Construct validity of an instrument to assess assertive feedback in initial teacher training

Validez de constructo de un instrumento para evaluar la retroalimentación asertiva en la formación inicial del profesorado

María de la Luz BERLANGA RAMÍREZ. Lecturer-Researcher. Escuela Normal Superior del Estado de Coahuila *(mtra.luzberlangarmz@gmail.com).*

Luis Gibran JUÁREZ HERNÁNDEZ, PhD. Lecturer-Researcher. Centro Universitario CIFE *(luisgibrancife@gmail.com).*

Abstract:

Feedback in the evaluation process has become more important in teaching practice since the start of the Covid-19 pandemic. The aim of the present study is to analyse the construct validity and reliability of the Socioformative Analytical Rubric for the Assessment of Assertive Feedback (RASERA). This instrument was applied to a sample of 525 students from normal schools in Mexico. Exploratory and confirmatory factor analysis were used to analyse its construct validity. Its reliability was analysed using Cronbach's alpha. The results of the first analysis revealed the formation of two factors; the first, we called execution of assertive feedback and the second, representativeness of assertive feedback. These two factors explained more than 65% of the variance and all of the items with significant factor loadings were found in them (FL > 0.50). For its part, the CFA revealed a good fit of this model (Ratio χ^2/df: 2.284; GFI: 0.909; RMSEA: 0.068; RMR: 0.035; CFI: 0.966; TLI: 0.955). For each factor, the average variance extracted, and the composite reliability were pertinent (AVE > 0.50 and CR > 0.70) and each item showed an adequate standardised factor load (SFL > 0.50). The reliability analysis gave optimal factor values (Cronbach's alpha and McDonald's omega > 0.85). We conclude that the RASERA instrument has adequate psychometric properties.

Keywords: factor analysis, assertiveness, evaluation, teacher training, feedback, validation.

Revision accepted: 2022-06-23.

This is the English version of an article originally printed in Spanish in issue 283 of the **revista española de pedagogía**. For this reason, the abbreviation EV has been added to the page numbers. Please, cite this article as follows: Berlanga Ramírez, M. L., & Juárez Hernández, L. G. (2022). Validez de constructo de un instrumento para evaluar la retroalimentación asertiva en la formación inicial del profesorado | *Construct validity of an instrument to assess assertive feedback in initial teacher training. Revista Española de Pedagogía, 80* (283), 583-600. https://doi.org/10.22550/REP80-3-2022-08

https://revistadepedagogia.org/ ISSN: 0034-9461 (Print), 2174-0909 (Online)

Resumen:

La retroalimentación en el proceso de evaluación cobró mayor relevancia en las prácticas docentes durante la pandemia de COVID-19. El objetivo del presente estudio fue realizar un análisis de la validez de constructo y confiabilidad del instrumento «Rúbrica Analítica Socioformativa para la Evaluación de la Retroalimentación Asertiva» (RASERA). El instrumento se aplicó a una muestra de 525 discentes de escuelas normales en México. Para el análisis de validez de constructo se empleó el análisis factorial exploratorio y confirmatorio, y el análisis de confiabilidad se efectuó mediante el Alfa de Cronbach. Los resultados del primer análisis mostraron la conformación de dos factores; el primero se denominó Ejecución de la retroalimentación asertiva y el segundo Representatividad de la retroalimentación asertiva. Ambos factores explicaron más 65 % de la varianza y en donde se encuentran incorporados todos los ítems con cargas factoriales significativas (CF>0.50). Por su parte, el AFC reveló un buen ajuste de este modelo (Razón $\chi 2$/gl: 2.284; GFI: 0.909; RMSAE: 0.068; RMR: 0.035; CFI: 0.966; TLI: 0.955). Se especifica que, para cada factor, la varianza media extraída y la confiabilidad compuesta fue pertinente (VME>0.50 y CC>0.70) y cada ítem manifestó una carga factorial estandarizada adecuada (CFE>0.50). Concerniente al análisis de confiabilidad, se obtuvieron valores óptimos por factor (Alfa de Cronbach y Omega de McDonald>0.85). Se concluye que el dispositivo RASERA posee propiedades psicométricas adecuadas.

Descriptores: análisis factorial, asertividad, evaluación, formación del profesado, retroalimentación, validación.

1. Introduction

Educational research has studied various aspects of evaluation: its evolution, defined through evaluation models, its application, and its methodology. During the Covid-19 pandemic, studies have centred on feedback resulting from evaluation in virtual education (Castro et al., 2020; Miguel, 2020; Temesio et al., 2021) owing to its importance in decision making by teachers to ensure that students meet learning targets. García-Jiménez (2015) defines feedback as the relevant information for learners to ensure that they are aware of their learning progress, can become conscious of it, and can take decisions about which metacognitive strategies to apply to improve their performance by improving evidence of learning. Furthermore, a range of types of feedback have been described: retroactive-proactive, intrinsic-extrinsic, corrective-indicative, assertive, specific-general, content-based, product-based, process oriented, self-regulation centred, and centred on the person itself, evaluative, descriptive and returnable, prescriptive, informative, confrontational, cathartic, catalytic and supportive, evaluative-descriptive, and negative-positive (Berlanga & Juárez, 2020a).

Of those listed, we focus here on assertive feedback. This entails consist-

ently, respectfully, and cordially evaluating the learner's performance throughout all of the learning process and, is given in an appropriate way for undertaking the corresponding improvements (Berlanga & Juárez, 2020b). This new feedback proposal has its genesis in socioformative evaluation, which includes stages of diagnosis, continuous appraisal, and feedback, the aim of which is to achieve the established goals by developing the learner's talent (Tobón, 2017).

In contrast, Sadler (1989) posits the importance of involving learners in feedback making available the information obtained to improve their performance. On the same lines, Bordas and Cabrera (2001) refer to forming evaluation, distinguishing it from formative evaluation because the feedback emanates from the learner's initiative and reflection.

This being so, assertive feedback corresponds to what is outlined in forming evaluation and its use offers a variety of benefits, such as fostering self-evaluation and self-regulation of learning because the learner has a more active role and can lead the evaluation, as it generates metacognitive mechanisms. Likewise, it facilitates consensual and participatory evaluation and quality information to guide learners in improving their learning (Berlanga & Juárez, 2020a). Another advantage of adding assertiveness to feedback, is that the teacher creates a positive classroom environment, and so helps provide a solid and integrated foundation for the learner's education (Monje et al., 2009; Triana & Velásquez, 2014).

Tobón (2013) has set out the requirements for implementing assertive feedback. The first is that it should be offered immediately while carrying out the evaluated activity; the second is that it is important to start by underlining achievements and positive aspects, with the purpose of accentuating motivation, as Mejía and Pasek de Pinto (2017) note. Other requirements are that the teacher must guide learners with reasoning and respect, while simultaneously giving them the opportunity to make improvements, so that feedback does not become a series of instructions.

Berlanga and Juárez (2020a) affirm that associating assertive feedback with feedforward (García-Jiménez, 2015) establishes a systematic and optimal process for evaluating learning. Nonetheless, it is advisable to incorporate the stages of the feedback for learning model proposed by Quezada and Salinas (2021): literacy, signification, construction, comparison, reworking, and visualisation.

Assertive feedback is understood from its perspectives as a dialogic and sustainable activity. The first perspective derives from the interaction between the teacher and the learner; while the second is because when it is combined with feedforward, there is an effort to improve the current evidence and identify future learning needs (Quezada & Salinas, 2021).

There are a number of instruments for examining feedback in higher education,

but despite the importance of feedback in education, we found no contributions for evaluating assertive feedback. In view of the above, Berlanga and Juárez (2020a) proposed the Socioformative Analytical Rubric for the Assessment of Assertive Feedback (RASERA), a teacher-training instrument that makes it possible to analyse various aspects of assertive feedback: its focus, execution, and representativeness in the initial training of teachers. It also provides data that give information for improving teachers' performance in evaluating, for example, establishing the degree of assertiveness in the evaluation process, because learners are more likely to feel confident in expressing doubts, confusions or concerns when this is given respectfully and cordially and so are more likely to be capable of leading their feedback themselves (Tobón, 2017), and identifying other types of feedback in teaching practices and identifying features they share with assertive feedback and how they differ from it (Tunstall & Gipps, 1996; Torrance & Pryor, 1998; Randall & Thornton, 2005; Hattie & Timperley, 2007; Farahman & Masoud, 2011; Evans, 2013; Jonsson, 2013; García-Jiménez, 2015; Contreras & Zuñiga, 2019).

Similarly, it provides information about the circumstances in which feedback is given with the aim of recognising which ones facilitate or hinder its realisation (Wiggins, 2011; Padilla & Gil, 2008; Wiliam, 2011; Martínez-Rizo, 2013; García, 2015; López & Osorio, 2016). When describing, analysing and evaluating a process of evaluation such as feedback in teaching practice, meta-evaluation is elicited, a key element of the socioformative focus, which involves reviewing the process of evaluation and its effects (Díaz, 2001).

The instrument considers the aspects of the focus, execution, and representativeness of assertive feedback. It includes 16 items and was constructed in the format of a socioformative analytical rubric, meaning that for each item there are levels of action, each of which comprises one descriptor. Socioformative analytical rubrics are tools that facilitate detailed evaluation of the performance of an individual in resolving a conflict that has to be solved in a given context. In accordance with socioformative evaluation, the levels that represent the progression of the competences go from the basic to the most complex. The specific and qualitative features of the activities to be evaluated in each domain are connected to the descriptors. For these elements the socioformative taxonomy was taken as a basis. This is a set of actions centred on meeting the challenges of the knowledge society, and so the levels of action of the socioformative focus were also considered: preformal, receptive, problem-solving, autonomous, and strategic (Tobón, 2017).

After the design of the rubric, it was subjected to a process of review by experts and expert judgement, which confirmed its face validity and content validity (Berlanga & Juárez, 2020a). This process established that the elements of the instrument are appropriate, rele-

vant, pertinent, and representative of the attribute or target construct (Connell et al., 2018; Koller et al., 2017). Carvajal et al. (2011) note that evaluating the psychometric properties of an instrument is an essential criterion for determining the quality of its measurement. Construct validity stands out among them and is regarded as the principal type of validity (Pérez-Gil et al., 2000; Messick, 1980) as it determines the link in the instrument between theory and the conceptualisation that supports the construct, and also verifies whether the structure of the instrument truly reproduces that of the proposed construct (Lagunes-Córdoba, 2017). This property is defined as the integral validation that "subsumes the relevance and representativeness of the content, as well as the relations with the criteria, as both give meaning to the scores of the tests" (Martínez, 1995, p.335, own translation). For its part, reliability refers to the capacity to obtain measurements with minimal error (Jabrayilov et al., 2016).

Consequently, as a result of the significance and relevance of the psychometric properties set out, the aim of the present work was to analyse the construct validity and reliability of the RASERA instrument for evaluating assertive feedback provided by teachers.

2. Material and Methods

2.1. Type of study

We carried out an instrumental study, which included the development of instruments and analysis of the psychometric properties of an instrument (Ato et al., 2013).

2.2. Process

The study of the validity and reliability of the instrument was done in the following phases:

1. Instrument. The RASERA instrument (Berlanga & Juárez, 2020a) comprises 16 items, which include aspects relating to the focus, execution, and representativeness of feedback. The rubric was first subjected to face validation through expert review and content validation through expert judgement. Pilot testing of the instrument was carried out, through which an initial analysis of the reliability and suitability of the instrument for the target population was performed (Berlanga & Juárez, 2020a). Consequently, the instrument was validated in terms of face and content validity, and through the piloting, it was determined that the understanding of instructions and items was optimal. In this pilot trial, the reliability was optimal (Cronbach's alpha: 0.906; 95% CI: 0.818 ± 0.963) (Berlanga & Juárez, 2020a).

After the process of review, expert judgements, and implementation of the pilot trial, the RASERA instrument was as shown in Table 1 (Berlanga & Juárez, 2020a).

TABLE 1. Socioformative Analytical Rubric for Evaluation of Assertive Feedback (RASERA).

Item
What is the purpose of the feedback the teacher provides?
What is the focus of the feedback received?
What type of feedback does the teacher provide?
Is the feedback received at a time that is appropriate and in line with the level of complexity of the learning outcome evaluated?
How often is feedback given?
Is feedback generated on the basis of an evaluation instrument?
What information does the teacher's feedback provide?
Is the process of feedback done to foster self-regulation?
Does the teacher propose evaluation targets?
Does the teacher set evaluation standards?
How is the communication by the teacher during the feedback?
What is the attitude of the teacher when a student disagrees with the evaluation or has doubts about it?
When the teacher provides me with feedback, what is my role?
As a student, how do I use the feedback provided by the teacher?
As a student, what relevance do I give the feedback?
Is the feedback meaningful for my teacher training?

Source: Berlanga & Juárez (2020a).

2. Selection of the sample population for application of the instrument. The instrument was applied in normal schools in the state of Coahuila, Mexico. The tool was applied to trainee teachers from ongoing semesters without prior notice, and so non-probability convenience sampling was used. A total of 525 respondents was obtained from the Escuela Normal Preescolar, Benemérita Escuela Normal de Coahuila, Escuela Normal Superior, Escuela Normal Regional de Especialización, and Escuela Normal de Educación Física. The sample comprised 77.5% women and 22.5% men, with a mean age of 26.95 years and with a mean of 2.8 years of study of teacher training. With regards to the place of origin of the sample, 390 learners were from the municipality of Saltillo in the state of Coahuila and 135 from other municipalities in the same state. There were also four learners who were from cities in other states of Mexico as a result of the student exchange in some of the country's normal schools. At the same time as the application of the rubric, the "Instrument satisfaction

questionnaire" (CIFE, 2018) was applied with the purpose of evaluating how easy the instrument is to complete. In order to comply with ethical research criteria, the participants were informed of the aim of the instrument, informed consent was sought, and personal data was protected (General Law on the Protection of Personal Data Held by Obligated Parties, 2017).

3. Construct validity and reliability analysis. Firstly, the fit of the items to the normal distribution was analysed by calculating skew and kurtosis with items with a value greater than ± 2 being eliminated (Bollen & Long, 1993). The item–test correlation was also examined with the aim of identifying items with a value lower than 0.20 or greater than 0.90, which were eliminated (Tabachnick & Fidell, 2001).

The sample was then divided into two equal parts to perform a cross validation (Brown, 2015). The first part was analysed using exploratory factor analysis (EFA) and the second with confirmatory factor analysis (CFA). It is important to note that the sample was divided by a process of randomisation. This was used to avoid any bias or pattern, using random numbers through an electronic spreadsheet. To proceed with the exploratory factor analysis, the relevance of the data was verified using the KMO index and Bartlett's test (Howard, 2016; Yong & Pearce, 2013). Having checked that the items fit the normal distribution, exploratory factor analysis was performed, selecting the maximum likelihood method (Howard, 2016; Yong & Pearce, 2013). The number of factors to retain is based on the Guttman-Kaiser criterion, scree plot, explained variance, and the eigenvalue > 1 criterion (Henson & Roberts, 2006). Following the analysis of the factor matrix, if factor complexity was found, the matrix was rotated using the most appropriate algorithm.

As with the factor structure obtained through EFA, CFA was carried out on the second subsample using the maximum likelihood estimation method. With regards to sample size, we followed the guidance of Kline (2015), who suggests a sample size of between 200 and 400 participants. Specifically, the goodness of fit of the model was evaluated using chi-squared, the chi-squared/degrees of freedom ratio (χ^2/df), and indices of fit (goodness) of fit index (GFI); root mean square error of approximation (RMSEA), root mean square residual (RMR); comparative fit index (CFI); and Tucker-Lewis index (TLI), considering the criteria proposed by Yuan (2005) and Blunch (2013). Subsequently, the average variance extracted and composite reliability were calculated in accordance with Fornell and Larcker (1981), taking as the threshold for the former values greater than 0.50 and for the latter 0.70 (Hair et al, 2014). Based on what Hair et al. (2014) state, for each factor we checked whether the standardised factor loadings by item were greater than 0.5, the average variance extracted greater than 0.5, and the composite reliability greater than 0.7. Finally, we calculated reliability using Cronbach's alpha coefficient (Cronbach, 1951)

with 95% confidence intervals (Koning & Frances, 2003), as well as the omega coefficient of reliability (McDonald, 1999). The criteria established by Taber (2018) were used for the values obtained from these coefficients.

The skew, kurtosis, item–test correlation, Cronbach's alpha coefficient, McDonald's omega coefficient, and the exploratory and confirmatory factor analysis were calculated using JASP version 0.11.1 (JASP Team, 2019) software.

3. Analysis and results

3.1. Analysis of construct validity and reliability

According to the analysis, as Table 2 shows, none of the items broke the criteria established for skew and kurtosis, and so the data can be assumed to have a normal distribution. No item had a value lower than 0.20 or greater than 0.90 on the item–test–correlation indicator, and so there was no need to eliminate any of them.

TABLE 2. Skew and kurtosis of the items.

Item	Skew	Kurtosis	Item–test correlation
1	0.473	-0.922	0.673
2	0.492	-1.041	0.609
3	0.509	-0.989	0.664
4	-0.005	-0.957	0.752
5	0.34	-1.037	0.681
6	0.462	-1.219	0.691
7	0.677	-0.816	0.556
8	0.53	-1.119	0.727
9	0.004	-1.526	0.666
10	0.062	-1.577	0.751
11	-0.206	-1.385	0.739
12	-0.069	-0.993	0.554
13	0.433	-1.156	0.618
14	0.037	-1.248	0.606
15	-0.17	-1.429	0.719
16	-0.339	-1.466	0.732

Source: Own elaboration.

The Kaiser–Meyer–Olkin test (KMO: 0.961) and Bartlett's test of sphericity ($X2$: 3751.286 df: 120; $p < 0.00001$) were used to test the relevance of the data to be analysed through EFA. The matrix of communalities represents all of the items within the factor model as Table 3 shows. The factor extraction found divergences with regards to the theoretical model, as two factors with an eigenvalue greater than 1 were found, which together explain more than 65% of the variance. It is worth noting that two items were found with factor loading in more than one factor, and so the matrix was rotated and the loadings were clarified. The factor structure obtained specifically indicates that items 1 to 12 were represented in factor one, and so we called this Execution of feedback. Factor two included items 13, 14, 15 and 16, and so we named it Representativeness of feedback. Table 3 shows the factor structure of both factors.

TABLE 3. Communalities and factor loadings.

Item	Communality	Factor Loading	
		Factor 1	Factor 2
1	0.629	0.820	.
2	0.565	0.674	.
3	0.622	0.631	.
4	0.718	0.661	.
5	0.687	0.750	.
6	0.839	0.834	.
7	0.573	0.763	.
8	0.818	0.892	.
9	0.638	0.771	.
10	0.771	0.657	.
11	0.638	0.608	.
12	0.447	0.709	.
13	0.611	.	0.583
14	0.573	.	0.776
15	0.686	.	0.719
16	0.707	.	0.732

Source: Own elaboration.

The confirmatory factor analysis revealed a good fit for the two-factor model. In particular, optimal values were apparent in the value of the ratio between chi-squared and degrees of freedom (c2/df: 2.284), goodness of fit index (GFI: 0.909), root mean square error of approximation (RMSEA: 0.068, 90% CI: 0.057 ± 0.080, p > 0.005), root mean square residual (RMR: 0.035), comparative fit index (CFI: 0.966), and Tucker-Lewis index (TLI: 0.955).

The composite reliability, average variance extracted by factor, and standardised factor loading by item are presented in the summary of the model in Table 4. Each factor fulfilled the condition of the standardised factor loading by item (SFL > 0.50), average variance extracted (AVE > 0.50), and composite reliability (CR > 0.70). The above is shown in Table 4 and Graph 1. Finally, the reliability (alpha and omega) by factor was optimal (Table 4).

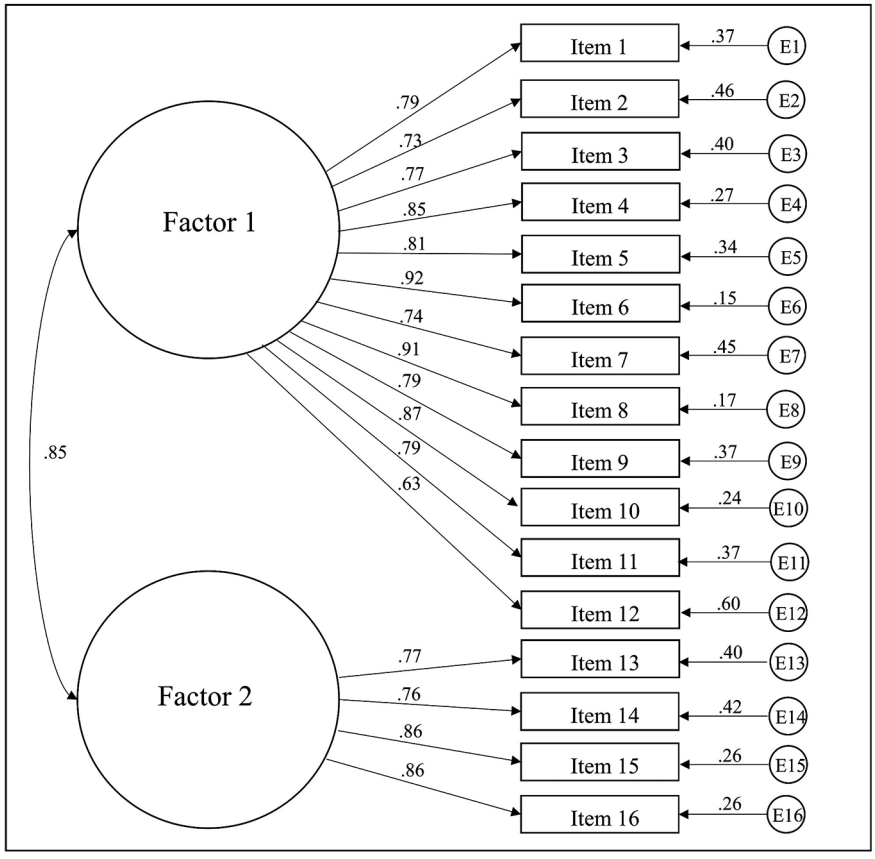
With regards to the analysis of feasibility of the instrument (Table 5), it can be seen that the perception of the respondents regarding comprehension of instructions and items and satisfaction with the instrument was evaluated as excellent.

TABLE 4. Summary of the model.

Factor	Number of items	Standardised factor loadings	Average variance extracted	Composite reliability	Cronbach's alpha (95% CI)	McDonald's omega
Execution of feedback	12	I1 (0.79), I2 (0.73), I3 (0.77), I4 (0.85), I5 (0.81), I6 (0.92), I7 (0.74), I8 (0.91), I9 (0.79), I10 (0.87), I11 (0.79), I12 (0.63)	0.646	0.955	0.95 (0.940 ± 0.958)	0.922
Representativeness of feedback	4	I13 (0.77), I14 (0.76), I15 (0.86), I16 (0.86)	0.662	0.886	0.88 (0.853 ± 0.902)	0.851

Source: Own elaboration.

GRAPH 1. Representation of the Confirmatory Factor Analysis of the Two-Dimensional Model.



Source: Own elaboration.

TABLE 5. Analysis of satisfaction with the instrument.

Questions	Low degree (%)	Acceptable degree (%)	Good degree (%)	Excellent degree (%)
Understanding of instructions	0.6	5.7	37.9	55.8
Understanding of items	0.4	6.9	40.2	52.6
Satisfaction with instrument	0.4	5.5	32.8	61.3

Source: Own elaboration.

4. Conclusions

Assertive feedback, defined through socioformative evaluation, makes it possible to evaluate the entire process of acquiring learning and developing competences, thus offering an opportunity to make the necessary adjustments to facilitate reorientation of learning (Tobón, 2017). This shows its importance in the education and learning process, since it enables the development of cognitive mechanisms in the student (Shute, 2008), which result in self-regulation of the learning process, thus making learners self-educating (García-Jiménez, 2015).

Evaluating assertive feedback in initial and continuing teacher training has a dual advantage on the basis of the results achieved. Theoretical-methodological elements complement one another and a taxonomy and instruments emerge that are more in line with the professional and disciplinary competences of the teacher for evaluating his or her performance from the socioformative focus. The other advantage lies in promoting assertive feedback in teacher training, given that it is one of the pillars for evaluating competences (Tobón, 2017).

As stated above, evaluation of the psychometric properties of an instrument is a fundamental criterion for determining the quality of its measurement (Carvajal et al., 2011). The proposed RASERA instrument was initially subjected to a process of review by experts and content validation by expert judgement, which firstly established the pertinence of the items to the phenomenon, their relevance, their wording, and whether they are understandable for the target population (Connell et al., 2018). The analysis of the content validity also showed that the items from the instrument cover the domain of content of the construct, indicating its pertinence, relevance, and representativeness (Koller et al., 2017).

This process was highly significant given that content validity is a fundamental component of construct validity (Messick, 1980). This property is regarded as fundamental because it determines the relevance of the items to the objective that has been designed as well as how much each of them represents the construct evaluated (Messick, 1980). In this respect, Furr (2020) asserts that construct validity is the degree to which the relations expected under the theory and the definitions that support the construct are confirmed.

Regarding the analysis of this property in the present work, we carried out a cross validation process, which corresponds with the classical, greatest relevance recommendation (Brown, 2015; Lloret-Segura et al., 2014). The first approach was done through exploratory factor analysis, which has the objective of identifying the factor structure underlying the items (Lloret-Segura et al., 2014) and observing the correspondence of the resulting factor structure with the theoretical proposal. The second approach was done with confirmatory factor analysis to validate the sustainability of the factor structure obtained with EFA and, consequently, the validity of the theoretical inferences made from it (Leyva, 2011).

The results of the EFA revealed discrepancies with what was theoretically proposed, as a one-dimensional model was initially considered but a model with two dimensions (factors) was found. This fit is functional in nature given that the theoretical elements listed directly and represented in the items were concentrated sequentially. For example, factor one (Execution of feedback) includes items that target aspects of focus (García-Jiménez, 2015; Tobón, 2017), information on the conditions in which it is created (Martínez -Rizo, 2013; Padilla & Gil, 2008), types (Tunstall & Gipps, 1996; Torrance & Pryor, 1998; Randall & Thornton, 2005; Hattie & Timperley, 2007; Farahman & Masoud, 2011; Evans, 2013; Jonsson, 2013; García, 2015; Contreras & Zuñiga, 2019), time (Tobón, 2017), frequency (García-Jiménez, 2015; Wiggins, 2011), implementation (Tobón, 2017), information created based on feedback (Jónsson et al., 2018; Anijovich & Cappelletti, 2017; García-Jiménez, 2015), self-regulation of the learner (García-Jiménez, 2015; Tobón, 2017; Quezada & Salinas, 2021), setting of evaluation targets, establishing evaluation rules, and interaction between teacher and learner during feedback (Tobón, 2017).

Meanwhile, factor two (Representativeness of feedback) addressed the role of the learner in the feedback, its use, and the importance given to it in their teacher training (Tobón, 2017). Both factors explained more than 56% of variance and there was representativeness and correspondence of 100% of the items from the tool, indeed, in all cases they displayed significant factor loadings (FL > 0.50), reflecting their significance and representativeness in the construct evaluated (Lagunes-Córdoba, 2017).

Evaluation through CFA provided elements of confirmation of the fit of the model to the data, which make it possible to validate the empirical sustainability of the proposed theoretical model (Herrero, 2010; Yuan, 2005) since the χ^2/df ratio, and the indices of fit used displayed optimal values. One significant aspect contributed by this analysis is the value of the standardised factor loadings (SFL > 0.50), composite reliability (CR > 0.70), and average variance extracted (AVE > 0.50). This set of results makes it possible to verify the empirical sustainability of the proposed model, and to note that the proposed indicators adequately measure this factor (Cheung & Wang, 2017; Fornell & Larcker, 1981).

Regarding the reliability analysis (Cronbach's alpha, McDonald's omega, and composite reliability), optimal values were obtained for each factor, indicating that at least 70% of the variance in the measurements in the instrument is free from error (Cho & Kim, 2015; Viladrich et al., 2017). Likewise, an optimal value was obtained with the pilot group (Cronbach's alpha: 0.906, 95% CI: 0.818 ± 0.963) (Berlanga & Juárez, 2020a). It is important to note that this calculation was done with 15 students and, as Charter (2003) states, the potential and consistency of the coefficient is determined by the sample size. On similar lines to the above, and with the sample size, in the present study, the value of the reliability was optimal and according to the confidence intervals, greater stability in the measurement is indicated.

With regards to this, it is appropriate to note the position of Jabrayilov et al. (2016) who argue that the reliability of an instrument refers to its ability to obtain measurements with minimal error and, shows the correlation between the items and the concept studied (Gliner et al., 2001).

One very important aspect analysed is the degree of satisfaction with the instrument or its feasibility (Carvajal et al., 2011), which considers elements associated with the instrument's ease, length, and time needed to complete it; the degree of comprehension of instructions and items; and the clarity of the wording (Conell et al., 2018; Halek et al., 2017). It is notable that the instrument presented here obtained an excellent evaluation regarding comprehension of instructions and items, as well as satisfaction with the instrument. Carvajal et al. (2011) underline that these aspects are relevant because they are threats to the validity, reliability, and precision of the instrument.

The methodological process followed means we can state that the RASERA mechanism is optimal given that its content and construct are valid and it is also reliable. The demonstration of the psychometric properties analysed in the present work means we can state that the instrument provides valid and reliable information (Mendoza-Mendoza & Garza, 2009). Based on this, we note that the proposed RASERA instrument provides a valid and reliable diagnostic tool.

We propose expanding its use owing to the benefits obtained in the execution and significance of assertive feedback, given that it has an influence on the improvement of the performance of the teacher and the learner. In relation to the action of the teacher, we propose its use because it would facilitate identification of the degree of assertiveness in the guidance given during evaluation, and because of its repercussion in the development of competences, the conditions in which it is provided: time, conduct, context (Canabal & Margalef, 2017), the implementation of evaluation, and the level of information (García-Jiménez, 2015; Jónsson et al., 2018). Regarding learner performance, when teachers incorporate assertive feedback into their educational praxis, this also has an impact on the cognitive and emotional realm of the learner.

Meanwhile, the evaluation of assertive feedback by learners is of value given that the information gathered comes into play in their performance and in the performance of the teacher. It is also important to note that knowledge of the teacher's perception would consolidate a complete meta-evaluation and would facilitate self-evaluation with the intention of contributing to the improvement of educational praxis relating to the process of evaluation. Consequently, a future line of research would be to construct an instrument that enables teachers to self-evaluate the degree of assertiveness in their feedback and the conditions in which it is executed. Furthermore, in the present work we did not consider the application of any other instrument that includes an aspect of assertiveness, whether as behaviour or communication, which limits understanding of the convergent and concurrent validity of the measurement instrument presented here.

References

Anijovich, R., & Cappelletti, G. (2017). *Más allá de las pruebas: la retroalimentación en la evaluación como oportunidad* [*Beyond testing: Feedback. in the evaluation as an opportunity*]. Paidós.

Ato, M., López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología [A classification system for research designs in psychology]. *Anales de Psicología, 29* (3), 1038-1059.

Berlanga Ramírez, M. L., & Juárez Hernández, L. G. (2020a). Diseño y validación de un instrumento para evaluar la retroalimentación asertiva en educación normal [Design and validation of an instrument to evaluate assertive feedback in Normal Education]. *IE Revista de Investigación Educativa de la REDIECH, 11*, e-791. https://doi.org/10.33010/ie_rie_rediech.v11i0.791

Berlanga Ramírez, M. L., & Juárez Hernández, L. G. (2020b). Paradigmas de evaluación: del tradicional al socioformativo [Evaluation paradigms: From the traditional to the socio-formative]. *Diálogos sobre educación. Temas actuales en investigación educativa*, *21*, 1-14. https://doi.org/10.32870/dse.v0i21.646

Blunch, N. (2013). *Introduction to structural equation modeling using IBM SPSS statistics and AMOS*. Sage.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Sage.

Bordas, M. I., & Cabrera, F. Á. (2001). Estrategias de evaluación de los aprendizajes centrados en el proceso [Learning assessment strategies focused on the process]. **revista española de pedagogía**, *59* (218), 25-48.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Canabal, C., & Margalef, L. (2017). La retroalimentación: la clave para una evaluación orientada al aprendizaje [The feedback: A key to learning-oriented assessment]. *Profesorado. Revista de Currículum y Formación de Profesorado, 21* (2), 149-170.

Carvajal, A., Centeno, C., Watson, R., Martínez, M., & Sanz Rubiales, A. (2011). ¿Cómo validar un instrumento de medida de la salud? [How to validate a health measurement instrument?]. *Anales del Sistema Sanitario de Navarra, 34* (1), 63-72.

Castro, S., Paz, L., & Cela, M. (2020). Aprendiendo a enseñar en tiempos de pandemia COVID-19: nuestra experiencia en una universidad pública [Learning to teach in times of the COVID-19 pandemic: Our experience at Universidad de Buenos Aires]. *Revista Digital de Investigación en Docencia Universitaria*, *14* (2), e1271.

Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method and the clinical implications of low reliability. *The Journal of General Psychology, 130* (3), 290-304. https://doi.org/10.1080/00221300309601160

Cheung, G. W., & Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: Issues and solutions. *Academy of Management Proceedings, 2017* (1), 12706. https://doi.org/10.5465/AMBPP.2017.12706abstract

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: wellknown but poorly understood. *Organizational Research Methods, 18* (2), 207-230. https://doi.org/10.1177/1094428114555994

CIFE (2018). *Instrumento «Cuestionario de satisfacción con el instrumento» [Instrument "Questionnaire of satisfaction with the instrument"].* Centro Universitario CIFE. https://docs.google.com/forms/d/e/1FAIpQLSc8-jOiWYwG64QbnhRyGAg1ElTggq2aP1XiSg45pyN9XLbXNQ/viewform

Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T., Barkham, M., Robotham, D., Rose, D., & Brazier, J. (2018). The importance of content and face validity in instrument development: Lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Quality of life research: an international journal of quality-of-life aspects of treatment, care, and rehabilitation, 27* (7), 1893-1902. https://doi.org/10.1007/s11136-018-1847-y

Contreras, G., & Zúñiga, C. G. (2019). Prácticas y concepciones de retroalimentación en formación inicial docente [Practices and conceptions of feedback in initial teacher training]. *Educação e Pesquisa*, *45*, 1-22. https://doi.org/10.1590/s1678-4634201945192953

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests [Coeficiente alfa y la estructura interna de pruebas]. *Psychometrika*, *16* (3), 297-334. https://doi.org/10.1007/BF02310555

Díaz, L. (2001). La metaevaluación y su método [Meta-evaluation and its method]. *Revista de Ciencias Sociales* (Cr), *II-III (*93), 171-192. https://www.redalyc.org/articulo.oa?id=15309314

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83* (1), 70-12. https://doi.org/10.3102/0034654312474350

Farahman, F., & Masoud, Z. (2011). A comparative study of EFL teachers' and intermediate High School students' perceptions of written corrective feedback on grammatical errors. *English Language Teaching, 4* (4), 36-48. http://dx.doi.org/10.5539/elt.v4n4p36

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18* (1), 39-50. https://doi.org/10.2307/3151312

Furr, R. M. (2020). Psychometrics in clinical psychological research. In *The Cambridge handbook of research methods in clinical psychology* (pp. 54–65). Cambridge University Press. https://doi.org/10.1017/9781316995808.008

García-Jiménez, E. (2015). La evaluación del aprendizaje: de la retroalimentación a la autorregulación. El papel de las tecnologías [Assessment of learning: From feedback to self-regulation. The role of technologies]. *Relieve: Revista Electrónica de Investigación y Evaluación Educativa, 21* (2), 1-24. http://dx.doi.org/10.7203/relieve.21.2.7546

General Law on the Protection of Personal Data Held by Obligated Parties. *Official Journal of the Federation,* 26 January 2017. Chamber of Deputies of the H. Congress of the Union. United Mexican States. https://www.diputados.gob.mx/LeyesBiblio/pdf/LGPDPPSO.pdf

Gliner, J. A., Morgan, G. A., & Harmon, R. J. (2001). Measurement reliability. *Journal of the American Academy of Child y Adolescent Psychiatry, 40*, 486-488. https://doi.org/10.1097/00004583-200104000-00019

Hair, Jr., William, C. B., Barry, J. B., & Anderson, R. (2014). *Multivariate data analysis*. Pearson

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77* (1), 81-112. https://doi.org/10.3102/003465430298487

Halek, M., Holle, D., & Bartholomeyczik, S. (2017). Development and evaluation of the content validity, practicability, and feasibility of the Innovative dementia-oriented Assessment system for challenging behaviour in residents with dementia. *BMC health services research*, *17* (1), 1-26.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, *66* (3), 393-416. https://doi.org/10.1177/0013164405282485

Herrero, J. (2010). El análisis factorial confirmatorio en el estudio de la estructura y estabilidad de los instrumentos de evaluación: un ejemplo con el cuestionario de autoestima CA-14 [Confirmatory factor analysis in the study of the structure and stability of assessment instruments: An example with the self-esteem questionnaire (CA-14)]. *Intervention Psychosocial, 19* (3), 289-300. https://doi.org/10.5093/in2010v19n3a9

Howard, C. M. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human–Computer Interaction*, *32* (1), 51-62. https://doi.org/10.1080/10447318.2015.1087664

JASP Team (2019). JASP (Version 0.11.1) [Computer software].

Jabrayilov, R., Emons, W. H. M., & Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment. *Applied Psychological Measurement, 40* (8), 559-572. https://doi.org/10.1177/0146621616664046

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14* (1), 63-76. https://doi.org/10.1177/1469787412467125

Jónsson, I. R., Smith, K., & Geirsdóttir, G. (2018). Shared language of feedback and assessment. Perception of teachers and students in three Icelandic secondary schools. *Studies in Educational Evaluation*, *56*, 52-58. https://doi.org/10.1016/j.stueduc.2017.11.003 Kline, P. (2015). *A handbook of test construction (psychology revivals). Introduction to psychometric design*. Routledge.

Koller, I., Levenson, M. R., & Glück, J. (2017). What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Frontiers in Psychology*, *8*, 126. https://doi.org/10.3389/fpsyg.2017.00126

Koning, A. J., & Frances, P. H. (2003). Confidence intervals for Cronbach's Coefficient Alpha values. *ERIM Report Series Reference* No. ERS-2003-041-MKT. http://hdl.handle.net/1765/431

Lagunes-Córdoba, R. (2017). Recomendaciones sobre los procedimientos de construcción y validación de instrumentos y escalas de medición en la psicología de la salud [Recommendations about procedures for construction and validation of scales in health psychology]. *Revista Psicología y Salud, 27* (1), 5-18.

Leyva, E. (2011). Una reseña sobre la validez de constructo de pruebas referidas a criterio [An overview of the construct validity of criterion-referenced tests]. *Perfiles Educativos*, *33* (131), 131-154. https://doi.org/10.22201/iisue.24486167e.2011.131.24238

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. & Tomas-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada [The exploratory factor analysis of the items: A practical guide, revised and updated]. *Anales de Psicología, 30* (3), 1151-1169. http://dx.doi.org/10.6018/analesps.30.3.199361

López, A., & Osorio, K. (2016). Percepciones de estudiantes sobre la retroalimentación formativa en el proceso de evaluación [Student perceptions about formative feedback in the evaluation process]. *Actualidades Pedagógicas, 68*, 43-64. http://dx.doi.org/10.19052/ap.2829.

Martínez, M. (1995). *Psicometría: teoría de los tests psicológicos y educativos* [*Psychometry: Theory of psychological and educational tests*]. Síntesis.

Martínez-Rizo, F. (2013). Dificultades para implementar la evaluación formativa: revisión de literatura [Difficulties in implementing formative assessment: Literature review]. *Perfiles educativos, 35* (139), 128-150. http://www.scielo.org.mx/pdf/peredu/v35n139/v35n139a9.pdf

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates, Inc.

Mejía, M., & Pasek de Pinto, E. (2017). Proceso general para la evaluación formativa del aprendizaje [General process for formative assessment of learning]. *Revista Iberoamericana de Evaluación Educativa, 10* (1), 177-193. https://doi.org/10.15366/riee2017.10.1.009

Mendoza-Mendoza, J., & Garza, J. B. (2009). La medición en el proceso de investigación científica: Evaluación de validez de contenido y confiabilidad [Measurement in the scientific research process: Content validity and reliability evaluation]. *Innovaciones de Negocios, 6* (11), 17- 32.

Messick, S. (1980). Test validity and ethics of assessment. *American Psychologist, 35* (11) 1012-1027. https://doi.org/10.1037/0003-066X.35.11.1012

Miguel, J. A. (2020). La educación superior en tiempos de pandemia: una visión desde dentro del proceso formativo [Higher education in times of pandemic: A view from within the training process]. *Revista Latinoamericana de Estudios Educativos*, *50* (ESPECIAL), 13-40. https://doi.org/10.48102/rlee.2020.50.ESPECIAL.95

Monje, V., Camacho, M., Rodríguez, E., & Carvajal, L. (2009). Influencia de los estilos de comunicación asertiva de los docentes en el aprendizaje escolar [Influence of teachers' assertive communication styles on learning in schools]. *Psicogente, 12* (21), 78-95.

Padilla, M. T., & Gil, J. (2008). La evaluación orientada al aprendizaje en la educación superior: condiciones y estrategias para su aplicación a la docencia universitaria [Learning-oriented assessment in higher education: conditions and strategies for its application in university education]. **revista española de pedagogía**, *66* (241), 467-486.

Pérez-Gil, J. A., Chacon-Moscoso, S., & Moreno-Rodríguez, R. (2000). Validez de constructo: el uso del análisis factorial exploratorio-confirmatorio para obtener evidencias de validez [Construct validity: The use of factor analysis]. *Psicothema, 12* (2), 442-446. https://www.psicothema.com/pdf/601.pdf

Quezada, S., & Salinas, C. (2021). Modelo de retroalimentación para el aprendizaje: Una propuesta basada en la revisión de literatura [Feedback model for learning: A proposal based on literature review]. *Revista mexicana de investigación educativa, 26* (88), 225-251.

Randall, M., & Thornton, B. (2005). *Advising and supporting teachers*. University Press.

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144. http://dx.doi.org/10.1007/BF00117714

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78* (1), 153-189. https://doi.org/10.3102/0034654307313795

Tabachnick, B. G., & Fidell, L.S. (2001). *Using multivariate statistics*. Allyn y Bacon.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48* (6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Temesio, S., García, S., & Pérez, A. (2021). Rendimiento estudiantil en tiempo de pandemia: percepciones sobre aspectos con mayor impacto [Student performance in times of pandemic: Perceptions of aspects with the greatest impact]. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, *28*, e45. https://doi.org/10.24215/18509959.28.e45

Tobón, S. (2017). *Evaluación socioformativa. Estrategias e instrumentos [Socioformative evaluation. Strategies and instruments]*. Kresearch. https://cife.edu.mx/recursos/wp-content/uploads/2018/08/LIBRO-Evaluaci%C3%B3n-Socioformativa-1.0-1.pdf

Tobón, S. (2013). *Evaluación de las competencias en la edcación básica* [*Evaluation of competencies in basic education*]. Santillana.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment. Teaching, learning and assessment in the classroom*. Open University Press.

Triana, A., & Velásquez, A. (2014). Comunicación asertiva de los docentes y clima emocional del aula en preescolar [Assertive teacher communication and emotional classroom climate in preschools]. *Voces y Silencios: Revista Latinoamericana de Educación, 5* (1), 23-41.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Journal, 22* (4), 389-404. https://doi.org/10.1080/0141192960220402

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Annals of Psychology, 33* (3), 755-782. https://doi.org/10.6018/analesps.33.3.268401

Wiggins, G. (2011). Giving students a voice: The power of feedback to improve teaching. *Educational Horizons, 89* (3), 23-26. https://doi.org/10.1177/0013175X1108900406

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, Solution Tree Press.

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology*, *9* (2), 79-94. https://doi.org/10.20982/tqmp.09.2.p079

Yuan, K. H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research, 40* (1), 115-148. https://doi.org/10.1207/s15327906mbr4001_5

Authors' biographies

María de la Luz Berlanga Ramírez. Doctoral candidate at the Centro Universitario CIFE. Professor-Researcher at the Escuela Normal Superior del Estado de Coahuila. Her current research interests are in the field of evaluation in teacher training.

https://orcid.org/0000-0001-9088-3991

Luis Gibran Juárez-Hernández. Doctor of Biological and Health Sciences from the Universidad Autónoma Metropolitana. Professor–Researcher at the Centro Universitario CIFE. His current research interests are the fields of evaluation instruments, sustainable development, and ecology.

https://orcid.org/0000-0003-0658-6818