Research Article https://doi.org/10.12973/eu-jer.11.4.2101

European Journal of Educational Research
Volume 11, Issue 4, 2101 - 2114.
ISSN: 2165-8714
http://www.eu-jer.com/

Teacher-Student Performance Criteria During Online Classes due to COVID-19: Self-Report by Postgraduate Students in Education

Aldo Bazán-Ramírez*
Cesar Vallejo University, PERU

Rolando Quispe-Morales
San Cristóbal de Huamanga National University, PERU

Carlos De La Cruz-Valdiviano
Federico Villarreal National University, PERU

Carmela Henostroza-Mota
Federico Villarreal National University, PERU

Received: March 23, 2022 ▪ Revised: July 6, 2022 ▪ Accepted: August 27, 2022

Abstract: During didactic interactions, teachers and students employ competencies and skills that correspond functionally to one another, and several models propose to typify didactic performances in higher education. For this study, we chose the didactic performance interbehavioral model as the substantive theory to identify six pairs of didactic performance criteria through self-evaluations from graduate students. These included: Competency exploration–precurrent learning behaviors, Criteria explicitness–Criteria identification, Illustration–Participation, Practice supervision–Relevant practice, Feedback–Improvement, and Evaluation–Application.309 Peruvian students from an in-person postgraduate course in Education Sciences filled out two scales (via Google Forms) to assess teacher-student performance criteria during online classes. Convergent and divergent validity were obtained separately for six teacher/student performance constructs through a confirmatory analysis using LISREL 8. Moreover, differences according to sex were only significant for the Illustration teacher criterion (favoring men) and the Application student criterion (favoring women). Students rated Explicitness of criteria and Illustration as the teacher performance criteria most frequently seen during online classes. On the other hand, Criteria identification ranked the highest in terms of the students' performance criteria, followed by Feedback–Improvement. Another conclusion was that the didactic performance interbehavioral model could be empirically supported by the two self-assessment questionnaires, which yielded good convergent and divergent validity of constructs.

Keywords: *Didactic, pandemic, performance, school lockdown, teaching.*

To cite this article: Bazán-Ramírez, A., Quispe-Morales, R., De La Cruz-Valdiviano, C., & Henostroza-Mota, C. (2022). Teacher-student performance criteria during online classes due to COVID-19: Self-report by postgraduate students in education. *European Journal of Educational Research*, *11*(4), 2101-2114. https://doi.org/10.12973/eu-jer.11.4.2101

Introduction

The spread of the COVID-19 pandemic called for the implementation of distance education models at all levels, when learning worldwide was primarily in-person. This new reality drawn by social distancing brought online education into a global educational landscape that continues to endure inequality in educational opportunities and access to basic technology and resources. COVID-19 forced numerous higher education institutions to switch from traditional face-to-face sessions to a digital approach (De Vincenzi, 2020; Gonzales-Zamora et al., 2020). Consequently, said institutions had to adapt pedagogical practices and evaluation from traditional in-person teaching to a virtual environment (De la Riva & Álvarez, 2020; García-Peñalvo, 2020), whereas teachers had to swiftly develop new digital competencies (Portillo et al., 2020).

The educational community, and especially the institutions, adopted different approaches to offer education services meant for on-site application, now in a virtual environment. As a result, new organizational strategies were developed to ensure the performance and well-being of teachers and administrative personnel working remotely (van Wyk et al., 2021) during the COVID-19 pandemic. The experimental suitability of several pedagogical strategies for online teaching, e.g., "flipped classroom" or "flipped learning" (Lee et al., 2022; Purwanti et al., 2022), has also been validated with positive results. Online classes provided teachers with new skills for academic work in virtual environments while favoring work satisfaction and improving their self-efficacy despite the greater workloads, physical health issues, new didactic behavior adaptation, and an emotional overload rarely experienced during traditional teaching (Szabó et al., 2022).

* **Corresponding author:**
Aldo Bazán-Ramírez, Cesar Vallejo University, Medicine School, Peru. ✉ abazanramirez@gmail.com

© 2022 The Author(s). **Open Access** - This article is under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

A widespread practice in assessing teacher performance has been using students' self-evaluation questionnaires (Boysen, 2016; Newton et al., 2019; Spooren et al., 2017). These are useful to identify trends or occurrence rates of certain behaviors exercised either by the teachers or students themselves during didactic interactions. Several self-reporting instruments have been developed for the assessment of teacher performance and didactic interactions by university students (Chan, 2018; Doménech-Betoret, 2018; Gitomer, 2018; Scherer & Gustafsson, 2015; Scherer et al., 2016; Spooren et al., 2017; Üstünlüoglu & Güngör-Culha, 2012). However, self-assessment questionnaires may well be supplemented with survey-sourced information and the use of qualitative or mixed methods (Ghislandi et al., 2020; Lee et al., 2022; Swart et al., 2019).

Assessing teacher performance criteria as forms of linguistic behavior during teaching-learning interactions in higher education is of utmost importance for improving the quality of teaching (Swart et al., 2019). This assessment can be conducted using quantitative and qualitative procedures. In doing so, the results obtained in these self-reports may be supplemented using qualitative approaches, such as those reported by Ghislandi et al. (2020) after analyzing the quality of teaching in an Italian university.

Although questionnaires to assess teacher performance criteria are deemed solid tools in educational evaluation and research, little has been studied about the students' self-reports on their performance during didactic interactions. Neither is there a consensus on what categories should be applied to analyze teacher performance during class or practices; such categories have indiscriminately included teacher behaviors during teaching, as well as strategies related to learning opportunities (Ramírez-Asís et al., 2020; van der Lans et al., 2018). For instance, van der Lans et al. (2018) included the following categories to measure teaching strategies: Safe learning environment, Classroom management, Clear instruction, Activation of teaching methods, Teaching learning-to-learn strategies, and Differentiation (student-centered teaching). Only two strategies from these six categories are deemed proper categories of teacher performance during classes or practices; Clear instruction and Teaching learning-to-learn strategies. The latter is referred to by the PISA test of the Organization for Economic Cooperation and Development (OECD), as Student-centered teaching (OECD, 2017).

Similarly, instruments and categories with limited content validity have also been used to assess teacher performance. For example, Ramírez-Asís et al. (2020) treat skill competencies as a category of teacher performance during teaching or practices: Instructional skill, Communication (referring to didactic interaction), Classroom Management, Instructional planning (a performance exercised by the teacher before the actual class or practice) and Networking (technologies). Furthermore, in the Teacher performance category, the referred-to authors mistake Providing feedback to the student (a teacher competency during class and practices) for other teacher practices in higher education, such as research development or administrative activities (management).

The Didactic Performance Interbehavioral Model

Throughout this study, the term "didactic performance" shall refer to the functional condition in which teaching and learning occur, involving various interactive areas or criteria that outline the teacher/student performance (Silva et al., 2014; Velarde-Corrales & Bazán-Ramírez, 2019). A seven-criteria didactic performance model based on the Functional Taxonomy of Behavior proposed by Ribes and López (1985) has been devised and agreed upon with slight variations: Didactic planning, Competency exploration, Criteria prescription, Illustration, Practice supervision, Feedback, and Evaluation (Carpio et al., 1998; Irigoyen et al., 2011; Reyna & Hernández, 2017; Silva, et al., 2014).

Applying these categories of didactic performance areas has led to several findings. Galindo et al. (2017) identified the effects of virtual observational training under controlled experimental conditions with psychology students. They found that the Illustration–Feedback pair was the most effective for learning and improving student performance in pre-and post-testing of the assessed skills and knowledge. Velarde-Corrales and Bazán-Ramírez (2019) validated a smaller model with five teacher performance criteria and five student performance criteria in didactic interactions during high school science classes using an observational record:

1. Competency exploration–Competencies needed for learning
2. Criteria explicitness–Identification of achievement and disciplinary criteria
3. Illustration–Adjustment to linguistic mode and illustration
4. Feedback–Making tasks and work available for assessment and feedback
5. Evaluation–Application

Based on the class video log of ten high school teachers of different subjects, Ávila (2020) validated a checklist to classify teacher actions into four teacher performance criteria or areas: Competency exploration, Criteria explicitness, Illustration, and Feedback. The author found a correlation between criteria seen in the recordings and those included in the instructional planning submitted in advance by each teacher. The author also acknowledged the importance of including the Evaluation criterion in the didactic performance analysis.

Based on the observational record categories reported by Velarde-Corrales and Bazán-Ramírez (2019) and supplementary to the didactic performance analysis, Bazán-Ramírez and Velarde-Corrales (2021) developed and validated a self-report instrument on said five pairs of teacher and student didactic performance criteria intended for psychology students from Mexico. Bazán-Ramírez et al. (2021) validated five criteria or areas of didactic performance of Peruvian psychology teachers as assessed by students and obtained invariance and construct validity according to sex and academic level.

Didactic Performance Criteria

This study takes into account six sets of teacher-student didactic performance criteria based on the theoretical model derived from the contributions of the interbehavioral psychologists (Irigoyen et al., 2011; Morales et al., 2017; Silva et al., 2014) and the studies for the empirical validation of categories and indicators derived from the teacher-student performance areas and criteria in the context of this didactic performance model (Ávila, 2020; Bazán-Ramírez et al., 2021, 2022; Bazán-Ramírez & Velarde-Corrales, 2021; Velarde-Corrales & Bazán-Ramírez, 2019). Table 1 summarizes this didactic performance model and outlines the six sets of didactic performance areas, maintaining a functional correlation between teacher and student behaviors. Said sets allow for the development of several instruments to identify such performances during didactic interactions.

Table 1. Teacher and Student Didactic Performance Criteria

Competency exploration: The teacher determines the competency level of the students at the beginning of the course or before discussing a class topic. Prior knowledge from students is assessed.	*Precurrent learning behaviors*: The student proves to possess the knowledge and potential skills needed to learn the topics that are being taught and learned.
Criteria explicitness: The teacher explains which parameters the student must fulfill according to the didactic criteria.	*Criteria identification*: The student mirrors/recreates the course/class criteria and inquires about which are and how to adjust to the thematic and disciplinary criteria.
Illustration: The teacher shows the student an efficient course of action to fulfill a given criterion based on the teacher's and the student's actions.	*Illustration - Participation*: The student behaves according to the criteria set forth by the teacher based on the didactic requirements and the appropriate linguistic mode.
Practice supervision: The teacher monitors the student's progress according to the achievement criteria and conditions related to what is and how to conduct a practice or activity. They also address the student's performance on a moment-by-moment basis.	*Pertinent practice*: The student's performance is outlined by the established achievement criterion, in correlation with what, how, when, and where the practice will be supervised.
Feedback: The teacher engages the student with their performance and shows them ways in which they can fulfill the criterion.	*Feedback – Improvement*: The student submits their completed tasks for evaluation and feedback from the teacher. They relate to their own performance and adjust their actions according to the criteria and observations made by the teacher or by themselves.
Assessment: The teacher compares the student's performance with the expected performance and its implementation when facing unfamiliar problems.	*Assessment – Application*: The student fulfills activities and exercises, solves problems, and implements procedures according to the achievement, effectiveness, and variability. They can apply their knowledge when facing unfamiliar problems and situations.

The following questions were prepared for this research study adapting the didactic model proposed by the interbehavioral psychologists and the six sets of didactic performance criteria, as shown in Table 1: a) What is the convergent and divergent construct validity level of two scales of teacher-student didactic performance during postgraduate classes online due to the COVID-19 pandemic? b) What didactic performance criteria of teachers and students were the most frequent according to Education Sciences postgraduate students? In addition, a third question was formulated: c) Are there any significant differences according to the students' sex and academic level among the twelve didactic performance criteria assessed?

Methodology

Participants

309 students from the education sciences postgraduate unit (198 men and 111 women; 222 enrolled in master's programs and 87 in the doctoral program) participated. they attended four master's programs, teaching learning and evaluation strategy, education management, university teaching, bilingual intercultural education, and a doctoral

program in education. both the master's and the doctoral courses are professional programs. Master's program students are mainly primary or secondary education teachers, whereas most doctoral students work as professors or administrative personnel at public universities. the age range for the master's program was between 23 and 40, and between 35 and 72 for the doctoral program.

The master's program has four academic terms or semesters, whereas the doctoral program has six terms or semesters. Each study term comprehends four sequential subjects, each lasting one month. Although designed as traditional in-person programs, all classes have been conducted online since June 2020 due to the COVID-19 pandemic.

Selection of participants: All students enrolled for the first and third terms of the four master's programs, and the first, third, and fifth terms of the doctoral program for the 2020-1 semester were emailed an invitation to participate in this study, where they were informed about its purpose and benefits for the institution and reminded that participation was voluntary. In case of acceptance, students would sign their informed consent electronically. They were also informed that there would be no penalties for anyone who did not wish to participate.

Description of Variables

Didactic Performance Criteria: "Didactic performance" refers to the skills and actions exercised by the professor and the student that enable didactic interactions at different moments or areas of the interactive processes during synchronous teaching-learning on-site or online. The following six teacher performance criteria were considered: Competency exploration, Criteria explicitness, Illustration, Practice supervision, Feedback, and Evaluation. Six student performance criteria were also included: Precurrent learning behaviors, Criteria identification, Participation, Relevant practice, Feedback–Improvement, and Evaluation–Application. Eight statements regarding certain situations throughout an academic term were included in each criterion or performance area: four for teacher performance and four for student performance.

Control variables: At the start of one of the questionnaires, participants were queried about two personal attributes to identify the variability in the assessment of teacher and student performances during didactic interactions. These variables were: sex and academic level.

Measurement Instruments and Materials

Two scales were designed to assess the graduate students' perception of the teacher and student performance criteria during online didactic interactions. These adapted several items (sentences) from the self-reports of Mexican high school students (in science subjects) and Psychology undergraduate students (Bazán-Ramírez & Velarde-Corrales, 2021; Velarde-Corrales, 2020) while also including newly devised items. Four sentences were assigned to each of the six performance criteria, totaling 24 for teacher performance and 24 for student performance. All items required the student to choose one of four options on a Likert scale: 0 = Never, 1 = Almost Never, 2 = Almost Always, and 3 = Always.

The first scale assesses the students' perception of their teacher's performance in six criteria: Competency exploration, Criteria explicitness, Illustration, Practice supervision, Feedback, and Evaluation (Appendix 1).

The second scale assesses the students' self-perception of their performance during the didactic interactions, also in six criteria: Precurrent learning behaviors, Criteria identification, Illustration–Participation, Pertinent practice, Feedback–Improvement, and Evaluation–Application (Appendix 2).

Procedure

The two self-report scales were uploaded via Google Forms upon project approval from a dedicated university research committee and after receiving consent from the students who agreed to participate in the study. The application was conducted from July 18 to August 9, 2020. Construct validity analyses were conducted for the two self-report scales through confirmatory factor analysis using the LISREL 8 software (Jöreskog & Sörbom, 2001). Descriptive and comparative analyses of the students' assessment of the online classes and the teacher and student didactic performance constructs were conducted using the SPSS 23 software.

Results

Validation of Didactic Performance Criteria

Figure 1 shows the result of the convergent and divergent construct validity of the teacher didactic performance during online classes as measured using self-reports. The resulting confirmatory factor analysis model (CFA) obtained good convergent and divergent validity for six teacher performance criteria constructs, and a good goodness-of-fit index as significant practical indicators were obtained: Chi-square = 281.39, gl =237; p=0.03; RMSEA = 0.03; NFI = 0.99; NNFI = 1.00; CFI = 1.00; RMR = 0.052; GFI = 0.99; AGFI = 0.99.

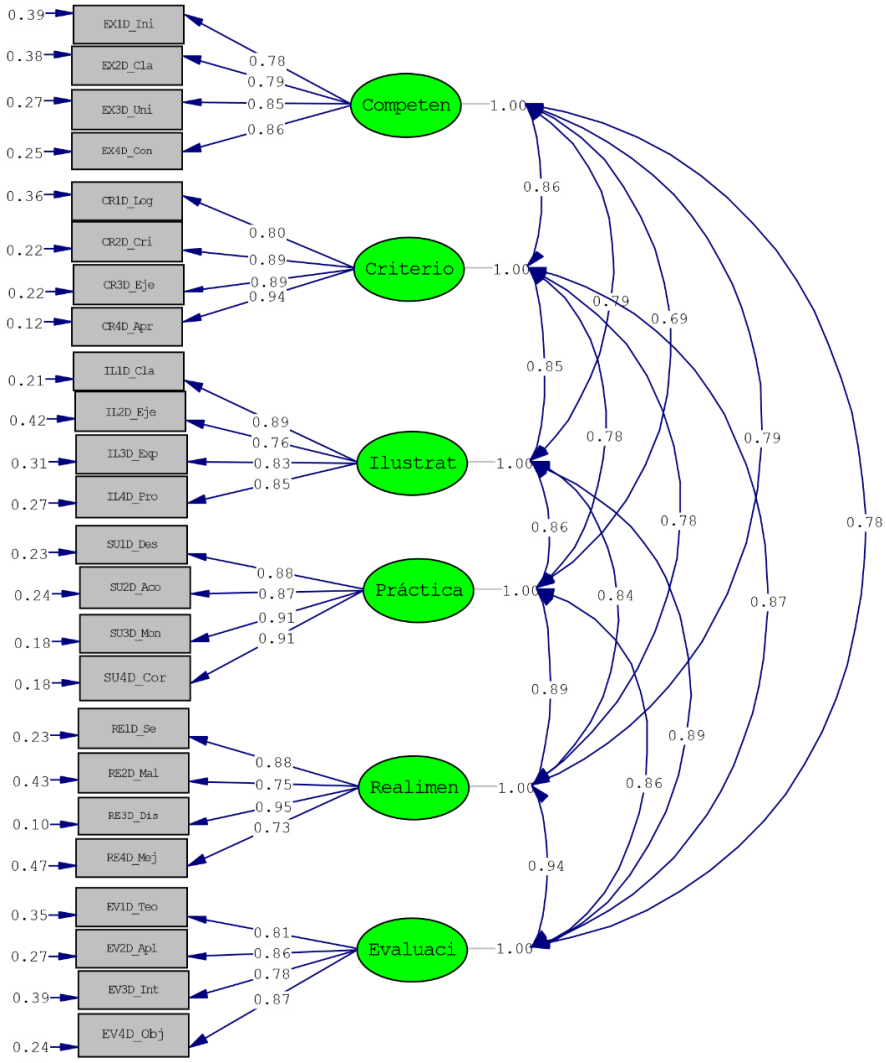

Figure 1. Teacher Performance Confirmatory Factor Analysis

Figure 2 shows the resulting model for the student performance confirmatory factor analysis through which the six student performance criteria constructs were confirmed, with a good divergent and convergent construct validity. However, provided that two of the indicators yielded very low factorial loads, convergent validity was not optimal for the last construct Application (Evaluation); therefore, the indicator with the lowest factor-indicator ratio was eliminated. Although the resulting model was not initially significant (Chi-Square=298.75, df = 215, P-value = 0.0001), it did obtain good goodness-of-fit practical indicators: RMSEA = 0.04; NFI = 0.98; NNFI = 0.99; CFI = 1.00; RMR = 0.03; GFI = 0.98; AGFI = 0.97.

Assessment of Didactic Performance During Online Classes

Table 2 shows the mean score ratios (ranging from 0 to 100) from the assessment of each didactic performance criterion. According to students, teacher didactic performances occur frequently (mean values ranging between 77.61 and 84.68). The Illustration and Criteria explicitness criteria had the highest overall rankings (84.68 and 84.66 respectively).

In terms of the didactic performance of students themselves, all criteria yielded high mean scores ranging between 76.05 and 83.79, except for the Evaluation criterion (Participation and Application), which had a regular assessment average (53%).

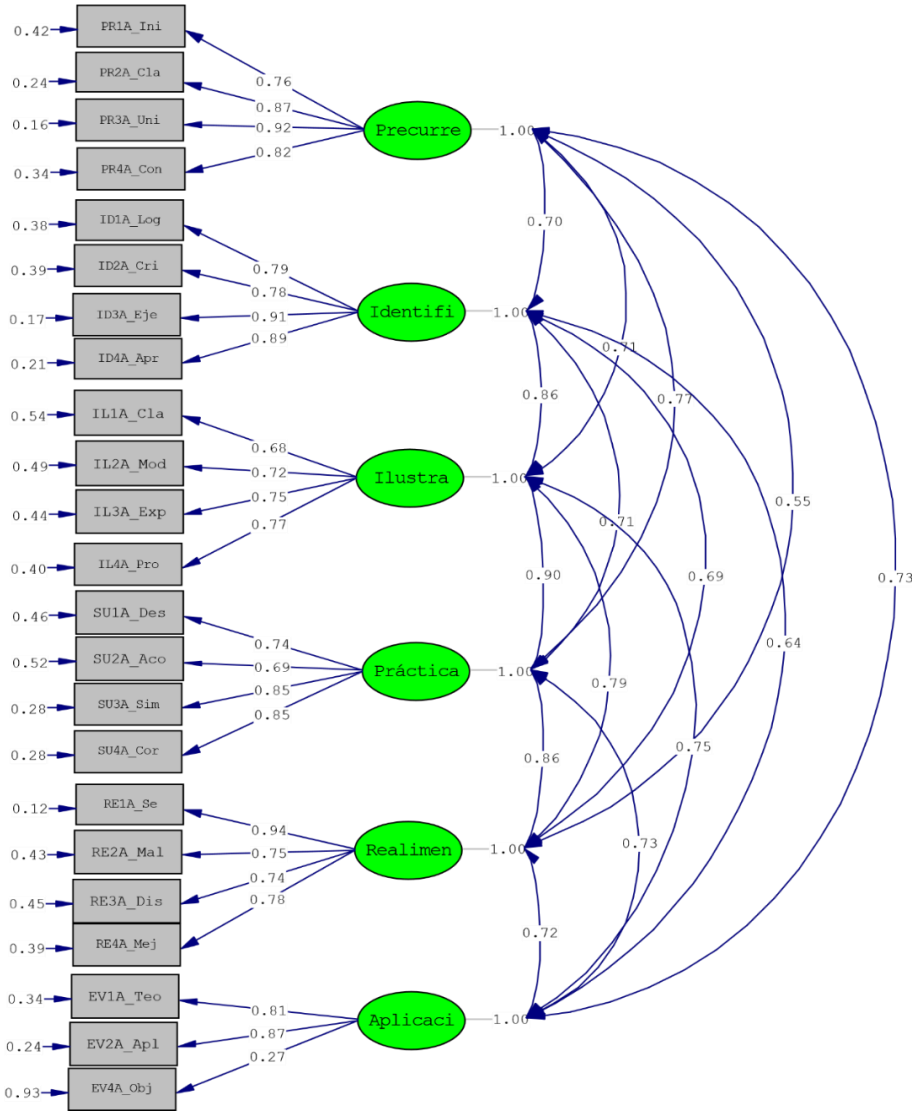

Figure 2. Student Performance Confirmatory Factor Analysis

Table 2. Mean Score Ratios for Didactic Performance Criteria (N = 309)

	Lowest value	Highest value	Mean	Std. deviation	Sex		Academic level	
					Male	Female	Master's	Doctor's
Teacher performance								
Competency exploration	.00	100	77.61	17.29	78.87	75.37	77.89	76.92
Criteria explicitness	33.33	100	84.66	13.99	85.22	83.63	84.27	85.63
Illustration	16.67	100	84.68	14.33	85.94	82.43	85.47	82.66
Practice supervision	16.67	100	80.79	14.42	81.10	80.25	80.22	82.28
Feedback	.00	100	77.56	16.44	78.12	76.58	77.70	77.20
Evaluation	.00	100	79.37	15.02	80.56	77.25	79.54	78.92
Student performance								
Precurrent learning behavior	16.67	100	76.05	13.82	76.01	76.12	13.46	14.76
Criteria identification	41.67	100	83.79	12.58	83.59	84.15	12.319	13.23
Illustration – Participation	50.00	100	80.53	11.75	80.30	80.93	11.57	12.10
Relevant practice	16.67	100	81.04	13.34	80.76	81.53	13.03	13.95
Feedback – Improvement	50.00	100	81.23	13.09	81.15	81.38	12.92	13.55
Evaluation – Application	25.00	75	52.94	10.10	51.51	55.48	9.66	11.13

Differences in Didactic Performances According to Sex and Academic Level

Differences between men and women in each teacher/student performance criterion were analyzed using a student's t-test for independent groups. Only two of the twelve constructs revealed significant differences between the sexes: Illustration (teacher performance) and evaluation–application (student performance). For the Illustration criterion (teacher performance), the difference favored men ($p = .04$). In contrast, evaluation–application (student performance) showed a significant deviation in favor of women ($p = .01$). Mean difference analyses by academic level (master's or doctor's) were also conducted. However, no significant differences were found for either didactic performance category.

Discussion

Teacher's Didactic Performance

Our results allowed for validating constructs of seven teacher performance categories based on the didactic performance model from an interbehavioral psychology standpoint as proposed by Carpio et al. (1998), Irigoyen et al., (2011), and Silva et al. (2014). According to these authors, such generic categories describing teachers' pedagogical practice can be used to describe teacher behaviors in psychology and related disciplines. However, empirical evidence was required to deem these outlooks suitable as interaction analysis categories during routine conditions in the classroom.

In contrast to theoretical propositions and approaches stemming from this didactic performance model (Carpio et al., 1998; Irigoyen et al., 2011; Morales et al., 2017; Reyna & Hernández, 2017; Silva et al., 2014), our findings are empirical evidences of the suitability of these categories to identify performance criteria according to postgraduate students' self-reports on their teachers. These findings also highlight that, because these are general didactic performance categories, they can also be applied to similar disciplinary areas, such as psychology teaching, education sciences teaching, and natural sciences teaching, among others.

Furthermore, our findings align with application and validation reports on these categories for analyzing the teachers' didactic performance from an interbehavioral psychology standpoint in high school classes in Mexico. In this line, Ávila (2020) reported using a direct observation checklist on the teacher's behavior during classes; Bazán-Ramírez et al. (2022) reported using an observational records & teacher-students performance self-reports hybrid study, and Velarde-Corrales and Bazán-Ramírez (2019) using observational records analysis on teacher-students didactic performance.

Moreover, our findings are similar to those of Bazán-Ramírez and Velarde-Corrales (2021), using psychology students' self-reports on their own and their teacher's didactic performance.

The similarity with the studies mentioned above stems from the validity and permanence of these constructs to identify teacher's and students' didactic performances in different higher education levels and disciplines. One difference, however, is the use of supplementary techniques and procedures to analyze didactic interactions, for instance, the behavioral checklist in the classroom, observational records and self-reporting.

As a second aspect to underscore, our study showed that the three best-ranked criteria of the teacher's didactic performance during virtual classes were criteria explicitness, illustration, and practice supervision. In this sense, these constructs that allow us to describe areas of didactic performance during online classes match the findings of other studies to outline actions and conditions of online learning (De la Riva & Álvarez, 2020; Portillo et al., 2020). Criteria explicitness, Illustration, and Practice supervision are three performance areas that refer to behaviors related to the development of classroom teaching.

On the one hand, these results correspond to findings on teacher performance as assessed by students (Chan, 2018; Scherer et al., 2016; Scherer & Gustafsson, 2015; Üstünlüoglu & Güngör-Culha, 2012). On the other hand, these students are teachers themselves who, due to the COVID-19 pandemic, were also teaching online to their own students at various academic levels. De la Riva and Álvarez (2020) indicate that the students' assessment of online learning processes is influenced by their experiences and teaching skills during this virtual period because of their roles as both graduate students and teachers giving online classes of traditional programs. Therefore, it is possible that giving their graduate teachers' performance higher scores for criteria inherent to teaching comes from their own pedagogical knowledge and from the shared experience adapting traditional teaching activities into online teaching (De Vincenzi, 2020).

Student's Didactic Performance

According to the student performance self-evaluations, Criteria identification was the top-ranking criterion. However, in general, Education Sciences graduate students gave high occurrence rates for five of the six areas, except for the Evaluation–Application area. On the one hand, these results confirm that student performance occurrence corresponds to teacher performance occurrence, with higher rates for five of the criteria. This result might indicate a functional correlation between teacher and student didactic competencies (Morales et al., 2017). On the other hand, this would also confirm that these categories of didactic performance criteria allow to identify teacher-student interaction trends as functional correlations between the teacher's and the students' behaviors during didactic interactions (Carpio et al., 1998; Irigoyen et al., 2011; Reyna & Hernández, 2017; Silva et al., 2014; Velarde-Corrales & Bazán-Ramírez, 2019).

Our results from verbal self-reports of Peruvian Education Sciences postgraduate students on their performance as a functional adjustment to the teacher's performance are in line with the findings on Mexican high school students in natural sciences from observational records (Velarde-Corrales & Bazán-Ramírez, 2019) and from an observational & self-reporting hybrid technique (Bazán-Ramírez et al., 2022). In addition, these data are consistent with the findings reported by Bazán-Ramírez and Velarde-Corrales (2021) about Mexican psychology undergraduate students self-reporting on their didactic performance as a functional adjustment to their teachers performance during Psychology classes and with the findings reported by Bazán-Ramírez et al. (2021) on Peruvian psychology students self-reporting on the teacher-students didactic performance during in-person classes. Regardless of the academic level, discipline, and geographical context, didactic performance constructs demonstrate their stability and validity. However, it is worth mentioning that earlier studies mentioned included only five of the seven criteria proposed on the didactic performance interbehavioral model. In contrast, this study encompassed six didactic performance constructs.

Lastly, Illustration occurrence ranked higher according to men. For this criterion, the teacher is expected to explain and bring the student into contact with effective strategies to fulfill a given criterion through the teacher's own actions and those of the student. Dissimilarities in the teacher's performance according to the sex of the students have been deemed significant and linked to the sex of the university teacher whose performance is being evaluated (Boring, 2017).

Furthermore, differences according to sex were found in the evaluation–application. This student performance criterion is where students are expected to perform activities and exercises, solve problems, apply procedures according to the expected achievement, and be able to apply their knowledge when facing unseen issues and situations. However, because further studies are needed on this construct (with indicators yielding better factorial convergence loadings), caution must be exercised regarding this significant difference according to sex in the self-assessment of the evaluation–application performance criterion. Once this pending issue has been solved, research can be supplemented by conducting a qualitative analysis with a gender perspective to explain better a potential difference by gender concerning the students' self-assessment, revisiting the term "study competencies" (Morales et al., 2017).

In general, our results provide a detailed account of the assessments made by Education postgraduate students regarding their teachers' (trainers of trainers) and their own didactic performance as students in online didactic interactions during the first six months of the pandemic in Peru. These results are consistent with earlier findings regarding teachers teaching virtually while also attending refresher or training courses during the pandemic (Lee et al., 2022; Purwanti et al., 2022; Szabó et al., 2022; van Wyk et al., 2021). Teachers may therefore have been able to adapt their performances from a face-to-face experience to these new virtual environments, especially after the first months of online teaching following the onset of the pandemic in 2020.

Conclusions

This study can be seen as the first approach to identifying teacher-student didactic performance criteria during online classes in a discipline other than psychology and using student self-reports based on the didactic performance model according to the interbehavioral perspective of psychology. These self-reports collect information on the occurrence rates of didactic performance indicators (Bazán-Ramírez et al., 2022) instead of using these categories to analyze class interactions (Velarde-Corrales & Bazán-Ramírez, 2019) with observational methods that allow assessing teaching and pursuing its improvement (Gitomer, 2018). However, one drawback of self-reporting a class on a finished term is the simultaneous occurrence of didactic interactions in other term subjects and the time elapsed between the events to be assessed and the fulfillment of the actual report. We may present five central conclusions:

a) An instrument with convergent and divergent construct validity for six performance criteria is available to identify teacher performances using student self-reports during didactic interactions in online classes of an in-person postgraduate program.

b) Graduate students filled out a self-report questionnaire to assess their didactic performance obtaining good divergent construct validity and moderate convergent validity for the Application criterion and good convergent validity for the remaining student didactic performance criteria.

c) Explicitness of criteria and Illustration were the teacher didactic performance criteria most frequently observed by the students attending online classes of the in-person Education Sciences postgraduate program. In other words, according to Education Sciences postgraduate students, their professors provide the specifics of the learning objectives and criteria to their students and demonstrate what to do and how to do it.

d) According to the Education Sciences postgraduate course students attending online classes, the Criteria identification criterion ranks the highest regarding their own performance, followed by Feedback–Improvement. This information implies that graduate students adjust to the explicitness of criteria and learning objectives from teachers to a greater extent. Likewise, they usually refer to their teachers' feedback when looking for ways to improve their performance.

e) Considering the twelve didactic performance criteria evaluated by the Education Sciences postgraduate students during didactic interactions in virtual classes, significant differences were only found for the Illustration teacher criterion, in favor of men, and for the Evaluation–Application student criterion, in favor of women.

In brief, new findings derived from this study include, among others:

a) Having obtained didactic performance measurements with convergent and divergent construct validity from theoretical categories using self-reports from Education Sciences postgraduate students,

b) Strategies to identify didactic performance criteria or categories with higher or lower occurrence during didactic interactions (according to the students' self-reports) to be able to create intervention proposals with teachers,

c) Strategies to identify differences in the didactic performance according to variables or characteristics of the research participants.

Recommendations

This study exemplifies how graduate professors and graduate students behave didactically according to the Education postgraduate students' self-perceptions during those first six months of online education, which began as a response to unprecedented events that would later bring fatal consequences worldwide. It is hard to ascertain if these didactic performances were different in a pre-pandemic context before March 2020 or how they would differ now, almost two years later and nearing the end of the third wave of the pandemic. Further studies are needed not only to retrospectively measure the various criteria or areas of didactic performance during online classes but also in terms of effectiveness, satisfaction, and empowerment of didactic strategies, among other variables (Lee et al., 2022; Purwanti et al., 2022; Szabó et al., 2022; van Wyk et al., 2021).

Limitations

An important limitation of this study is the significance of having invariance indicators for the measurements of both the didactic performance of the postgraduate teacher and the self-evaluations of graduate students on their didactic performance (with six performance criteria constructs in both cases). Two specific limitations are underscored: 1. Factorial invariance by sex in both didactic performance instruments and 2. To determine the factorial invariance of both instruments, consider the participants' academic level (Master's degree or Doctorate). According to the Education Science postgraduate students' evaluations, these two aspects are essential to establish more precise differences in teacher and student didactic performances.

Authorship Contribution Statement

Bazán-Ramírez: Conceptualization, design, conducted the investigation, analysis, writing. Quispe-Morales: Design, Conducted the investigation, Editing/reviewing. De La Cruz-Valdiviano: Writing, review of the APA style and formal aspects for the publication. Henostroza-Mota: Review of the theoretical foundations, punctuation marks and structure of the manuscript.

References

Ávila, M. A. (2020). *Evaluación de interacciones didácticas mediante ámbitos del desempeño docente en profesores de preparatoria en Cananea* [Evaluation of didactic interactions through areas of teaching performance in high school teachers in Cananea] [Master's thesis, Universidad de Sonora, México]. CRIS-UNISON. https://bit.ly/3PQa8bv

Bazán-Ramírez, A., Pérez-Morán, J. C., & Bernal-Baldenebro, B. (2021). Criteria for teaching performance in psychology: invariance according to age, sex, and academic stage of Peruvian students. *Frontiers in Psychology, 12*, 1-13. https://doi.org/10.3389/fpsyg.2021.764081

Bazán-Ramírez, A., & Velarde-Corrales, N. M. (2021). Autoreporte del estudiantado en criterios de desempeño didáctico en clases de Psicología [Students Self-report within didactic performances criteria in Psychology classes]. *Journal of Behavior, Health & Social Issues, 13*(1), 22-35. https://bit.ly/3uedxJt

Bazán-Ramírez, A., Velarde-Corrales, N. M., Rodríguez-Pérez, M. E., Guerrero-Barrios, J., & Anaya-González, R. B. (2022). Observational record and self-report of teacher-student performance in high school lessons. *International Journal of Educational Methodology, 8*(3), 479-491. https://doi.org/10.12973/ijem.8.3.479

Boring, A. (2017). Gender biases in student evaluation of teachers. *Journal of Public Economics,145*, 27-41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Boysen, G. A. (2016). Using student evaluations to improve teaching: Evidence-based recommendations. *Scholarship of Teaching and Learning in Psychology, 2*(4), 273-284. https://doi.org/10.1037/stl0000069

Carpio, C., Pacheco, V., Canales, C., & Flores, C. (1998). Comportamiento inteligente y juegos de lenguaje en la enseñanza de la psicología [Intelligent behavior and language games in the teaching of psychology]. *Acta Comportamentalia, 6*(1), 47-60. https://bit.ly/3O6FcE2

Chan, W. M. (2018). Teaching in higher education: Students' perceptions of effective teaching and good teachers. *Social Sciences and Education Research Review, 5*(1), 40-58. https://bit.ly/3zHbTDy

De la Riva, M., & Álvarez, G. (2020). Artefactos de inscripción digitales en la formación docente de posgrado [Digital inscription artifacts in postgraduate teacher education]. *Transdigital, 1*(2), 1-26. https://doi.org/10.56162/transdigital32

De Vincenzi, A. (2020). Del aula presencial al aula virtual universitaria en contexto de pandemia de COVID-19. Avances de una experiencia universitaria en carreras presenciales adaptadas la modalidad virtual. [From the university face-to-face classroom to the virtual classroom during the COVID-19 pandemic. Developments of a university experience in on-site courses adapted to the virtual modality]. *Debate Universitario, 8*(16), 67-71. https://bit.ly/3Tk6utk

Doménech-Betoret, F. (2018). The educational situation quality model: Recent advances. *Frontiers in Psychology, 9*, 328. https://doi.org/10.3389/fpsyg.2018.00328

Galindo, L., Silva, H., Serrano, V., Rocha, E., & Galguera, R. (2017). Aprendizaje por observación de interacciones didácticas de ilustración y retroalimentación [Interbehavioral survey about learning by observation in didactic interactions of illustration and feedback]. *Interacciones, 3*(3), 131-140. https://doi.org/10.24016/2017.v3n3.71

García-Peñalvo, F. J. (2020). Modelo de referencia para la enseñanza no presencial en universidades presenciales [Reference model for virtual education at face-to-face universities]. *Campus Virtuales, 9*(1), 41-56. https://bit.ly/3Rwh13d

Ghislandi, P., Raffaghelli, J., Sangrà, A., & Ritella, G. (2020). The street lamp paradox: Analysing students' evaluation of teaching through qualitative and quantitative approaches. *Journal of Educational, Cultural and Psychological Studies*, *21*, 65-85. https://doi.org/10.7358/ecps-2020-021-ghis

Gitomer, D. H. (2018). Evaluating instructional quality. *School Effectiveness and School Improvement*, *30(*1), 68-78. https://doi.org/10.1080/09243453.2018.1539016

Gonzales-Zamora, J. A., Alave, J., De Lima-Corvino, D. F., & Fernandez, A. (2020). Videoconferences of Infectious Diseases: An educational tool that transcends borders. A useful tool also for the current COVID-19 pandemic. *Le Infezioni in Medicina, 28*(2), 135-138. https://cutt.ly/CyjLpWf

Irigoyen, J., Acuña, K., & Jiménez, M. (2011). Interacciones didácticas en educación superior. Algunas consideraciones sobre la evaluación de desempeño [Didactic interactions in higher education. Some considerations about

performance evaluation]. In J. Irigoyen, K. Acuña, and M. Jiménez (Eds.), *Evaluación de Desempeños Académicos* (pp. 73–96). Universidad de Sonora. https://bit.ly/3n3R79r

Jöreskog, K., & Sörbom, D. (2001). *LISREL 8: User's reference guide.* Scientific Software International.

Lee, Y., Davis, R., & Li, Y. (2022). Implementing synchronous online flipped learning for pre-service teachers during COVID-19. *European Journal of Educational Research, 11*(2), 653-661. https://doi.org/10.12973/eu-jer.11.2.653

Morales, G., Peña, B., Hernández, A., & Carpio, C. (2017). Competencias didácticas y competencias de estudio: Su integración funcional en el aprendizaje de una disciplina [Didactic competencies and competencies of study: its functional integration in the learning of a discipline]. *Alternativas en Psicología 21*(37), 24–35. https://bit.ly/3csDfUL

Newton, G., Poung, K., Laila, A., Bye, Z., Bettger, W., Cottenie, K., Dawson, J., Graether, S. P., Jacobs, S., Murrant, C., & Zettel, J. (2019). Perception of biology instructors on using student evaluations to inform their teaching. *International Journal of Higher Education, 8*(1), 133-147. https://doi.org/10.5430/ijhe.v8n1p133

Organisation for Economic Cooperation and Development. (2017). *PISA 2015 Assessment and Analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving. Revised edition*. https://doi.org/10.1787/9789264281820-en

Portillo, J., Garay, U., Tejada, E., & Bilbao, N. (2020). Self-perception of the digital competence of educators during the COVID-19 pandemic: A cross-analysis of different educational stages. *Sustainability*, *12*(23), 1-13. https://doi.org/10.3390/su122310128

Purwanti, I. T., Suryawati, E., & Eliwarti. (2022). Video lectures in online EFL flipped-classroom: Effectiveness, students' evaluation and experiences. *European Journal of Educational Research, 11*(2), 885-898. https://doi.org/10.12973/eujer.11.2.885

Ramírez-Asís, E., Espinoza, M., Esquivel, S., & Naranjo-Toro, M. (2020). Inteligencia emocional, competencias y desempeño del docente universitario: Aplicando la técnica mínimos cuadrados parciales SEM-PLS [Emotional intelligence, competencies and performance of the university professor: Using the SEM-PLS partial least squares technique]. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, *23*(3), 99-114. https://doi.org/10.6018/reifop.428261

Reyna, W., & Hernández, M. (2017). Enseñanza-aprendizaje de la psicología: Reflexiones desde la matriz científica interconductual [Teaching-learning of psychology: Reflections from interbehavioral scientific matrix]. *Interacciones: Revista de Avances en Psicología, 3*(3), 171-182. https://doi.org/10.24016/2017.v3n3.67

Ribes, E., & López, F. (1985). *Teoría de la conducta: Un análisis de campo y paramétrico* [Behavior theory: A field and parametric analysis]. Trillas.

Scherer, R., & Gustafsson, J. E. (2015). Student assessment of teaching as a source of information about aspects of teaching quality in multiple subject domains: An application of multilevel bifactor structural equation modeling. *Frontiers in Psychology, 6,* 1-15. https://doi.org/10.3389/fpsyg.2015.01550

Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating individual students' perceptions of instructional quality: An investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Frontiers in Psychology, 7,* 1-16. https://doi.org/10.3389/fpsyg.2016.00110

Silva, H. O., Morales, G., Pacheco, V., Camacho, A. G., Garduño, H. M., & Carpio, C. A. (2014). Didáctica como conducta: Una propuesta para la descripción de las habilidades de enseñanza [Didactic as behavior: A proposal for the description of teaching skills]. *Revista Mexicana de Análisis de la Conducta, 40*(3), 32-46. https://doi.org/10.5514/rmac.v40.i3.63679

Spooren, P., Vandermoere, F., Vanderstraeten, R., & Pepermans, K. (2017). Exploring high impact scholarship in research on student's evaluation of teaching (SET). *Educational Research Review, 22*, 129-141. https://doi.org/10.1016/j.edurev.2017.09.001

Swart, F., Knezic, D., Onstenk, J., & de Graaff, R. (2019). Evaluating and improving teacher educators' language oriented performance in content-based teaching. *International Journal of Educational Methodology, 5*(1), 71-86. https://doi.org/10.12973/ijem.5.1.71

Szabó, E., Kóródi, K., Szél, E., & Jagodics, B. (2022). Facing the inevitable: The effects of coronavirus disease pandemic and online teaching on teachers' self-efficacy, workload and job satisfaction. *European Journal of Educational Research, 11*(1), 151-162. https://doi.org/10.12973/eu-jer.11.1.151

Üstünlüoglu, E., & Güngör-Culha, D. (2012). Investigating student evaluation of teachers by using latent class analysis: A case study at a tertiary level. *International Journal of Education, 4*(3), 147-159. https://doi.org/10.5296/ije.v4i3.1811

van der Lans, R. M., van de Grift, W., & van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education, 86*(2), 247-264. https://doi.org/10.1080/00220973.2016.1268086

van Wyk, M. M., Kotze, C. J., Tshabalala, S. L., & Mukhati, F. (2021). The responsiveness of teacher education managers at an ODel college to resilience and the well-being of staff working from home during COVID-19. *International Journal of Educational Methodology, 7*(4), 623-635. https://doi.org/10.12973/ijem.7.4.623

Velarde-Corrales, N. (2020). *Análisis de los criterios de desempeño didáctico en docentes y estudiantes de educación media superior como componentes de la interacción didáctica* [Analysis of the didactic performance criteria in high school teachers and students as components of didactic interaction] [Doctoral dissertation, Autonomous University of the State of Morelos, Mexico]. DSpace UAEM. https://bit.ly/3ARqDzS

Velarde-Corrales, N., & Bazán-Ramírez, A. (2019). Sistema observacional para analizar interacciones didácticas en clases de ciencias en bachillerato [Observational system to analyze didactic interactions in science classes in bachelor]. *Revista de Investigación en Psicología, 22*(2), 197-216. https://doi.org/10.15381/rinvp.v22i2.16806

Appendices

Appendix 1. Teacher Performance Assessment

Competency exploration

At the beginning of the course, the teacher assessed my prior knowledge orally and/or through a written questionnaire or task.

At the beginning of each class, the teacher inquired (explored) about our skills and knowledge related to the topic to be discussed.

At the beginning of each instructional unit, the teacher confirmed our level of proficiency, according to the competencies indicated in the syllabus.

At the beginning of each new topic, the teacher inquired about related concepts before explaining the topic.

Criteria explicitness

The teacher explained what the expected unit achievement was at the beginning of each Instructional Unit.

The teacher explained what criteria students had to meet to conduct a certain activity or task.

The teacher explained what criteria and requirements I had to meet to conduct a class exercise or a course practice.

In each class, the teacher explained clearly what achievement criteria students had to meet to acquire the class contents.

Illustration

The professor explained the topic of each class in detail.

The teacher showed examples of how to conduct a task or practice. The professor provided examples.

The teacher described how an expert with a postgraduate degree would solve a complex problem relevant to the education field.

The professor solved problems in front of the students according to the class topic.

Practice supervision

The teacher oversaw my performance during course practices.

All my colleagues received assistance from the teacher during the course activities.

The teacher established the practice conditions and oversaw their efficient fulfillment.

The teacher remediated our performance during the course activities.

Feedback

The professor remediated our performance during class activities, pointing out rights and wrongs.

The professor underscored my mistakes and shared viable solutions.

The professor taught me different ways in which I can meet the criteria for their class activities.

The professor graded the requested tasks and exercises and provided me with indications to correct or improve my work.

Evaluation

The professor conducted periodical evaluations of my theoretical knowledge or subject fundamentals.

The professor conducted applied evaluations and solutions to practical problems related to the subject.

The professor assessed my capacity to integrate knowledge from other courses into their class.

The professor evaluated the students according to the learning objectives shown at the beginning of the term, which can also be found in the syllabus.

Appendix 2. Student Performance Assessment

Precurrent learning behaviors

I demonstrated my prior knowledge at the beginning of the course, according to the examination conducted by the teacher.

At the beginning of each class, I expressed and showed my skills and knowledge related to the class topic to the teacher.

I expressed my prior competencies and skills for each instructional unit, according to the syllabus.

I answered the teacher's questions and/or asked questions about concepts related to a new topic or instructional unit.

Criteria identification

I identified what achievement (learning objectives) the teacher expected me to reach during that Instructional Unit.

I completed the activity or task while fulfilling the criteria explained by the teacher.

I met the criteria established by the professor when conducting class exercises or course practices.

I applied the achievement criteria established by the teacher for my class learning.

Illustration–Participation

I had no issues whatsoever comprehending the topic explained by the professor during class.

I completed class exercises or homework following the model devised by the professor.

I would solve a complex problem relevant to the education field, as close as possible to how an expert with a postgraduate degree would do it

I solved the problems or tasks devised by the professor during class.

Relevant practice

I conducted my activities under the supervision of the teacher.

I participated in learning activities controlled by the teacher.

I participated efficiently in activities mirroring problems that I might face when holding a Master's/Doctor's degree in Education.

I understood and applied the corrections provided by the teacher during the course activities.

Feedback–Improvement

I improved my performance to conduct class activities appropriately and according to the required criteria.

I identified my mistakes during class and adopted remedial suggestions from the teacher.

I conducted the exercises relying on the different methods taught by the teacher.

I adjusted my actions to improve my work and meet the criteria set for the task.

Evaluation–Application

I showed proficiency in the theoretical principles of the relevant topic.

I made contributions for practical application and problem-solving.

I applied the expected knowledge and competencies according to the course syllabus.