Teacher's Achievement Verbalizations Questionnaire in Oral Exams

Cuestionario de Verbalizaciones de Logro Docentes en Exámenes Orales

Javier Sánchez-Rosas * [1] , Sergio Alexis Dominguez-Lara [2] , Luis Alberto Furlan [3] ,
Mirai Okinishi [3]

1 - Catholic University of Temuco, Temuco, Chile.
2 - Private University Norbert Wiener, Lima, Perú.
3 - National University of Cordoba, Córdoba, Argentina.

Abstract

At the moment, there is no measurement instrument to assess verbal expressions and phrases of feedback about students' achievement issued by teachers in the presence of one or more students during an oral exam. This article reports the design, structural validity, reliability and external validity of the Teacher Achievement Verbalizations in Oral Exams Questionnaire (TAVQ), which assesses several teachers' verbalizations from the perspective of students in oral exams. The structural validity, reliability, and external validity were evaluated in a sample of university students (N = 252) from Argentina. Several plausible measurement models were specified based on the dimensions of valence, object focus, and temporal frame, which were tested through confirmatory factor analysis and bifactor analysis. Two scales that measure with very good reliability positive and negative verbalizations related to achievement expressed by teachers during oral exams were validated. These verbalizations showed appropriate relationships with achievement emotions and academic performance. The need for future studies and practical implications are discussed.

Keywords: test anxiety, teacher behavior, feedback, oral exams, achievement emotion

Resumen

Hasta el momento, no existe un instrumento de medición que evalúe las expresiones verbales y comentarios sobre el logro de los alumnos que emiten los docentes en presencia de uno o varios estudiantes durante un examen oral. Este artículo informa la construcción, validez estructural, confiabilidad y validez externa del cuestionario de verbalizaciones de logro docentes en exámenes orales (TAVQ). El mismo evalúa varias verbalizaciones del docente desde la perspectiva de los estudiantes en exámenes orales. La validez estructural, confiabilidad y validez externa fueron evaluadas en una muestra de estudiantes universitarios (N = 252) de Argentina. Se especificaron varios modelos de medición plausibles basados en las dimensiones de valencia, foco y marco temporal, que fueron testeados mediante análisis factorial confirmatorio y bifactor. Se validaron dos escalas que miden con muy buena confiabilidad verbalizaciones de logro positivas y negativas expresadas por docentes durante los exámenes orales. Estas verbalizaciones mostraron relaciones apropiadas con emociones de logro y rendimiento académico. Se discute la necesidad de estudios futuros e implicancias prácticas.

Palabras clave: ansiedad ante los exámenes, comportamiento docente, comentarios, exámenes orales, emoción de logro

Introduction

Oral exams are frequent in higher level education and are designed to assess students' understanding of a subject, as well as the ability to articulate ideas and knowledge effectively (Hazen, 2020; Theobold, 2021). Whoever faces an oral exam must develop skills to respond adequately to a double challenge in a stressful context that can activate several achievement emotions (Furlan & Sánchez-Rosas, 2018). On the one hand, the evaluated content must be communicated in a clear, fluent and organized way, using the specific terms of the subject in an appropriate manner. On the other hand, the student must sustain an asymmetric interpersonal interaction, while processing information related to his/her performance received through the gestural and verbal language of his/her evaluator (Burić, 2015; Gardner & Giordano, 2023; Guo et al., 2022; Puertas-Molero et al., 2022; Putwain et al., 2017). The messages related to achievement can affect control and value appraisals of the ongoing activity and their outcomes, and promote increased achievement emotions (Goetz et al., 2018).

Achievement emotions are defined as emotions that are directly linked to achievement activities or achievement outcomes (Pekrun, 2018, 2021). Achievement emotions are present daily in the academic environment and have the ability to affect students' thoughts, motivation and actions in evaluative situations (Furlan & Sánchez-Rosas, 2018; Pekrun, 2018; Pekrun et al., 2023; Rojas-Torres et al., 2022; Sánchez-Rosas & Furlan, 2017). These emotions are activated by control-value appraisals and the learning context would contribute to their activation by affecting these appraisals (Pekrun, 2018, 2021).

Beyond the existence of individual causes, there is a growing interest in knowing the environmental characteristics that impact students' emo-tional experiences (Dewaele et al., 2018, 2019; Lei et al., 2018; Pekrun et al., 2023; Raccanello et al., 2018; Sánchez-Rosas et al., 2019; Ventura-León et al., 2022; Yang et al., 2021). In this regard, teachers, through their behaviors in the classroom, (Sánchez-Rosas et al., 2016; Sánchez-Rosas & Esquivel, 2016), would play a crucial role in oral exam situations by affecting control-value appraisals of exams (Burić, 2015; Putwain et al., 2022; Reilly & Sánchez-Rosas, 2021).

One specific case on teaching behaviors is teachers' achievement verbalizations in oral exams, which can increase the demands of the task, provide clarity and structure to the exam, or transmit messages related to achievement in terms of success or failure. These achievement verbalizations could impact students' emotions and other important aspects of student performance (Pekrun et al., 2023). When teachers use negative, judgmental, or frightening language, it can generate negative emotions in students that affect students' motivation and academic behavior (Apto et al., 2017; Putwain et al., 2017, 2023). In contrast, if teachers use positive and motivating language, it can lead to less negative emotional, motivational, and behavioral consequences.

Although it is possible to envision the effects of these positive or negative achievement verbalizations on achievement emotions, instruments that allow measuring teachers' achievement verbalizations in oral exams are still nonexistent. To address this gap and provide for related research, a study that seeks to develop a tool to assess various teacher achievement verbalizations in oral exams and to analyze some of their psychometric properties is reported. Specifically, the dimensional structure of a set of items is analyzed as well as the reliability of the measurements made by the resulting scales, and the relationship that teachers' achievement verbalizations have with achievement emotions during oral exams is also tested.

When developing the instrument, an empirical-rational strategy was followed, paying attention to the ability of achievement verbalizations to affect control and value appraisal and, consequently, to activate achievement emotions (Pekrun et al., 2023). In addition, various measurement models were analyzed based on possible combinations according to the valence and object focus of the emotion that achievement verbalizations could activate. Also, although a large pool of items was developed, the final instrument is a brief and practical version that will shorten the length of research protocols in studies with many variables and avoid the presence of less representative items of the construct.

Achievement Emotions and Teacher Behaviors

The CVT offers a frame of reference to define teachers' achievement verbalizations that activate achievement emotions in oral exams, while building scales and validating instruments (Pekrun, 2018, 2021; Pekrun et al., 2023). It is important to note that the CVT establishes that control (e.g., self-efficacy) and value (e.g., task value) appraisals are the direct causes of the activation of achievement emotions. This means that emotions are induced when the individual feels in control of, or out of control of, activities and outcomes that are subjectively important. In turn, the emotions activated in achievement situations, such as oral exams, can be distinguished by their valence (positive vs. negative) or by the object focus of the emotion (activity or outcomet); emotions can even be distinguished as current (activity), prospective (future outcomes), or retrospective (past outcomes) emotions.

The control-value theory postulates that the affective impact of social environments is mediated by control and value appraisals. Ac-

cordingly, it is assumed that the features of environments that deliver information related to controllability and academic values are of critical importance for students' emotions. Important variables include quality of instruction, induction of values, autonomy support, goal structures and achievement-related expectancies of significant others, as well as feedback and consequences of achievement (Pekrun, 2018).

Understanding the role of immediate environmental factors in achievement situations, such as teacher behavior or teacher feedback (Apto et al., 2017; Awad-Igbaria et al., 2022; Frenzel et al., 2021; Narciss et al., 2022) is important because it allows us to understand how they influence the students' constitution of achievement beliefs and expectations in oral exams. Various teaching behaviors have been considered when analyzing their relationship with control-value appraisals and emotions (Becker et al., 2014; Lazarides & Buchholz, 2019; Lei et al., 2018; Sánchez Rosas et al., 2016; Westphal et al., 2018). Some of these behaviors can be categorized as non-verbal behaviors (Derakhshan et al., 2023; Guo et al., 2022; Juma et al., 2022; Puertas-Molero et al., 2022) and their influence can sometimes be ambiguous depending on students' interpretation: space management, gestures, body language, position and body orientation, gaze, facial expression, and paralinguistic features such as tone of voice and rhythm.

Teacher's achievement verbalizations in oral exams

Teachers' achievement verbalizations are messages that have a much more precise capacity to transmit information related to achievement than non-verbal behaviors and, consequently, their effect on appraisals and emotions is clear-

er (Putwain et al., 2017, 2022, 2023). Teachers' achievement verbalizations in oral exams are understood here as verbal expressions, phrases, and feedback on achievement issued by teachers in the presence of one or more students (Narciss et al., 2022; Putwain et al., 2017, 2022, 2023). If, for instance, before starting the exam the teacher warns that the exam will be difficult, this enhances the perceived difficulty of the exam and can elicit anxiety (Putwain et al., 2017). If, by contrast, the teacher gives negative feedback on the current performance (He/she says: *How could you not know that?*), it can activate shame (Apto et al., 2017). By contrast, if the teacher gives positive feedback about current performance or provides support to continue responding, pride and enjoyment can be activated (Pekrun et al., 2023).

As it can be seen, it is possible to think that some verbalizations are more associated with one emotion than others or that they are even related to several emotions simultaneously. For example, arbitrary questions or humiliating expressions about skill level or knowledge can clearly mobilize ideas of unfair treatment and anger, but also anxiety and hopelessness by inducing loss of control. Some verbalizations can even anticipate success before the exam starts, which can increase hope regarding the possibility of obtaining a positive result and activate enjoyment of the situation.

Items development of the (TAVQ)

Item construction was based on the aforementioned theoretical aspects and a preliminary exploratory study (Sánchez-Rosas, 2016). This exploratory study analyzed the occurrence of various teachers' achievement verbalizations in oral exams associated with some achievement emotions.

Seven sets of items were developed, each one initially thought to be closely related to one of seven possible emotions (enjoyment, anger, pride, hope, anxiety, shame and hopelessness). In this way, one set of items would evaluate verbalizations that could activate enjoyment (e.g., *The teacher affirms that the exam is one more instance of learning*), while another item would evaluate hopelessness, (e.g., *The teacher says that even if I try hard, I will not be able to improve my exam performance*). The selection of these items associated with these seven emotions would cover verbalizations that were thought to be associated with frequent emotions in exams and with positive-negative emotions of activity (enjoyment and anger) and of past or future outcomes (hope, anxiety and hopelessness, pride, and shame). It is important to note that these items would not evaluate emotions but verbalizations that would be believed to be discernible from one another and that their grouping could be due to their possible association with a specific emotion. However, it must be recognized that the same verbalization could simultaneously lead to experiencing emotions that are empirically difficult to separate due to their valence, object focus, or time frame (e.g., anxiety, hopelessness; Pekrun et al., 2023; Sánchez-Rosas, 2016). The item design contemplated the specific achievement context (oral exams) and the temporal nature (before/beginning, during, after/end) predominant in the achievement context that triggers each emotion (Pekrun, 2018, 2021). This is because emotions can be activated before, during and after the achievement activity (questions, problem situations) or before and after the achievement outcomes (anticipations, feedback). However, some emotions are more prospective (hopelessness), others more retrospective (pride), and others more concurrent (anger). Therefore, when writing the items, we sought to give relevance to the temporal aspect of the achievement. Thus, for example, an item that was thought to be associated with anxiety included verbalizations by the teacher mentioning the diffi-

culty of the situation at the beginning of the exam or conveying uncertainty about the appropriateness of the responses during the exam. In contrast, to evaluate a verbalization associated with shame, an item was written in which the teacher explained to other people that the student failed to respond.

Method

Participants

Two hundred fifty-two students in different academic programs (72% were studying psychology) from the National University of Córdoba, Argentina, participated. The participants were between the ages of 18 and 55 (M = 27.70, SD = 8.81), in their first to fifth year of studies (5th year = 40%), and 84% of the participants were women.

Instruments

Teacher's achievement verbalizations in oral exams. For the assessment of teacher's achievement verbalizations in oral exams, 56 preliminary items were used, which were distributed according to the measurement model tested. For example, if oblique model of seven dimensions B was tested, items were distributed in groups of eight items depending on whether they had been designed to evaluate their association with enjoyment (*The teacher greets at the beginning of the exam*), anger (*The teacher makes questions about topics that are not on the syllabus*), anxiety (*Before starting, the teacher warns that the exam will be difficult*), shame (*At the end of the exam, the teacher says: I thought you had prepared better*), hope (*Before starting, the teacher assures me that I have the enough ability to pass the exam*), hopelessness (*Before starting, the teacher assures me that the exam will be very difficult*) and pride (*When the exam is over, the teacher congratulates me on my performance*). The final version of the instrument includes eight items of positive verbalizations (*When the exam is over, the teacher compliments the way I prepared myself for the exam*) and eight items of negative verbalizations (*While I take the exam, the teacher reproaches: You should have already known that!*). Each of the items is answered on a 5-point Likert scale (1 = *almost never*, 5 = *almost always*) to describe the frequency with which teachers make various comments or verbalizations related to achievement in oral exams. Although the response scale used evaluated the typical experience in oral exams, it is possible to evaluate the experience in a particular exam with slight modifications in the response instruction.

Achievement emotions. Seven items were used to assess the emotions of enjoyment (*I enjoy taking the exam*), anger (*I get angry*), anxiety (*I am very nervous*), shame (*I feel ashamed*), hope (*I am very confident*), hopelessness (*I feel hopeless*) and pride (*I am very pleased with myself*). Each of the items is answered on a 5-point Likert scale (1 = *almost never*, 5 = *almost always*) to describe the frequency with which students experience these emotions in oral exams (Sánchez-Rosas, 2015).

Academic performance. Students' performance was measured by assessing their average grades attained over the academic career.

Procedures

The set of items was included in an online survey to which questions about gender, age, career, and current academic year were added. The survey included an invitation to participate, the objectives of the study and an informed consent form. It was conducted through a platform that takes online surveys and the invitation was made through social networks.

Data Analysis

Measurement models of teachers' achievement verbalizations in oral exams. Considering the valence, object focus, and temporality of emotions, seven alternative measurement models of teacher achievement verbalizations in oral exams were specified: (a) uni-dimensional model A, in which all items load on a single factor; (b) oblique model of seven dimensions B, in which the seven sets of designed items load their respective latent factors with specific emotions; (c) bi-dimensional model C, in which the items of positive verbalizations and negative verbalizations load their respective latent factor; (d) bi-dimensional model D, in which the items of activity (enjoyment + anger) and outcomes (anxiety + shame + hopelessness + pride) load their respective latent factor; (e) tri-dimensional model E, in which the items of activity (enjoyment + anger), of positive (hopelessness + pride) and negative (anxiety + shame) outcomes load their latent factor respectively; (f) tri-dimensional model F, in which the items of activity (enjoyment + anger), prospective (anxiety + hopelessness + hope) and retrospective (shame + pride) outcomes load their latent factor respectively; (g) bifactor model G, in which all the items load on a general factor and the items with positive and negative verbalizations load on their respective latent factor.

Preliminary analysis. The approximation to univariate normality was analyzed through the magnitude of the skewness and kurtosis of the items, and they were considered acceptable if they were less than 2 and 7, respectively (Finney & DiStefano, 2006). In turn, for multivariate normality, the Mardia multivariate kurtosis coefficient was used (G2 < 70; Pérez et al., 2013).

Assessment of measurement models. The different measurement models were assessed based on the magnitude of the fit indices such as the CFI (Comparative Fit Index > .90; McDonald & Ho, 2002), the RMSEA (Root Mean Square Error of Approximation < .08; Browne & Cudeck, 1993), and the WRMR (Weighted Root Mean Square Residual Index < 1.00; DiStefano et al., 2018). In turn, factorial loads (> .50; Dominguez-Lara, 2018) and interfactorial correlations were also considered in order to detect potential cases of factorial redundancy (ϕ > .80; Brown, 2015).

If the bifactor model obtains a favorable fit, complementary indicators will be assessed to evaluate the representativeness of the general factor (ω_h, ω_{hs}, and ECV; Rodriguez et al., 2016), while values less than .30 for ω_{hs} reaffirm it.

Short version of the TAVQ. Based on the criteria mentioned above, the best measurement model was chosen, with which a brief version was designed, selecting the items whose factor loadings are greater than .60 by progressively eliminating those that were below that requirement. After designing it, the equivalent with the extended version was analyzed using the corrected Pearson correlation coefficient, since both versions share items, and a corrected correlation above .70 was expected to conclude the equivalence of the versions.

Reliability

Finally, the reliability of the brief version was evaluated both at the score level (α > .70; Ponterotto & Charter, 2009) and the construct level (ω > .70; Hunsley & Marsh, 2008).

Estimation and software

A series of confirmatory factor analyses were implemented in order to assess the measurement models proposed in the introductory section. The weighted least squares means and variance adjusted estimation method (WLSMV) and

Table 1
Descriptive statistics of the items.

	M	SD	g_1	g_2		M	SD	g_1	g_2
Item 1	1.93	1.18	0.97	-0.20	Item 29	1.84	1.04	0.89	-0.38
Item 2	2.07	1.23	0.80	-0.59	Item 30	1.28	0.78	3.18	10.21
Item 3	2.31	1.27	0.57	-0.82	Item 31	2.05	1.22	0.80	-0.51
Item 4	3.15	1.27	-0.05	-1.03	Item 32	1.32	0.75	2.53	6.24
Item 5	1.78	1.00	1.18	0.66	Item 33	4.25	1.00	-1.19	0.53
Item 6	2.65	1.20	0.24	-0.75	Item 34	2.65	1.08	0.06	-0.56
Item 7	1.81	1.06	1.11	0.21	Item 35	3.04	1.29	-0.17	-1.06
Item 8	1.53	0.94	1.79	2.56	Item 36	2.86	1.21	0.08	-0.88
Item 9	1.96	1.16	0.92	-0.31	Item 37	2.67	1.08	-0.01	-0.64
Item 10	2.54	1.26	0.23	-0.95	Item 38	2.67	1.05	0.17	-0.45
Item 11	2.05	1.36	0.86	-0.73	Item 39	2.32	1.17	0.29	-0.97
Item 12	2.53	1.24	0.29	-0.88	Item 40	2.52	1.09	0.03	-0.91
Item 13	2.36	1.30	0.50	-0.93	Item 41	2.33	1.20	0.48	-0.68
Item 14	1.86	1.15	1.25	0.62	Item 42	2.00	1.04	0.81	-0.03
Item 15	2.28	1.09	0.43	-.62	Item 43	1.56	0.91	1.61	2.01
Item 16	1.86	1.05	1.07	0.37	Item 44	2.23	1.23	0.47	-1.10
Item 17	2.13	1.12	0.53	-0.81	Item 45	2.31	1.10	0.32	-0.74
Item 18	2.31	1.18	0.37	-0.97	Item 46	1.59	0.89	1.33	0.76
Item 19	1.69	0.99	1.28	0.79	Item 47	1.94	1.06	0.83	-0.27
Item 20	1.44	0.84	1.99	3.38	Item 48	1.90	1.03	0.89	-0.08
Item 21	1.81	1.08	1.22	0.60	Item 49	1.82	1.12	1.23	0.58
Item 22	1.91	1.17	1.01	-0.16	Item 50	1.45	0.87	2.37	5.80
Item 23	2.44	1.25	0.41	-0.88	Item 51	1.39	0.84	2.40	5.81
Item 24	2.23	1.20	0.60	-0.67	Item 52	1.71	1.01	1.33	0.98
Item 25	1.69	1.02	1.33	0.82	Item 53	2.20	1.10	0.46	-0.72
Item 26	1.84	1.13	1.17	0.41	Item 54	1.80	1.01	1.08	0.46
Item 27	1.87	1.14	1.08	-0.02	Item 55	1.72	1.04	1.28	0.62
Item 28	1.91	1.16	0.93	-0.31	Item 56	2.42	1.08	0.28	-0.59

Note. M= Mean; SD= Standard Deviation; g_1= Asymmetry; g_2= Kurtosis; items 1 to 8= anxiety; items 9 to 16= shame; items 17 to 24= anger; items 25 to 32= hopelessness; items 33 to 40= enjoyment; items 41 to 48= pride; items 49 to 56= hope. The numbering corresponds to the database ordered for analytical purposes. The items were presented to the participants in order.

the polychoric correlation matrix between items were used. For this purpose, the software Mplus v. 7 (Muthén & Muthén, 2012).

Results

Evidence of validity related to internal structure: Preliminary analysis

Most of the items present magnitudes of skewness and kurtosis that are close to univariate normality (e.g., item 1), but in other cases they far exceed the established limits (e.g., item 30) (Table 1). Likewise, G2 is above what is suggested to conclude on multivariate normality (G2 = 225.530).

Evaluation of measurement models

Regarding the analysis of the internal structure, four models (A, D, E and F) presented fit indices with unacceptable magnitudes. Likewise, the

bifactor model (model G) did not achieve a better fit than the other models, so the analysis of the complementary indicators was not continued.

Those that had a more acceptable adjustment were the model of seven oblique factors (model B) and two factors (model C), although in the case of the first, the interfactorial correlation was high among those factors that corresponded to the same type of verbalization, either negative (ϕ_{mean} = .949) or positive (ϕ_{mean} = .830). In this sense, the two-factor model was considered the most parsimonious one and the one which best represents the evaluated construct.

Short version of TAVQ

The short version was prepared based on the two-factor model (negative verbalizations and positive verbalizations) by gradually discarding the items whose factorial loads were less than .60 in their respective factors.

Table 2
TAVQ measurement models.

	CFI	RMSEA	IC_{RMSEA} 90%	WRMR
Model A	.661	.103	.100, .106	2.468
Model B	.835	.072	.069, .076	1.785
Model C	.831	.073	.070, .076	1.846
Model D	.696	.108	.105, .111	2.336
Model E	.753	.097	.094, .101	2.099
Model F	.665	.103	.100, .106	2.455
Model G	.749	.092	.089, .095	2.062

Note. Model A= uni-dimensional model; Model B= oblique model of seven dimensions; Model C= bi-dimensional model of positive and negative factors; Model D= bi-dimensional model of activity and outcomes factors; Model E= tri-dimensional model of activity and positive and negative outcomes factors; Model F= tri-dimensional model of activity and prospective and retrospective factors; Model G= bifactor model.

Table 3
TAVQ Factorial Loadings and Correlations.

Item		Factorial loadings
	Negative verbalizations	
13	While I am taking the exam, the teacher reproaches: You should have already known that!	.78
	Mientras rindo, reprocha ¡Eso ya lo deberías saber!	
19	The teacher makes negative comments about my skill.	.78
	Hace comentarios negativos sobre mi capacidad.	
20	The teacher makes fun of what I say.	.77
	Se burla de lo que digo.	
27	While I am taking the exam, the teacher says: So far, you could not answer anything right.	.86
	Mientras rindo dice: hasta el momento no pudiste responder nada bien.	
28	While I am taking the exam, the teacher assures me that my level of knowledge is not enough to pass.	.84
	Mientras rindo asegura que mi nivel de conocimiento es insuficiente para aprobar.	
29	The teacher assures me that, with the knowledge I have, I will not be able to pass the exam.	.84
	Asegura que con lo que sé no podré aprobar el examen.	
31	The teacher maintains that it makes no sense to keep asking me.	.78
	Sostiene que no tiene sentido seguir preguntándome.	
32	The teacher says that even if I make an effort, I won't be able to improve my performance.	.73
	Dice que aunque me esfuerce, no podré mejorar mi desempeño en el examen.	
	Positive verbalizations	
41	When the exam is over, the teacher congratulates me on my performance.	.82
	Al terminar el examen, me felicita por mi desempeño.	
42	When the exam is over, the teacher compliments the way I prepared myself for the exam.	.79
	Al terminar el examen, elogia la forma en que me preparé.	
44	When the exam is over, the teacher encourages me to keep on the same track.	.83
	Al terminar el examen, me alienta a seguir así.	
47	The teacher says that my answer is excellent.	.71
	Dice que mi respuesta es excelente.	
51	Before we start the exam, the teacher asserts that he/she has confidence in me.	.60
	Antes de comenzar, dice que confía en mí.	
52	Before we start the exam, the teacher says I will do great just like everyone else.	.64
	Antes de comenzar, asegura que me va a ir bien como a todos.	
54	The teacher asserts that, by what I seem to know, for sure I will do well.	.64
	Afirma que, por el nivel de conocimiento que demuestro, seguro me va a ir bien.	
56	The teacher thinks that I have a good understanding of the topics.	.70
	Considera que tengo un buen dominio de los temas.	
	Factor Correlations	-.36

Regarding the equivalence between the long and short version of the negative verbalizations factor, the initial correlation was .91, and a post correction of .88 was obtained; and in the case of positive verbalizations, initially it reached .92 initially, and when implementing the correction, it decreased to .86. In both cases, it is concluded that the long and short versions are equivalent.

Reliability

Finally, in relation to reliability, adequate values were obtained both at the construct level and for scores of the negative verbalizations dimension ($\omega = .936$; $\alpha = .895$) and positive verbalizations ($\omega = .898$; $\alpha = .856$).

Evidence of validity by association with other variables

The association of verbalizations with achievement emotions and academic performance was significant in almost all cases (Table 4), with moderate and low results.

Discussion

This study presented the development of an instrument to evaluate teachers' achievement verbalizations in oral exams and its psychometric properties assessed through internal structure, reliability and criterion related variables. Altogether, the results provide psychometric evidence for an instrument that allows to evaluate teachers' achievement of positive and negative verbalizations in oral exams. These verbalizations include, among others, messages that anticipate difficulty or uncertainty in the exam, indicate poor performance, imply an arbitrary attitude or promote control of the exam. In addition, these verbalizations are related to the activation of various emotions and academic performance.

The progress presented here constitutes the focus on an important area of research on academic emotions, a topic that is not much studied at the local level (specifically, in the research on teachers' verbalizations that affect the activation of student's emotions during oral exams) (Narciss et al., 2022; Putwain et al., 2017, 2022, 2023).

Table 4
Association of teachers' verbalizations with achievement emotions and academic performance.

	1	2	3	4	5	6	7	8	9	10
1. NV	-									
2. PV	-.27***	-								
3. enjoyment	-.28***	.43***	-							
4. pride	-.30***	.39***	.56***	-						
5. anger	.27***	-.08	-.08	-.18**	-					
6. anxiety	.07	-.12*	-.41***	-.24***	.15*	-				
7. shame	.20**	-.24***	-.41***	-.42***	.19**	.43***	-			
8. hope	-.15*	.26***	.46***	.58***	-.09	-.25***	-.42***	-		
9. hopelessness	.33***	-.27***	-.39***	-.55***	.25***	.23***	.51***	-.41***	-	
10. GPA	-.29***	.29***	.16**	.35***	-.14*	-.04	-.25***	.16**	-.32***	-

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. GPA= Grade Point Average; NV= negative verbalizations; PV= positive verbalizations.

Item development and evidence of validity and reliability

The item design intended to incorporate the dimensions of the valence, object focus and temporal frame (Pekrun, 2018, 2021) of the emotions to which teachers' achievement verbalizations would be associated. In turn, the items were designed considering aspects that would be core to each achievement emotion, for example, uncertainty about the domain of the exam (Putwain et al., 2017), exposure of the error (Apto et al., 2017) or certainty about a positive result (Pekrun et al., 2023).

Considering the dimensions and these central aspects, plausible models for measuring teachers' achievement verbalizations in oral exams were specified. Seven measurement models were evaluated and the most favorable adjustments were obtained for a model with seven oblique but highly related dimensions and a more parsimonious model with two dimensions that grouped positive and negative verbalizations. As anticipated, the same verbalization could simultaneously activate various emotions that are empirically difficult to separate due to their valence, emotional focus, or time frame (e.g., anxiety and hopelessness; Pekrun et al., 2023; Sánchez-Rosas, 2016). Some of the evaluated models considered the time frame of the verbalizations; however, they did not show a good fit to the data. This difficulty in discriminating the time frame may be due to the fact that the items refer to verbalizations that are performed at the beginning and at the end of the exam, instead of being performed before or after it. In this way, all the items would have been constructed as concurrent to the examination situation and not as prospective or retrospective. Added to this, the items failed to discriminate between verbalizations related to the activity in progress or their outcomes. From the data, it can be deduced that the items evaluate teachers' verbalizations whose content provides information regarding the distinguishable achievement only by their valence (Pekrun, 2018).

In contrast, a large number of items were designed with the intention of arriving at a brief instrument with the best items that facilitate and make their application more practical. After identifying the items with the best factor loadings, two scales of eight items each with very good internal consistency were retained.

Teachers' achievement verbalizations can affect control and value appraisals and, through these, activate achievement emotions and academic performance (Goetz et al., 2018). This research provided favorable evidence on the relationship of the scores of each scale with theoretically related variables (Apto et al., 2017; Pekrun et al., 2023; Putwain et al., 2017). Low correlations with achievement emotions were obtained, although the association seems to be greater between verbalizations and emotions of equal valence. There were also positive and negative relationships between the positive and negative verbalizations, respectively, with the academic average. The magnitudes of these relationships, although low, seem consistent with the idea that control-value appraisals would be mediating this association (Pekrun, 2018, 2021).

Limitations and further studies

Although the instrument developed presents some good initial psychometric properties, the results should be taken with caution and further studies should be carried out. First, the items express verbalizations made during oral exams and fail to capture the verbalizations that are anticipated several days before, for example, in a class situation or that are even carried out a few

days later. This distinction, not achieved with our development of items, could have differential effects not only on achievement emotions, but also on, for example, behavioral avoidance in oral exams (Furlan & Sánchez-Rosas, 2018). Verbalizations before or after an exam would affect the postponement of the evaluation instances, while those that are said during the exam would have a greater impact on inhibition during the oral exam.

Conversely, the clear gender bias in the sample has not made it possible to verify whether the measurements remain invariant based on sex. This would be important to analyze since, if achievement emotions related to negative outcomes seem to be more frequent in women (Reilly & Sánchez-Rosas, 2019; Sánchez-Rosas, 2015), verbalizations, considering their environmental antecedents, could also vary depending on this categorical variable.

Finally, although the constructed items clearly refer to verbal expressions on achievement, it would be convenient to consider an analysis of convergence or divergence through correlation with instruments that assess feedback or fear appeals (Putwain et al., 2017, 2023), instructional teaching quality in class (Becker et al., 2014; Lazarides & Buchholz, 2019; Narciss et al., 2022; Sánchez-Rosas et al., 2016), teacher support (Apto et al., 2017; Lei et al., 2018) or the inclusion of non-verbal behaviors (Derakhshan et al., 2023; Guo et al., 2022; Juma et al., 2022; Puertas-Molero et al., 2022).

Practical implications

Through the TAVQ, the measurement of teachers' verbalizations during oral exams makes it possible to broaden and enrich the functional analysis of emotional dysregulation problems and avoidance behaviors in evaluative contexts (Furlan & Sánchez-Rosas, 2018). The messages related to achievement during an oral exam offer feedback on the control and value appraisals that precede the emotional responses of the people evaluated and that later activate their coping behaviors (Pekrun, 2018, 2021). The information processing related to performance during the exam is part of executive control tasks when goal-directed behaviors are implemented (Zeidner & Matthews, 2005). In this way, the information provided by teachers will be processed by each student according to their beliefs and appraisals and will lead to behaviors that tend to regulate their emotional state, using the set of strategies that can be accessed (Rojas-Torres et al., 2022). For this reason, it is valuable to have an evaluation tool that allows one to reflect on the nature and effects of verbal messages, noting their relevance in students' performances in oral exams.

Conclusions

An instrument is provided with two scales that evaluate positive and negative teachers' achievement verbalizations in oral exams with evidence of validity and reliability. In addition to the optimal reliability values, the relationship of the scores of each scale with the achievement emotions and academic performance is demonstrated. In short, we count with a useful instrument for the assessment of verbal expressions, phrases and comments about students' achievement that a teacher emits in the presence of one or more students during an oral exam.

References

Apto, M. S., Pesqueira, N. G., Vilca, B., & Sánchez-Rosas, J. (2017). Efecto y contribución explicativa del apo-

yo e inhibición, interacción-ilustración, vergüenza, disfrute y amenazas a la evitación de la búsqueda de ayuda académica. *Anuario de Investigaciones de la Facultad de Psicología, 3*(1), 247-263. https://revistas.unc.edu.ar/index.php/aifp/article/view/18670

Awad-Igbaria, Y., Maaravi-Hesseg, R., Admon, R., & Karni, A. (2022). Only tomorrow: Delayed effects of teachers attitude on motor skill learning. *Learning and Instruction, 82*. https://doi.org/10.1016/j.learninstruc.2022.101681

Becker, E. S., Goetz, T., Morger, V., & Ranellucci, J. (2014). The importance of teachers' emotions and instructional behavior for their students' emotions – An experience sampling analysis. *Teaching and Teacher Education: An International Journal of Research and Studies, 43*(1), 15-26. https://doi.org/10.1016/j.tate.2014.05.002

Brown, T. (2015). *Confirmatory factor analysis for applied research (2ⁿᵈ ed.*). The Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Sage.

Burić, I. (2015). The role of social factors in shaping students' test emotions: A mediation analysis of cognitive appraisals. *Social Psychology of Education, 18*(4), 785-809. https://doi.org/10.1007/s11218-015-9307-9

Derakhshan, A., Zhang , L. J., & Zhaleh, K. (2023). The effects of instructor clarity and non-verbal immediacy on Chinese and Iranian EFL students' affective learning: The mediating role of instructor understanding. *Studies in Second Language Learning and Teaching, 13*(1), 71-100. https://doi.org/10.14746/ssllt.31733

Dewaele, J.-M., Magdalena, A. F., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal, 103*(2), 412-427. https://doi.org/10.1111/modl.12555

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research, 22*(6), 676-697. https://doi.org/10.1177/1362168817692161

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 453-466. https://doi.org/10.1080/10705511.2017.1390394

Dominguez-Lara, S. (2018). Propuesta de puntos de corte para cargas factoriales: Una perspectiva de fiabilidad de constructo. *Enfermería Clínica, 28*(6), 401-402. https://doi.org/10.1016/j.enfcli.2018.06.002

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 269-314). Information Age.

Frenzel, A. C., Daniels, L., & Burić, I. (2021). Teacher emotions in the classroom and their implications for students. *Educational Psychologist, 56*(4), 250-264. https://doi.org/10.1080/00461520.2021.1985501

Furlan, L., & Sánchez-Rosas, J. (2018). Evidencias de validez y confiabilidad de una Escala de Evitación Conductual en Exámenes Orales en estudiantes universitarios. *Ansiedad y Estrés, 24*(2-3), 90-98. https://doi.org/10.1016/j.anyes.2018.05.001

Gardner, D. E., & Giordano, A. N. (2023). The challenges and value of undergraduate oral exams in the physical chemistry classroom: A useful tool in the assessment toolbox. *Journal of Chemical Education, 100*(5), 1705-1709. https://doi.org/10.1021/acs.jchemed.3c00011

Goetz, T., Lipnevich, A. A., Krannich, M., & Gogol, K. (2018). Performance feedback and emotions. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge Handbook of Instructional Feedback* (pp. 554-574). Cambridge University Press. https://doi.org/10.1017/9781316832134.027

Guo, H., Gao, W., & Shen, Y. (2022). L2 Enjoyment of English as a foreign language students: Does teach-

er verbal and non-verbal immediacy matter? *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.897698

Hazen, H. (2020). Use of oral examinations to assess student learning in the social sciences. *Journal of Geography in Higher Education, 44*(4), 592-607. https://doi.org/10.1080/03098265.2020.1773418

Hunsley, J., & Marsh, E. J. (2008). Developing criteria for evidence-based assessment: An introduction to assessments that work. In J. Hunsley & E. J. Marsh (Eds.), *A guide to assessments that work* (pp. 3-14). Oxford University Press.

Juma, O., Husiyin, M., Akhat, A., & Habibulla, I. (2022). Students' classroom silence and hopelessness: The impact of teachers' immediacy on mainstream education. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.819821

Lazarides, R., & Buchholz, J. (2019). Student-perceived teaching quality: How is it related to different achievement emotions in mathematics classrooms? *Learning and Instruction, 61*, 45-59. https://doi.org/10.1016/j.learninstruc.2019.01.001

Lei, H., Cui, Y., & Chiu, M. M. (2018). The relationship between teacher support and students' academic emotions: A meta-analysis. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02288

McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64-82. https://doi.org/10.1037/1082-989X.7.1.64

Muthén, L. K., & Muthén, B. O. (2012). Mplus User's Guide. Muthén & Muthén Press.

Narciss, S., Prescher, C., Khalifah, L., & Körndle, H. (2022). Providing external feedback and prompting the generation of internal feedback fosters achievement, strategies and motivation in concept learning. *Learning and Instruction, 82*. https://doi.org/10.1016/j.learninstruc.2022.101658

Pekrun, R. (2018). Control-value theory: A social-cognitive approach to achievement emotions. In G. A. D. Liem & D. M. McInerney (Eds.), *Big theories revisited 2:*

A volume of research on sociocultural influences on motivation and learning (pp. 162-190). Information Age Publishing.

Pekrun, R. (2021). Self-appraisals and emotions: A generalized control-value approach. In T. Dicke, F. Guay, H. W. Marsh, R. G. Craven, & D. M. McInerney (Eds.), *Self - A multidisciplinary concept* (pp. 1-30). Information Age.

Pekrun, R., Marsh, H. W., Elliot, A. J., Stockinger, K., Perry, R. P., Vogl, E., Goetz, T., van Tilburg, W. A. P., Lüdtke, O., & Vispoel, W. P. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology, 124*(1), 145-178. https://doi.org/10.1037/pspp0000448

Pérez, E., Medrano, L., & Sánchez-Rosas, J. (2013). El Path Analysis: Conceptos básicos y ejemplos de aplicación. *Revista de la Asociación Argentina de Ciencias del Comportamiento, 5*(1), 52-66. https://revistas.unc.edu.ar/index.php/racc

Ponterotto, J., & Charter, R. A. (2009). Statistical extensions of Ponterotto and Ruckdeschel's (2007) Reliability Matrix for estimating the adequacy of internal consistency coefficients. *Perceptual and Motor Skills, 108*(3), 878-886. https://doi.org/10.2466/PMS.108.3.878-886

Puertas-Molero, P., Zurita-Ortega, F., González-Valero, G., & Ortega-Martín, J. L. (2022). Design and validation of the Non-Verbal Immediacy Scale (NVIS) for the evaluation of non-verbal language in university professors. *International Journal of Environmental Research and Public Health, 19*(3), 1159. http://dx.doi.org/10.3390/ijerph19031159

Putwain, D. W., Nakhla, G., Liversidge, A., Nicholson, L. J., Porter, B., & Reece, M. (2017). Teachers use of fear appeals prior to a high-stakes examination: Is frequency linked to perceived student engagement and how do students respond? *Teaching and Teacher Education, 61*, 73-83. https://doi.org/10.1016/j.tate.2016.10.003

Putwain, D. W., Nicholson, L. J., & Kutuk, G. (2023). Warning students of the consequences of examination

failure: An effective strategy for promoting student engagement? *Journal of Educational Psychology, 115*(1), 36-54. https://doi.org/10.1037/edu0000741

Putwain, D. W., Pekrun, R., Rainbird, E., & Roberts, C. (2022). Cognitive-behavioural intervention for test anxiety: Could teachers deliver the STEPS Program and what training would they require? In L. R. V. Gonzaga, L. L. Dellazzana-Zanon, & A. M. Becker da Silva (Eds.), *Handbook of Stress and Academic Anxiety* (pp. 381-399). https://doi.org/10.1007/978-3-031-12737-3_25

Raccanello, D., Hall, R., & Burro, R. (2018). Salience of primary and secondary school students' achievement emotions and perceived antecedents: Interviews on literacy and mathematics domains. *Learning and Individual Differences, 65*, 65-79. https://doi.org/10.1016/j.lindif.2018.05.015

Reilly, P., & Sánchez-Rosas, J. (2019). The achievement emotions of English language learners in Mexico. *Electronic Journal of Foreign Language Teaching, 16*(1), 34-48. http://e-flt.nus.edu.sg/v16n12019

Reilly, P., & Sánchez-Rosas, J. (2021). Achievement emotions and gender differences associated with second language testing. *International Journal of Instruction, 14*(4), 825-840. https://doi.org/10.29333/iji.2021.14447a

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150. https://doi.org/10.1037/met0000045

Rojas-Torres, L., Furlan, L. A., Sánchez-Rosas, J., & Rojas-Rojas, G. (2022). Construcción y validación de la Escala de Afrontamiento de la Ansiedad durante los Exámenes. *Ansiedad y Estrés, 28*(2), 81-90. https://doi.org/10.5093/anyes2022a9

Sánchez-Rosas, J. (2015). The Achievement Emotions Questionnaire-Argentine (AEQ-AR): Internal and external validity, reliability, gender differences and norm-referenced interpretation of test scores. *Evaluar, 15*(1), 41-74. https://doi.org/10.35670/1667-4545.v15.n1.14908

Sánchez-Rosas, J. (2016). Avances preliminares en la identificación de causas docentes de la ansiedad, vergüenza y disfrute en exámenes orales [Comunicaciones libres]. *Revista Argentina de Ciencias del Comportamiento, Suplemento I (Agosto)*, 108-109. https://revistas.unc.edu.ar/index.php/racc

Sánchez-Rosas, J., Correa, P., & Díaz, I. (2019). Revisión de las intervenciones que mejoran la utilidad percibida del aprendizaje de los estudiantes. *Revista Digital de Investigación en Docencia Universitaria, 13*(2), 41-52. https://revistas.upc.edu.pe/index.php/docencia

Sánchez-Rosas, J., & Esquivel, S. (2016). Instructional teaching quality, task value, self-efficacy, and boredom: A model of attention in class. *Revista de Psicología, 25*(2), 1-20. https://doi.org/10.5354/0719-0581.2016.44966

Sánchez-Rosas, J., & Furlan, L. A. (2017). Achievement emotions and achievement goals in support of the convergent, divergent and criterion validity of the Spanish-Cognitive Test Anxiety Scale. *International Journal of Educational Psychology, 6*(1), 67-92. https://doi.org/10.17583/ijep.2017.2268

Sánchez-Rosas, J., Takaya, P. B., & Molinari, A. V. (2016). The role of teacher behavior, motivation and emotion in predicting academic social participation in class. *Pensando Psicología, 12*(19), 39-53. https://doi.org/10.16925/pe.v12i19.1327

Theobold, A. S. (2021). Oral exams: A more meaningful assessment of students' understanding. *Journal of Statistics and Data Science Education, 29*(2), 156-159. https://doi.org/10.1080/26939169.2021.1914527

Ventura-León, J., Caycho-Rodríguez, T., Sánchez-Villena, A. R., Peña-Calero, B. N., & Sánchez-Rosas, J. (2022). Academic inspiration: Development and validation of an instrument in higher education. *Electronic Journal of Research in Education Psychology, 20*(58), 635-660. https://doi.org/10.25115/ejrep.v20i58.5599

Westphal, A., Kretschmann, J., Gronostaj, A., & Vock, M. (2018). More enjoyment, less anxiety and bore-

dom: How achievement emotions relate to academic self-concept and teachers' diagnostic skills. *Learning and Individual Differences, 62*, 108-117. https://doi.org/10.1016/j.lindif.2018.01.016

Yang, Y., Gao, Z., & Han, Y. (2021). Exploring Chinese EFL learners' achievement emotions and their antecedents in an online English learning environment. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.722622

Zeidner, M., & Matthews, G. (2005). Evaluation Anxiety: Current Theory and Research. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 141-163). Guilford Publications.