Academic co-creation: development and validation of a short scale

José Ventura-León[1]*, Andy Rick Sánchez-Villena[1],
Tomás Caycho-Rodríguez[2] and Michael White[3]

[1]Facultad de Ciencias de la Salud, Universidad Privada del Norte, Lima, Peru, [2]Facultad de Psicología,
Universidad Científica del Sur, Lima, Peru, [3]Facultad de Ciencias Humanas y Educación, Universidad
Peruana Unión, Lima, Peru

Introduction: Given the profound changes caused in higher education by the COVID-19 pandemic, affecting 1.6 billion students and 63 million educators globally, there arises the necessity for quantifiable measures that capture the essence of academic co-creation. This study aimed to develop and validate a short scale that measures academic co-creation (AC-S) in a sample of higher education students.

Methods: A total of 3,169 students from three Peruvian cities participated in the study (Mean Age = 25.77 years old; SD = 8.92 years); 1889 were female (59.60%) and 1,280 (40.40%) males. Qualitative and quantitative procedures were used for test construction. Item response theory (IRT) under the two-parameter graded response model (GRM-2PL) and test information function were used to examine reliability; additionally, a brief measure of academic satisfaction was used to provide evidence of relationship with another variable.

Results: The AC-S displayed strong fit and reliability, assessed through the test information function and standard error. It also showed a moderate correlation with academic satisfaction, bolstering its validity by linking with a pertinent variable. Its brevity enhances its practicality for education and research, efficiently fitting explanatory models and educational contexts. Despite substantial sample size and advanced psychometric methods, the study acknowledges limitations in sample representativeness and cross-sectional design. In conclusion, IRT and SEM techniques compellingly support the AC-S's reliability and validity.

Conclusion: The scale's one-dimensionality, local independence, reliability, and academic satisfaction relationship form a foundation for future exploration of co-creation-based educational models. Further studies should evaluate its performance across diverse cultural contexts.

KEYWORDS

academic co-creation, validation, development, short scale, higher education

1. Introduction

The COVID-19 pandemic has induced significant changes in higher education, affecting 1.6 billion students and 63 million teachers globally (United Nations, 2020). This scenario has triggered a set of far-reaching structural changes in higher education (Tilak and Kumar, 2022). One of them is the incorporation of new study variables as motivation for research (Esteban et al., 2022), satisfaction with virtual courses (Ventura-León et al., 2022a), blended learning (Singh et al., 2021), and inspiration (Ventura-León et al., 2022b). Therefore, the university can be conceptualized as an entity in which various elements associated with the process of knowledge acquisition converge (Dollinger et al., 2018).

In line with this, the university has a strong responsibility to society because it seeks to provide highly trained professionals (Schlesinger et al., 2015; Yin and Wang, 2016) by developing in its student's team collaboration, decision-making, problem-solving, knowledge generation, and innovation competencies (Irigoyen et al., 2011; Yamada et al., 2012). In addition, another motivation of the university is to increase students' participation in their learning process, and interactions among them prove to be a useful tool for collaborative learning, where students have similar opportunities to contribute to an academic topic (Cook-Sather et al., 2014; Bovill, 2020).

In the framework of these participatory actions, the concept of Co-creation is born, which involves collaborative participation between students and teachers and emphasizes active learning that contributes to bonding with the teacher (Ryan and Tilbury, 2013; Bovill et al., 2016). The term Co-creation involves interactive learning among students (Bovill, 2020) and implies cooperation, coordination, and collaboration in the sense of building relationships, planning, making decisions, and engaging in the whole learning process (Bentzen, 2020). Recently, this term is spreading in academia due to teacher-student interaction (Bovill, 2019). As a result of this dynamic, knowledge is built, focused on the student (Dollinger et al., 2018; Lystbæk et al., 2019), who elaborates a final learning product based on an interactive experience (Kaminskiene et al., 2020). In this context, Academic Co-creation can be defined as the interaction between the student, their peers, or the teacher, which favors the active teaching-learning process through the joint realization of academic actions (The justification of this method appears in Table 1).

In higher education, there are some factors associated with Co-creation, such as academic engagement (Tarı and Mercan, 2020), satisfaction along with positive experiences during classes (Lubicz-Nawrocka, 2018), motivation, quality of communication, an adequate teacher-student relationship, creativity, flexibility (Kaminskiene et al., 2020) and the short duration of lectures (Bovill, 2020). In turn, teachers must have the ability to create activities that take students out of their comfort zone and create new learning spaces (Kaminskiene et al., 2020), such as the development of negotiation skills and an ability to manage frustrating situations (Iversen and Pedersen, 2017) that may arise during teamwork in the classroom.

Co-creation in education breaks with traditional teacher-centered models, giving protagonist to the student, who plays an active role in their learning process (Kaminskiene et al., 2020). This strengthens their sense of belonging, commitment, positive experiences, and closeness with the teacher, increasing the value of learning and its outcomes (Dollinger et al., 2018; Lystbæk et al., 2019). This method, rooted in the principles of constructivism, entails the student acquiring knowledge through engagement with an instructor who evaluates the disparity between the student's existing knowledge and their capacity for learning (Rodríguez, 2008). It has garnered attention in higher education over an extended period (González et al., 2011), resulting in a shift where educators take on roles as facilitators and mediators in their interactions with students (Serrano and Pons, 2011).

Currently, there are no specific measurement instruments for co-creation in the educational context. However, we can refer to instruments developed in the business field as a guide, especially in determining key factors such as information sharing, tolerance, support, feedback, dialogue, and interpersonal connection. These factors have been previously highlighted by theorists (Yi and Gong, 2013; Taghizadeh et al., 2016; Merz et al., 2018). This emphasis is because the study focuses its attention on the Co-creation of students, as they are the ones who share experiences, commit to certain tasks, and achieve objectives in their educational work (Dollinger et al., 2018; Lystbæk et al., 2019; Kaminskiene et al., 2020). In fact, the exchange of knowledge among students allows for the development of solid competencies for the professional field (Dante, 2018), because professionals work in teams and must collaboratively carry out work-related projects.

In this context, this study aims to develop and provide evidence of content validity, internal structure, and reliability of a brief Co-creation scale in the context of higher education through quantitative and qualitative methods.

2. Materials and methods

2.1. Participants

The participants were 3,169 university students from three Peruvian cities (77.40% from Lima, 13.5% from La Libertad, and 9.15% from Cajamarca), whose ages ranged from 16 to 56 years old (Mean = 25.77 years old; SD = 8.92 years); 1889 were female (59.60%) and 1,280 (40.40%) male. Students participated form a total of six fields of study: Architecture and Design (8.46%), Communications (3.72%), Law (9.88%), Engineering (30%), Business (32.50%), and Health (15.50%). The sample size was estimated using the "semPower" library (Moshagen and Erdfelder, 2016) establishing *a priori* 2 degrees of freedom [k(k-3)/2]; RMSEA = 0.05; power of 0.95 and an alpha of 0.05, giving a total of 2,774 observations; thus, the recommended minimum was exceeded. Sampling was purposive, because participants were deliberately selected (Maxwell, 2012).

2.2. Instruments

Academic co-creation short scale (AC-S), which is a unidimensional measure composed of four items. Its psychometric properties are the subject of this study.

Academic Satisfaction Scale [Escala de Satisfacción Académica] (ESA) developed by Lent et al. (2007) with the Spanish version by Medrano and Pérez (2010). This is a unidimensional measure consisting of eight items with three response alternatives (0 = Never [Nunca], 1 = Sometimes [A veces], 2 = Often [A menudo], 3 = Always [Siempre]). The evidence of validity is through factor analysis presenting factor loadings above 0.40 and an explained variance of 49%. Reliability is considered good (α = 0.84).

2.3. Academic co-creation short scale construction procedure

First, a search for information was conducted in specialized databases. Then, a qualitative analysis method, proposed by Ventura-León (2021a), was implemented to provide evidence based on the content of the items. This method included the following stages: (a) Familiarization: the collected information was placed in a table and read and reread (see Table 1); (b) Segmentation: relevant information

TABLE 1 Comparative table of different definitions of co-creation.

Familiarization and segmentation			
MacMillan online dictionary (2020)	Oertzen et al. (2018, p. 642)	Bovill et al. (2016, p. 197)	Bovill (2020, p. 1026)
General	General	Educación	Educación
"*a way of* working together where people from different backgrounds are invited to jointly produce *a product or service that will benefit them all.*" ["*una forma de* trabajar en conjunto donde se invita a personas de diferentes orígenes a producir conjuntamente *un producto o servicio que los beneficiará a todos*"]	"Co-creation has *its roots in the verb to create, which is defined as to bring something into existence*, to make something happen as a result of one's own actions, and in *co-,* meaning together with another or others." ["La Co-creación tiene *sus raíces en el verbo crear, que se define como dar existencia a algo,* hacer que algo ocurra como resultado de las propias acciones, y en *co-,* que significa junto con otro u otros."]	"one way to conceptualize co-creation is to occupy the space between *student engagement and partnership, to suggest* meaningful collaboration between students and staff*, with students becoming more active participants in the learning process, *building* understanding and resources with academic staff." ["una forma de conceptualizar la co-creación es ocupar el espacio entre el *compromiso de los estudiantes y la asociación, para sugerir una* colaboración significativa entre los estudiantes y el personal*, con los *estudiantes convirtiéndose en participantes más* activos en el proceso de aprendizaje, *construyendo* la comprensión y los recursos con el personal académico."]	"Co-creation involves the development of *deeper relationships between student and* teacher*, and between students and other students. *Education is seen as* a shared endeavor in *which* learning and teaching *are done with students*, not for them." ["La Co-creación implica el desarrollo de *relaciones más profundas entre estudiante y* el profesor*, *y entre los estudiantes* y otros estudiantes. *La educación se percibe* como un esfuerzo compartido *en el que el* aprendizaje y la enseñanza *se realizan con los estudiantes*, no para ellos."]
Categorization			
a. Collaborative behavior: A joint, shared action among students for a defined purpose b. Student: An event that happens among students* c. Learning: promotes active learning			
Correspondence			
Instruction: *I learn best…* [*Yo aprendo mejor…*] (Expression indicating learning improvement) 1. ...exchanging ideas with my classmates. […intercambiando ideas con mis compañeros de aula.] (Collaborative behavior, students) 2. ...collaborating on group projects with my classmates. […colaborando en proyectos grupales con mis compañeros de aula.] (Collaborative behavior, students) 3. ...contributing ideas in group work with my classmates. […aportando ideas en los trabajos grupales con mis compañeros de aula.] (Collaborative behavior, students) 4. ...actively listening to my classmates in group work. […escuchando activamente a mis compañeros de aula en los trabajos grupales.] (Collaborative behavior, students)			

Note: Italics denote the relevant segment (segmentation). The different types of underlining indicate the similarity of the information segments (Categorization). *: It could also happen between student and teacher, but this form of evaluation could involve another test or other forms of evaluation. The expressions General and Education denote aspects of the definitions taken. For example, the first two are broad definitions and the last two are specifically in the field of higher education.

segments were identified; (c) Categorization: information segments were organized by common aspects; and (d) Correspondence: the relationship between the items and the previously generated categories was explained. This procedure led to the generation of a few items, which were revised prior to their mass application. The decision to formulate only a few items was related to the goal of constructing a brief test, i.e., one with fewer than 10 items (Ziegler et al., 2014). The response alternatives did not include the expression "Never" ["Nunca"] because a previous pilot study had shown that this alternative was not selected by the participants. A similar situation occurred with the term "Always" ["Siempre"], which was changed to "Very often or always is my case" ["Con mucha frecuencia o siempre es mi caso"]. All these changes improved the metric properties of the scale.

The pilot study was conducted using a virtual form shared through the virtual platform of a higher education institution. In this regard, recommendations for online administration, as outlined by Hoerger and Currell (2011), were followed. An informed consent form was utilized, explaining the key aspects of the study, such as its objectives, the anonymity of responses, and the treatment of collected information. The virtual form was distributed to students through their online classrooms. These actions were carried out as part of an institutional study with the support of the university's academic department.

2.4. Data analysis

Statistical analyses were performed using the R programming language implemented in its RStudio environment (RStudio Team, 2022). Specifically, the "mirt" library (Chalmers, 2012) was used for the estimation of the IRT model; "jrt" (Myszkowski, 2021) in combination with "ggplot2" (Wickham et al., 2020) for the IRT graphics, "semPlot" (Epskamp, 2015) for the explanation diagram and "dplyr" for data manipulation (Wickham et al., 2021).

The analysis was conducted in stages: preliminarily, descriptive statistics were calculated through the reporting of response rates because it is a Likert scale that is on an ordinal scale. Second, a latent variable approach such as item response theory (IRT) was used. It has been shown that IRT has notable advantages over classical test theory (CTT). In particular, IRT offers parameter invariance independent of the sample studied and provides an estimate of reliability through the test-item information function, where test accuracy along the latent trait can be observed (Zickar and Broadfoot, 2009). Specifically, a Graded Response Model (GRM; Samejima, 1997) was used, which proved to be the best model after a comparison with other IRT models (e.g., PCM and GPCM) based on a lower value of Bayesian information criterion (BIC; Schwarz, 1978). This indicator has been reported as one of the most accurate methods in polytomous IRT (Kang et al., 2009). Prior to the implementation of the IRT model, two

assumptions were checked; (a) local independence, through the $Q3^*$ statistic whose values below 0.20 are indicators of the presence of the assumption (Christensen et al., 2017); (b) monotonicity, by inspecting the characteristic curve of the categories.

IRT was performed considering a two-parameter model (2PL). The discrimination parameter (α) indicates the ability of the test to differentiate between people with low and high ability (θ); it is often located between −3 to 3; however, values higher than 1 are indicators of high discrimination. The location parameter (β) suggests the value on the θ scale where the person is likely to respond between one and the other response alternative. The algorithm for determining the dimensional reduction was MCEM (MonteCarlo EM estimation). The global adjustment of the model was performed using the recommendations by Maydeu-Olivares (2013): Log likelihood, comparative index (CFI $\geq$ 0.95), Tucker-Lewis index (TLI $\geq$ 0.95) and root mean square error of approximation (RMSEA $\leq$ 0.05).

The impact of differential item functioning (DIF) by gender was calculated using the Expected Score Standardized Difference (ESSD), based on Meade (2010) expected scores. ESSD values were interpreted following Cohen (1988) effect size categories, where ESSD >0.30 indicates a small effect, ESSD >0.50 is medium, and ESSD >0.80 is large.

Reliability was estimated through the test-item information function, together with the empirical reliability (r_{xx}) which is a ratio between *factor scores* and model estimates (Du Toit, 2003).

The evidence based on the relationship with other variables was carried out using structural equation modeling (SEM). For such purposes, the WLSMV estimator was employed, which is specifically used for observed categorical data (Likert scale). In this approach, the normality of the data is not assumed, nor is the continuity property considered necessary (Li, 2016). In this sense, the model fit was first evaluated through the Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA) whose values above 0.95 and below 0.08, respectively, are indicators of an optimal fit (Hu and Bentler, 1999). The interpretations of the relationships were carried out considering the recommendations of Cohen (1988): $\rho \geq 0.10$ small; $\rho \geq$. 30 mediums; $\rho \geq 0.50$ large. It is worth mentioning that these interpretations are pertinent when there is little information about the relationship between variables (Ventura-León, 2021b).

Finally, it is good to mention that during the whole data analysis process statistical significance tests were not used for two reasons: (a) they require the implementation of random sampling (Hirschauer et al., 2020) and (b) they are sensitive to large sample sizes (Lin et al., 2013).

3. Results

3.1. Preliminary analysis

Table 2 shows the descriptive statistics. It is observed that the highest values occur in response alternative 2 ("Sometimes it is my case" ["A veces es mi caso"]) and 3 ("Many times it is my case" ["Muchas veces es mi caso"]). In addition, the calculation of the ESSD index shows that in all items there is a small effect size (ESSD <0.50).

3.2. Item response theory

Preliminarily, it was examined which IRT model best fits the data. This procedure was performed using the BIC index, which

TABLE 2 Descriptive statistics and model parameters.

Item	Response rates				Item statistics				ESSD
	1	2	3	4	α	β_1	β_2	β_3	
1	13	32	32	22	3.34	−1.21	−0.16	0.81	0.42
2	14	30	34	21	4.20	−1.10	−0.18	0.82	0.47
3	14	25	39	22	4.88	−1.09	−0.30	0.78	0.49
4	12	28	37	23	2.70	−1.36	−0.32	0.85	0.42

Note. β, localization parameter; ESSD, Expected Score Standardized Difference (DIF).

shows that the GRM (BIC = 25715.739) is the best model for the data when compared to other polytomous models such as PCM (BIC = 26652.749) and GPCM (BIC = 26389.386). The $Q3^*$ index obtained a value of 0.15, below the established cut-off ($Q3^* \leq 0.20$). The discrimination parameters (α) were high for each of the items (i.e., > 1.0) and the localization values (β) showed a monotonic increase (see Table 2; Figure 1). Finally, the goodness of fit of the model was optimal: M2(2) = 8.23; RMSEA = 0.031; SRMR = 0.048; TLI = 0.998, CFI = 0.999.

3.3. Reliability

The reliability obtained through the empirical coefficient demonstrates the presence of good internal consistency at the peak of the trait assessment (r_{xx} = 0.89). These data are supported by the information and standard error function of the test which points to a maximum value of 15.36 (SE = 0.26) when the trait level at θ is equal to −0.26. This suggests that the instrument is more accurate at medium levels of the latent trait (see Figure 2).

3.4. Validity in relation to another variable

The literature suggests a relationship between Co-creation and academic satisfaction. The goodness of fit of the model was optimal: $\chi^2(42) = 686.281$; CFI = 0.992; TLI = 0.989; RMSEA = 0.070; SRMR = 0.044. The measurement model items were correctly represented. Figure 3 shows that there is a moderate relationship between Co-creation and academic satisfaction ($r = 0.43$). These results indicate that the AC-S scores show convergent validity with the ESA scores, as expected theoretically. It is worth mentioning that the errors of items 5 and 6 on the academic satisfaction scale had to be correlated, improving the instrument's fit as per modification index suggestions. The connection between the errors could indicate additional measurements beyond the academic satisfaction factor due to the expressions used or the relationship between expectations and personal preferences. Finally, the reliability obtained using the omega coefficient (which assumes a congeneric model) for the construct's Co-creation ($\omega = 0.89$) and Academic Satisfaction ($\omega = 0.90$) are good.

4. Discussion

Contemporary educational models assign an active role to learners within the teaching-learning process (Kaminskiene et al., 2020). Consequently, students are actively engaged in developing their

FIGURE 1
Characteristic curves of the categories.

FIGURE 2
Item and test information curves.
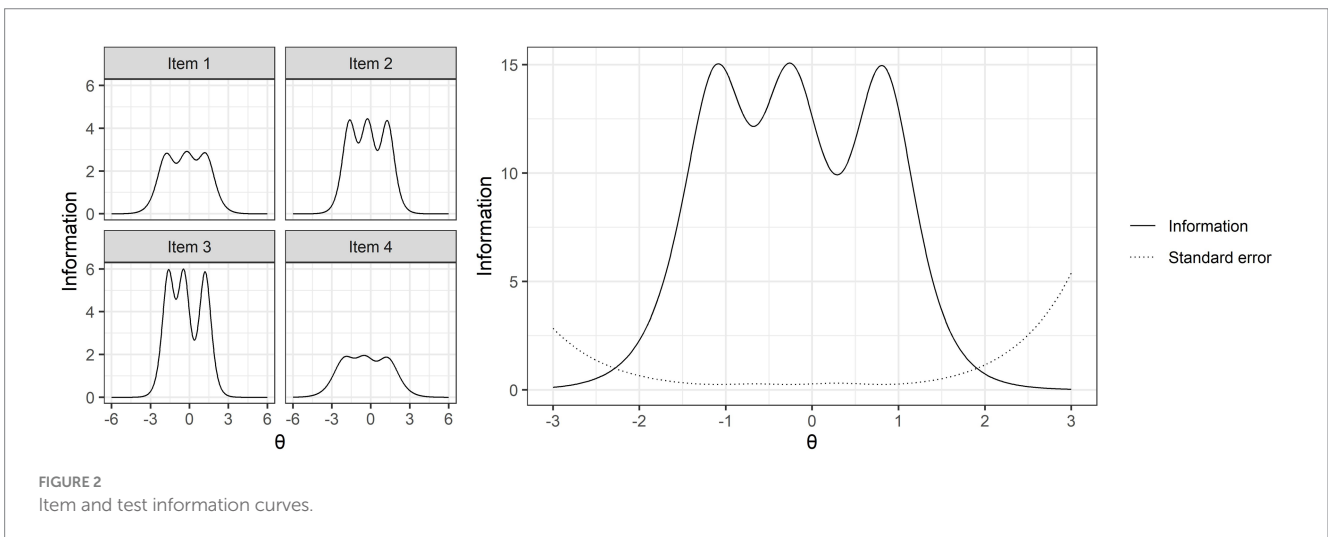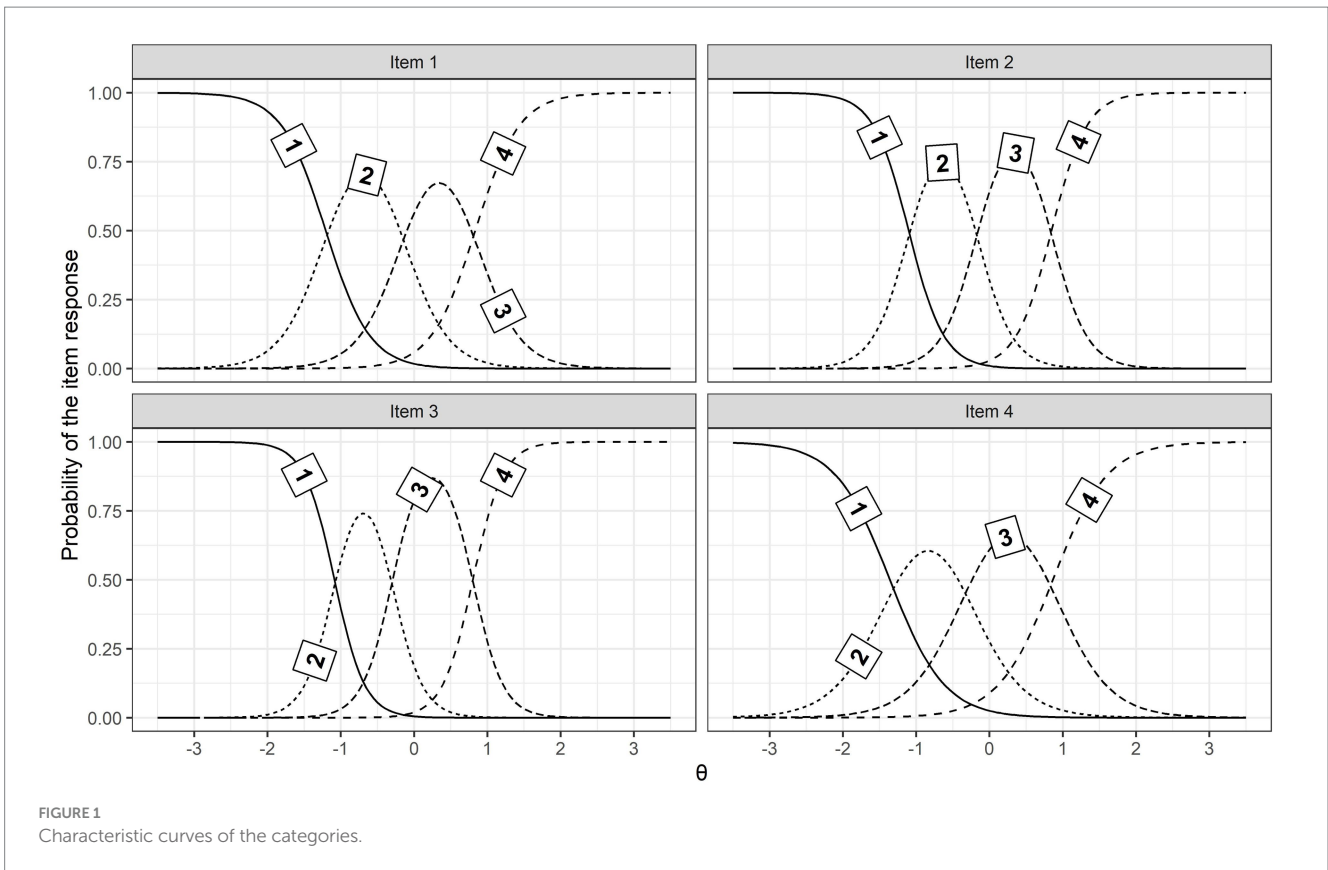
competencies (Irigoyen et al., 2011; Yamada et al., 2012). In this context, the concept of co-creation gains significance in higher education, emphasizing students' capacity to collaborate and work in teams to enhance academic learning (Bovill, 2020). Nevertheless, there is currently a lack of instruments designed to measure co-creation in the context of higher education, as existing tools are primarily designed from a business perspective (Yi and Gong, 2013; Taghizadeh et al., 2016). Consequently, this study aims to create and validate a concise instrument for assessing academic co-creation among university students, utilizing modern analysis methods (IRT) and qualitative techniques.

Initially, for the development of the instrument, a qualitative methodology was employed to establish the theoretical consistency between the items and the concept of Co-creation. Broadly speaking, Co-creation is understood as an active and collaborative engagement involving students, peers, and teachers (Ryan and Tilbury, 2013; Bovill et al., 2016). This dynamic promotes stronger interpersonal relationships among students, as well as greater commitment, planning, decision-making, and learning (Bentzen, 2020). In this context, we conceptualize academic co-creation as the interaction between students, their peers, and/or the teacher, which facilitates an active teaching-learning process through joint actions (Bovill et al.,
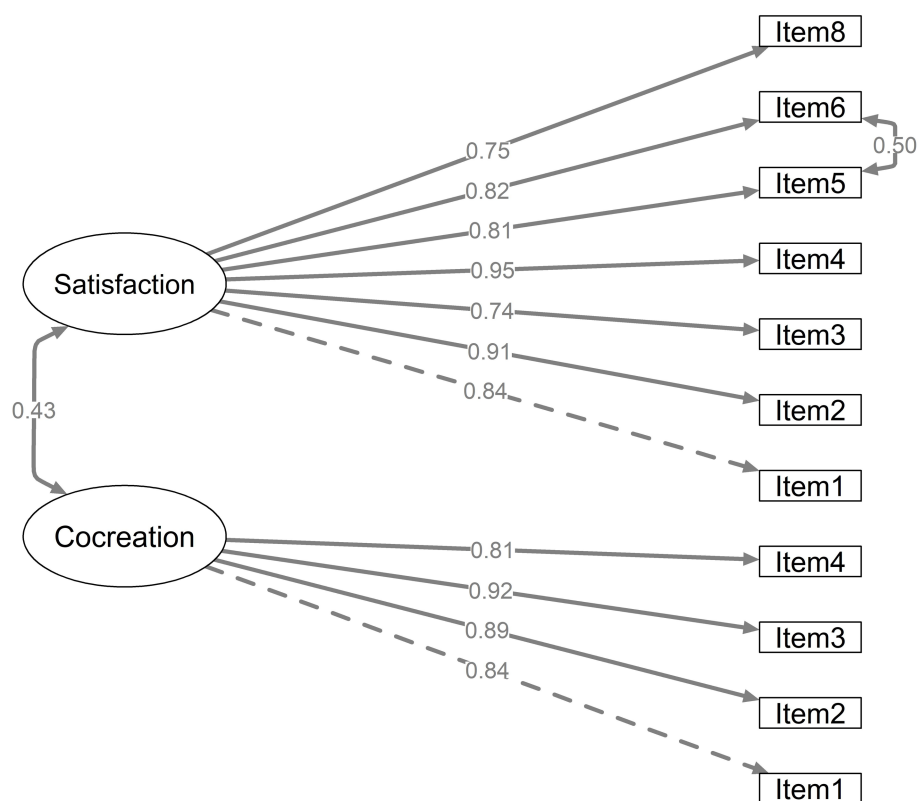
FIGURE 3
Explanatory model between the academic co-creation (AC-S) and the Escala de Satisfacción Académica (ESA).

2016; Oertzen et al., 2018; Bovill, 2020; MacMillan online dictionary, 2020), as defined through thematic analysis (Ventura-León, 2021b). This approach allowed us to construct items that are theoretically coherent. Moreover, this prompts us to contemplate that the notion of co-creation within education may be examined through a social constructivist lens (Rodríguez, 2008). This perspective posits that co-creation arises through interactions with the environment, wherein individuals shape their own comprehension of reality and juxtapose it with the perspectives of those in their vicinity (Helou et al., 2018). This approach has demonstrated its worth in higher education settings (González et al., 2011). Although Co-creation can involve other educational stakeholders, such as teachers, our focus in this study has been primarily on students. This is because there is a continuous exchange of experiences and interactions among students (Dollinger et al., 2018; Lystbæk et al., 2019; Kaminskiene et al., 2020) that contributes to the development of professional competencies that are valuable in their future careers (Dante, 2018).

Secondly, the items were administered to university students through virtual forms. The data collected were subjected to a descriptive analysis to determine the response trend, where it was found that the highest response rates were options 2 ("Sometimes it is my case") and 3 ("Many times it is my case"). It is worth mentioning that "Sometimes is my case" was chosen as the lowest expression, because the expression "Never" in a previous pilot was not selected by any of the participants. This may occur because, being a construct that denotes something positive, there is a tendency to choose at least the expression "Seldom." A similar situation occurred with the expression "Always" collapsing the alternative into "Very often or always is my case." After these changes, an improvement in the metric properties of the AC-S was observed.

The psychometric properties of the instrument were analyzed using Item Response Theory (IRT) due to its advantages over CTT (Zickar and Broadfoot, 2009). Specifically, the Graded Response Model (GRM) was chosen as the most suitable option among various IRT models (such as PCM and GPCM) because it demonstrated a lower BIC (Schwarz, 1978) and was well-suited for the ordinal nature of the Likert scale (Samejima, 1997). Regarding the discrimination and location parameters, the results indicated that the AC-S performed well in both aspects of metric properties. Furthermore, it exhibited an excellent goodness of fit, surpassing the minimum cut-off points established by Maydeu-Olivares (2013).

As Co-creation is a relatively new variable, there is limited existing information regarding differences between genders. Nevertheless, a Differential Item Functioning (DIF) analysis was conducted using the effect size measured by the ESSD index (Meade, 2010). The results indicate the presence of minimal differences in expected scores between men and women. Consequently, measurement invariance by gender is supported, affirming that the test is equitable (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014). However, it should be noted that some items (2 and 3) showed differences close to the moderate category. Therefore, it is recommended to further investigate the invariance of the AC-S in future studies.

Regarding the reliability of the scores, they were estimated using the test information function and the standard error (Du Toit, 2003), which indicate that the test performs well overall. Specifically, the items demonstrate similar performance. However, it is worth noting that item 3 ("I learn best by contributing ideas in group work with my classmates") exhibited a different pattern of behavior, with a

somewhat pointed distribution. This might be attributed to the presence of high discrimination parameters related to sampling variability (Feuerstahler, 2020).

Regarding the validity based on the relationship with other variables, the Co-creation scale was shown to have moderate and direct correlation with academic satisfaction. This is compatible with a previous study that points out that having positive experiences in the classroom fosters collaboration with peers (Lubicz-Nawrocka, 2018). In addition, it is important to mention that co-creation is related to several academic variables such as creativity, flexibility, new learning spaces (Kaminskiene et al., 2020), negotiation and management of frustrating situations (Iversen and Pedersen, 2017). These aspects will be important to consider in future studies.

One of the strengths of this study was the large sample size. Additionally, modern psychometric techniques, such as the IRT and SEM models, were employed in constructing the scales. From a practical standpoint, having a concise instrument simplifies the measurement of co-creation for both educational and research purposes, providing a tool that can be incorporated into explanatory models. As a result, less time is required to administer this scale, without sacrificing the information needed to assess the construct. Moreover, the AC-S can be easily applied in educational settings, particularly in higher education, to assess the level of co-creation among students, with lower scores indicating a lower degree of this phenomenon.

Despite its merits, it is essential to recognize certain limitations. Initially, even though the study boasted a substantial sample size, it may not provide a comprehensive representation of Peruvian university students due to the non-random sampling method employed. Therefore, future research using this type of sampling is needed (Bentzen, 2020). Additionally, the cross-sectional design presents another limitation, as it does not allow us to assess reliability through the test–retest method (Herting et al., 2018). Hence, it is recommended that future studies employ longitudinal approaches to examine the scale's stability over time.

5. Conclusion

The application of the IRT model and SEM have provided evidence of adequate reliability and validity of the AC-S. It is important to mention that the analyses indicate one-dimensionality and that the items are locally independent, reliable and have an acceptable fit. In addition, they show a relationship with academic satisfaction. These findings can further future research that aims to evaluate educational models based on Co-creation. Further research

is invited to investigate the performance of the scale in other cultural contexts.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Universidad Privada del Norte Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JV-L conceptualized and designed the study, data preparation, statistical analysis, and interpretation of data. AS-V wrote the first draft of the manuscript. TC-R: critical review of the first draft and suggestions for substantial improvement of the manuscript. MW critical review of the manuscript in the English language. All authors have access to the data and accept responsibility for data integrity and reporting accuracy. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing. 7th ed.* Washington, DC: American Educational Research Association.

Bentzen, T. Ø. (2020). Continuous co-creation: how ongoing involvement impacts outcomes of co-creation. *Public Manag. Rev.* 24, 34–54. doi: 10.1080/14719037.2020.1786150

Bovill, C. (2019). A co-creation of learning and teaching typology: what kind of co-creation are you planning or doing? *Int. J. Students as Partners* 3, 91–98. doi: 10.15173/ijsap.v3i2.3953

Bovill, C. (2020). Co-creation in learning and teaching: the case for a whole-class approach in higher education. *High. Educ.* 79, 1023–1037. doi: 10.1007/s10734-019-00453-w

Bovill, C., Cook-Sather, A., Felten, P., Millard, L., and Moore-Cherry, N. (2016). Addressing potential challenges in co-creating learning and teaching: overcoming resistance, navigating institutional norms and ensuring inclusivity in student–staff partnerships. *High. Educ.* 71, 195–208. doi: 10.1007/s10734-015-9896-4

Chalmers, R. P. (2012). Mirt: a multidimensional item response theory package for the R environment. *J. Stat. Softw.* 48, 1–29. doi: 10.18637/jss.v048.i06

Christensen, K. B., Makransky, G., and Horton, M. (2017). Critical values for Yen's Q 3: identification of local dependence in the Rasch model using residual correlations. *Appl. Psychol. Meas.* 41, 178–194. doi: 10.1177/0146621616677520

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. Abingdon.* United Kingdom: Routledge.

Cook-Sather, A., Bovill, C., and Felten, P. (2014). *Engaging students as partners in learning and teaching: a guide for faculty*. San Francisco: Jossey Bass.

Dante, G. P. (2018). Diseño de una auditoría del conocimiento organizacional orientada hacia los procesos principales y el desarrollo profesional. *Rev. Cuba. Inf. en Ciencias la Salud* 29, 1–12. Available at: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132018000300007

Dollinger, M., Lodge, J., and Coates, H. (2018). Co-creation in higher education: towards a conceptual model. *J. Mark. High. Educ.* 28, 210–231. doi: 10.1080/08841241.2018.1466756

Du Toit, M. (2003). *IRT from SSI: Bilog-MG, multilog, parscale, testfact*, United States: Scientific Software International.

Epskamp, S. (2015). semPlot: unified visualizations of structural equation models. *Struct. Equ. Model. Multidiscip. J.* 22, 474–483. doi: 10.1080/10705511.2014.937847

Esteban, R. F. C., Mamani-Benito, O., Huancahuire-Vega, S., and Lingan, S. K. (2022). Design and validation of a research motivation scale for Peruvian university students (MoINV-U). *Front. Educ.* 7, 1–7. doi: 10.3389/feduc.2022.791102

Feuerstahler, L. M. (2020). Metric stability in item response models. *Multivar. Behav. Res.* 57, 94–111. doi: 10.1080/00273171.2020.1809980

González, A. D., Rodríguez, A. D. L. Á., and Hernández, D. (2011). El concepto zona de desarrollo próximo y su manifestación en la educación médica superior Cubana. *Rev. Cuba. Educ. Medica Super.* 25, 531–539. Available at: http://scielo.sld.cu/pdf/ems/v25n4/ems13411.pdf

Helou, M. M., Newsome, L. K., and Ed, D. (2018). Original paper application of lev Vygotsky's sociocultural approach to Foster students' understanding and learning performance. *J. Educ. Cult. Stud.* 2, 347–355. doi: 10.22158/jecs.v2n4p347

Herting, M. M., Gautam, P., Chen, Z., Mezher, A., and Vetter, N. C. (2018). Test-retest reliability of longitudinal task-based fMRI: implications for developmental studies. *Dev. Cogn. Neurosci.* 33, 17–26. doi: 10.1016/j.dcn.2017.07.001

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C., and Jantsch, A. (2020). Can p-values be meaningfully interpreted without random sampling?. *Stat. Surv.* 14, 71–91. doi: 10.1214/20-SS12

Hoerger, M., and Currell, C. (2011). "Ethical issues in internet research" in *APA handbook of ethics in psychology, Vol 2: Practice, teaching, and research* eds. S. J. Knapp, M. C. Gottlieb, M. M. Handelsman and L. D. VandeCreek (eds.), (Washington, D.C: American Psychological Association), 385–400 doi: 10.1037/13272-018

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Irigoyen, J., Jiménez, M., and Acuña, K. (2011). Competencias y educación superior. *Rev. Mex. Investig. Educ.* 16, 243–266. Available at: https://www.redalyc.org/pdf/140/14015561011.pdf

Iversen, A., and Pedersen, A. (2017). "Co-creating knowledge: students and teachers together in a field of emergence" in *Co-creation in higher education*. eds. T. Chemi and L. Krogh (Aalborg: Brill Sense), 15–30.

Kaminskiene, L., Žydžiunaite, V., Jurgile, V., and Ponomarenko, T. (2020). Co-creation of learning: a concept analysis. *Eur. J. Contemp. Educ.* 9, 337–349. doi: 10.13187/ejced.2020.2.337

Kang, T., Cohen, A. S., and Sung, H.-J. (2009). Model selection indices for Polytomous items. *Appl. Psychol. Meas.* 33, 499–518. doi: 10.1177/0146621608327800

Lent, R. W., Singley, D., Sheu, H.-B., Schmidt, J. A., and Schmidt, L. C. (2007). Relation of social-cognitive factors to academic satisfaction in engineering students. *J. Career Assess.* 15, 87–97. doi: 10.1177/1069072706294518

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav. Res. Methods* 48, 936–949. doi: 10.3758/s13428-015-0619-7

Lin, M., Lucas, H. C.Jr., and Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Inf. Syst. Res.* 24, 906–917. doi: 10.1287/isre.2013.0480

Lubicz-Nawrocka, T. (2018). From partnership to self-authorship: the benefits of co-creation of the curriculum. *Int. J. Students as Partners* 2, 47–63. doi: 10.15173/ijsap.v2i1.3207

Lystbæk, C. T., Harbo, K., and Hansen, C. H. (2019). Unboxing co-creation with students. *Nord. J. Inf. Lit. High. Educ.* 11, 3–15. doi: 10.15845/noril.v11i1.2613

MacMillan online dictionary. (2020). MacMillan online dictionary. [electronic resource]. Available at: https://www.macmillandictionary.com/.

Maxwell, J. A. (2012). *Qualitative research design: an interactive approach*. California State, US: Sage publications.

Maydeu-Olivares, A. (2013). Goodness-of-fit assessment of item response theory models. *Meas. Interdiscip. Res. Perspect.* 11, 71–101. doi: 10.1080/15366367.2013.831680

Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *J. Appl. Psychol.* 95, 728–743. doi: 10.1037/a0018966

Medrano, L. A., and Pérez, E. (2010). Adaptación de la escala de satisfacción académica a la población universitaria de Córdoba. *Summa Psicológica UST* 7, 5–14. doi: 10.18774/448x.2010.7.117

Merz, M. A., Zarantonello, L., and Grappi, S. (2018). How valuable are your customers in the brand value co-creation process? The development of a customer co-creation value (CCCV) scale. *J. Bus. Res.* 82, 79–89. doi: 10.1016/j.jbusres.2017.08.018

Moshagen, M., and Erdfelder, E. (2016). A new strategy for testing structural equation models. *Struct. Equ. Model. A Multidiscip. J.* 23, 54–60. doi: 10.1080/10705511.2014.950896

Myszkowski, N. (2021). Development of the R library "jrt": automated item response theory procedures for judgment data and their application with the consensual assessment technique. *Psychol. Aesthet. Creat. Arts* 15, 426–438. doi: 10.1037/aca0000287

Oertzen, A.-S., Odekerken-Schröder, G., Brax, S. A., and Mager, B. (2018). Co-creating services—conceptual clarification, forms and outcomes. *J. Serv. Manag.* 29, 641–679. doi: 10.1108/JOSM-03-2017-0067

Rodríguez, H. (2008). Del constructivismo al construccionismo: implicaciones educativas. *Rev. Educ. y Desarro. Soc.* 2, 71–89. Available at: https://dialnet.unirioja.es/servlet/articulo?codigo=2692738

RStudio Team. (2022). *RStudio: Integrated development for R*. RStudio, PBC, Boston, MA.

Ryan, A., and Tilbury, D. (2013). Flexible pedagogies: New pedagogical ideas. 41. Available at: http://www.heacademy.ac.uk/resources/detail/flexible-learning/flexiblepedagogies/new_ped_ideas/report?utm_medium=email&utm_source=The+Higher+Education+Academy&utm_campaign=4074096_140506&utm_content=New-pedagogical-ideas-report.

Samejima, F. (1997). "Graded response model" in *Handbook of modern item response theory*, van der Linden, W.J., Hambleton, R.K. (eds) (New York: Springer), 85–100. doi: 10.1007/978-1-4757-2691-6_5

Schlesinger, W., Cervera, A., and Iniesta, M. Á. (2015). Key elements in building relationships in the higher education services context. *J. Promot. Manag.* 21, 475–491. doi: 10.1080/10496491.2015.1051403

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Serrano, J. M., and Pons, R. M. (2011). El constructivismo hoy: enfoques constructivistas en educación. *Rev. electrónica Investig. Educ.* 13, 1–27. Available at: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1607-40412011000100001

Singh, J., Steele, K., and Singh, L. (2021). Combining the best of online and face-to-face learning: hybrid and blended learning approach for COVID-19, post vaccine, & post-pandemic world. *J. Educ. Technol. Syst.* 50, 140–171. doi: 10.1177/00472395211047865

Taghizadeh, S. K., Jayaraman, K., Ismail, I., and Rahman, S. A. (2016). Scale development and validation for DART model of value co-creation process on innovation strategy. *J. Bus. Ind. Mark.* 31, 24–35. doi: 10.1108/JBIM-02-2014-0033

Tarı, B., and Mercan, H. (2020). Co-creating positive outcomes in higher education: are students ready for co-creation? *J. Mark. High. Educ.* 32, 73–88. doi: 10.1080/08841241.2020.1825031

Tilak, J. B. G., and Kumar, A. G. (2022). Policy changes in global higher education: what lessons do we learn from the COVID-19 pandemic? *High. Educ. Policy* 35, 610–628. doi: 10.1057/s41307-022-00266-0

United Nations. (2020). Policy brief: Education during COVID-19 and beyond. Available at: https://unsdg.un.org/resources/policy-brief-impact-covid-19-children.

Ventura-León, J. (2021a). Instrumentos breves: Un método para validar el contenido de los ítems. *Andes Pediatr.* 92:812. doi: 10.32641/andespediatr.v92i5.3961

Ventura-León, J. (2021b). Una misma talla para todo: Repensando los tamaños del efecto de Cohen. *Educ. Médica* 22:445. doi: 10.1016/j.edumed.2020.07.002

Ventura-León, J., Caycho-Rodríguez, T., Mamani-Poma, J., Rodriguez-Dominguez, L., and Cabrera-Toledo, L. (2022a). Satisfaction towards virtual courses: development and validation of a short measure in COVID-19 times. *Heliyon* 8:e10311. doi: 10.1016/j.heliyon.2022.e10311

Ventura-León, J., Caycho-Rodríguez, T., Sánchez-Villena, A. R., Peña-Calero, B. N., and Sánchez-Rosas, J. (2022b). Academic inspiration: development and validation of an instrument in higher education. *Electron. J. Res. Educ. Psychol.* 20, 635–660. doi: 10.25115/ejrep.v20i58.5599

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., et al. (2020). ggplot2: Create elegant data Visualisations using the grammar of graphics (version 3.3. 0)[computer software]. Retrievd from https://CRAN.R-project.org/package=ggplot2.

Wickham, H., François, R., Henry, L., and Müller, K. (2021). Dplyr: A grammar of data manipulation. R package version 1.0.7. Available at: https://cran.r-project.org/web/packages/dplyr/dplyr.pdf.

Yamada, G., Castro, J., and Rivera, M. (2012). *Educación Superior en el Perú: Retos para el Aseguramiento de la Calidad*. Lima: Sistema Nacional de Evaluación, Acreditación y Certificación de la Calidad Educativa.

Yi, Y., and Gong, T. (2013). Customer value co-creation behavior: scale development and validation. *J. Bus. Res.* 66, 1279–1284. doi: 10.1016/j.jbusres.2012.02.026

Yin, H., and Wang, W. (2016). Undergraduate students' motivation and engagement in China: an exploratory study. *Assess. Eval. High. Educ.* 41, 601–621. doi: 10.1080/02602938.2015.1037240

Zickar, M. J., and Broadfoot, A. A. (2009). The partial revival of a dead horse? Comparing classical test theory and item response theory. *Stat. Methodol. Myth. urban Legend. Doctrin. Verit. fable Organ. Soc. Sci.*, 37–59.

Ziegler, M., Poropat, A., and Mell, J. (2014). Does the length of a questionnaire matter?: expected and unexpected answers from generalizability theory. *J. Individ. Differ.* 35, 250–261. doi: 10.1027/1614-0001/A000147

Appendix A

Academic co-creation short scale (AC-S).

Instrucciones: A continuación, encontrarás un conjunto de preguntas acerca de tu aprendizaje en la UNIVERSIDAD.

Pocas veces es mi caso	A veces es mi caso	Muchas veces es mi caso	Con mucha frecuencia o siempre es mi caso			
1	2	3	4			
Preguntas						
Yo aprendo mejor…						
1. …intercambiando ideas con mis compañeros de aula.			1	2	3	4
2. …colaborando en proyectos grupales con mis compañeros de aula.			1	2	3	4
3. …aportando ideas en los trabajos grupales con mis compañeros de aula.			1	2	3	4
4. …escuchando activamente a mis compañeros de aula en los trabajos grupales			1	2	3	4