495.059, gl=199, GFI= .996, CFI= .981, TLI= .978, RMSEA=.060, 90% CI [.058- .061], p-value RMSEA=<.001, SRMR=.054). In the case of AFE, the proportion of variance explained by the three factor solutions was 71%, 49% and 60%, respectively.

## Procedure

A census application was carried out through a digital platform developed for this purpose. The invitation to teachers was sent via email through the Baja California Ministry of Education and was open from 3 to 27 May 2022. The questionnaire included an initial section dedicated to explaining the confidentiality and anonymity of the information collected. In this space, each teacher was asked to give consent for the use of data, citing the regulations governing the processing of personal data in Mexico.

## Data analysis

To answer the research question regarding what teacher profiles are identified in mathematics assessment practices, we used ACL, a statistical technique belonging to the family of mixed models that allows the identification of underlying classes of individuals based on differences in their responses to test items (Nylund-Gibson and Choi, 2018; Ondé and Alvarado, 2019). In this procedure, subjects are grouped by modal assignment, i.e. in the class where they have the highest probability of belonging according to their response patterns (Bauer, 2022)

For the above, the R software package *poLCA* (Linzer and Lewis, 2011) was used, which uses the Expectation-Maximisation (EM) algorithm to maximise the log likelihood function of the model. Since results vary due to the random initialisation of this algorithm and in order to counter the risk of identifying a local rather than a global maximum in the likelihood function, each model was estimated with at least 50 sets of random initial values and the number of iterations was set to 3,000, which ensured convergence on similar solutions and thus stability of the classification.

To determine the number of profiles, both statistical and theoretical criteria were considered. In relation to the statistical criteria, the lowest values were observed in three fit indices: Bayesian information criterion (BIC), its sample size adjusted variant (SABIC) and the Akaike information criterion (AIC), as reported by Weller et al. (2020). From the selected model, entropy and average posterior probabilities were calculated for each class to corroborate that appropriate values were obtained.

Subsequently, in order to answer the question "Which variables allow us to differentiate the teaching profiles associated with assessment practices?", the Student's t-test - together with the calculation of Cohen's $d$ - was used to observe the differences in terms of the 11 factors on teaching assessment practices and the age variable. In addition, the chi-square test was used for the remaining socio-demographic variables of an ordinal or nominal nature. These analyses were carried out in SPSS 27 software, and the calculation of the effect size of the chi-square test using the G*Power programme.

## Result

Regarding the identification of teacher profiles associated with assessment practices, the two latent class model was selected since, following Bauer (2022) and Nylund-Gibson and Choi (2018), it is where the inflection point is located that shows the most prominent decrease of values in the BIC, SABIC and AIC criteria, registering a decrease with marginal values for each additional class (Figure 1).

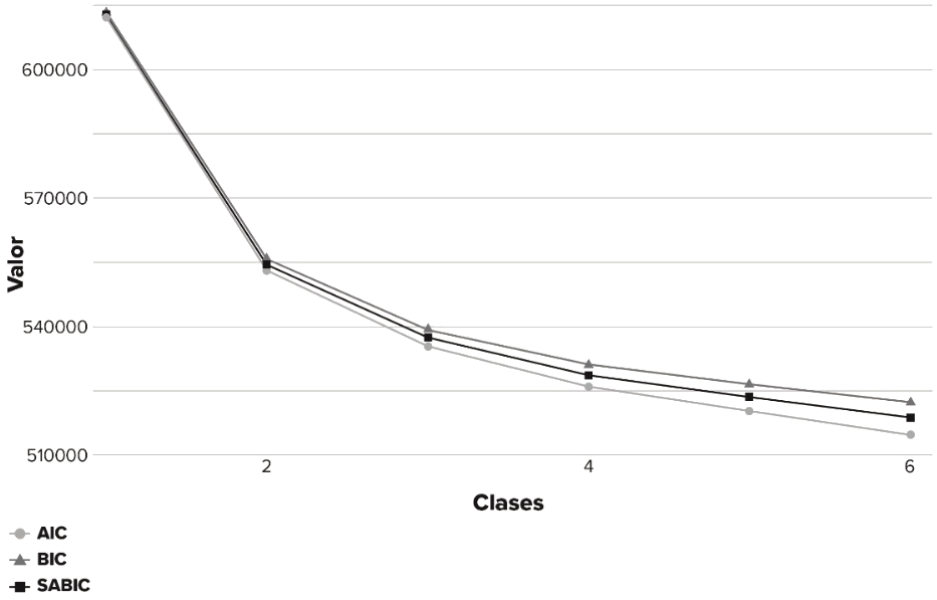
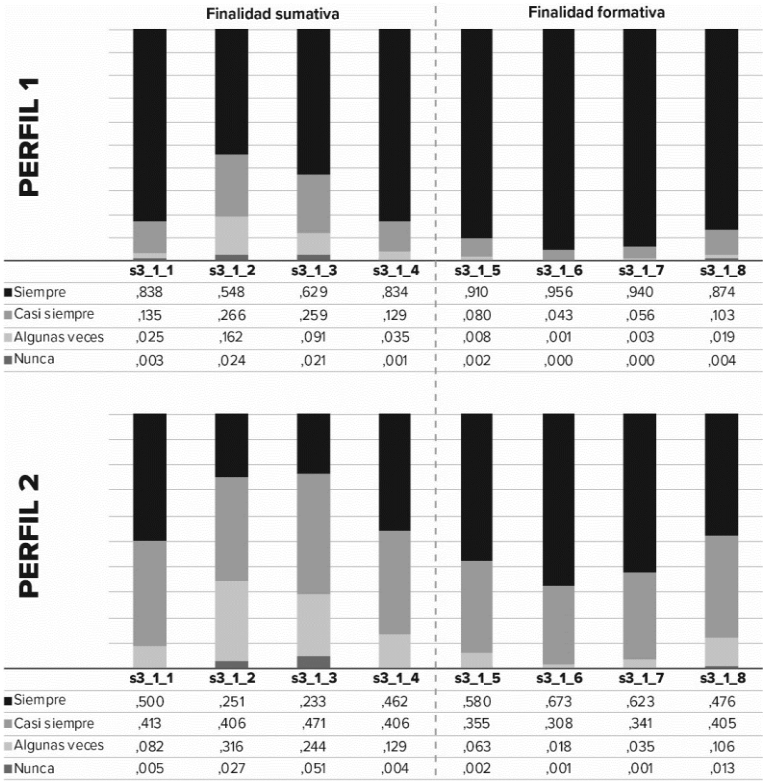



*Figure 1.* Elbow diagram with BIC, SABIC and AIC values for models with one to six latent classes. *Note.* BIC=Bayesian information criterion; SABIC=Sample size adjusted Bayesian information criterion and AIC=Akaike information criterion.

For the two-class model, entropy and average posterior probabilities were calculated for each class. Regarding entropy, minimum values of .80 are expected (Bauer, 2022; Weller et al., 2020), obtaining a result of .90, which indicates that the two latent classes are clearly separated. The average posterior probabilities of each class were .99, which provided additional evidence for this classification, since according to Masyn (2013) values higher than .70 are desirable.

The two latent classes obtained, called *teacher profiles associated with assessment practices in mathematics*, divided the sample into 55% of teachers belonging to profile 1 (2,558) and 45% to profile 2 (2,116). From the analysis of conditional response probabilities, it was observed that teachers belonging to profile 1 generally have higher probabilities of responding to the categories of *always* and *almost always* than those belonging to profile 2.

Figure 2 shows these results, where from the items associated with the purpose of the assessment, it can be seen that profile 1 teachers have a probability of over 80% of answering in the *always* category, except in the case of the assessment to assign a grade and level the group with 55% and 63%, respectively, where the *always* category continues to have the highest probability. In profile 2, the probabilities were distributed between the categories *always* and *almost always,* and for the items related to the assessment to assign a grade and group levelling, the category *sometimes* also obtained outstanding probability values.



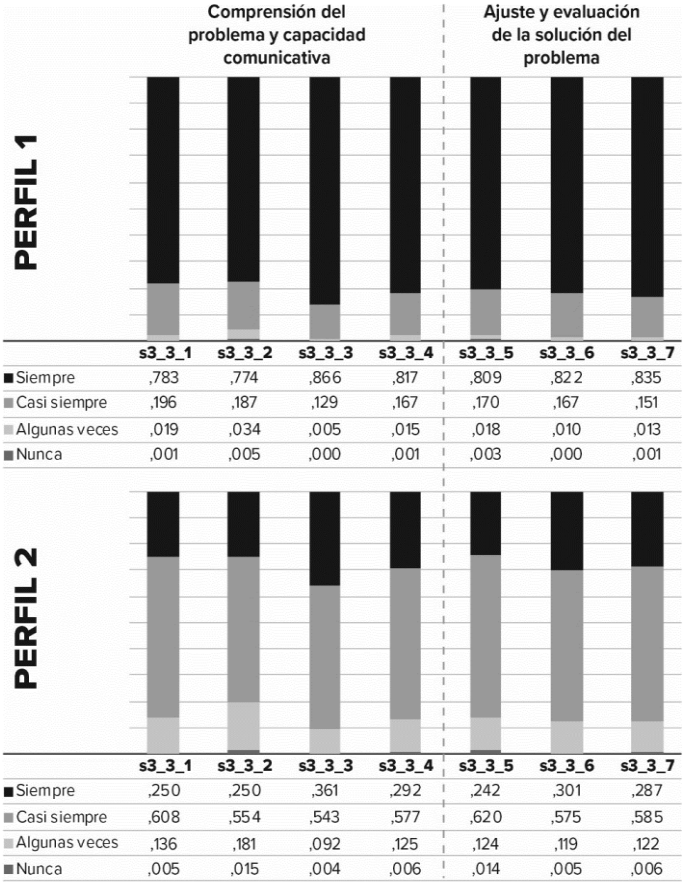

| PERFIL 1 | s3_1_1 | s3_1_2 | s3_1_3 | s3_1_4 | s3_1_5 | s3_1_6 | s3_1_7 | s3_1_8 |
|---|---|---|---|---|---|---|---|---|
| Siempre | ,838 | ,548 | ,629 | ,834 | ,910 | ,956 | ,940 | ,874 |
| Casi siempre | ,135 | ,266 | ,259 | ,129 | ,080 | ,043 | ,056 | ,103 |
| Algunas veces | ,025 | ,162 | ,091 | ,035 | ,008 | ,001 | ,003 | ,019 |
| Nunca | ,003 | ,024 | ,021 | ,001 | ,002 | ,000 | ,000 | ,004 |

| PERFIL 2 | s3_1_1 | s3_1_2 | s3_1_3 | s3_1_4 | s3_1_5 | s3_1_6 | s3_1_7 | s3_1_8 |
|---|---|---|---|---|---|---|---|---|
| Siempre | ,500 | ,251 | ,233 | ,462 | ,580 | ,673 | ,623 | ,476 |
| Casi siempre | ,413 | ,406 | ,471 | ,406 | ,355 | ,308 | ,341 | ,405 |
| Algunas veces | ,082 | ,316 | ,244 | ,129 | ,063 | ,018 | ,035 | ,106 |
| Nunca | ,005 | ,027 | ,051 | ,004 | ,002 | ,001 | ,001 | ,013 |

*Figure 2.* Conditional probabilities of the items associated with the evaluation purpose

*Note.* S3_1_1_1=Measure their learning; S3_1_2=Assign a grade; S3_1_3=Level the group; S3_1_4=Communicate their children's progress to their families; S3_1_5=Identify their prior knowledge; S3_1_6=Guide them in their learning; S3_1_7=Adapt my teaching strategies; S3_1_8=Identify my own needs for teacher updating.

A similar situation was observed in the items of the factors associated with the object of assessment (Figure 3), the formative assessment strategies factor (Figure 4), those related to the assignment of marks (Figure 5), as well as those related to the communication of results (Figure 6).




| PERFIL 1 | s3_3_1 | s3_3_2 | s3_3_3 | s3_3_4 | s3_3_5 | s3_3_6 | s3_3_7 |
|---|---|---|---|---|---|---|---|
| Siempre | ,783 | ,774 | ,866 | ,817 | ,809 | ,822 | ,835 |
| Casi siempre | ,196 | ,187 | ,129 | ,167 | ,170 | ,167 | ,151 |
| Algunas veces | ,019 | ,034 | ,005 | ,015 | ,018 | ,010 | ,013 |
| Nunca | ,001 | ,005 | ,000 | ,001 | ,003 | ,000 | ,001 |

| PERFIL 2 | s3_3_1 | s3_3_2 | s3_3_3 | s3_3_4 | s3_3_5 | s3_3_6 | s3_3_7 |
|---|---|---|---|---|---|---|---|
| Siempre | ,250 | ,250 | ,361 | ,292 | ,242 | ,301 | ,287 |
| Casi siempre | ,608 | ,554 | ,543 | ,577 | ,620 | ,575 | ,585 |
| Algunas veces | ,136 | ,181 | ,092 | ,125 | ,124 | ,119 | ,122 |
| Nunca | ,005 | ,015 | ,004 | ,006 | ,014 | ,005 | ,006 |

*Figure 3.* Conditional probabilities of the items associated with the evaluation object

*Note.* S3_3_1=Identify the relationships between explicit and implicit data provided in a problem statement; S3_3_2=Explain what a problem is about in their own words.S3_3_3_3=Perform the steps they will follow before solving a problem; S3_3_3_4=Be able to explain each step they perform when solving a problem; S3_3_3_5=Adjust or modify the steps to solve a problem when they encounter difficulties; S3_3_3_6=Check that the results obtained actually solve the problem; S3_3_7=Identify other ways to solve the same problem.

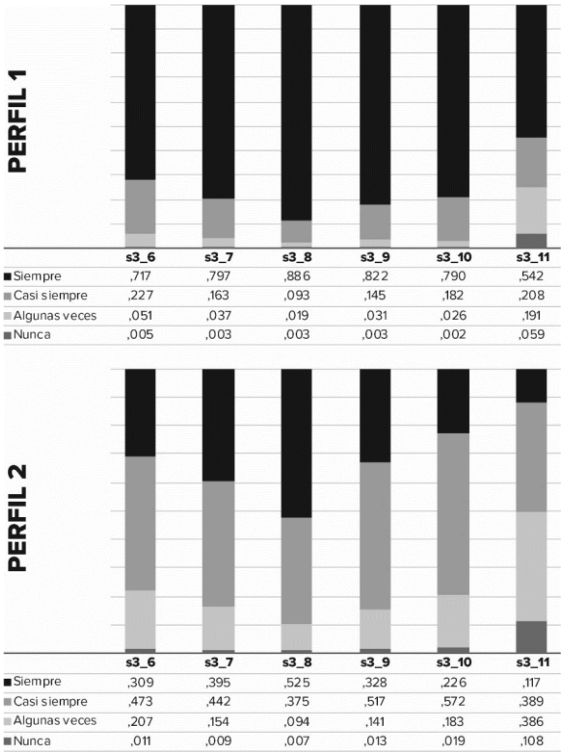
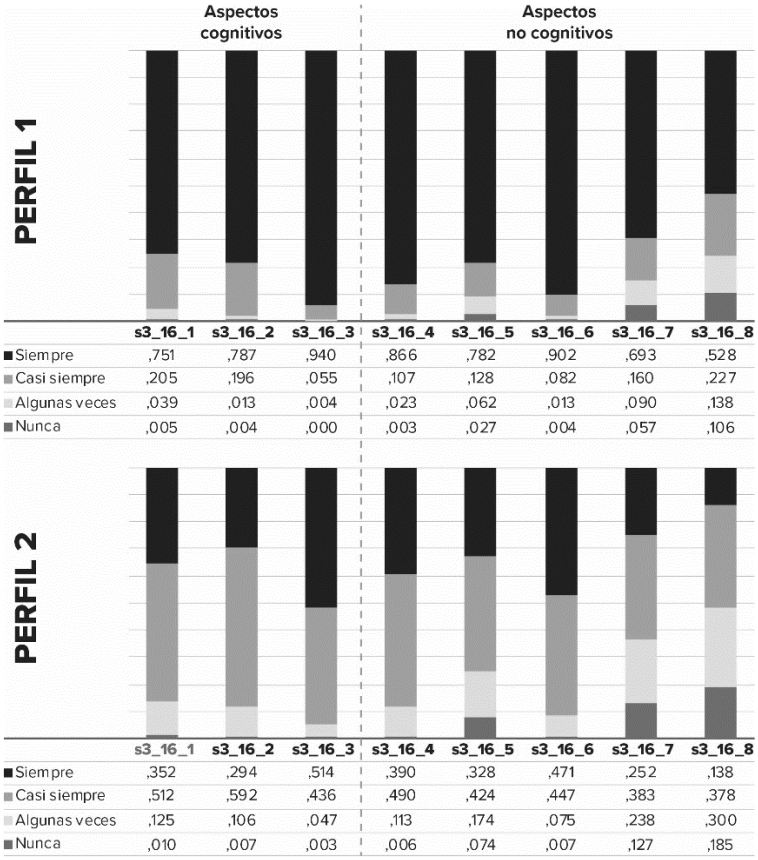



*Figure 4.* Conditional probabilities for items associated with formative assessment strategies

*Note.* S3_6=I communicate to my students what they are expected to learn in each thematic block; S3_7=I explain to my students how they will be assessed; S3_8=I inform families how their children will be assessed; S3_9=I ask my students to review their work or homework to identify their strengths and weaknesses before providing feedback; S3_10=I ask my students to reflect on their strengths and opportunities for learning; S3_11=I ask my students to reflect on their strengths and opportunities for learning; S3_11=I ask my students to reflect on their strengths and opportunities for learning; S3_12=I ask my students to reflect on their strengths and opportunities for learning; S3_10=I ask my students to reflect on their strengths and learning opportunities; S3_11=I ask my students to review their peers' work or assignments to identify strengths and areas of opportunity.

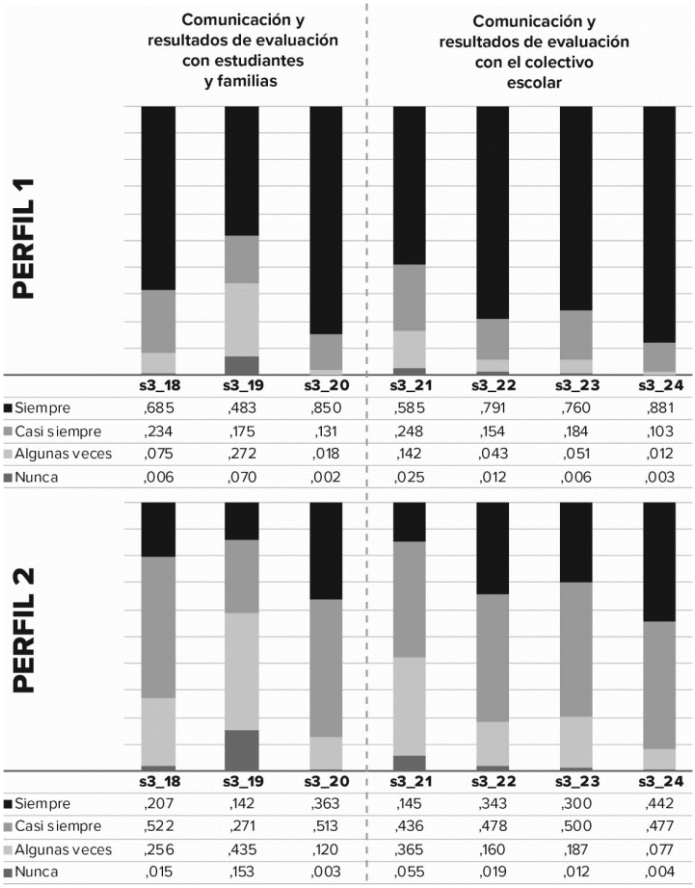



| PERFIL 1 | s3_16_1 | s3_16_2 | s3_16_3 | s3_16_4 | s3_16_5 | s3_16_6 | s3_16_7 | s3_16_8 |
|---|---|---|---|---|---|---|---|---|
| Siempre | ,751 | ,787 | ,940 | ,866 | ,782 | ,902 | ,693 | ,528 |
| Casi siempre | ,205 | ,196 | ,055 | ,107 | ,128 | ,082 | ,160 | ,227 |
| Algunas veces | ,039 | ,013 | ,004 | ,023 | ,062 | ,013 | ,090 | ,138 |
| Nunca | ,005 | ,004 | ,000 | ,003 | ,027 | ,004 | ,057 | ,106 |

| PERFIL 2 | s3_16_1 | s3_16_2 | s3_16_3 | s3_16_4 | s3_16_5 | s3_16_6 | s3_16_7 | s3_16_8 |
|---|---|---|---|---|---|---|---|---|
| Siempre | ,352 | ,294 | ,514 | ,390 | ,328 | ,471 | ,252 | ,138 |
| Casi siempre | ,512 | ,592 | ,436 | ,490 | ,424 | ,447 | ,383 | ,378 |
| Algunas veces | ,125 | ,106 | ,047 | ,113 | ,174 | ,075 | ,238 | ,300 |
| Nunca | ,010 | ,007 | ,003 | ,006 | ,074 | ,007 | ,127 | ,185 |

*Figure 5.* Conditional probabilities for items associated with aspects making up the rating

S3_16_1=The correct answers in exams, class activities or homework; S3_16_2=The solution process followed in exams, class activities or homework; S3_16_3=The participation in class; S3_16_4=The performance in team work; S3_16_5=The attendance in class; S3_16_6=The delivery of work and homework on time; S3_16_5=The attendance in class.S3_16_4=Performance in teamwork; S3_16_5=Class attendance; S3_16_6=Delivery of assignments and homework in due time and form; S3_16_7=Discipline/behaviour in the classroom; S3_16_8=Interest and enjoyment of mathematics.




| PERFIL 1 | s3_18 | s3_19 | s3_20 | s3_21 | s3_22 | s3_23 | s3_24 |
|---|---|---|---|---|---|---|---|
| Siempre | ,685 | ,483 | ,850 | ,585 | ,791 | ,760 | ,881 |
| Casi siempre | ,234 | ,175 | ,131 | ,248 | ,154 | ,184 | ,103 |
| Algunas veces | ,075 | ,272 | ,018 | ,142 | ,043 | ,051 | ,012 |
| Nunca | ,006 | ,070 | ,002 | ,025 | ,012 | ,006 | ,003 |

| PERFIL 2 | s3_18 | s3_19 | s3_20 | s3_21 | s3_22 | s3_23 | s3_24 |
|---|---|---|---|---|---|---|---|
| Siempre | ,207 | ,142 | ,363 | ,145 | ,343 | ,300 | ,442 |
| Casi siempre | ,522 | ,271 | ,513 | ,436 | ,478 | ,500 | ,477 |
| Algunas veces | ,256 | ,435 | ,120 | ,365 | ,160 | ,187 | ,077 |
| Nunca | ,015 | ,153 | ,003 | ,055 | ,019 | ,012 | ,004 |

*Figure 6.* Conditional probabilities for the items associated with reporting of results

*Note.* S3_18=I take time to talk with my students about their grades; S3_19=I give a written report (other than a report card) to families; S3_20=I inform families how they can help their children improve their learning; S3_21=I get feedback from my teaching colleagues to improve my students' learning; S3_22=I share assessment results with the teacher who will teach my group in the next grade; S3_23=I share assessment results in school staff meetings; S3_24=The learning outcomes of my students are taken into account in school planning.

However, in the case of the items that make up the factors of *alternative assessment techniques and instruments* and *examinations as assessment instruments*, it is observed that the highest probabilities do not occur for all cases in the same categories (Figure 7). For example, in the case of written tests, profile 1 teachers have a higher probability than 40% of to answer that they *always* use both open and closed written tests, while profile 2

teachers will have a higher probability of answering that they *almost always* use open-ended tests, but only *sometimes* use closed ones.

For the instruments grouped under the name of alternatives, even though profile 1 presents higher probabilities in the *always* category than profile 2, similar probabilities are observed within each item between categories, as in the case of debates or discussions, where probabilities range from 27.1% to 35% for *sometimes*, *almost always* or *always*. On the other hand, the tendencies observed in profile 2 vary, where the use of portfolios of evidence, debates or discussions, observation guides and rubrics are mainly located in the categories *sometimes* and *almost always*, while questions during class are applied with a higher probability *almost always* and *always*. The instruments in which both profiles coincide are the notebook, textbook, supplementary books or support guides, where the category *always* has the highest probability of response.

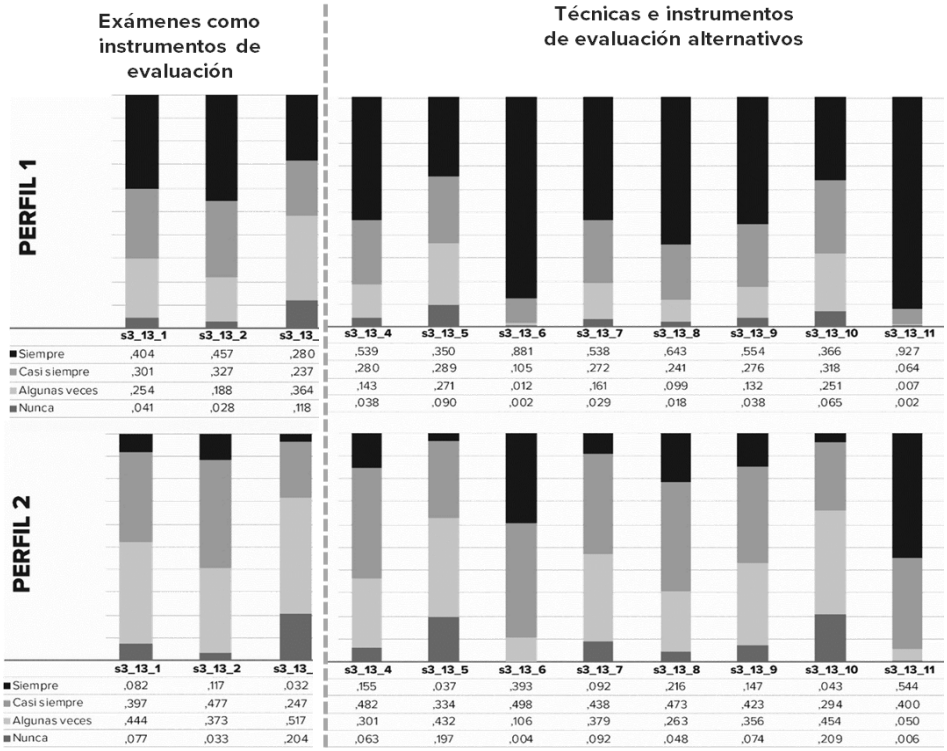



*Figure 7.* Conditional probabilities of items associated with assessment techniques and instruments

*Note.* S3_13_1=Written examinations with open-ended questions; S3_13_2=Written examinations with closed or multiple choice questions; S3_13_3=Oral examinations. S3_13_4=Portfolios of evidence; S3_13_5=Debates or discussions on mathematical concepts and procedures; S3_13_6=Questions during class; S3_13_7=Observation guides; S3_13_8=Checklists; S3_13_9=Rubrics; S3_13_10=Attitude scales; S3_13_11=Portfolios, textbook, supplementary books or support guides.

The results presented are corroborated by the Student's t-test, where profile 1 scored above the mean in all factors, while profile 2 scored below the mean. However, the tests report significant differences in all cases, with the exception of the factors corresponding to the application of assessment instruments, whether alternative or exam-type. The differences found present large effect sizes, ranging from .884 to 1.724, according to Cárdenas and Arancibia (2014) (see Table 2 and Figure 8).

From the behaviour observed in the two teacher profiles, it is clear that their differences lie more in the frequency than in the type of evaluation practices. In this sense, it was decided to add the label *intensive* to profile 1 and *moderate* to profile 2, since the latter, although registering a lower frequency, place a large part of their responses in the category *almost always*.

Table 2

*Results of the t-student test between the assessment practices variables*

| | **Profile 1 Intensive (n= 2558)** | | | **Profile Moderate (n= 2116)** | | |
|---|---|---|---|---|---|---|
| **Factors** | **M** | **DE** | **Mean standard error** | **M** | **DE** | **Mean standard error** |
| p1 | 0.55 | 0.81 | 0.02 | -0.67 | 0.78 | 0.02 |
| p2 | 0.57 | 0.72 | 0.01 | -0.69 | 0.84 | 0.02 |
| p3 | 0.42 | 0.68 | 0.01 | -0.51 | 1.09 | 0.02 |
| p4 | 0.43 | 0.72 | 0.01 | -0.51 | 1.05 | 0.02 |
| p5 | 0.59 | 0.76 | 0.01 | -0.72 | 0.76 | 0.02 |
| p6 | 0.02 | 1.00 | 0.02 | -0.02 | 1.00 | 0.02 |
| p7 | 0.01 | 1.00 | 0.02 | -0.01 | 1.00 | 0.02 |
| p8 | 0.60 | 0.73 | 0.01 | -0.73 | 0.77 | 0.02 |
| p9 | 0.56 | 0.79 | 0.02 | -0.68 | 0.78 | 0.02 |
| p10 | 0.37 | 0.79 | 0.02 | -0.44 | 1.05 | 0.02 |
| p11 | 0.39 | 0.87 | 0.02 | -0.47 | 0.94 | 0.02 |

| Factors | t | gl | p | Difference in averages | Standard error difference | *d* | LI | LS |
|---|---|---|---|---|---|---|---|---|
| p1 | 52.71 | 4,562.59 | >.001 | 1.22 | 0.02 | 1.54 | 1.48 | 1.61 |
| p2 | 53.93 | 4,186.01 | >.001 | 1.26 | 0.02 | 1.61 | 1.54 | 1.67 |
| p3 | 34.41 | 3,400.03 | >.001 | 0.93 | 0.03 | 1.05 | 0.99 | 1.12 |
| p4 | 34.96 | 3,604.36 | >.001 | 0.94 | 0.03 | 1.06 | 1.00 | 1.13 |
| p5 | 58.64 | 4,497.60 | >.001 | 1.31 | 0.02 | 1.72 | 1.66 | 1.79 |
| p6 | 1.42 | 4,672.00 | 0.156 | 0.04 | 0.03 | 0.04 | -0.02 | 0.10 |
| p7 | 0.53 | 4,672.00 | 0.597 | 0.02 | 0.03 | 0.02 | -0.04 | 0.07 |
| p8 | 59.94 | 4,415.13 | >.001 | 1.33 | 0.02 | 1.77 | 1.70 | 1.84 |
| p9 | 53.92 | 4,543.90 | >.001 | 1.24 | 0.02 | 1.58 | 1.52 | 1.65 |
| p10 | 29.30 | 3,840.35 | >.001 | 0.81 | 0.03 | 0.88 | 0.82 | 0.94 |
| p11 | 32.54 | 4,672.00 | >.001 | 0.86 | 0.03 | 0.96 | 0.90 | 1.02 |

*Note*. N=4,674. p1=Summative purpose; p2=Formational purpose; p3=Problem understanding and communicative ability; p4=Adjustment and evaluation of problem solving; p5=Formational assessment strategies; p6=Examinations as assessment instruments; p7=Alternative assessment techniques and instruments; p8=Cognitive aspects that make up the rating; p9=Non-cognitive aspects that make up the rating; p10=Communication with the school community; p11=Communication with students and families; gl=degrees of freedom; *d*= Cohen's d; LI=Lower limit of effect size; LS=Upper limit of effect size.

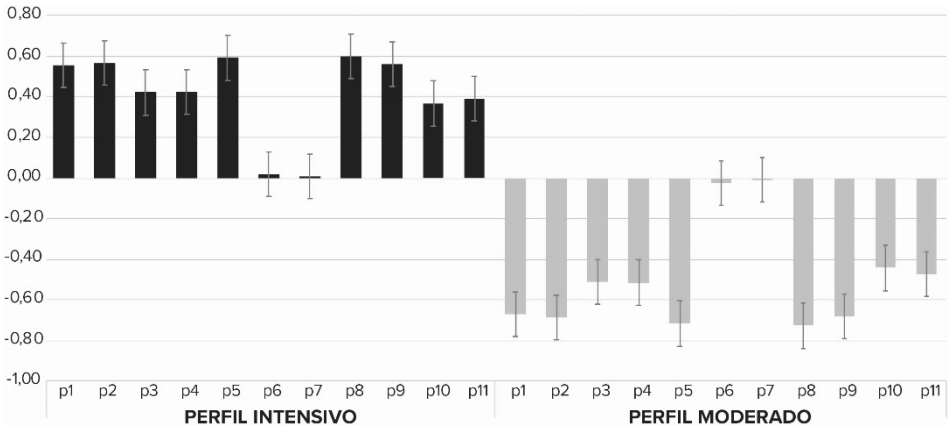



*Figure 8*. Behaviour of evaluation practices variables

*Note*. N=4,674. p1=Summative purpose; p2=Formational purpose; p3=Problem understanding and communicative ability; p4=Adjustment and evaluation of problem solving; p5=Formational assessment strategies; p6=Exams as assessment instruments; p7=Alternative assessment techniques

and instruments; p8=Cognitive aspects that make up the grade; p9=Non-cognitive aspects that make up the grade; p10=Communication with the school community; p11=Communication with students and families.

On the other hand, when exploring the differences between the *intensive* and *moderate* profiles, according to the socio-demographic variables, we find that age is a variable with statistically significant differences between them, where teachers in the moderate profile (M=39.00, SD=9.81) are slightly older than those in the intensive profile (M=37.12; SD=9.52), although the effect size indicates that the difference is small (d=0.20, 95% CI [0.14- 0.25]).

In terms of gender, significant differences were also found, since a greater number of female teachers belong to the *intensive profile*, equivalent to a proportion of 79.67% within this group. Similarly, a higher proportion of male teachers are in the *moderate profile* (25.43% versus 20.33% in the *intensive profile*). However, the effect size is basically zero (Table 3).

Table 3

*Frequencies and results of the chi-square test between teacher profiles and contextual variables*

| Variables | | Profile intensive | | Moderate profile | | $X^2$ | gl | p | w |
|---|---|---|---|---|---|---|---|---|---|
| | | N | % | N | % | | | | |
| Sex | Woman | 2,038 | 79.67% | 1,578 | 74.57% | 17.18 | 1 | **<0.001** | 0.060 |
| | Man | 520 | 20.33% | 538 | 25.43% | | | | |
| Initial training institution | Normal School | 1,419 | 55.47% | 1,344 | 63.52% | 36.45 | 2 | **<0.001** | 0.090 |
| | Public university | 517 | 20.21% | 393 | 18.57% | | | | |
| | Private university | 622 | 24.32% | 379 | 17.91% | | | | |
| Last grade of studies completed | Bachelor's degree | 2,074 | 81.8% | 1,742 | 82.33% | 5.79 | 3 | 0.122 | - |
| | Speciality | 76 | 2.97% | 41 | 1.94% | | | | |
| | Master | 379 | 14.82% | 314 | 14.84% | | | | |
| | PhD | 29 | 1.13% | 19 | 0.90% | | | | |
| Participation in continuous training spaces [a] | Yes | 2,135 | 83.46% | 1,583 | 74.81% | 53.29 | 1 | **<0.001** | 0.055 |
| | No | 423 | 16.54% | 533 | 25.19% | | | | |
| Years of experience | 0-5 years | 966 | 37.76% | 658 | 31.10% | 48.83 | 5 | **<0.001** | 0.103 |
| | 6-10 years | 436 | 17.4% | 287 | 13.56% | | | | |
| | 11-15 years | 345 | 13.49% | 343 | 16.21% | | | | |
| | 16-20 years | 317 | 12.39% | 320 | 15.12% | | | | |
| | 21-25 years | 302 | 11.81% | 296 | 13.99% | | | | |
| | Over 26 years old | 192 | 7.51% | 212 | 10.02% | | | | |

*Note.*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Grade where you teach | 1st and 2nd grade | 917 | 35.85% | 693 | 32.75% | 5.98 | 3 | 0.113 | - |
| | 3rd and 4th grade | 758 | 29.63% | 673 | 31.81% | | | | |
| | 5th and 6th grade | 813 | 31.78% | 682 | 32.23% | | | | |
| | Multigrade | 70 | 2.74% | 68 | 3.21% | | | | |
| Sector to which your current school belongs | Public | 2,095 | 81.90% | 1,901 | 89.84% | 58.86 | 1 | **<0.001** | 0.118 |
| | Private | 463 | 18.10% | 215 | 10.16% | | | | |
| Type of school you work in according to the type of organisational structure | Complete organisation | 2,432 | 95.7% | 2,019 | 95.42% | 0.30 | 1 | 0.585 | - |
| | Multigrade | 126 | 4.93% | 97 | 4.58% | | | | |
| Type of school you work in by type of school day | Regular day | 2,212 | 86.47% | 1,834 | 86.67% | 0.04 | 1 | 0.842 | - |
| | Extended working day | 346 | 13.53% | 282 | 13.33% | | | | |
| Type of school settlement | Urban | 1,540 | 60.20% | 1,160 | 54.82% | 13.96 | 2 | **<0.001** | 0.055 |
| | Rural | 432 | 16.89% | 396 | 18.71% | | | | |
| | Marginalised | 586 | 22.91% | 560 | 26.47% | | | | |

N=4,674. Variables that were statistically significant are highlighted in bold. Assessment of the effect size *w* for the chi-square test result based on Cárdenas and Arancibia (2014).

[a] Refers to continuous training opportunities on topics related to learning assessment in the last five years.

Table 3 also shows that, with respect to the initial training institution, there are significant differences, although with smaller effect sizes, with a higher proportion of teachers who studied a degree at a public or private university in the composition of the *intensive profile* than in the *moderate profile*. On the other hand, teachers who graduated from teacher training colleges[1], although in absolute frequencies they are also concentrated in the *intensive profile*, represent a higher proportion in the integration of the *moderate profile* (63.52%).

On the other hand, the *intensive profile* brought together a higher number of teachers who have participated in continuous training activities on topics related to learning assessment in the last five years. In this respect, 83.46% of the teaching staff in the *intensive profile* have participated in courses, seminars, workshops, diploma courses or other training options in this area, as opposed to 74.81% in the *moderate profile*.

In relation to the socio-demographic variables of an occupational nature, the years of service as primary school teachers represent significant differences, where it is possible to distinguish that in the integration of the *intensive profile* there is a higher proportion of teachers from 0 to 10 years than in *the moderate profile*. On the other hand, teachers with more than 11 years of experience are proportionally more likely to belong to the *moderate profile* than to the *intensive profile*

With regard to the sector to which the school belongs, although a greater number of teachers from both public and public schools make up the *intensive profile*, it is possible to detect differences within each profile. In this sense, in the *intensive profile* there is a higher proportion of teachers from public schools than in the *moderate profile*, while in the *moderate profile* there is a higher proportion of teachers working in public schools.

The type of school settlement is a variable that also showed significant differences, with the *intensive profile* comprising a higher proportion of teachers working in an urban context and a lower proportion of teachers from rural or marginalised areas compared to the *moderate profile*.

Finally, it is worth noting that the variables corresponding to the last grade of school completed, the school grade in which classes are taught, the organisational structure of the school and the type of school day did not show significant differences between profiles.

## Discussion

Within the sample of primary school teachers in Baja California, Mexico, two profiles associated with assessment practices in mathematics were identified, referred to as the *intensive profile* and the *moderate profile*

The *intensive profile* is similar to the enthusiastic profile described in the studies by Sayac and Veldhuis (2021), Veldhuis and Van Den Heuvel-Panhuizen (2014) and Zhao et al. (2018), where this group is above average in the frequency with which they perform the different assessment practices documented here. One aspect to note is that, in these

studies, the enthusiastic profile was not the largest group in the sample of teachers, unlike in the present research where teachers with the intensive profile represent slightly more than half of the participating teachers.

In turn, according to the results of the studies referred to above, the *moderate profile* is associated with a low enthusiasm profile, as the teaching staff register below-average scores in the evaluation practices under analysis. However, for the purposes of this study, we chose to use the term *moderate* because, although the response patterns register a lower frequency in the behaviours explored, they are far from being considered infrequent.

It is worth noting that the group of teachers with an intensive profile, although they report a higher frequency of assessment practices associated with a formative approach, also tend to consider non-cognitive criteria in the awarding of school grades, which has been documented as an undesirable practice with reduced impact on the achievement of expected learning (Chappuis et al., 2012; Dagdag and Dagdag, 2020, Paderes and Ramos, 2024). Nevertheless, these results reinforce the importance of developing strategies to support teachers in clearly determining the grades they assign to their students from a pedagogical perspective that supports learning.

Furthermore, it is relevant in interpretative terms that, although there are differences basically in all the dimensions involved in the assessment process - purpose, object, formative assessment strategies, assignment of marks and communication of results - teachers did not report differentiated practices in terms of the tools they use to retrieve evidence of their students' learning, at least in relation to the techniques and instruments included in the questionnaire.

On the other hand, it was possible to characterise the profiles based on socio-demographic variables, although these results should be taken with caution, given that the effect sizes were low, it is relevant that seven of these variables showed significant differences. In this sense, the *intensive profile* was proportionally composed, compared to the *moderate profile*, of more female teachers, younger, university graduates, teachers who have participated in training schemes associated with evaluation in the last five years, with less than 10 years of experience, working in privately supported primary schools and located in urban areas.

In this respect, there are discrepancies and similarities with previous studies in terms of the differences found between profiles in terms of socio-demographic variables. In this regard, in the research by Veldhuis and Van Den Heuvel-Panhuizen (2014), carried out in the Netherlands, gender, age, number of years of experience, professional qualification and the school grade in which they teach represented significant differences, which coincides with this study with the exception of school grade. In the case of Sayac and Veldhuis (2021), who investigated teachers in France, differences were found in some variables according to years of experience, but conversely age, gender and previous education were not related to assignment to any of the groups. Finally, in Zhao et al.

(2018), gender, years of experience also represented differences between profiles as in the sample of teachers in this research.

The results obtained lead to the partial acceptance of the hypothesis that socio-demographic variables allow us to characterise teacher profiles associated with assessment practices in mathematics, which can guide the implementation of differentiated strategies to support teaching.

## Conclusions

The aim of this research was to identify teacher profiles associated with mathematics assessment practices and their differentiation in terms of socio-demographic variables. In a large sample of primary school teachers in Baja California, Mexico, intensive and moderate profiles were found, with clear differences in the different dimensions of the assessment process, with the exception of the techniques and instruments that teachers reported applying to assess their students. In turn, a series of socio-demographic variables such as gender, age, years of experience, initial training, in-service training, school maintenance and type of settlement made it possible to differentiate, albeit with low effect sizes, these two groups of teachers who apply assessment in the classroom to a greater or lesser extent.

The possibility of identifying teacher profiles with different practices, in addition to contributing to the theoretical understanding of the object of study, provides input for the development of differentiated strategies to support teaching practice. Furthermore, in methodological terms, the present study demonstrated that LCA is an appropriate technique for addressing research objectives of this nature.

As a prospective approach, it is recommended that future studies consider linking assessment profiles with student performance levels, in order to identify whether a greater presence of assessment in the classroom has an impact on student learning. It is also desirable to collect information from other educational actors, such as students, families and educational authorities, in order to draw a clearer and more complete map of the evaluation process.

Notes.

1. Normalista training is that received within a teacher training institution, which in the Mexican context is known as Escuelas Normales. Teachers who fall into this category have studied a curriculum focused exclusively on teaching at primary level.

## References

Acar-Erdol, T. and Yildizli, H. (2018). Classroom assessment practices of teachers in Turkey. *International Journal of Instruction*, *11*(3), 587-602.

Bauer, J. (2022). A Primer to Latent Profile and Latent Class Analysis. In M. Goller, E. Kyndt, S. Paloniemi, and C. Damşa (Eds.), *Methods for Researching Professional Learning and Development: Challenges, Applications and Empirical Illustrations* (pp. 243-268). Springer International Publishing. https://link.springer.com/book/10.1007/978-3-031-08518-5

Black, P. and Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*, 5-31.

Braund, H. and DeLuca, C. (2018). Elementary students as active agents in their learning: an empirical study of the connections between assessment practices and student metacognition. *Australian Educational Researcher*, *45*, 65-85. https://doi.org/10.1007/s13384-018-0265-z

Cárdenas, M. and Arancibia, H. (2014). Statistical power and effect size calculation in G*power: complements to statistical significance testing and its application in psychology. *Health & Society*, 5(2), 210-224.

Chan, K. T. (2021). Embedding formative assessment in blended learning environment: The case of secondary Chinese language teaching in Singapore. *Education Sciences*, *11*(7), 360. https://doi.

Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2012). *Classroom Assessment for Student Learning. Doing it right-Using it well* (2nd ed.). Pearson.

Dagdag, J. M. H. and Dagdag, J. D. (2020). Constructivism and the mathematics classroom assessments of elementary teachers. *Journal of Critical Reviews*, *7*(12), 816-823. https://doi.org/10.31838/jcr.07.12.144

Demosthenous, E., Christou, C., and Pitta-Pantazi, D. (2021). Mathematics classroom assessment: A framework for designing assessment tasks and interpreting students' responses. *European Journal of Investigation in Health, Psychology and Education*, *11*(3), 1088-1106.

Ebel, R. L. and Frisbie, D.A. (1986). *Essentials of Education Measurement* (4th ed.). Prentice-Hall.

George, D. and Mallery, P. (2003). *SPSS for Windows step by step: A Simple Guide and Reference* (4th ed.). Allyn and Bacon.

Herbert, S. (2021). Overcoming Challenges in Assessing Mathematical Reasoning. *Australian Journal of Teacher Education*, *46*(8), 17-30. https://doi.org/10.14221/ajte.2021v46n8.2

INEE (2018). *Planea Resultados nacionales 2018, 6to primaria. Language and communication, Mathematics*. Plan Nacional de Evaluación de los aprendizajes, Instituto Nacional para la evaluación de la Educación. https://www.

Lara, L. and Saracostti, M. (2019). Effect of parental involvement on children's academic achievement in Chile. *Frontiers in Psychology*, *10*, 1464.

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., and Loyens, S. (2021). Formative assessment as practice: the role of students' motivation. *Assessment and Evaluation in Higher Education*, *46*(2), 236-255. https://doi.org/10.1080/02602938.2020.1765228

Linzer, D. A. and Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software, 42*(10), 1-29.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). The exploratory factor analysis of items: A practical, revised and updated guide. *Anales de Psicología*, *30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361.

Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods, Vol. 2: Statistical Analysis* (pp. 551-611). Oxford University Press.

Montoya, O. (2007). Application of factor analysis to market research. Case study. *Scientia Et Technica, 13*(35), 281-286. https://www.redalyc.org/pdf/849/84903549.pdf

Navarro-Mosquera, N. G., Falconí-Asanza, A. V., and Espinoza-Cordero J. (2017). Improving the evaluation process of students in basic education. *Universidad y Sociedad, 9*(4), 58-69. http://scielo.

Nylund-Gibson, K. and Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science, 4*(4), 440- 461. https://doi.org/10.1037/tps0000176

Ondé, D. and Alvarado, J. M. (2019). Latent class analysis as a typology identification technique. *INFAD Journal of Psychology. International Journal of Developmental and Educational Psychology-, 5(1), 251-260...* https://doi.org/10.17060/ijodaep.2019.n1.v5.1641

Organisation for Economic Co-operation and Development (OECD) (2023). *PISA 2022 Results (Vol I and II) - Country Notes: Mexico.* https://www.oecd.org/publication/pisa-2022-results/country-notes/mexico-519eaf88/#chapter-d1e11

Paderes, B. M. and Ramos, C. (2024). Teachers' Decision-Making in Grading Students at Risk of Academic Failure during the Implementation of the Basic Education Learning Continuity Plan. *FEU Journal of Graduate Students' Research, 1*(1), 1-17.

Polya, G. (1965). *How to pose and solve problems* (J. Zugazagoitia, Transl.). Editorial Trillas.

Ravela, P., Picaroni, B., and Loureiro, G. (2017). *How to improve classroom assessment? Reflections and work proposals for teachers* (2nd ed.) Grupo Magro Editores.

Sayac, N. and Veldhuis, M. (2021). Mathematics Assessment Practices of Primary School Teachers in France. *International Journal of Science and Mathematics Education, 20*, 1595-1610. https://doi.org/10.1007/s10763-021-10229-3

Schellekens, L. H., Bok, H., de Jong, L. H., van der Schaaf, M. F., Kremer, W., and Van der Vleuten, C. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, *71*, 101094.

Sheikh, A. B. M. and Manap, M. N. A. (2024). Classroom Assessment Practices. *Open Journal of Social Sciences*, *12*(3), 239-258. https://doi.org/10.4236/jss.2024.123018

Schildkamp, K., Van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, *103*, 101602.

Tariq, E., Asif, A., and Nazirullah (2023). Assessment Practices and Assessment Skills of Teachers at Secondary School Level: Factorability and Associative Perspective. *Journal of Peace, Development and Communication, 7*(01), 101-116. https://doi.org/10.36968/JPDC-V07-I01-10

Veldhuis, M. and Van Den Heuvel-Panhuizen, M. (2014). Primary school teachers' assessment profiles in mathematics education. *PLoS ONE*, *9*(1), e86817.

Wafubwa, R. N. and Csíkos, C. (2021). Formative Assessment as a Predictor of Mathematics Teachers' Levels of Metacognitive Regulation. *International Journal of Instruction*, *14*(1), 983-998.

Weller, B. E., Bowen, N. K., and Faubert, S. J. (2020). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4), 287-311. https://doi.

Zhao, X., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2018). Chinese Primary School Mathematics Teachers' Assessment Profiles: Findings from a Large-Scale Questionnaire Survey. *International Journal of Science and Mathematics Education*, *16*(7), 1387-1407. https://doi.org/10.1007/s10763-017-9841-3.

Zulliger, S., Buholzer, A., & Ruelmann, M. (2022). Observed Quality of Formative Peer and Self-Assessment in Everyday Mathematics Teaching and its Effects on Student Performance. *European Journal of Educational Research*, *11*(2), 663-680. https://doi.org/10.12973/eu-jer.11.2.663



Date received: 4 February, 2024.
Review date: 13 March, 2024.
Acceptance date: 12 November, 2024