

Elaboration of quality perception instrument of remote teaching amidst COVID-19 pandemics in a University of Northern Brazil.

Rafael da Silva Fernandes (✉ rafasfer2@ufra.edu.br)

Federal Rural University of the Amazon <https://orcid.org/0000-0002-3035-8025>

Rosana Maria do Nascimento Luz

Federal Rural University of the Amazon <https://orcid.org/0000-0003-2650-4875>

Daniela Castro dos Reis

Federal Rural University of the Amazon <https://orcid.org/0000-0002-9505-4516>

Marcos Augusto Lima da Luz

Federal Rural University of the Amazon <https://orcid.org/0000-0002-7076-0633>

Gabriella Vitorino Guimarães

Federal Rural University of the Amazon

Research Article

Keywords: Psychometric assessment, Teaching assessment, Measuring Instrument, Student perspectives, Quality of teaching

Posted Date: January 31st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1308160/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

COVID-19 (SARS-CoV-2) pandemics changed social dynamics in several aspects, the introduction of containment measures to ensure social distancing was one of these changes, it made education institutions adopt the remote learning system. In order to understand the impact on the learning process, this paper aims at devising and developing a measurement instrument to quantify students' perception regarding the adaptation to the remote education provided by the Federal University of Amazonas – UFRA, over classroom-based learning, which was adopted previously to the COVID-19 context, that being the latent trait to be measured. For this purpose, theoretical procedures from the Construction Model for Psychological Tests and Measures and the psychological evaluation incorporated multiple perspectives. To validate the content, the study set three phases: 1) initial construction of the measurement instrument and literature review; 2) establishment of student focus groups; 3) judge's analyses to outline factors and potential items. Finally, the instrument presented to the content validity index an average acceptance of 94.16% per item, and acceptance values of over 93.43% to factors and domains.

Introduction

COVID-19 pandemics produced changes in social interactions and social distancing intended on reducing these interactions amongst people, causing education institutions to adopt the emergency remote teaching format. In this national higher education context, a survey carried out by the College of Deans of Undergraduate Studies (COGRAD-2020), linked to the National Association of Directors in Federal Institutions of Higher Education (ANDIFES), demonstrated that most of the IFES (Federal Institutions of Higher Education), planned complementary academic calendars that were completely remote or hybrid (Aguilera-Hermida, 2020; Ali, 2020; Buttler et al., 2021a; Chiou, 2020; Sarah I Hofer et al., 2021; Secundo et al., 2021a).

Due to the long duration of the suspension of in-person activities, the proper fulfillment of the academic calendar for the years 2020.1 and 2020.2 was disrupted. This situation became even worse because of the uncertainty regarding the resumption of academic in-person activities on the IFES. In addition, on April 28th 2020, CNE/MEC (National Council of Education) approved the opinion N° 05, which covers the reorganization of the school calendar and the possibility of validation of non-face-to-face activities in order to meet the minimum annual class load.

Another aspect that required readjusting was the evaluation process, since it was conducted contemplating classroom teaching characteristics, the search for more precise indicators became necessary on the remote teaching assessment.

In the COVID-19 context, loss of teaching quality is projected due to the conditions for carrying out a semester entirely online in 2020. The present study aimed at devising and developing a measurement instrument to quantify students' perception regarding the adaptation to the remote education provided by the Federal University of Amazonas – UFRA, over classroom-based learning.

Psychometrics can be understood as a theory or technique to assess psychological processes, usually applied in the areas of Psychology and Education (Pasquali, 2009). In this paper, the construct to be assessed is the perception of the students which represents a latent trait, such as its objectives, we have the quality evaluation of emergency remote teaching in Brazilian universities and the explanation of the meaning that individuals bestow to the answers on a variety of tasks through items. Thus, we expect psychometrics to assist the comprehension of mental processes through the quantitative method, which is an important aspect in science when seeking to depict the understanding of nature with greater precision.

The methodology applied was based on the creation of new instruments revised by (Coluci et al., 2015) and on mental health assessment principles (Gorenstein & Yuan-Pang, 2016), which present the theoretical procedures of design and construction of the psychological measurement proposed in (Pasquali, 2010). It is based on three poles: theoretical procedures, empirical procedures (experimental) and analytical procedures (statistical).

It is worth noticing that the multi-criteria proposed by (Nickel et al., 2010) was applied as complementary basis to the construction of the measuring instrument. Hence, it is necessary the identification of the requirements/needs of the students, definition of the requirements to compose the instrument and the definition of the goal-specifications that relate to the validation of the instrument by specialists. Finally, we sought to create an assessment instrument for the emergency remote teaching to validate the quality of this teaching and also, subsidize the various assessments and educational examinations in a large scale, such as: Basic Education System (SAEB), National Exam of Students Performance (ENADE), External on Site Evaluation of Higher Education Institutions, National High School Exam (ENEM) and Programme for International Student Assessment (PISA). Next, the bibliographic review is going to be demonstrated, the methodological steps to the construction and validation of the perception instrument about the quality of remote learning.

Literature review

Some studies about remote teaching were performed, among them the practice of psychological assessments, reinforcing the importance and need of researches that demonstrate accuracy, ethics and safety (Marasca et al., 2020), the investigation of students' perception about the emergency online teaching (adoption, use and acceptance) and its quality (Aguilera-Hrmida, 2020; Sarah I. Hofer et al., 2021; Iglesias-Pradas et al., 2021; Moser et al., 2021; Secundo et al., 2021b; Shim & Lee, 2020) . Other articles indicate the psychological impacts experienced by higher education students in the new reality which is a resultant of the fear of COVID-19 (Buttler et al., 2021b; Cavalheiro & Sticca, 2020; Faro et al., 2020; Shim & Lee, 2020).

There were few publications that analyzed the emergency remote teaching in Brazilian Universities (De la Fuente et al., 2021; Quispe-Pietro et al., 2021; Ribeiro & Corrêa, 2021), this can be based on two assumptions: 1) It is the first time the universities are experiencing this type of drastic change on teaching setup, called Emergency Remote Teaching (ERT), which differs from the Distance Education, offering a temporary instructional support and developed with a reduced planning time (Hodges et al., 2020; Mohammed et al., 2020; Whittle et al., 2020). 2) It is observed that the research focused on the impact of ERT, and directed to the Education literature, has not used psychometric procedures with an approach focused on latent trait theory and without validation performed with the support of mental health measurement fundamentals.

Moreover, papers or research measuring the quality of the adjustment to the remote teaching in the country were not identified, neither content where the methodology that had aspects related to psychometrics, method applied only in other countries using structural equation modeling (Buttler et al., 2021b). Therefore, the studies cited provided a bibliometric longitudinal framework, with themes related to: psychometrics, measurement, construction of measurement instruments and their validation, latent trait, quality of remote teaching and COVID-19, which finally supported and guided the construction of a measurement instrument with validation for this research.

Present study

Practical motivating situation

The concern and relevance relative to the exceptional moment the world is experiencing in the face of COVID-19 pandemics reinforces the finding that it is necessary for universities to create alternatives in all the sectors they operate.

Thus, after the suspension of the Academic Calendar (AC) 2020.1 on March 19, 2020 and aiming on finding alternatives to face the situation, the Federal University of Amazonas – UFRA, located in the Northern region of Brazil, with its head office in Belém/PA, and five other campuses located in the cities of Parauapebas/PA, Paragominas/PA, Tomé-Açu/PA, Capitão Poço/PA and Capanema/PA, started the planning process for teaching.

Through Technical Note No. 01/2020 - PROEN/UFRA, dated March 26, 2020, the institution encouraged teachers to perform academic activities through the SIGAA platform, in order to maintain virtual contact with their students, such as the recommendation of topics for reading and activities.

Considering Resolution No. 567-Consepe/2020, of July 1, 2020, the UFRA, through Ordinance No. 943 of July 6, 2020, created a special committee. This committee proposed an alternative academic calendar, called Supplementary Academic Period (SAP) offering curricular components in emergency and temporary digital format, most importantly, seeking to keep the university active, notably from an

undergraduate teaching perspective, where students and professors would maintain social, cognitive, and emotional interaction.

The Supplementary School Period (PLS - 2020.5), consisted of a reduced period of 8.5 weeks, which intended on minimizing the impacts of the long duration of the suspension of academic activities due to the COVID-19 pandemic, enabling the provision of curricular components and other academic activities in non-face-to-face format, synchronously or asynchronously, employing information and communication technologies (ICT). It is emphasized that PLS – 2020.5 did not reduce the load of curricular components but relaxed the shift and the offer period. The evaluation procedures are part of a branch that aims on evaluating the knowledge acquired by the students throughout their education. In contrast, subjective aspects and several factors directly impact the student's learning beyond the course's own knowledge, particularly when concerning higher education.

Over the years, some factors have been investigated, such as reasons relating to evasion, social and economic impacts, home study environments, amongst others associated to the student's experience in the academic environment (Ambiel, 2015; Buttler et al., 2021b; Newton, 2017; Sindhu et al., 2019).

Latent trait measured and assessment design

This study analyzes the students' perspective and constitutes a property for evaluating the quality of education. Specifically, the object or psychological system to be represented is the latent trait "perception of students regarding the adequacy to remote teaching offered by the Federal University of Amazonas (UFRA), over the face-to-face teaching adopted before the COVID-19 pandemics context". The properties or attributes of an object are the several aspects that characterize it and must be defined to the conception of an instrument.

Thus, given the latent trait, this can be measured by the adequacy attribute, and it is associated to a negative perception, neutral or positive, considering the following aspects of social and academic systems regarding: class dynamics, its own adjustment to remote teaching, structure and regulations of remote teaching, its own resources, study environment and external incentives, and finally, their own effort, relationships and personal motivations. As an example, the following model is presented in Figure 1:

This psychological attribute can be understood as a meter of latent trait of the evaluation of quality and adequacy of the remote teaching offered by the UFRA, which, in turn, it is not something directly observable or identifiable. Hence, it arises the need for the elaboration of a construct that manifests the reality observed indirectly, through other variables or factors that can be observed and are related to the subject of interest to be measured (Pasquali, 1997).

The adequacy concept or perception of adequacy, in the paper, refers to the idea of a search for criteria that measure the similarity between its knowledge and a corresponding object. Consequently, considering the remote teaching environment, in which the student is inserted and incorporating the academic and social integration aspects of the longitudinal model of (Tinto, 1975), the hypothesis of the conceptual model proposed (Figure 1) is that there is a set of factors that exert direct influence in this environment, and that, in turn, generates a negative perception, neutral or positive perception associated with the quality of remote teaching. Thus, the evaluation of the student, given the impact of these factors, will produce a worse adequacy perception, equivalent or better than face-to-face teaching, aiming at measuring the adequacy quality of the emergency remote teaching.

Based on the fact that the students' perceptions about the educational environment influence directly their learning, (Roff et al., 1997) define educational "environment" or "climate" as any learning experience that influences the motivation of students to learn, affecting their attitudes, values and behaviors related to a learning task. Hence, measuring teaching quality from the students' perspective corroborates as justification and contribution of this paper to the literature.

It is known from (Burlison & Thoron, 2014), that the premises of a good educational environment must include base-factors that are able to meet the following aspects: 1) the physical needs of students and associates regarding ergonomic conditions; 2) students' safety and trust regarding the handling of the teaching/learning activities; 3) the engagement, the feeling of belonging in their development process as humans and the importance of performing professional activities to promote social, political, economic, cultural and environmental changes during the school term; 4) the self-esteem, linked to the importance of providing constructive feedbacks, given the context of the COVID-19 pandemics.

Multiple perspectives

Within this framework, and prior to the determination of the facts that compose this construct, the main motivation of this evaluation and measurement comes from the reports of students and teachers regarding the challenges of adjustment to the remote teaching methods applied. This exemplifies the importance of considering multiple perspectives on the evaluation purposes, thus, aiming to explicitly distinguish the perspectives, and in order to help the design of the evaluation (Newton, 2017), improve the way of designing the performance evaluation on the premises of jointly assess the perspectives of the concerned parties. Hence, it is possible to classify these perspectives in three different types, which in this paper are presented as: teaching perspective, adequacy perspective and engagement perspective.

This way, since its pure intent idealizes the principle that the *design* of the measurement instrument must be guided by a single measurement purpose, that is, solely and exclusively to measure the latent trait, we have in this unity a test *design* that begins with the consideration of the expected interpretations to the intended uses of the scores obtained by the tests. The content and form of the test are then specified to provide evidence and support interpretations for the intended uses. When the scores of a single test are interpreted differently, to make different types of decisions, it is admitted that each intended interpretation

must be independently validated (Association, 2018). In other words, a single perspective of the evaluation purposes is informed, and from a single point of perspective, their results are applied for different purposes.

The concept of multiple perspectives about the objectives of evaluation guides the test design to the intended uses of the scores, however it does not explicitly discuss how decision-oriented and content-oriented approaches intersect (Newton, 2017). The premise is that the instrument needs to provide information from multiple perspectives, in other words, the students have different perspectives, hence, they are capable of generating information that can be used by them, the teachers and the institution. The flowchart presented in Figure 2 exemplifies the relationship between multiple perspectives and factors associated with the latent trait.

The first perspective focuses on teaching and using the evaluation results in order to support, for example, the regulatory committee with reliable information, backed by quantitative methods and validated by evidences that aim on supporting the interpretation of the scores obtained from the tests. The *designer* is focused on the experience that must be attained by the students regarding their academic integration.

The second perspective focuses on the adequacy of the structure, regulations and technological resources of the UFRA, which is then, perceived by the students. The premise is that several students needed to adjust to the remote format, and by implication, derives from an institutional integration and its determinations. The *designer* is focused on the experience acquired by the students while adjusting to remote teaching, and from this perspective, the problem of how to specify the evaluation construct arises from the question of how to determine specific answers, that is, the instrument needs to minimize disturbances from other answers, and, most importantly, not to confuse aspects of the teaching perspective.

Finally, the third perspective of engagement focuses on the motivational consequences of the adequacy to remote teaching. The engagement operates mostly through social integration and the student's agreement on being evaluated in terms of this new teaching structure. Therefore, it is assumed that the current situation of the COVID-19 pandemics tends to motivate students to study harder, or at least raise the comprehension of the need in keeping education active, that otherwise would have dedicated themselves to achieving learning results.

The *designer* is focused on capturing the engagement of the students and, given social and career aspects, the instrument needs to provide *insights* and some type of self-assessment of the student. So, the construct refers to the search of a set of answers that relates to the context of the pandemics, that is, the student will respond not as a direct comparison between teaching formats but context weighted.

According (Newton, 2017) researches that focus on performance evaluations and question the persistency of the purism of purpose need the coordination of the multiple perspectives of evaluation. In

this sense, the construct can be defined and thus, determined as theoretical bases of the factors that compose it.

Finally, it is necessary to conceptualize in detail the construct, which will be done under two products: constitutive definitions and operational definitions. On the constitutive definition, it is understood that all the theoretical foundation, since the definition of the psychological object and the base-factors along with the three presented perspectives, provided subsidies for the creation of a conceptual model and aims at theoretical and abstract understanding, as well as the semantic validation of the instrument.

From an operational point of view, many instruments are available to assesses the educational environment, such as, in the field of medical sciences and related fields it is commonly evaluated applying DREEM (Roff et al., 1997) which reflects the overall construction of the teaching environment. Hence, given the specificity of the context and a target audience originated from a single higher education institution, this paper presents a procedure for the operationalization of the measurement instrument.

Research Methodology

Operationalization and validation procedure

The procedure of creating, adapting and validating the items was based on the multi-criteria model (Nickel et al., 2010). The step of operationalization and analysis of the items are described in detail as follows:

- Step 01: Elaboration and initial description of the items: a set of items is elaborated based solely on experiences and suggestions of the authors of this article. This process refers to the identification of the requirements of the customers of the product, that is, it is related to the compilation of the characterization needs of the multiple perspectives, factors, and latent trait. It is worth noticing that those needs were somehow mentioned or indicated by students and teachers throughout remote teaching terms.
- Step 02: Grouping and classification of items: refers to the distribution of the needs in primary elements of factor specification, that is, the process of directing the items to the respective factor. When there is a larger set of items presented, grouping and classifying those needs is useful for verifying similar items, eliminating repetitions and irrelevant needs.
- Step 03: Definition and hierarchization of factors: refers to the understanding of the researchers of this study regarding the criteria of clarity, relevance or representativeness and scope and it is called internal validation. Thus, through a Value tree, it is possible to determine the necessary requirements for the factors in order to have a deeper and more accurate understanding of what should be of concern in this situation.
- Step 04: Target question and conversion of items: as a simple explanation for intuitive items, when a satisfactory answer to a difficult question is not quickly found, it is necessary to establish a target question that is the evaluation that the research tends to produce and turn it into an adequacy (or

empirical) question that, besides its framing to the answer's category, it is a simpler question that is answered as a substitute of the target question. The process of conversion of the items, after grouping and sorting, obtains clearer and simpler descriptions, that is, measurable questions or expressions. In addition, this conversion consists of a set of objectives, each of which can be seen as a means to an end or the cause of an effect.

- Step 05: External validation of the items: after the definition and hierarchization of the items, it was necessary its external validation, which was performed in three separate ways:
 1. Literature review of the validated items in other measurement instruments and that, following previous steps, can be incorporated into the pilot measurement instrument.
 2. Execution of rounds of focus groups, with the participation of about 5 students from the institution, in order to comprehend the difficulty of understanding the items by the target interviewees, that is, the items ask or empirically approach the target question.
 3. Application of a form to more than five specialists with working experience in the education and/or psychometrics area, in order to size the minimum criteria required for the set of items of each factor of the measurement instrument. The domains are initially verified, considering clarity, relevance or representativeness and scope.
 4. Next, each of the factors and items are evaluated according to the same criteria to verify whether each domain is adequately covered by the set of factors and items. In addition, we consider the evaluation of the organizational structure of the instrument, the aspects of social and academic integration and impact on the student's academic environment.
- Step 06: Pre-test procedure: The pre-test intends on verifying that all items are comprehensible for a small sample and not necessarily representative of the population to which the instrument is intended. It also aims at capturing information regarding the necessary time to answer.

At the end of this activity, there is a pilot instrument that can be applied to the entire student population and will be ready to have its psychometric properties evaluated. The first and last procedures need to follow the order as described above, however the other ones were not executed in the described order.

Technical issues

Typically, it is expected that in the construction of the constructs, the suggested items present heuristics and biases that can create validation problems of the final instrument. Thus, we present three important technical problems to understand the framing of this process. As presented by (Kahneman, 2012, pp. 165-185) and arising from an *availability heuristic*, the first problem refers to the ease with which occurrences come to mind, which in turn, can hide important items, though not initially suggested.

Another problem, that implicated on the initial proposition of the instrument and directly impacted the regulation of the 2020.1 remote teaching, is the WYSIATI Principle (*What You See Is All There Is*).

As several students and professors presented difficulties adapting to the remote teaching, they generated a greater number of complaints and requests to adjust the remote teaching model. Thus, the dissatisfaction or negative perception about the teaching method presented a representation bias, that is, everything that was seen as information to regulation, largely came from a negative perception. Furthermore, this problem implies a false sense of causality, when in fact the general idea about the students' perceptions is the result of a bias of representativeness and/or availability of information. The concern with this situation was used as a criterion for not using the attribute "satisfaction" and choosing the attribute in question.

Finally, there is still a third problem known as the *Priming Effect* and it is a suggestion that selectively evokes compatible evidence. Therefore, both the instrument of measure and the regulation itself may have been influenced by the suggestions of one or more people. For example, the regulation was strongly influenced by teachers, as they reported the dissatisfactions of their students which evoked negative indications from other teachers. This situation may have caused a difficulty in approving the academic calendar, since there is a latent pressure of condensing the calendar, that is, students from the last terms of their courses are the most interested in condensing the academic calendar, mainly due to their concerns about graduation, exerting greater representation influence and evoking similar feelings on students in the beginning of their studies. This situation resulted in a high amount of content in a shorter time and, consequently, in several dissatisfactions from beginner students.

Those problems embody great importance in the theoretical validation of this measurement instrument, that is, it is necessary to resist this large set of potential availability biases, which require a great effort to reconsider the initial impressions and intuitions. At first it may appear it is considered a process of standard scientific method or at least required in research, however, through literature review, a set of studies that constantly presented several biases was identified.

Results And Discussion

Step 01: Elaboration and Initial Description of the Items

In a preliminary way, we sought to elaborate an instrument of measurement, based on experiences collected through reports in the remote teaching of the school term PLS - 2020.5 and PL - 2020.1 and/or discussions held by the Regulatory Committee of Remote Education of UFRA.

Therefore, this work, in addition to understanding the potential biases of the construction of the items, initially provided a set of 24 items, which were elaborated following the diagram in Figure 3. It is worth noting that in the diagram there is a total of 41 items, the additional of which comes from the application of item (i) step 05.

Step 02: Grouping and Classification of Items

Note that the diagram in Figure 3 covers the process of grouping and classification of items. This step, by checking similar items, eliminating repetitions and irrelevant needs, contributes by minimizing the problem originated from the availability heuristic. Moreover, it is observed in some items the specification of comparison between elements of face-to-face teaching and elements of remote teaching, for example, the item "Own Tools" on Factor 04 refers to an adequacy and/or comparison between the use of notebooks, printed materials and/or computer labs in face-to-face teaching over the adequacy and use of notebooks, smartphones, and private internet connection in the remote teaching. So, a response that adapted much worse or worse in this item is associated with an average of other associated aspects, namely, about the quality of this resource (NTB / cellular) and on the quality of their own internet connection.

Consequently, this grouping and classification process conducts a discussion and determines if an item should be kept, altered, divided, or merged. This analysis, along with step 05 (i), analyzes item-to-item relationships with its factor and associated domains.

Step 03: Definition and hierarchization of factors

The definition of the construct requirements, particularly, the minimum requirements that each factor should have, were obtained, and presented in Figure 4. Note that the requirements are a direct derivation of the factors and domains and constitute relevant aspects of an academic environment.

The process of hierarchization of the requirements in this study is only the indication of each item to its respective requirement. To exemplify the contribution of this process, we have item 32, contained in Factor 4 and described as "Dedication to evaluative activities (Face-to-face versus Remote)", being indicated more strongly by requirement 4 "Concentration/Attention".

The tree-shaped structuring helped clarify the decision context, validation, and representativeness determination for each topic of the content, that is, it becomes possible to evaluate whether the item really belongs to one factor or the other.

Step 04: Target question and conversion of items

After grouping and defining the requirements of the items, these can be called item core and represent the essence that the final question of the questionnaire should have. Hence, the workflow for converting items to the response categories is presented in Figure 5, exemplifying factor 1.

It is observed that the flow begins in the initial elaboration of the items, description of the nucleus, which is transformed into one or more target questions and, finally, converted to the category of answers. This process counted on 91 target questions, which are somehow being represented by the 41 items to be validated.

Finally, it is important to emphasize that this step is a heuristic and it is a simple procedure that helps to find appropriate answers, even though usually flawed, to difficult questions (Kahneman, 2012).

Step 05: External validation of the items

Literature Review

The literature review associated 32 items validated in other measurement instruments, to the 41 items of this instrument. This review also counted on over 20 papers that presented instruments, in addition to what had already been cited in this work.

Focus Group

For the focus group, the study recruited 3 to 5 students for each of the 3 focus groups held at different times. The intentional sampling method with indication of teachers from other campuses was used based on four criteria, namely: (1) each group will be composed of students from different campuses and courses; (2) the participant is an active student of UFRA and is enrolled in a discipline offered in Emergency Remote Teaching; (3) the participant is a student with previous experience on learning through face-to-face teaching according to the formal curriculum; and (4) ideally, students with lower academic performance index (API) will be preferred, on the premise that they will present greater difficulties in understanding the items to be analyzed and, consequently, they will provide better information regarding the purpose of the focus group.

The items discussed after each one of the first two focus groups were readjusted, i.e., focal group 02 will have a set of items already validated and altered according to the considerations of focus group 01. This procedure is presented in Figure 6.

It is important highlighting that after holding the focus groups, in addition to the modifications made to the construct items, this process allowed a visualization of implicit target questions incorporated. Therefore, after this process there was an internal revalidation and a new adjustment of the items that compose each factor.

Judges Validation

Invitations were sent to 10 specialists, obtaining 7 responses. Recruitment was conducted nominating specialists with working experience in the education and/or psychometrics area. The procedure for judging the validity of the content was structured and delivered to the judges on a form to evaluate the instrument considering clarity, relevance or representativeness and scope.

Regarding the representation diagram of the multiple perspectives, the content, representativeness, and influence/impact of each factor, all related to Figure 1, presented 100% agreement of the judges. On the question that the factors represent aspects of social and academic integration, there were three disagreements in factor 03, being in fact understood and justified, because this factor intends on representing an institutional integration.

Regarding the scope, representativeness of the domains and how much each factor expresses the content of the domains, all related to Figure 2, there was 100% agreement of the judges. When asking whether each domain factor should be kept, altered, divided, or merged, for factors 3, 4 and 5, 100% believe that it should be kept. As for factors 01 and 02, there was a suggestion for them to be altered and/or divided, but without comments from the judges regarding the changes. In this case, taking into account comments from other judges that the subdivision could make it more difficult to frame the issues and reviewing the literature, we consider that factors 01 and 02 should be kept.

In a second item by item analysis of the construct, we are interested in quantifying the content validity index. The response rate for the 5 questions asked to the judges was an average of 99.65% expressiveness of the core item, 99.30% of representativeness of the content of the final question, 95.82% that the item should continue the way it was presented to the judges, 82.58% it is clear and understandable and 93.43% that the items (through their final question) are representative and relevant to the concept explored. Rates are expressed in Table 1. The items were considered a 4-point ordinal likert scale for the criteria presented, and answers 3 and 4 were acceptable and indicative that the item should continue in the measurement instrument and answers 1 and 2 as an indication for the investigation of the item. In addition, the content validity index was calculated by taking the average of each set of answers from Table 1, obtaining as the item with the worst index the items 09 and 10 with 85.71%. According to the comments, disagreement is associated with an alteration and not the removal of these items, however, this is an aspect that was validated by the students in the focus group phase and will be maintained as presented. Rewriting was explicitly suggested only for item 19 of the Factor 3 and it was adopted for application and final validation of the pre-test.

Table 1 - Item content validity analysis.

Factors	Items	Does every core factor express the content?	Does every item (final question) express the content of the core factor?	Does every item (final question) of the factor must be kept, altered, divided or merged?	Is every item (final question) of the factor evident? Comprehensible?	Is every item (final question) of the factor representative to the explored concept? Is it relevant?
Factor 1	Item 01	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 02	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 03	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 04	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 05	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 06	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 07	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 08	100,00%	100,00%	100,00%	71,43%	85,71%
	Item 09	85,71%	85,71%	100,00%	71,43%	85,71%
	Item 10	100,00%	85,71%	85,71%	71,43%	85,71%
Factor 2	Item 11	100,00%	100,00%	71,43%	100,00%	100,00%
	Item 12	100,00%	100,00%	71,43%	85,71%	100,00%
	Item 13	100,00%	100,00%	85,71%	85,71%	100,00%
	Item 14	100,00%	100,00%	85,71%	85,71%	100,00%
	Item 15	100,00%	100,00%	85,71%	85,71%	100,00%
	Item 16	100,00%	100,00%	85,71%	100,00%	100,00%
	Item 17	100,00%	100,00%	85,71%	85,71%	100,00%
Factor 3	Item 18	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 19	100,00%	100,00%	71,43%	71,43%	100,00%
	Item 20	100,00%	100,00%	100,00%	85,71%	87,71%
	Item 21	100,00%	100,00%	100,00%	85,71%	100,00%

	Item 22	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 23	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 24	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 25	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 26	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 27	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 28	100,00%	100,00%	100,00%	85,71%	100,00%
Factor 4	Item 29	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 30	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 31	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 32	100,00%	100,00%	100,00%	85,71%	100,00%
	Item 33	100,00%	100,00%	100,00%	85,71%	100,00%
Factor 5	Item 34	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 35	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 36	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 37	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 38	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 39	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 40	100,00%	100,00%	100,00%	85,71%	85,71%
	Item 41	100,00%	100,00%	100,00%	85,71%	85,71%

Step 06 – Pre-test Procedure

The form was applied online, in order to facilitate access and participation of the people invited, it was sent through a link, generated through two free tools offered by Google: Google Forms and Google Meet.

Thirty-seven students attended this step. The survey was available to be filled out for a stipulated time of 30 minutes with the assistance of the mediators via Google Meet, and later the students were interviewed.

We sought to observe the difficulties of understanding the measurement instrument as well as verifying if there could still be any alteration of the instrument. At the end of the process, it is reported that the students understood the purpose of the instrument and raised questions observed in the survey. In summary, the measurement instrument was consistent, and some of the expected responses were confirmed, even though obtained from a non-representative sample.

Finally, in order to obtain a more universal answer, in addition to the 41 items, 3 questions about the domain and 5 more questions about the factors were added, compiling a total of 49 items in the pilot instrument. As an example, right after the 10 items referring to Factor 01, the following question was asked: "As for the dynamics of the classes (teacher + teaching), you consider:". Thus, it is possible to evaluate whether the factor is sufficiently adequate by the items that compose it.

Conclusions

Considering the objectives of this study, which were elaborating and developing a measurement instrument to measure the perception of students and demonstrating the methodological steps for the construction and content validation of the instrument of perception about the quality of remote teaching, it can be said that the goals were achieved in a way that could contribute to the evaluation of the teaching quality using psychometric measures.

These methodological steps are ways of building this type of instrument to assist the measurement of psychological constructs, especially the perception of the students from the Federal University in the Amazon region. These steps can contribute to the construction of new instruments in other university or regions.

It can be said that this research is an attempt on apprehending some aspects that concern remote teaching. Regarding the limitations, it can be reported that some items for the evaluation of remote teaching were not contemplated, aspects that can be referenced in other studies with the same object of study.

As a perspective of new research, it is indicated that the instrument is improved, following all the steps described in this manuscript, inserting new evaluation items. It is also suggested the application of the instrument to a number of university students in a way that includes more regions and more universities. This attempt is an essential step for universities to build scientifically validated indicators through psychometrics.

Declarations

Acknowledgements

The authors thank the Federal Rural University of Amazonas for assisting the preparation of this manuscript, to the students involved in the focal groups and the execution of the pre-test and the judges for their contributions.

Disclosure statement

This study was approved by the National Research Ethics Committee (CONEP – Comissão Nacional de Ética em Pesquisa) and all parties signed the Free Written Informed Consent to the terms.

Competing Interests

The authors declare that they have no competing interests.

References

- Aguilera-Hermida, A. P. (2020). College students' use and acceptance of emergency online learning due to COVID-19. *International Journal of Educational Research Open*, 1, 100011.
- Aguilera-Hrmida, A. P. (2020). College students' use and acceptance of emergency online learning due to Covid-19. *International Journal of Educational Research Open*, 1, e100011.
- Ali, W. (2020). Online and remote learning in higher education institutes: A necessity in light of COVID-19 pandemic. *Higher education studies*, 10(3), 16-25.
- Ambiel, R. A. M. (2015). Construção da escala de motivos para evasão do ensino superior. *Avaliação Psicológica*, 14(1), 41-52.
- Association, A. E. R. (2018). *Standards for educational and psychological testing*. American Educational Research Association.
- Burleson, S. E., & Thoron, A. C. (2014). Maslow's Hierarchy of Needs and Its Relation to Learning and Achievement. *Agricultural Education and Communication Department, UF/IFAS Extension*, 3.
- Buttler, T., George, D., & Bruggemann, K. (2021a). Student input on the effectiveness of the shift to emergency remote teaching due to the COVID crisis: Structural equation modeling creates a more complete picture. *International Journal of Educational Research Open*, 2, 100036.
- Buttler, T., George, D., & Bruggemann, K. (2021b). Student input on the effectiveness of the shift to emergency remote teaching due to the COVID crisis: Structural equation modeling creates a more complete picture. *International Journal of Educational Research Open*, 2(2), e100036.
- Cavalheiro, F. R. S., & Sticca, M. G. (2020). Adaptation and Validation of the Brazilian Version of the Fear of COVID-19 Scale. *Int J Ment Health Addiction*.

- Chiou, P. Z. (2020). Learning cytology in times of pandemic: an educational institutional experience with remote teaching. *Journal of the American Society of Cytopathology*, 9(6), 579-585.
- Coluci, M. Z. O., Alexandre, N. M. C., & Milani, D. (2015). Construção de instrumentos de medida na área de saúde. *Ciência & Saúde Coletiva*, 20(3), 925-936.
- De la Fuente, C. I., Guadagnin, E. C., Kunzier, M. R., & Carpes, F. P. (2021). Programming course for health science as a strategy to engage students during the coronavirus pandemic. *Advances in Physiology Education*, 45, 53-58.
- Faro, A., Silva, L. d. S., Santos, D. N. d., & Feitosa, A. L. B. (2020). Adaptation and validation of The Fear of COVID-19 Scale.
- Gorenstein, C., & Yuan-Pang, W. (2016). Fundamentos de Mensuração em Saúde Mental (*Instrumentos de Avaliação em Saúde Mental* (pp. 1-30). Artmed.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause review*, 27, 1-12.
- Hofer, S. I., Nistor, N., & Scheibenzuber, C. (2021). Online teaching and learning in higher education: Lessons learned in crisis situations. *Computers in Human Behavior*, 121, 106789.
- Hofer, S. I., Nistor, N., & Scheibenzuber, C. (2021). Online teaching and learning in higher education: Lessons learned in crisis situations. *Computers in Human Behavior*, 121, e106789.
- Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, 119, e106713.
- Kahneman, D. (2012). *Rápido e devagar: duas formas de pensar*. Objetiva.
- Marasca, A. R., Yates, D. B., Schneider, A. M. d. A., & Feijó, L. P. (2020). Avaliação psicológica online: considerações a partir da pandemia do novo coronavírus (COVID-19) para a prática e o ensino no contexto a distância. *Revista Estudos em Psicologia*, 37, e200085.
- Mohammed, A. O., Khidhir, B. A., & Vijayan, V. J. (2020). Emergency remote teaching during Coronavirus pandemic: the current trend and future directive at Middle East College Oman. *Innovative Infrastructure Solutions*, 5, 72.
- Moser, K. M., Wei, T., & Brenner, D. (2021). Remote teaching during COVID-19: Implications from a national survey of language educators. *System*, 97, e102431.
- Newton, P. E. (2017). There Is More to Educational Measurement Than Measuring: The Importance of Embracing Purpose Pluralism. *Educational Measurement: Issues and Practice Summer*, 36(2), 5-15.

- Nickel, E. M., Ferreira, M. G. G., Forcellini, F. A., Santos, C. T. d., & Silva, R. A. Á. (2010). Modelo multicritério para referência na fase de Projeto Informacional do Processo de Desenvolvimento de Produtos. *Gestão e Produção*, 17(4), 707-720.
- Pasquali, L. (1997). *Psicometria: Teoria e Aplicações*. Universidade de Brasília.
- Pasquali, L. (2009). Psychometrics. *Revista da Escola de Enfermagem da USP*, 43, 992-999.
- Pasquali, L. (2010). Testes Referentes a Construto: Teoria e Modelo de Construção (*Instrumentação psicológica: fundamentos e práticas* (pp. 165-198). Artmed.
- Quispe-Pietro, S., Cavalcanti-Bandos, M. F., Caipa-Ramos, M., Paucar-Caceres, A., & Rojas-Jiménez, H. (2021). A Systemic Framework to Evaluate Student Satisfaction in Latin American Universities under the COVID-19 Pandemic. *systems*, 9(1), 15.
- Ribeiro, H. C. M., & Corrêa, R. (2021). ESTRATÉGIAS DE ENSINO PRATICADAS NAS INSTITUIÇÕES DE ENSINO SUPERIOR PRIVADA DE UM GRUPO EDUCACIONAL DO BRASIL FRENTE A PANDEMIA DO COVID-19. *Revista Gestão Organizacional*, 14(1), 333-355.
- Roff, S., McAleer, S., Harden, R. M., Al-Qahtani, M., Ahmed, A. U., Groenen, G., & Primparyon, P. (1997). Development and validation of the Dundee Ready Education Environment Measure (DREEM). *Medical Teacher*, 19(4), 295-299.
- Secundo, G., Gioconda, M., Del Vecchio, P., Gianluca, E., Margherita, A., & Valentina, N. (2021a). Threat or opportunity? A case study of digital-enabled redesign of entrepreneurship education in the COVID-19 emergency. *Technological forecasting and social change*, 166, 120565.
- Secundo, G., Gioconda, M. E. L. E., Del Vecchio, P., Gianluca, E. L. I. A., Margherita, A., & Valentina, N. D. O. U. (2021b). Threat or opportunity? A case study of digital-enabled redesign of entrepreneurship education in the COVID-19 emergency. *Technological Forecasting and Social Change*, 166, e120565.
- Shim, T. E., & Lee, S. Y. (2020). College students' experience of emergency remote teaching due to COVID-19. *Children and Youth Services Review*, 119, e105578.
- Sindhu, I., Daudpota, S. M., Bakhtyar, M., Baber, J., & Nurunnabi, M. (2019). Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation. *IEEE Access*, 7, 108729-118741.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research Winter*, 45(1), 89-125.
- Whittle, C., Tiwari, S., Yan, S., & Williams, J. (2020). Emergency remote teaching environment: a conceptual framework for responsive online teaching in crises. *Information and Learning Sciences*, 121(5/6), 311-319.

Figures

Figure 1

Conceptual Model of the Latent Trait.

Figure 2

Conceptual Model of Multiple Perspectives.

Figure 3

Item elaboration diagram based on the factors.

Figure 4

Definition of the requirements of each factor.

Figure 5

Construction flow and adequacy of the items. Example for factor 1.

Figure 6

Flowchart of the item validation process. Example for factor 1.