UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

THOMAZ EDSON VELOSO DA SILVA

EDUCOMETRICS:
FROM THEORY TO APPLICATION

FORTALEZA - BRASIL
2017

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE
DEPARTMENT OF FOOD SCIENCE

THOMAZ EDSON VELOSO DA SILVA

EDUCOMETRICS:
FROM THEORY TO APPLICATION

FORTALEZA - BRAZIL
2017

THOMAZ EDSON VELOSO DA SILVA

EDUCOMETRICS:
FROM THEORY TO APPLICATIONS

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia. Área de concentração: Sinais e Sistemas.

Orientador: Prof. Dr. João César Moura Mota (UFC)
Orientador: Prof. Dr. Rasmus Bro (UCPH)
Coorientador: Prof. Dr. Wagner Bandeira Andriola
Coorientador: Prof. Dr. André Lima Férrer de Almeida

FORTALEZA

2017

THOMAZ EDSON VELOSO DA SILVA

EDUCOMETRICS:
FROM THEORY TO APPLICATIONS

A doctoral thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD) from the PhD School of Science at the University of Copenhagen.

Supervisor: Professor João César Moura Mota (UFC)
Supervisor: Professor Rasmus Bro (UCPH)
Co-supervisor: Associate Professor Wagner Bandeira Andriola
Co-supervisor: Adjunct Professor André Lima Férrer de Almeida

FORTALEZA
2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

S584e Silva, Thomaz Edson Veloso da.
Educometrics: from theory to application / Thomaz Edson Veloso da Silva. – 2017.
82 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2017.
Orientação: Prof. Dr. João César Moura Mota.
Coorientação: Prof. Dr. Wagner Bandeira Andriola.

1. Educometrics. 2. Multivariate statistics. 3. Multilinear analysis. 4. Educational evaluation. 5. Data analysis. I. Título.

CDD 621.38

THOMAZ EDSON VELOSO DA SILVA

EDUCOMETRICS:

FROM THEORY TO APPLICATION

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia. Área de concentração: Sinais e Sistemas.

Aprovada em: 07 / 12 / 2017.

BANCA EXAMINADORA

Prof. Dr. João César Moura Mota (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. José Manuel Amigo Rubio
University of Copenhagen (UCPH)

Prof. Dr. Fátima N. Sombra de Medeiros
Universidade Federal do Ceará (UFC)

Prof. Dr. Charles Casimiro Cavalcante
Universidade Federal do Ceará (UFC)

Prof. Dr. João Marcos Travassos Romano
Universidade Estadual de Campinas (Unicamp)

Prof. Dr. Renato da Rocha Lopes
Universidade Estadual de Campinas (Unicamp)

Prof. Dr. André Lima Férrer de Almeida
Universidade Federal do Ceará (UFC)

Prof. Dr. Wagner Bandeira Andriola
Universidade Federal do Ceará (UFC)

I would like to dedicate my thesis to my beloved parents, Edson and Edna, and to my brother and sister, Diego and Nayra.

ACKOWLEDGMENTS

First of all, I would like to thank God for the opportunity given to me, the greatest challenge I have faced in my professional career so far. I thank my parents, Virgilio Edson Amaral da Silva and Edna Denize Veloso da Silva, for all the patience, affection and dedication that has been given to me throughout my life and especially in these almost five years of doctorate's degree. To my brother and sister, Diego Veloso da Silva and Nayra Veloso da Silva, thank you so much for the encouragement given in moments of doubts and for all the effort to keep me on the right track.

I also would like to thank Professors João César and Rasmus Bro for the briliant supervision, giving me very good insights of my work, as well as Professors Wagner Andriola and André Férrer for their support in educational evaluation and tensor decompositions fields, respectively.

Thank you to all my friends Albano, Herbert, Aparecida, Paulo, Vanessa, Márcio, Franzé, Nicola, Carol, Viola, Nunzia, Joe, José, Rodrigo, Gonçalo, Helena and Manu, for supporting me in all kinds of moments, both academic- and beer- wise. Guys, you are the best.

"Science is not only a disciple of reason but,
also, one of romance and passion."
Stephen Hawking

RESUMO

Atualmente, no contexto educacional, foi dada ênfase à coleta e análise de dados por cientistas de diversas áreas do conhecimento, tais como: psicologia e economia. Esses profissionais analisam dados e seus resultados podem ser usados, por exemplo, para ajudar na tomada de decisões de uma política pública. No entanto, as medidas educacionais tornaram-se muito populares e podem abranger toda a multidimensionalidade contida no processo educacional, desde o ensino e a aprendizagem até a interação social na sala de aula. A compreensão da análise de dados em uma sala de aula precisa ser feita por professores e pedagogos que conheçam exatamente o significado empírico da variabilidade de uma determinada variável medida. Nesse sentido, esta tese conceitua, discute, define e aplica a Educometria, a qual se reconhece ser uma área de conhecimento que faz uso de modelos de estatística multivariável para análise de dados relacionados a contextos educacionais. Depois de estabelecer o conceito de Educometria, aplicamos alguns modelos matemáticos no contexto da avaliação do contexto no ensino à distância. Uma amostra de 791 alunos respondeu o questionário QEOn atualizado para três cursos e a estrutura fatorial do questionário foi válida a partir da aplicação da análise fatorial. A análise dos componentes principais e o Parafac2, sendo modelos bilinear e multilinear, respectivamente, foram aplicados e capazes de identificar comportamentos intrínsecos em relação às 34 assertivas contidas no questionário QEOn. Como conclusão, a aplicação de modelos que permitem a intervenção pedagógica na sala de aula tendo em vista que essa ação é a chave do suporte fornecido pela educometria desenvolvida ao longo desta tese.

Palavras-chave: Educometria. Estatística Multivariada. Análise Multilinear. Avaliação Educacional. Análise de dados.

ABSTRACT

Nowadays, in the educational context, emphasis has been placed on the collection and analysis of data by scientists from several areas of knowledge, such as: psychology and economics. These professionals analyze data and their results can be used, for example, to aid in the decision making of a public policy. However, educational measures have become very popular and can encompass all the multidimensionality contained in the educational process, from teaching and learning to social interaction in the classroom. Understanding of data analysis in a classroom needs to be done by teachers and pedagogues who know exactly the empirical significance of the variability of a particular measured variable. In this sense, this thesis conceptualizes, discusses, defines and applies Educometrics, which is recognized as an area of knowledge that makes use of multivariate statistical models to analyze data related to educational contexts. After establishing the concept of Educometrics, we apply some mathematical models in the learning context of teaching in distance learning. A sample of 791 students answered the updated QEOn questionnaire for three courses and the factorial structure of the questionnaire was valid from the application of factorial analysis. The principal component analysis and the Parafac2, bilinear and multilinear models, respectively, were applied and able to identify intrinsic behaviors in relation to the 34 statements contained in the QEOn questionnaire. As a conclusion, the application of models that allow pedagogical intervention in the classroom as it is the key of the support provided by the educometrics developed throughout this thesis.

Keywords: Educometrics. Multivariate Statistics. Multilinear Analysis. Educational Evaluation. Data Analysis.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

CAPES	Coordination for Improvement of Higher Education Personnel
CFA	Confirmatory Factor Analysis
COMFOR	Management Committee for Initial and Continuing Education of Basic Educational Professionals
CONAFOR	Management Committee of the National Policy for Initial and Continuing Education of Basic Educational Professionals
EA	Environmental Education
EFA	Exploratory Factor Analysis
EJA	Youth and Adult Education
FA	Factor Analysis
FNDE	and the National Fund for the Development of Education
GDE	Gender and Diversity in School
IFES	Federal Institutions of Higher Education
IFET	Federal Institutions of Professional, Scientific and Technological Education
MEC	Ministry of Education
PC	Principal Component
PCA	Principal Component Analysis
QEOn	Quality of Distance Education Teaching Questionnaire
RM	Reference Matrix
RMSECV	Root Mean Square Error of Cross-Validation
SEEQ	Students' Evaluation of Educational Quality
SET	Students' Evaluation of Teaching
SINAES	National System for the Evaluation of Higuer Education
SVD	Singular Value Decomposition
UFC	Federal University of Ceará
VLE	Virtual Learning Environment

Notation

Scalar:	a	Lower-case letters.
Vector:	$\mathbf{a}$	Boldface lower-case letters.
Matrix:	$\mathbf{A}$	Boldface capital letters.
Transpose:	$\mathbf{a}^T, \mathbf{A}^T$	Transpose of $\mathbf{a}$ and $\mathbf{A}$, respectively.
Tensor:	$\underline{\mathbf{X}}$	Underline boldface capital letters.
Kronnecker Product:	$\otimes$	
Hadammard Product:	$\diamond$	
Khatri-Rao Product:	$\odot$	
Outer product:	$\circ$	
Mode 1 product:	$\times_1$	

TABLE OF CONTENTS

1 INTRODUCTION

In this chapter we will introduce this thesis, highlighting some early conceptions of Education, as well as the comprehension of how this discipline is related to others. This association will be made, especially, with educational psychology which brings the concept of measures in the evaluative context, being a key part in the development of this manuscript. The problem of data use and analysis in the educational context will also be addressed, assigning indicatives for the development of a field of investigation called Educometrics. This thesis originates some scientific publications that are also highlighted in this chapter. Finally, the structure of the thesis will be presented in order to give the reader a general understanding of the themes discussed throughout this study.

1.1 General Context

Education is a multidisciplinary discipline that aggregates concepts and relations of social, behavioral and political sciences and also neuroscience and pedagogy. "Educate" comes from the Latin *educare*, in turn linked to *educere*, verb composed of the prefix *ex* (outside) + *ducere* (lead), and literally means 'lead for' be, preparing the individual for the world. With the joining of several sciences, education ends up having a hyper hybrid context with nuances and characteristics related to each one of the aforementioned sciences. Directly, education can be understood as a process that facilitates the acquisition of knowledge and learning trough teaching methodology and didactics. Thus, among the possible methods to formally educate a person based on the curriculum proposed, teaching, storytelling and discussion can be highlighted.

Between 1913-1914, Edward L. Thorndike wrote the three-volume "Educational Psychology" and the books were based on experimental and statistics analysis (Lorge, 1949). Considered the "father of educational measurements", Thorndike opened a new branch of investigation at the beginning of the XX century. Nowadays, we know that the data measured at the educational context can become information and knowledge, but only if it is well handled and collected by specialists. It is also important to highlight that realible data is required for data analysis, no further analysis can be done if we have poor/inaccurate quality data. In this sense, educational research has performed experiments and procedures, especially with the application of psychometric tools that were developed in the behavioral sciences, supporting, e.g., the personality analysis and intelligence quotient tests by assisted data collection.

Although the application of these psychometric tests is used in educational contexts, there are still many other areas of education that are not contemplated. It is still necessary to fully understand and go deeper on the studies in teacher workload, teaching effectiveness, curriculum design effectiveness, relationship among stakeholders, textbook effectiveness and so on. Researches associated with each of these topics are important but still needs to be integrated. Marsh and Bailey (1993) emphasize that education is a multidimensional phenomenon and needs to be analyzed and evaluated as a whole, attesting that evaluation made only by parts of this process is precarious and weakly performed.

Searching for robust and new statistical methods to apply in education we faced a paper published by Tepper (2006), and the following question comes up at the end of it: "Is Educometrics a new field?". Then, an intriguing question was raised: What would be educometrics about?

To the best of our knowledge, the first time the term "educometrics" came up was in a book written by Lulla (1980). Although the title of this book can be easily found on the internet, we couldn't access the manuscript itself even by request to the editor or library. Since then, no further discussion has been conducted in order to establish educometrics as a field of investigation.

This thesis is based on an overlapping of educational evaluation and applied statistics so called Educometrics. Statistical tools present in other "metrics" will be used to analyze data related to the quality of teaching in blended learning courses.

1.2 What is Educometrics?

Educometrics can be undertood as an application of socio- , econo- and psychometrics tools as well as the basic statistical applications in educational contexts. Based on this concept, we may consider any interrelational and measurable construct among teaching, learning and the educational context, both quantitative or qualitative, as an input for educometrics analysis, having the teacher/professor/instructor, the students and the educational environment as part of an integrated process.

An example of application of educometrics is the analysis of the grades obtained by the students in school evaluations. These grades bring with them the information about the knowledge learned by the student in a given subject. Furthermore, analysing the grades, the teacher understands the skills and competences learned by the students. Then we can ask two questions: Is there any (non-)linear relationship between the grade obtained and the teaching method? How do we measure this relationship? We believe that educometrics is the key to thinking deeper and more intriguing questions that can be, for example, answered from the data collected in the classroom.

As can be observed in Figure 1, the focus of educometrics is the pedagogical intervention. As stated in the previous section, educational data has been extensively collected and analyzed, but only for public policy purposes. It is important to fill this gap and give subsidies to teachers to improve their pedagogical practice and optimize their student's learning in the areas that can be evaluated. In this thesis, the Learning Context

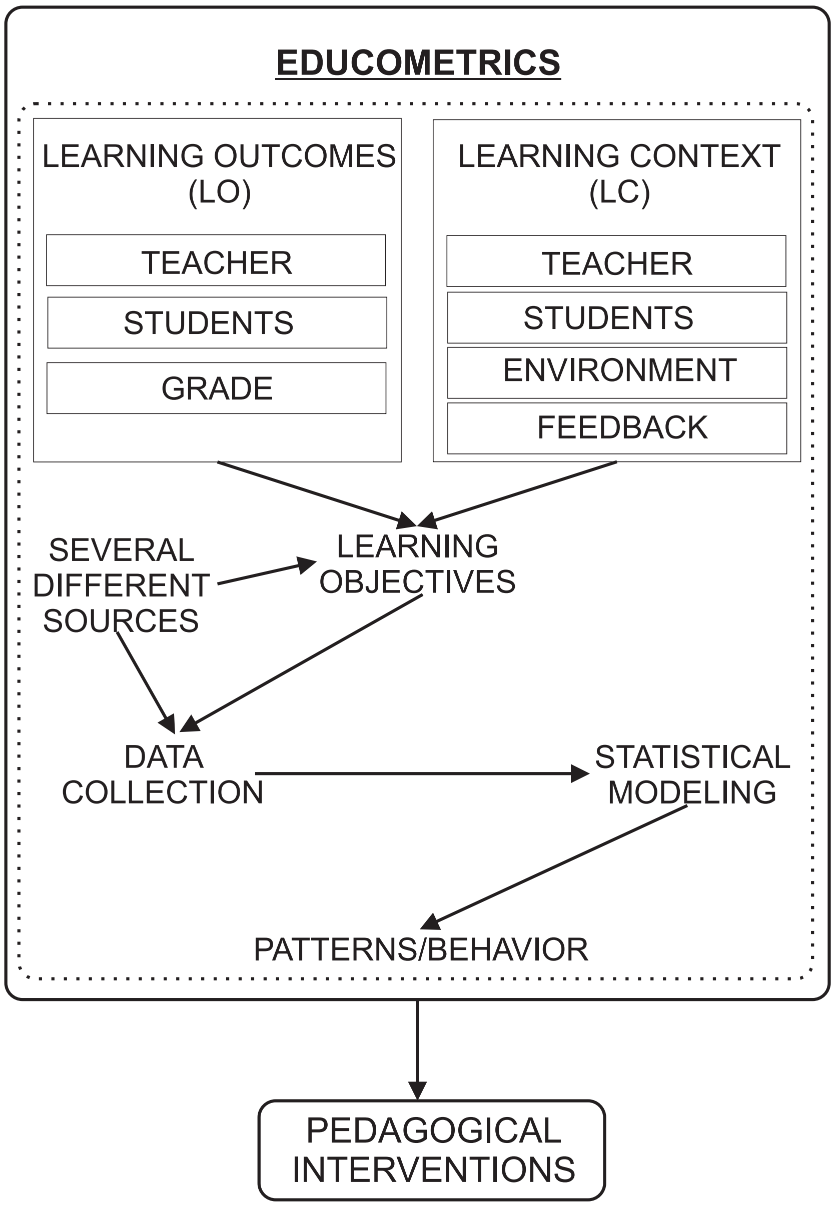


Figure 1: Educometrics organizational concept.

of three blended learning courses are going to be analyzed using classical and advanced statistical thecniques in order to understand possible intrinsic information into the data collected.

Establishing relationships, finding, or proposing explanatory laws, is the main purpose of science. For this, it is necessary to control, manipulate and measure the variables that are considered relevant for the comprehension of the phenomenon. There are many issues on how the information obtained can be translated into knowledge, especially when it comes from a very complex and multidimensional scenario, such as education.

1.3 Contributions

The major contributions of this thesis has been summarized as follows:

- Understand the intrinsic conceptual relations obtained by an educational evaluation and assessment and build up a bridge between educational concepts and the findings of applied multivariate statistics well establishing the Educometrics' concept, especially filling in the gaps non-covered by educational psychologists;
- Develop an assessment questionnaire for the evaluation of teaching quality on blended learning context based on the Students' Evaluation of Teaching (SET) methodology;
- Make use multidimensional and multivariate tools to comprehend how the use of educometrics would benefit the understanding of intrinsic patterns among bi- and multi-linear data.

1.4 List of Publications

Follow the list of publications of this thesis and of some scientific collaboration throughtout the doctorate program.

1.4.1 Edited book

Thomaz Edson Veloso da Silva, Germano de Oliveira Ribeiro, Ismar Frango Silveira, Francisco Herbert Lima Vasconcelos: Assessment and Evaluation in Online Education: Theory and Applications (in portuguese). 1 12/2015; Editora Imprima., ISBN: 978-85-64778-24-5.
DOI:10.13140/RG.2.1.3566.4724

1.4.2 Book chapter

Germano de Oliveira Ribeiro, **Thomaz Edson Veloso da Silva**, Albano Oliveira Nunes, Francisca Aparecida Prado Pinto, Francisco Herbert Lima Vasconcelos: Institucional Evaluations: A Longitudinal Study on Blended Learning Courses at UFC

(in portuguese). Avaliação em EAD: Teoria e Prática, edited by Thomaz Edson Veloso da Silva, Germano de Oliveira Ribeiro, Ismar Frango Silveira, Francisco Herbert Lima Vasconcelos, 12/2015; Editora Imprima., ISBN: 978-85-64778-24-5.
DOI:10.13140/RG.2.1.1469.3202

1.4.3 Journal papers

Thomaz Edson Veloso da Silva, João César Moura Mota, Wagner Bandeira Andriola, Rasmus Bro, André Lima Férrer de Almeida: Educometrics: Principles, Issues and Possible Applications. Scientometrics. *Submitted.* September, 2017.

Thomaz Edson Veloso da Silva, Francisco Herbert Lima Vasconvelos: Students' Evaluation of Teaching (SET) Methodology: Possibilities for Bleended Learning Education (in portuguese). Revista Sustinere. *Accepted for publication.* December, 2017.

Thomaz Edson Veloso da Silva, Germano de Oliveira Ribeiro, Albano Oliveira Nunes, Francisco Herbert Lima Vasconvelos, Wagner Bandeira Andriola, João César Moura Mota: QEOn Questionnaire for Assessing Experiences in Virtual Learning Environments. IEEE Latin America Transactions 05/2017; 15(6):1197 - 1204.
DOI:10.1109/TLA.2017.7932709

1.4.4 Conference papers

Thomaz Edson Veloso da Silva, João César Moura Mota, Wagner Bandeira Andriola, André Lima Férrer Almeida: Data Mining for Open Educational Governmental Data: The Case Study of Brazilian Higher Education. International Technology, Education and Development Conference; 03/2017.
DOI:10.21125/inted.2017.2211

Germano Oliveira Ribeiro, **Thomaz Edson Veloso da Silva**, Albano Oliveira Nunes, Francisco Herbert Lima Vasconcelos: Analysing the Influence of External Factor Associated to the Quality of Online Teaching (in portuguese). Proceedings of XXII Workshop de Informática na Escola (WIE 2016), Uberlândia; 10/2016.
DOI:10.5753/cbie.wie.2016.101

Thomaz Edson Veloso da Silva, Germano Oliveira Ribeiro, Albano Oliveira Nunes, Francisco Herbert Lima Vasconcelos, Wagner Bandeira Andriola: Assessment of Online Teaching Quality Indicators: A Case Study (in portuguese). Proceeding of Workshops do CBIE 2015, Maceió-Alagoas; 10/2015.
DOI:10.5753/cbie.wcbie.2015.503

1.4.5 Other publications by the author

Francisco Herbert Lima Vasconcelos, **Thomaz Edson Veloso da Silva**, João César Moura Mota: Multilinear Educational Data Analysis for Evaluation of Engineering Education. IEEE Latin America Transactions 11/2015; 13(8).
DOI:10.1109/TLA.2015.7332163

Francisco Herbert Lima Vasconcelos, **Thomaz Edson Veloso da Silva**, João César Moura Mota: The Context and Outcomes of Learning in Educational Evaluation an Engineering Course. IEEE Latin America Transactions 09/2015; 13(7).
DOI:10.1109/TLA.2015.7273811

Albano Oliveira Nunes, **Thomaz Edson Veloso Da Silva**, João César Moura Mota, André Lima Férrer de Almeida, Wagner Bandeira Andriola: Validation of the academic management evaluation instrument based on principal component analysis for engineering and technological courses. Ingeniería e Investigación 07/2015; 34(2).
DOI:10.15446/ing.investig.v35n2.47369

Albano Oliveira Nunes, **Thomaz Edson Veloso da Silva**, João César Moura Mota, André Lima Férrer de Almeida, Wagner Bandeira Andriola: Developing an Instrument for Assessment of Academic Management in Engineering Courses. IEEE Latin America Transactions 01/2015; 13(1).
DOI:10.1109/TLA.2015.7040657

Katiuscia C. B. Teixeira, **Thomaz Edson Veloso da Silva**, João César Moura Mota, Natália Cordeiro Barroso, Eduardo V. O. Teixeira: Peer instruction methodology for linear algebra subject: A case study in an engineering course. 2015 IEEE Frontiers in Education Conference, FIE 2015, El Paso; 10/2015.
DOI:10.1109/FIE.2015.7344346

Alberto Lima, Wagner Andriola, Neuman de Souza, **Thomaz Edson Veloso da Silva**, Zarathon L. Viana: A Mixed Pedagogical Method to Improve Teaching and Learning in Brazilian Computing Area Undergraduate Courses. 2015 IEEE Frontiers in Education Conference, FIE 2015, El Paso; 10/2015.
DOI:10.1109/FIE.2015.7344399

Igor P. da Silva, Alberto S. Lima, Neuman De Souza, Flávio R. C. Sousa, Lincoln S. Rocha, **Thomaz Edson Veloso da Silva**: Improving group decision-making in IT service management by the use of a consensus-based MCDM method. Network Operations and Management Symposium (LANOMS), 2015 Latin American, João Pessoa - Paraíba; 08/2015.
DOI:10.1109/LANOMS.2015.7332678

1.5 Organization

This thesis is structured in six chapters, including the introduction. Following a brief content of each one of the five remaining chapters.

Chapter 2: The main concepts of educational measures and their impacts on some types of educational evaluation will be presented in this chapter, as well as a review of the literature on how this data collection has been carried out over the last few years.

Chapter 3: Some statistical models have been widely used in the educational context, the presentation of these models and others that can potentially be used will be shown in this chapter.

Chapter 4: The application context of this thesis, the updated QEOn questionnaire, the audience that compose the analyzed sample and the organization of the data will be presented in this chapter.

Chapter 5: The validation of the updated QEOn questionnaire will be presented in this chapter, making possible a better understanding of the existing relationships between the actors that compose the blended learning courses in the context of the Federal University of Ceará. Multivariate and multi-linear tools are also applied in order to find intrinsic information hidden in the latent variables.

Chapter 6: After presenting the results obtained, in this chapter we will highlight the final considerations regarding the impact of the use of educometrics in several contexts and the need for specialists, both educational and statistical, whose are able to make data analysis for pedagogical intervention.

2 EDUCATIONAL MEASUREMENTS

Educational measures are, for the most part, linked to the collection of data on student achievement. In this chapter, we will work on the concept of educational measure as any measurement associated with the student/learner, both with respect to their learning outcomes and learning context.

2.1 Educational Evaluation and Assessment

One of the scientific fields that re-unites ethical theory of psychological measurement and behavioral aspects which includes the measure of characteristics such as knowledge, intellectual, cognitive, attitudes, personality traits and educational evaluation is psychometrics. In general, the achievement of such measures is done through the construction and validation of information gathering instruments, such as questionnaires, surveys, tests, personality assessments, among others.

In addition to the field of psychology, psychometrics has also been widely used in educational assessment by obtaining, analyzing and interpreting educational measures and indicators using mathematical tools. However, based on information processing in the branch of cognitive and contextual development, by making them a fundamental part of the analysis and interpretation of "data collected in educational context" a step further need to be carried by Educometrics. In this way, this research work will be based on its principle of conception.

According to Tyler (2000), educational evaluation and measurement are distinct processes but often based on descriptive analysis with low mathematical formalism. The educational measure is very important in the evaluation process, however the evaluation can not be limited to the quantitative analysis of the data (Andriola, 1999). In the model proposed by Tyler (2000), the evaluation objectives become key pieces to guide the evaluation process It should be emphasized, therefore, that the evaluation, in general, allows a critical analysis of what is being evaluated, enabling actions that seek to correct errors or to potentiating correctness in the educational process (Andriola, 2001). The educational evaluation is based on a strong description of processes and by its results, without any necessary statistic or probabilistic background.

Among the various forms and objectives of the evaluating act, we highlight the Students' Evaluation of Teaching (SET), as a methodology of real applicability to improve the quality of the courses in a general way (Romero and Ventura, 2010; Andriola, 2002).

The SET is supported by a strong factorial structure (Marsh and Bailey, 1993), which allows us to guarantee its scales and factors as being coherent and stable to investigate what the methodology is proposed for. In this sense, the Students' Evaluation of Educational Quality (SEEQ) questionnaire was developed (Marsh and Bailey, 1993), which aims to measure the nine factors that are considered essential for the evaluation of course quality (Andriola, 2002).

The SEEQ instrument has already been applied in several educational contexts and in various formats (Silva *et al.*, 2012; Vasconcelos *et al.*, 2013; Franco *et al.*, 2013). In all cases, its factor structure was confirmed, generating results that allow the improvement of the quality of the evaluated courses.

Thus, the SET methodology has a great potential to capture information related to the context of students' learning, allowing the distribution of not direct observed variables in daily life at school and enabling improvement of teaching and learning based on administrative and/or pedagogical intervention.

2.2 Measurements in Education

Hodges and Stanton (2007) and Ryan, Anderson, and Birchler (1980) indicate that evaluations in courses, disciplines, curricular components or teaching sessions performed by students are source of concern for the faculty and may lead to displeased students (Franklin and Theall, 1989). Such attitudes occur because teachers believe, while activities are in progress, assessments can be biased (Eiszler, 2002) (Feldman, 1976). Hence, students are not competent and mature assessors (Ryan, Anderson, and Birchler, 1980; Nasser and Fresko, 2002), and most of the time their opinions are influenced by expectations in relation to their performance and results obtained at the end of the teaching process (Baldwin and Blattner, 2003). These issues lead teachers and managers to question the general validity of student assessments and their use, as well as their feasibility (Beran, Violato, and Kline, 2007; Ory, 2001) to influence decision making (Sproule, 2000; Ryan, Anderson, and Birchler, 1980; Nasser and Fresko, 2002). However, according to Franklin (2001), the discomfort of teachers with assessments and classifications are related to the poor quality of their classroom teaching practices. These negative perceptions of assessments may lead teachers to disregard their importance, which may hinder the teaching and development of efforts to improve an activity or discipline, and, according to Aleamoni (1999) and Ory (2001), induces teachers and managers to have misperceptions about course evaluations. Several authors point out that the faculty in general do not consider the evaluation carried out by the students and do not approve the evaluation instruments/questionnaires used (Nasser and Fresko, 2002; Abrami, 2001; Wachtel, 1998; Theall and Franklin, 2001; Centra, 1993)

Other surveys (Beran, Violato, and Kline, 2007; Beran *et al.*, 2005) found that there are teachers who believe that assessment data cannot be used properly by managers academics.

For managers, mostly teachers have a positive attitude towards evaluation data and find it a useful source of information for decision making (Beran *et al.*, 2005; Campbell and Bozeman, 2008). However, there are others who have concerns about the validity of these results and instruments, and according to Franklin (2001); Abrami (2001), this is due to the lack of knowledge and lack of familiarity with the classification tools and the fundamentals of the research under evaluation. As far as students' opinion is concerned, there is little research in this regard and is often limited to case studies in institutions (Beran *et al.*, 2005; Campbell and Bozeman, 2008) which indicate how students perceive the process of collecting their opinions as valid and useful feedback. In addition, they also believe that students can be effective evaluators of teaching. However, other authors indicate that students are not always aware of how institutions use the collected data (Wachtel, 1998; Beran *et al.*, 2005; Campbell and Bozeman, 2008), and do not understand the impact that evaluations generate for change and decision-making and also do not believe that their opinion is used and evaluated (Wachtel, 1998), thus making no use of the data collected in the institution.

Faced with these perceptions and works that indicate the hierarchy of teachers, managers and students, we can still consider that these evaluational tools bring common characteristics, as Algozzine *et al.* (2004) and Clayson (2009) stated. These assessment tools in general education provide student feedback on the effectiveness of teaching using some grading scale and have a number of common characteristics such as open and closed questions about course content and the effectiveness of the teaching process. At least one item of overall effectiveness, written comments on course content and teacher effectiveness are also requested. Another characteristic is that the anonymity of the answers is guaranteed and assured; the answers and the questionnaires are applied and obtained at the end of the term, when the students have already finished their activities, and in the absence of the teacher. In addition, responses to items and scales may be useful for teachers, managers, departments, and faculties assessing the effectiveness of the educational process undertaken and can be used to make various professional development decisions within educational management and administration. Spooren *et al.* (2017) have shown that SET is still conducted and very reliable for the evaluation of teaching effectiveness.

Typically, the items related to activity or discipline assessments seek information about course design, posture, and teacher behavior. According to Sproule (2000), there are elements that commonly appear in these types of evaluations: questions about activities or discipline and content; questions about the teacher's communication skills; questions about teacher-student interaction; questions about the difficulty of the course and the workload; questions about current assessment practices and, finally, student self-assessment questions.

2.2.1 Learning Context

Researches also point out that the evaluation of efficacy of learning context is multidimensional and some teaching elements can be highlighted, categorized and identified (Algozzine *et al.*, 2004; Marsh and Roche, 1997; Marsh, 1987). The Students' Evaluation of Educational Quality (SEEQ) questionnaire, proposed by Marsh (1987), presents categories of questions about teaching behaviors and if carried out seriously guarantees the evaluation of the teaching effectiveness. This questionnaire is formed by the following categories: learning/value; teacher enthusiasm; organization; individual relationship; group interaction; range of coverage; the examinations/classification; readings; and the workload/difficulty.

In other works proposed by Braskamp and Ory (1994) and Centra (1993), similar measures of teaching effectiveness have been identified in the assessment system, called the Individual Development and Educational Assessment (IDEA), which includes categories such as course organization and planning, clarity/communication skills, student and teacher interaction and their relationships, difficulty/workload, student grading and exams, and self-assessment.

Another research instrument on the quality of higher education is the Course Perceptions Questionnaire (CPQ) and it was used to measure the experiences of British students. This instrument contained 40 items on eight scales and was proposed by Ramsden (1981) in a survey of 2,208 students in a total of 66 academic departments of engineering, physics, economics, psychology, history, and English. A factorial analysis of this questionnaires allowed to identify eight scales of characteristics and highlighted two dimensions, the first referring to the positive evaluation of teaching and programs, and the second referring to the use of formal teaching methods with emphasis on training with professional relevance. According to Gibbs, Habeshaw, and Habeshaw (1988), CPQ could be used for teaching assessment and course evaluation, although the correlations obtained in Ramsden (1981) between students' perceptions and their approaches to the study were relatively weak. Similar results were found by another researcher (Parsons, 1988) and this raised doubts about the adequacy and validation of CPQ as a research tool (Meyer and Muller, 1990).

From the CPQ, a new revised instrument, called Course Experience Questionnaire (CEQ), was also proposed in Ramsden (1991), which aimed to obtain indicators to monitor the performance of teaching quality in academic programs and was used in institutions in Australia (Ramsden, 1991; Linke, 1991). The instrument had 30 items on five scales and aimed to identify different perceptions of the quality dimensions of teaching.

A common and present element in all these evaluations is the feedback generated, that is, the possible practical returns related to the results obtained. However, according to Beran, Violato, and Kline (2007), while institutions believe in their teaching assessments, they rarely actually practice the results of their own ongoing assessments for decision-making. Studies show that the lack of financial resources and the lack of robust mathematical and statistical tools for interpretation, identification of teaching strategies and the interpretation of results, generate problems that arise when the results are misunderstood (Beran *et al.*, 2005; Wagenaar, 1995).

An overview of several questionnaires can be found in Richardson (2005).

2.2.2 Learning Outcomes

From another perspective, knowledge assessment activities that express students' performance results often have an important influence on learning. The way in which the evaluation is developed and applied by the teacher and the results or expectations of the students, can generate different learning profiles (Sternberg and Zhang, 2005). Struyven, Dochy, and Janssens (2005) argue that the way in which assessment is carried out in higher education has an important influence on student learning. In this context, a student's expectation regarding the evaluation methodologies or procedures that will be used by the teachers in a course, establish a direct relation of how students deal with academic tasks and get ready for exams, assignments, tests and other activities in which they will be evaluated. Equally, strategies and procedures regarding how to study and learning itself on the part of the students are generally strongly influenced by evaluation experiences provided previously with appropriate feedback.

In education, defining what learning goals we want to achieve means structuring the educational process in ways that allow for changes in thinking and behavior. The educator may have expectations and guidelines for the teaching process that are not clear but will be part of the learning assessment process. Learning assessment is a key factor in determining whether learning objectives have been achieved. It is clear that it is easier to achieve goals when they are well defined, but it is more difficult for students to reach the level of cognitive development because they do not know exactly what is expected of them during and after the teaching process. Learning occurs simultaneously and interactively in, at least, three main domains: affective, cognitive and psycho-motor.

The learning evaluation experiences, besides being important elements in the curriculum, directly influence the students' behaviour, since they can be determinants in their academic development, as well as the students' plan of studies, assigning even priority and meaning to the diverse academic tasks. Research indicates (Rickards and Friedman, 1978; Nolen and Haladyna, 1990) that the way the student observes the lesson, watches and even carries out his notes is linked to the expectations as to the forms of evaluation that will be used by the teachers. Thus, overloading of academic activities and tasks in a discipline or even evaluation methods dissociated from the level at which content is approached can lead to superficial learning about certain content (Ramsden, 1981, 1991).

Both teachers and their modes of assessment and consequent student performance influence learning approaches; however, depending on the learning context they may be modifiable (Struyven, Dochy, and Janssens, 2005). In this context, if an institution wants to develop critical and creative thinking for its students and increase its problem-solving capacity, it must devise structural change. This change must take into account not only a new design of its curricular structure, contents of disciplines, or even the competences and abilities of the area under study or of the pedagogical proposal of the course, but broad aspects that involve classroom methodologies and the development of scientific and technological skills and mainly evaluation practices of excellence that reflect not only on students' learning but also the effectiveness of the teaching done by the teachers in the opinion of the students.

Learning outcomes should outline the most central and essential elements of a course or curriculum. They may also shape the evaluation proposal of an institution. As such, the process of developing learning outcomes provides an opportunity for reflection on what is most needed to help students acquire certain knowledge and skills. We can also consider the following elements to characterize Learning Objectives: a) core ideas for the course, b) desired types of learning, and c) the context in which the knowledge and skills acquired in the course will be used, including possible applications, providing a basis for the development of learning outcomes.

The central ideas assume that to begin the process of developing learning outcomes, it may be useful to consider some central ideas of the programmatic content and generalizable skills taught in the course. In addition to information about the context and types of learning, we have to consider the following concerns that can be expressed in the following questions: What do students need to know in order to succeed in the course or discipline? What should students do to succeed in the course or discipline? What knowledge or skills should students bring to the classroom to take as a basis? What knowledge or skills will be new to students during a course or discipline? What other areas of knowledge are linked to the work of the course or subject being studied?

Types of learning arise from learning taxonomies to help make learning outcomes more accurate. Identifying learning levels can also help develop appropriate methods for better learning outcomes in assessing a course. One of the examples is Bloom's Educational Objectives Taxonomy (Bloom *et al.*, 1956; Bloom, 1944), which is particularly useful because it associates verbs and specific words with each level of learning. Although the Bloom Taxonomy is hierarchized, each type of learning may be an important aspect of a course (Anderson *et al.*, 2000; Bloom, Hastings, and Madaus, 1971; Bloom, 1986). Ultimately, however, learning outcomes should focus on higher order thinking found at the highest levels of Taxonomy through features such as analyzing, evaluating, and creating.

2.3 Assessment in Distance Education

With the advancement of distance education in Brazil and worldwide, defining a coherent evaluation process that takes into account the peculiar aspects of this type of teaching has been difficult (Laguardia, Portela, and Vasconcelos, 2007).

These difficulties arise due to the complexity of e-learning, as a result of several variables that compose it, such as: teaching material quality, curriculum, teaching and learning process, accessibility of the Virtual Learning Environment (VLE), students and online instructor's interaction, among others (Sales *et al.*, 2011; Andriola and Loureiro, 2005).

In order to propose improvements in the process of evaluation of distance learning courses, some researches (Sales *et al.*, 2011; Liaw, 2008; Hervas, García, and Penalvo, 2005) point out to the development of tools that aid the student's learning evaluation, whose objective is to improve and systematize the evaluation process. These tools are characterized by choosing parameters, which, according to their proponents, have a greater impact on the quality of the course. Adding to this the mathematical bias to seek to establish the proof of its reliability.

In the context of data analysis or data mining, recent research points to the use of data mining techniques in data extracted from VLE (Romero and Ventura, 2006). The VLE allows storing many kinds of data that can be used to extract patterns and characteristics, such as dropout behaviour or lany kind of students' profiles.

Standards can help identify possible dropouts. Such a prediction will contribute not only to reducing the high rates of dropouts, but also the improvement of the satisfaction ratings and credibility of the e-learning courses as a solid teaching modality for the current contemporary society.

2.4 Summary

Throughout this chapter, we briefly review on some concepts and definitions of educational evaluation and the importance of educational measures in this process. We exemplify the evaluation of contexts and learning as forms of evaluation that can be quantified by instruments that are already widely disseminated by the scientific community. In the following chapter, we will present some statistical models that are used in some "metrics" fields and can be replicated in the context of educometrics.

3 ”METRICS” MODELS: AN OVERVIEW

In contemporary society, we are increasingly using data collection and analysis in several areas of knowledge. Each dataset is related to a field of knowledge, with nuances and specific features for that field. Thus, it is necessary that specialists in these areas have statistical knowledge to analyze this data. Moreover, they have the possibility of developing models to solve specific problems of each area. In this chapter, we will raise the theoretical discussion about the already well established ”metrics” in the scientific community, as well as the description of some statistical models that will be carried out throughout this thesis.

3.1 Theoretical Discussion

Nowadays, the ”metrics” are becoming widely used by researchers, accepting that the information brought by the data is powerful and was not well treated in the past. We may highlight some of the well stablished metrics as:

- Sociometrics is an analytical tool for studying interactions among social groups;
- Econometrics is the branch of Economics that, starting from the general economic theory, analyzes the data provided by Statistics, through the application of mathematical methods, expressing economic laws in mathematical language;
- Psychometrics is a specialized branch of psychology that deals with the study and development of psychological assessment tests and the development and application of statistical knowledge and other mathematical processes linked to psychology;
- Medicometrics is the science of integrating different sources of measurements related to a pathological system;
- Chemometrics emerged from the need to extract chemical information from the profusion of data resulting from modern instrumentation;
- Biometrics refers to metrics related to human characteristics;
- Bibliometrics is one of the key ways of measuring the impact of scholarly publications.

As can be seen in the list above, the ”metrics” have in common the use of statistical models applied to their respective areas of knowledge. In this sense, educometrics emerges as a ”metric” responsible for bringing together data related to the teaching and learning process. Sleeping ”in the bosom of” other metrics, especially psychometry, educometrics could show off its usefulness and start to be seen as a fruitful field of investigation such as the well established psychometrics which already has some applications in educational contexts; The link between psychometrics and educometrics comes with the belief that every measurement in education describes the educational development of the personality and the cognitive development in the same way that height and weight describe physical growth.

According to Professor de Leeuw (2006):

> If Foo is a science then Foo often has both an area Foometrics and an area Mathematical Foo. Mathematical Foo applies mathematical modelling to the Foo subject area, while Foometrics develops and studies data analysis techniques for empirical data collected in Foo. Each of the social and behavioural sciences has a form of Foometrics, although they may not all use a name in this family.

Additionally, he also complemented that "Psychometrics and Educometrics have been around for a long time, at least since Galton, and their development has been very closely linked and often the two have been indistinguishable". It is time to split up both and start to think of educometrics as a whole branch of nuances and specific features.

In the following sections, some models of the multivariate and multi-linear data analysis will be presented, highlighting those that will be in the scope of this research.

3.2 Multivariate Analysis

With the modern world's need to provide better quality, low-cost products and services, effective knowledge management is needed. We are surrounded by information that is often not collected, thus not becoming useful knowledge. Nowadays, this information is collected and stored with greater ease, considering the technological advances that have been developed so far.

From the storage and collection of this information, a step forward was given: the mining process of these data. Multivariate analysis is understood as a set of statistical techniques that simultaneously use multiple variables Carroll, Green, and Chaturvedi (1997). Thus, if we have more than two variables in the same statistical model we can state that we are dealing with a multivariate analysis.

Many multivariate techniques are extensions of univariate and bivariate analysis. Thus, we do not have to apply a single model for each variable, and it is possible to apply the statistical model to a set of variables. Other techniques are unique to the multivariate analysis, such as Factor Analysis (see Section 3.4.1.1), which distinguishes groups based on linear combinations between the analyzed variables.

In general, whenever a decision needs to be made, a large number of factors must be taken into account. Obviously, not all of these factors weigh the same way at certain conditions. Sometimes, by making an intuitive decision, these factors or variables are not identified in a systematic way, that is, the variables that affect decision making are not identified.

When analyzing the world around us, all events, whether cultural or natural sciences, involve a large number of variables. Science intents to know reality, and to interpret events and phenomena based on the knowledge of the intervening variables, considered important in these events.

There are several methods of multivariate analysis, with very different purposes among them. Therefore, one goes back to the following questions: Which knowledge de we intend to generate? Or rather, what kind of hypothesis we want to generate regarding the data and the methods are chosen according to the aims of the research? Since it is known that the multivariate analysis can be seen as an exploratory data analysis, lending itself to generate hypotheses, and not to confirm them, or using confirmatory techniques, as in the hypothesis tests, in which one has an confirmation on the sample under study.

3.3 Multidimensional Analysis

Nowadays, we live in a data deluge that has being generated every minute from several computational systems around the world. Coupled with this generated data, the extraction of useful information appears to be fundamental to understand the set of standards and patterns associated with this data. Data processing techniques (clustering/discovery, regression/ classification/prediction) became essential to comprehend all the information brought by the collected data. However, some of the data has multiple entries and then goes beyond the classical bilinear representation.

Multidimensional techniques are becoming new trends in several topics and disciplines in the contemporary society. The idea is that all data collected by several sources are somehow correlated and deserve to be treated and analysed according to its multidimentional structure. In this thesis, multidimensional arrays are called as "tensors" such as described by Smilde, Bro, and Geladi (2004), Cichocki *et al.* (2009) and Kroonenberg (2008).

The study of tensors revealed a new way of analyzing and understanding the data. Although the beginning of studies of these techniques dates back to the 1970s (Carroll and Chang, 1970), many areas of knowledge, nowadays, use tensor decomposition techniques in an effective way: signal processing (Almeida, Favier, and Mota, 2008), chemometrics (Smilde, Bro, and Geladi, 2004), psychometrics (Kroonenberg, 2008), data mining (Morup, 2011) and many others. A survey of many unsupervised multiway applications can be found in Acar and Yener (2009).

3.4 General Models in Data Analytics for Education

As the concept of educometrics is being discussed and formalized in this thesis, we only have indications of tools that are adequate to be allocated within its context. Kroonenberg (2008) makes use of student performance measures in different courses over five years. This example refers to the multidimensionality of the act of evaluating this, often, the uni or multivariate analysis is performed.

In this sense, the matrix and tensor decompositions appear as a powerful tool to analyze the intrinsic relations of the variables analyzed. Studies presented in Kolda and Bader (2009) become important to draw a parallel with the proposal of consolidation of the application of educometrics as a field of investigation.

In addition, two doctoral theses present the application of the principal component analysis (PCA) for educational management (Nunes, 2016) and of parallel factor analysis (Parafac) for the analysis of learning outcomes and learning context as a whole integrated system (Vasconcelos, 2016). Both are making a direct citation to the works that present the term educometrics as a research area.

3.4.1 Multivariate Models

In this section we will present the multivariate models Factorial Analysis and Principal Component Analysis that will be used throughout this research.

3.4.1.1 Factorial Analysis model

Factorial Analysis (FA) stands out as being a set of statistical techniques that has as one of its main objectives the dimensionality reduction of the variables with low information loss. In general, FA seeks to establish standards that best represent the original variables, which can be grouped and classified (Cronbach, 1951).

The latent variables, formed by the original data, provide information not observed initially from the data analysis. FA is one of the classic techniques for data mining, being widely applied in problems in the field of psychometry (Romero and Ventura, 2010), mainly in the development of evaluation instruments, and is currently widely used in pattern recognition and signal processing for noise and redundant information elimination, among others (Romero and Ventura, 2010) .

The extraction of patterns, from the AF, is based on the matrix of correlation or covariance of the data related to the original variables. Several mathematical models can represent the FA, however, in the following a linear decomposition is presented to describe the model, regarding the equation below (Gorsuch, 1997):

$$\mathbf{M}_{L\times C} = \mathbf{A}_{L\times P}(\mathbf{B}_{C\times P})^T + \mathbf{E}_{L\times C}, \tag{1}$$

in which $\mathbf{M}_{L\times C}$ is the original data matrix with $L \times C$ dimensions, and, by convention, the number of lines (L) is treated as the number of observed samples and the number of columns (C) the number of variables analyzed. The $\mathbf{A}_{L\times P}$ and $\mathbf{B}_{C\times P}$ are matrices called components matrices, which are related to the observations and the original variables, respectively, and $\mathbf{E}_{L\times C}$ is the information considered not relevant from the system of linear equations. Thus, to be a non-zero matrix ($det(\mathbf{M} \neq 0)$), the number of extracted factors (P) must be less than the number of original variables (C).

It is important to emphasize that the matrices $\mathbf{A}_{L\times P}$ and $\mathbf{B}_{C\times P}$ are orthogonal to each other. This information guarantees that the data generated by the FA are linearly uncorrelated, ensuring that the vector projections in this new coordinate system are made using only the intrinsic information of each original variable.

For each column vector of the matrix $\mathbf{B}_{C\times P}$, we will find values that will represent the influence of each original variable on the latent variable, these values are called loadings.

Some prior procedures are necessary to apply the FA and ensure the applicability of the technique to the data obtained, which are (Gorsuch, 1997; Silva *et al.*, 2012; Majors and Sedlacek, 2001):

- Use of the Kaiser-Meyer-Olkin (KMO) and Bartlett Tests to verify if the sample is suitable for the application of FA;
- Selection of the number of components of the model (if it is a procedure for model validation);
- Verification of the internal consistency of the factors analyzed through Cronbach's α;
- Verification of the representativeness analysis of each original variable in FA through the commonality of each observed variable in the new latent structure;
- Rotation matrix is also an option to be considered to better comprehend the new data generated by the model.

It is worth noting that, depending on the application, FA may be confirmatory or exploratory (Fabrigar *et al.*, 1999). Confirmatory Factorial Analysis (CFA) is used when the researcher seeks to attest or validate a factorial structure established by previous work. In relation to Exploratory Factorial Analysis (EFA), is a process of data mining that identify patterns among the observed variables analyzed.

3.4.1.2 Principal Component Analysis Model

Principal Component Analysis (PCA) is one of the most classic statistical methods for multivariate data analysis. Its purpose is to allow an analysis of the data in order to minimize the internal correlation of a set of variables and to minimize the experimental noise obtained during the preparation of the data set. This is achieved through a (linear) transformation of the original data into a new set of uncorrelated data, called the principal components (PC), so that the first components of this new set of variables concentrate the greater variability of the original variables and noise can be minimized through an ideal selection of the principal components.

Geometrically, the determination of the principal components occurs through a change of coordinate axes in such a way that in a certain ordering, the first axes carry with them a greater information of the data and such axes are orthonormal. This transformation of referential axes can be seen as a linear transformation whose matrix associated with such a transformation on the canonical basis is $\mathbf{V} \in \mathbb{R}^{m\times m}$ and is applied over a set of m, n-dimensional vectors represented by a $m \times n$ matrix, $\mathbf{X}$, in the canonical basis. The new vector coordinates on the new axes determined by $\mathbf{V}$ are given by a $m \times n$ matrix, $\mathbf{Y} = \mathbf{XV}$. In such a way, the product $\mathbf{XV}$ corresponds to a projection of the old variables in the new coordinated axes determined by $\mathbf{V}$. As we shall see, the matrix $\mathbf{V}$ must be orthogonal. Therefore, finding the major components is equivalent to determining a factorization of the original data matrix in the product $\mathbf{X} = \mathbf{YV}^T$. The matrix $\mathbf{Y}$ is called the principal components matrix, score matrix, and the matrix $\mathbf{V}$ is called loading matrix.

In fact, it is important to point out that the ideal objective is to mathematically model the data as complex information from a complete set of independent elementary information, in which each elementary information has an indivisible characteristic, and whose algebraic combination faithfully reproduces each complex information, unless resilient observation errors and representative without content of the sources are presented in the data. Therefore, such errors are independent of sources, with no representativity identified in the set of elementary information.

The identification of the independence between elementary information being a difficult and exhausting task, leads us to seek, in most cases, a description of the phenomena from attributes associated with their elementary information. Thus, the choice of independent references is directed to a choice of orthogonal references, which are chosen without having a biunivocal relation with the elementary information. On the contrary, the choice of the systems formed by orthogonal dimensions is carried out from criteria that consider the data without necessarily extracting from them their elementary characteristics, but only to aligning them to each one of the attributes. In this way, the elementary characteristics remain all present in each attribute. The choice of references is associated with the choice of attributes, which are supposed to be associated with criteria that incorporate orthogonality structures.

From the point of view of the multivariate analysis, we can think of the matrix $\mathbf{X}$, $m \times n$, as a set of m samples given in function of n attributes. Initially, these n attributes may have some internal relationships. Determining the principal components would be to determine new uncorrelated attributes or components such that the first components describe as much information as possible, the second describing as much information as possible of the part that the former could not describe, and so on, up to the n-th component.

Ideally, the matrix $\mathbf{X}$ represents a data set with a mean equal to 0 and standard deviation equal to 1, i.e., the data is centered around the origin and normalized. In this case, we will say that the data is standardized. This practice prevents any discrepancy between values, often due to choices of units of inconvenient measures.

There are two classical methods for calculating the principal components. One is based on the Eigenvalues Decomposition (EVD) of the covariance (or correlation) matrix of the original data and the other method is based on the Singular Values Decomposition (SVD) of the original data matrix. The main difference is if we consider a matrix $\mathbf{X}$ with $m \times n$ dimentions, the EVD can be used if and only if $m = n$, and for $m \neq n$ the SVD should be used. In this thesis, we will demonstrate the method based on SVD.

In order to find the principal components through SVD, we will base the theorem below, which is known as the Singular Value Decomposition Theorem. Proof of this theorem can be found in Smilde, Bro, and Geladi (2004).

Considering $\mathbf{X}$ a real value full rank matrix with $m \times n$ dimentions. Then, there is an orthogonal real matrix $\mathbf{U}$, $m \times m$, and another orthogonal real matrix $\mathbf{V}$, $n \times n$, such as:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{2}$$

where $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \sigma_2, ..., \sigma_p)$, $p = \min(m, n)$, is a diagonal matrix $m \times n$ uniquely determined, such that the elements of its main diagonal are nonnegative real numbers satisfying $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_p \geq 0$.

Furthermore, the matrix $\mathbf{U}$, called the matrix of the left singular vectors, will be the matrix of the eigenvectors of $\mathbf{X}\mathbf{X}^T$, the matrix $\mathbf{V}$, called the matrix of the right singular vectors, will be the matrix of the eigenvectors of $\mathbf{X}^T\mathbf{X}$ and the real numbers σ_i will be equal to $\sqrt{\lambda_i}$, for $i = 1, 2, ..., p$, where λ_i is the i-th eigenvalue of the matrix $\mathbf{X}\mathbf{X}^T$, which also corresponds to the i-th eigenvalue of the matrix $\mathbf{X}^T\mathbf{X}$.

Given a matrix $\mathbf{X}$ of a given dataset (preferably standardized), the principal components matrix $\mathbf{Y}$ can be determined directly from the SVD computation of the matrix $\mathbf{X}$. Let $\mathbf{U}$,$\mathbf{V}$ and $\mathbf{\Sigma}$, respectively, the matrix of the left singular vectors, the matrix of the right singular vectors and the matrix of the singular values, all referring to the matrix $\mathbf{X}$. The matrix of the principal components can be determined as follows:

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma} \tag{3}$$

In this way, we have that the matrix $\mathbf{V}$ determined by the SVD of $\mathbf{X}$ is the loading matrix, i.e., the matrix that gives us the new coordinated axes.

The variance explained ν_i of the i-th component can be determined through the i-th singular value (σ_i) of $\mathbf{X}$:

$$\nu_i = \frac{\sigma_i^2}{\operatorname{tr}(\Sigma^T\Sigma)} = \frac{\sigma_i^2}{\sum_{j=1}^{p} \sigma_j^2} \tag{4}$$

The cumulative variance μ_i of the first i components is given by:

$$\mu_i = \sum_{k=1}^{i} \nu_k = \frac{\sum_{k=1}^{i} \sigma_k^2}{\sum_{j=1}^{p} \sigma_j^2} \tag{5}$$

The difference of variance explained and cumulative variance is that the variance explained is related to the retained singular variance of each principal component extracted and the cumulative variance is the sum of the variance explained of the principal components used in the model.

A reasonable choice of the number of components is given by choosing r ($r << i$) such a that the number $\mu_r = \sum_{i=1}^{r} \mu_i$ is relatively close to 1. In some applications, for example, it is assumed that $\mu_r \approx 0.70$ is a reasonable value. For instance, another criteria for rank (components) selection is based on the division of each eigenvalue by the highest one and this percentage must be less than 10%. But there is no general method to determine which minimum value of μ_r is ideal, since, in general, such value and the analysis made on the components that determine this value depends on the geometry of the problem.

3.4.2 Multidimensional Models

Historically, the first work on the subject of multidimensional analysis was introduced by Cattell (1944). Based on the Thurstone principle, arguing that a simple structure could be found to describe a data matrix or its correlation matrix with the help of factors, Cattell proposed the simultaneous analysis of several matrices and the use of the principle of "proportional parallel profiles" (Cattell, 1944). That is, from an arrangement of matrices, find a common set of factors amog them. He defined objects, circumstances/time, attribute, scale, and observer as the five inputs to an idealized multidimensional arrangement and for practical reasons, reduced them to a three-input arrangement with people, attributes, and circumstances.

The decomposition of a three-input arrangement was first presented by Tucker (1966). This decomposition consists of finding loading matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ and a three-way core-tensor $\underline{\mathbf{G}}$, which were introduced with a hypothetical example of 12 individuals, 9 treatments and 5 observers. In another independent research, Lathauwer, Moor, and Vandewalle (2000) had shown a similarity between the core-tensor and the singular value matrix in the singular value decomposition (SVD).

According to Cichocki *et al.* (2009), the matrix factorizations such as Principal Component Analysis and Singular Value Decomposition (SVD) are important tools for dimensionality reduction, noise reduction and data mining. However, these factors have only a two-dimensional representation, such as space and time, making their limited use in data structures requiring more than two dimensions (Cichocki *et al.*, 2009).

3.4.2.1 Basic Principles

According to Kolda and Bader (2009), a tensor is a multidimensional array that can vary according to its order. For instance, a scalar number is a zero order tensor, a vector is a first order tensor, a second order tensor is a matrix and higher order tensors are those with three or higher dimensions.

Some matrix products and unfolding matrices forms will be presented below.

3.4.2.1.1 Matrix Products

Matrix products are important for the algebraic development of tensor factorizations.

The Kronecker product ($\otimes$) of two matrices $\mathbf{A}$ $(I \times J)$ and $\mathbf{B}$ $(K \times M)$ can be defined as:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1J}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & \dots & a_{IJ}\mathbf{B} \end{pmatrix}$$

The dimension of the product $\mathbf{A} \otimes \mathbf{B}$ is $IK \times JM$. The Kronecker product is also defined by two matrices where the regular matrix product does not exist (if $J \neq K$).

Another important product is the Hadamard product ($\diamond$), which can be defined by the two matrices $\mathbf{A}$ and $\mathbf{B}$ with dimensions $I \times J$:

$$\mathbf{A} \diamond \mathbf{B} = \begin{pmatrix} a_{11}b_{11} & \dots & a_{1J}b_{1J} \\ \vdots & \ddots & \vdots \\ a_{I1}b_{I1} & \dots & a_{IJ}b_{IJ} \end{pmatrix}$$

where a_{IJ} and b_{IJ} are elements of $\mathbf{A}$ and $\mathbf{B}$, respectively. Then, the Hadamard product can be seen as an element-wise product.

The third product is the Khatri-Rao product ($\odot$), that can be used by the computation of matrices with the same number of K columns, defined as:

$$\mathbf{A} \odot \mathbf{B} = (\mathbf{A}_1 \otimes \mathbf{B}_1 \dots \mathbf{A}_K \otimes \mathbf{B}_K)$$

Some of the very useful properties of the products presented above can be found in Smilde, Bro, and Geladi (2004).

3.4.2.1.2 Unfolding

In matrices it is useful to organize their input elements as vector-rows or vector-columns. Tensors, or three-way arrays, can also have their elements organized in two-dimensional sections, as can be arranged in flat horizontal, lateral or frontal slices, according to Figure 2. It is worth noting that in this case, elements of tensors are located geometrically at the meeting point of the three planes perpendicular intersection.

As many computational resources are not able to manipulate tensors of order

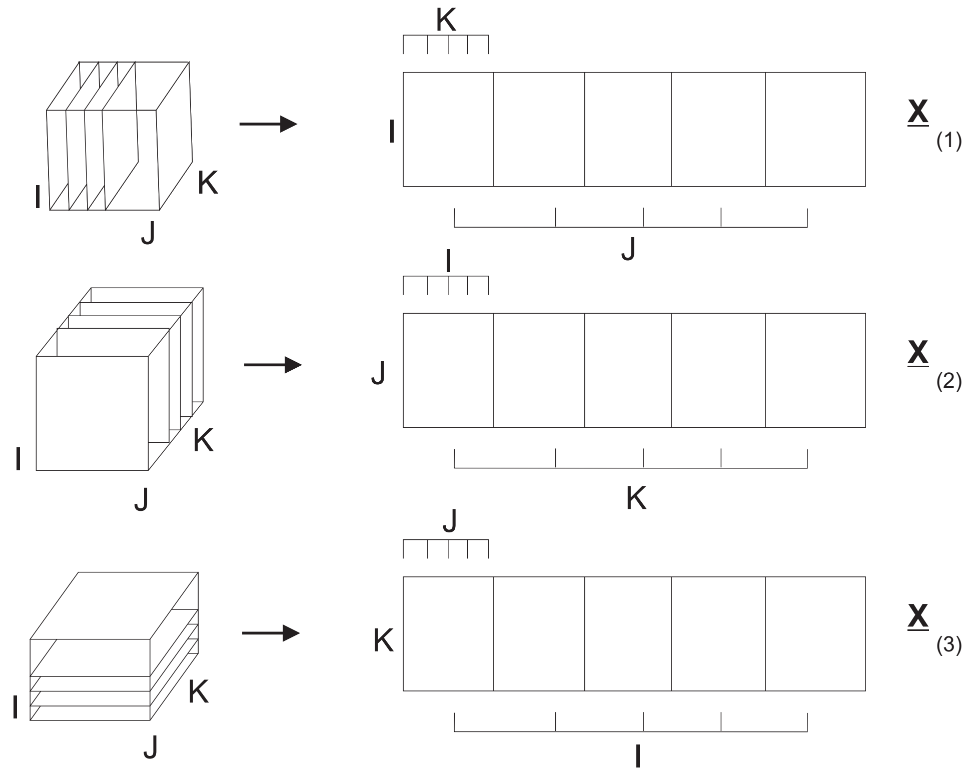


Figure 2: Unfolding of a tensor in each mode.

higher than 2, it is convenient to transform them into matrices in each mode, $\mathbf{X}_{(1)}$, $\mathbf{X}_{(2)}$ and $\mathbf{X}_{(3)}$, in order to facilitate its computational manipulation. For this, the concept of slice becomes fundamental to understand how the unfolding process works. The slices of the tensor can be considered as cuts in the tensor and its modes.

3.4.2.2 Parafac

The Parafac model can be considered a straightforward extension of the PCA model. In the Parafac method Smilde, Bro, and Geladi (2004); Kolda and Bader (2009), a trilinear model is found to minimize the sum of squares of the residues, e_{ijk}, according to $x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk}$, where R is the number of components used in the Parafac and a_{ir},b_{ir} and c_{ir} indicate weights of the r-th component and e_{ijk} denotes the entry's unexplained information.

Connections between PCA and Parafac can be further visualized via tensor product notation: PCA approximates $\mathbf{X}$ as a sum of R rank-1 second-order tensors (i.e., outer-products), and Parafac approximates $\underline{\mathbf{X}}$ as a sum of R rank-1 third-order tensors (Figure 3). In Parafac, a three-way array $\underline{\mathbf{X}}$ is decomposed into a sum of triple products of vectors (triplets).

As can be seen in Equation 6, the Parafac model decomposes the tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1, I_2, \ldots, I_N}$ in a vector-product $\mathbf{a}_r^{(1)}, \mathbf{a}_r^{(2)}, \ldots, \mathbf{a}_r^{(N)}$, contained in the matrices $\mathbf{A}_r^{(1)}, \mathbf{A}_r^{(2)}, \ldots, \mathbf{A}_r^{(N)}$, which can be represented as linear combinations as shown in Equation 6.

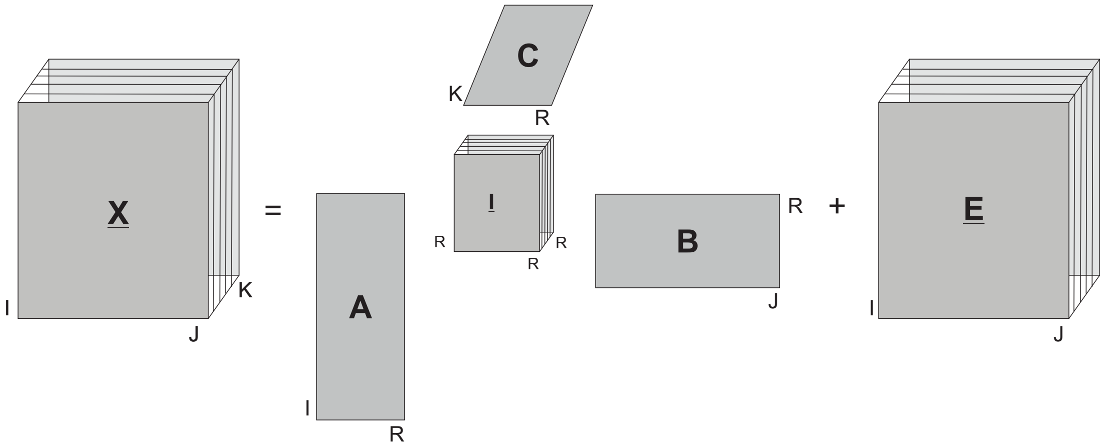


Figure 3: Parafac pictorical representation.

$$\begin{aligned}\underline{\mathbf{X}} &\approx \textstyle\sum_{r=1}^{R} \lambda_r [\mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \ldots \circ \mathbf{a}_r^{(N)}] \\ &\approx \underline{\mathbf{I}} \times_1 \mathbf{A}_r^{(1)} \times_2 \mathbf{A}_r^{(2)} \ldots \times_N \mathbf{A}_r^{(N)} \end{aligned} \tag{6}$$

where $\underline{\mathbf{I}} = diagtensor[\lambda_1, \lambda_2, \ldots, \lambda_R] \in \mathbb{R}^{R,R,\ldots,R}$ is a superdiagonal core tensor with $\lambda_r \neq 0$ values on its main diagonal, R is the number of components extracted by the model, $\times_N$ is the n-mode product and the external products of the triads $(\mathbf{a}_r^{(1)}, \mathbf{a}_r^{(2)}, \ldots, \mathbf{a}_r^{(N)})$ are the rank-1 tensors.

When compared to matrix models, the Parafac model brings with it a very important characteristic: uniqueness. The uniqueness of a model, ensures that there is only one solution to the problem, regardless of the rotation of the resulting data model.

An important step in using the Parafac method consists of estimating the number of components or latent variables (R) in the dataset. Several criteria can be used to determine this parameter, such as the explained variance and core consistency diagnostic (Cordondia) that was used in this work (Smilde, Bro, and Geladi, 2004). The Parafac algorithm used in this work was implemented in MATLAB via PLS Toolbox.

3.4.2.3 Tucker3

Originally, the Tucker3 model is considered a generalized Parafac model, since in the Parafac there is no iteration between the resulting vector-components of the model, ensuring that the core tensor is a superdiagonal tensor ($\underline{\mathbf{G}} = \underline{\mathbf{I}}$).

The Tucker3 model (see Figure 4) can be established based on a tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1,I_2,\ldots,I_N}$ in which the components $J_1, J_2, \ldots, J_N << I_1, I_2, \ldots, I_N$ are extracted by the factorization, guaranteeing the existence of a core tensor $\underline{\mathbf{G}} \in \mathbb{R}^{J_1,J_2,\ldots,J_N}$ which contains all the interactions among all the column-vectors at the loading matrices (see Equation 7).

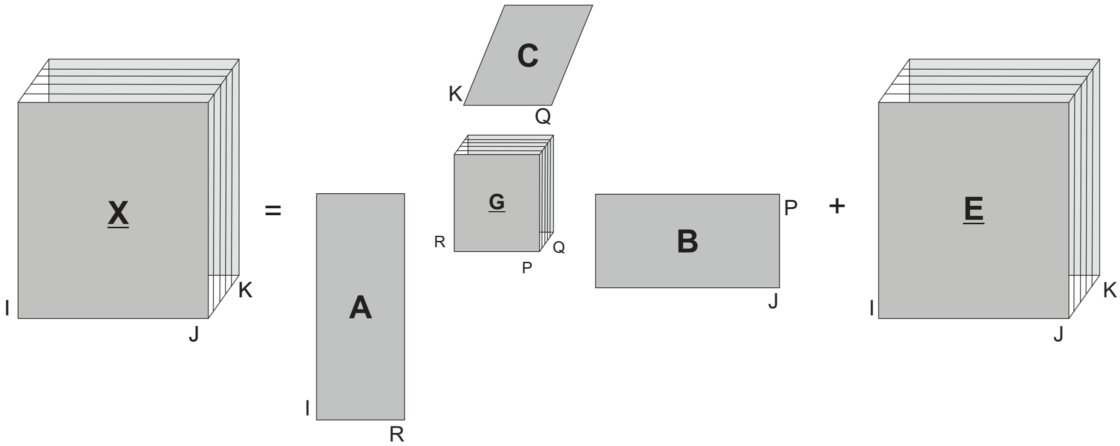


Figure 4: Tucker3 pictorical representation.

$$\begin{aligned}\underline{\mathbf{X}} &\approx \textstyle\sum_{j=1}^{J_1} \sum_{j=2}^{J_2} \ldots \sum_{j=N}^{J_N} g_{j_1, j_2, \ldots, j_N} [\mathbf{a}_{j_1}^{(1)} \circ \mathbf{a}_{j_2}^{(2)} \circ \ldots \circ \mathbf{a}_{j_r}^{(N)}] \\ &\approx \underline{\mathbf{G}} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \ldots \times_N \mathbf{A}^{(N)} \end{aligned} \tag{7}$$

In the model presented in the Equation 7, an important feature can be highlighted: the non-uniqueness of the model. This characteristic can be observed in the development of Equation 8. To exemplify, a third-order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I_1, I_2, I_3}$ is adopted and the matrices resulting from the factorization $(\mathbf{A}_{I_1 \times J_1}^{(1)}, \mathbf{A}_{I_2 \times J_2}^{(2)}, \mathbf{A}_{I_3 \times J_3}^{(3)})$ are multiplied by three invertible matrices $(\mathbf{B}_{J_1 \times J_1}^{(1)}, \mathbf{B}_{J_2 \times J_2}^{(2)}, \mathbf{B}_{J_3 \times J_3}^{(3)})$, such as:

$$\begin{aligned}\underline{\mathbf{X}} &\approx (\underline{\mathbf{G}} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \times_3 \mathbf{B}^{(3)}) \times_1 (\mathbf{A}^{(1)} \mathbf{B}^{(1)^{-1}}) \times_2 (\mathbf{A}^{(2)} \mathbf{B}^{(2)^{-1}}) \times_3 (\mathbf{A}^{(3)} \mathbf{B}^{(3)^{-1}}) \\ &\approx \hat{\underline{\mathbf{G}}} \times_1 \hat{\mathbf{A}}^{(1)} \times_2 \hat{\mathbf{A}}^{(2)} \times_3 \hat{\mathbf{A}}^{(3)}, \end{aligned} \tag{8}$$

we only observe the rotation of the initial matrices $(\mathbf{A}_{I_1 \times J_1}^{(1)}, \mathbf{A}_{I_2 \times J_2}^{(2)}, \mathbf{A}_{I_3 \times J_3}^{(3)})$.

The deduction imposed by the Equation 8 brings with it some important implications. The Tucker3 model does not guarantee a single solution for the factorization, even imposing, for example, orthogonality or orthonormality constraints on the matrices $\mathbf{B}_{J_1 \times J_1}^{(1)}, \mathbf{B}_{J_2 \times J_2}^{(2)}$ e $\mathbf{B}_{J_3 \times J_3}^{(3)}$. Thus, the traditional Tucker3 model does not apply to contexts where the search for a single optimal solution is the key problem.

In the Tucker3 and Parafac models, some constraints are widely used for model development, such as: non-negativity, orthogonality, linear independence, orthonormality, among others. These constraints are useful when analyzing each type of context to be studied, for example, if we are working with real positive data, it is important that the model of factorization brings with it this characteristic, as it occurs in the applications of matrix factorizations (Smilde, Bro, and Geladi, 2004).

Applications of tensor factorizations are expanding in different contexts and the classical Parafac and Tucker3 models need to be "adjusted" to these new situations, initiating the development of new types of factorizations. The Parafac2 decomposition, to be discussed in the next section, is an example of a model derived from Parafac decomposition and it will be used in this thesis.

3.4.2.4 Parafac2

Considering trilinearity as a linear relationship between all of the three modes in a three-way tensor, the Parafac2 model does not follow this concept like Parafac does. Parafac2 does not assume that the shape or length of the factors in one of the modes must be the same for each sample. A pictorial representation of the Parafac2 model can be seen in Figure 5.

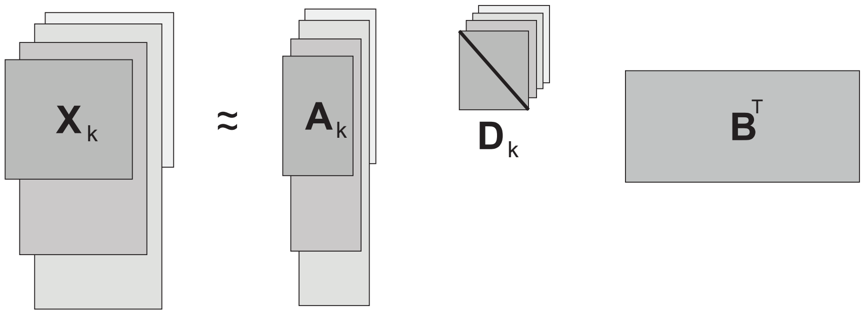


Figure 5: Parafac2 pictorical representation.

The three-way Parafac2 can be perceived as a Parafac model with no trilinearity restriction. For this, the Parafac can also be called Parafac1. Assuming that $\mathbf{X}_k$ is a tensor with k-samples with diferent sizes, we may have an individual $\mathbf{A}$ for each k, so called $\mathbf{A}_k$. Then, the model is given by:

$$\mathbf{X}_k = \mathbf{A}_k\mathbf{D}_k\mathbf{B}^T + \mathbf{E}_k$$

where $\mathbf{X}_k$ is a slab of data $(I_k \times J_k)$ and I varies with k. $\mathbf{D}_k$ is a diagonal matrix that holds the k-th row of $\mathbf{C}$ in its diagonal. Restricting that the cross-product $\mathbf{A}_k^T\mathbf{A}_k$ remains constant. Then, $\mathbf{A}_k$ is modeled as the product $\mathbf{P}_k\mathbf{H}$, in which $\mathbf{P}_k$ is an orthogonal matrix, keeping the sizes od $\mathbf{A}_k$, and $\mathbf{H}$ is a small quadratic matrix with dimension equal to the number of components. Then, the alternating least square algorithm minimize the following cost funcion (Kiers, ten Berge, and Bro, 1999):

$$FIT_{err} = \sum_{k=1}^{K} ||\mathbf{X}_k - \mathbf{A}_k\mathbf{D}_k\mathbf{B}^T||^2$$

The structure of the data collected in this study, we conduct the use of Parafac2 decomposition to extract the latent components of the data. In our case, each slab $(\mathbf{X}_k)$ will change its number of $k-$respondents.

3.5 Summary

In this chapter, multivariate and multidimensional analysis concepts were presented. In total, five models were described, in which three will be used in this research. In the following chapter, all methodological procedures for data collection, research context, the updated QEOn questionnaire and the way the data were structured for several analyzes will be presented.

4 EDUCOMETRICS ON TEACHING ASSESSMENT IN DISTANCE EDUCATION

This chapter presents the methodological aspects of this thesis. The context of application, data structure and materials and methods are presented. Also, an updated version of the QEOn questionnaire (da Silva et al., 2017) will be presented, including the indicator Self-Assessment.

4.1 Context of Application

The process of social development experienced in recent years in Brazil has provided series of investments, such as significant increase in access to education. This has required a higher level of formation by the professionals responsible for training these people. In this sense, Decree No. 8.752, dated May 6th, 2016, formalized the National Policy for the Training of Basic Education Professionals (MEC, 2016), and the Ministry of Education (MEC) in November 8th (MEC, 2013), established the Management Committee of the National Policy for Initial and Continuing Education of Basic Educational Professionals (CONAFOR), defining its general guidelines for the creation of an Institutional Management Committee for Initial and Continuing Education of Basic Educational Professionals (COMFOR) in the Federal Institutions of Higher Education (IFES) and in the Federal Institutions of Professional, Scientific and Technological Education (IFET's). The CONAFOR is responsible for formulating, coordinating and evaluating the actions and programs of the MEC, the Coordination for Improvement of Higher Education Personnel (CAPES) and the National Fund for the Development of Education (FNDE).

COMFOR acts in the coordination of the training programs and courses, setting up a dialogue space for the distribution of resources already allocated to the IFES' budget through the 20RJ resolution, and the COMFOR, together with the coordination of the training courses proposed at the IFES, assessing and evaluating both the quality of the course offered and the teaching and learning process (MEC Ordinance No. 1,105, November 08, 2013, MEC Ordinance No. 1,087, Of August 10, 2011 and CONAFOR Resolution No. 1 of August 17, 2011) (MEC, 2013; 2011b; 2011a).

In this context, the Federal University of Ceará (UFC), through its COMFOR, has been developing actions and continuing education courses in partnership with several secretariats of the MEC. In this thesis, training courses for teachers and other professionals from all over the world, Public educational networks and the municipalities of the State of Ceará are going to be analyzed.

4.2 Methods and Materials

In this section, the methods and data collection procedures will be presented, the updated QEOn questionnaire that was used to collect the information will also be discussed.

4.2.1 Exploratory Factor Analysis, Principal Component Analysis and Parafac2

This research made use of Factor Analysis (FA), Principal Component Analysis (PCA) and Parafac2 to obtain the intrinsic relations related to the statements described in Table 5. Since this instrument is brand new and the psycometric validation has never been performed to ensure its factorial structure, FA has been performed in order to explore and identify the factorial structure of the questionnaire.

From the FA, we can observe the clusters related to a set of statements that are statistically related. In the analysis of the database, the software Matrix Laboratory - MATLAB® with the PLS Toolbox were used.

In addition to the clusters formed by the linear combinations described in Section 3.4.1.1, the FA also highlights the importance of each component extracted, which is given here as the name of the indicator, and it is done by calculating the explained variance associated with each eigenvalue obtained in the matrix decomposition.

Cronbach's α index (Cronbach, 1951) was used to verify the internal consistency of the indicators obtained, that is, if the statement's responses converged to a central tendency, then the index tends to be close to 1, pointing out that its variability ranged from 0 to 1. In other words, it consists in verifying if the specific variance of the items is low and variance of the set of items is high: if so, α is close to 1. All the processes presented here followed the recommendations for validation of research instruments (Gorsuch, 1997; Majors and Sedlacek, 2001; Silva *et al.*, 2012).

PCA and Parafac2 are going to be applied to explore the intrinsic relationship among the observed variables.

4.2.2 Sampling

Since 2012, the Federal University of Ceará (UFC), in partnership with SECADI/MEC, has been promoting extension courses for teacher training in the areas of Youth and Adult Education (EJA), Gender and Diversity in School (GDE) and Environmental Education (EA) using the blended learning education to reach people from all over the state of Ceará - Brazil.

In the coursers' structure that served as samples, none are totally online, and these are characterized as blended-learning courses, hence the need for statements related to the presential instructor that are at the location the courses has been taken.

The extension courses have between 180 and 200 hours distributed in modules, lasting an average of four months. Composed of presential and online presences, mediated by a Virtual Learning Environment (VLE), through which the students are encouraged to identify a concrete problem in their school community and then propose an intervention project.

In 2014, a total of 1639 teachers were certified by these continuing teacher training courses in GDE, EJA and EA. All the concluding teachers were able to participate in this research, aiming to contribute with their perception of the course, so as to improve the quality of future offers.

At the end of each course, all the certified students were invited, by e-mail, to respond to the online version of the updated QEOn questionnaire. The sample obtained consisted only of those respondents who, in fact, fully answered the questionnaire. Thus, for the analysis of the results, the data provided by 791 students, or 48.5% of the total target audience, were considered.

In addition to the indicators extracted by QEOn, some initial questions were conducted in order to understand how the sample were characterized.

Table 1: Characterizing the gender in the courses analyzed.

Course	Male Abs[%]	Female Abs[%]	Total Abs[%]
EA	48(25,8%)	138(74,2%)	186(100,0%)
EJA	80(22,7%)	272(77,3%)	352(100,0%)
GDE	70(27,7%)	183(72,3%)	253(100,0%)
Total	198(25,0%)	593(75,0%)	791(100,0%)

As can be seen in Table 1, in all courses there is a majority of the female audience (> 70%). Notably, the EJA course is the one that presents a greater percentage of females (77.3%).

Table 2: Characterizing the age range in the courses analyzed.

Course	20 - 30 years old Abs[%]	30 - 40 years old Abs[%]	40 - 50 years old Abs[%]	50 - 60 years old Abs[%]	Above 60 years old Abs[%]	Total Abs[%]
EA	49(26,3%)	60(32,3%)	46(24,7%)	29(15,6%)	2(1,1%)	186(100,00%)
EJA	63(17,9%)	125(35,5%)	105(29,8%)	52(14,8%)	7(2,0%)	352(100,00%)
GDE	83(32,8%)	89(35,2%)	66(26,1%)	15(5,9%)	0(0%)	253(100,00%))
Total	195(24,7%)	274(34,6%)	217(27,4%)	96(12,1%)	9(1,1%)	791(100,00%)

The classification of the participants regarding its age range is presented in Table 2. Although, in the three courses, the majority of students are between 30 - 40 years old (34.6%), the students classified as 40-50 years old and 20-30 years old right behind.

One of the issues we want to analyze is how students who have had previous experiences in blended learning respond to the indicators raised by the QEOn question-

Table 3: Characterizing how many blended courses each students have taken.

Course	First Course Abs[%]	1 - 2 courses Abs[%]	3 - 4 courses Abs[%]	5 - 6 courses Abs[%]	Above 6 courses Abs[%]	Total Abs[%]
EA	36(19,4%)	63(33,9%)	51(27,4%)	10(5,4%)	26(14,0%)	186(100,00%)
EJA	99(28,1%)	98(27,8%)	93(26,4%)	26(7,4%)	36(10,2%)	352(100,00%)
GDE	60(23,7%)	84(33,2%)	66(26,1%)	17(6,7%)	26(10,3%)	253(100,00%))
Total	195(24,7%)	245(34,6%)	210(26,5%)	53(6,7%)	88(11,1%)	791(100,00%)

naire. According to Table 3, approximately one quarter (24.7%) of the attendees are newcomers to blended learning courses. The EJA course is the one that presents the highest percentage (28.1%) of students who are taking a course in this modality for the very first time.

Table 4: Characterizing the obtained grade (range) by the students in the courses analyzed

Course	0,0 - 4 Abs[%]	4,1 - 6 Abs[%]	6,1 - 8 Abs[%]	8,1 - 10 Abs[%]	Total Abs[%]	Abs[%]
EA	1(0,5%)	6(3,2%)	63(33,9%)	10(62,4%)	186(100,00%)	
EJA	3(0,9%)	5(1,4%)	112(31,8%)	26(65,9%)	352(100,00%)	
GDE	1(0,4%)	10(4,0%)	83(32,8%)	17(62,8%)	253(100,00%))	
Total	4(0,6%)	21(2,7%)	258(32,6%)	53(64,1%)	791(100,00%)	

As we did not have access to the final grade obtained by the student at the end of the courses, we requested that they provide us with this information. The vast majority (64.1%) of the students who took part in this research obtained a final grade between 8.1 - 10.

In the following section, the updated QEOn questionnaire will be presented.

4.2.3 Updated QEOn Questionnaire

With the current need to evaluate the quality of the courses being offered by the UFC, the Online Teaching Quality Assessment (QEOn) questionnaire was developed (da Silva *et al.*, 2017). This questionnaire is based on the Students' Evaluation of Teaching (SET) methodology, proposed initially by Marsh (1987). In SET, several evaluation instruments were developed, where the Students' Evaluation of Educational Quality (SEEQ) questionnaire is the most popular and commonly used (Richardson, 2005). The SEEQ evaluates the quality of courses and programs (Marsh and Bailey, 1993). Although it evaluates aspects related to the quality of the courses, it does not apply to the reality of blended-learning education since it does not take into account the specific characteristics of such an educational model. Thus, there was a need for the development of another instrument, the QEOn.

From the development of the QEOn questionnaire (da Silva *et al.*, 2017), it was proposed to improve it with the insertion of the new factor "Self-Assessment". This factor aims to identify the ability of the experiment participants to evaluate the object under study, then it proposes to act as a kind of meta-evaluation.

According to Marsh and Bailey (1993), the quality evaluation of courses was divided into nine factors, whose data were obtained through the application of the SEE-Q questionnaire, but from a re-reading of Marsh's studies, this thesis listed six factors as potential identifiers of the quality of blended learning courses. The updated QEOn questionnaire has 34 statement (see Table 5), in which the respondents need to assign a degree of agreement to each of these assertions using a 5-point Likert scale (1- Strongly Disagree, 2- Partially Disagree, 3 - Indifferent, 4- Partially agree and 5- Strongly Agree).

Table 5: QEOn Questionnaire in its original version (tested) and the english version.

	Original version (Portuguese)	English version
Q1	Você considera o curso intelectualmente desafiador e estimulante.	You consider the course intellectually challenging and stimulating.
Q2	Você aprendeu algo que considera pertinente.	You have learned something that you consider pertinent.
Q3	O seu interesse sobre o tema cresceu como consequência do curso.	Your interest in the subject grew as a consequence of the course.
Q4	Você compreendeu os conteúdos do curso.	You have understood the contents of the course.
Q5	O tutor à distância mostrou entusiasmo ao ministrar o curso.	The online instructor showed enthusiasm throughout the course.
Q6	O tutor à distância foi dinâmico e energético na condução do curso.	The online instructor was dynamic and energetic in driving the course.
Q7	O tutor à distância melhora a apresentação dos conteúdos com sugestões de sites e vídeos.	The online instructor improves the presentation of the content with suggestions of websites and videos.
Q8	O tutor à distância apresenta interesse pelo aprendizado do aluno.	The online instructor is interested in student learning.
Q9	O tutor à distância elucida as indagações.	The online instructor elucidates the inquiries and questions.
Q10	Os materiais do curso foram bem preparados e cuidadosamente transmitidos.	The course materials were well prepared and carefully transmitted.
Q11	Os objetivos propostos estão de acordo com o que foi ensinado durante o curso.	The proposed objectives are in accordance with what was taught during the course.

Continued on next page

Table 5 – continued from previous page

	Original version (Portuguese)	English version
Q12	O tutor à distância propôs leituras complementares que facilitam a obtenção de nota.	The online instructor proposed complementary reading materials that make it easier to obtain a grade.
Q13	Os cursistas são encorajados a participarem das discussões no fórum.	The students are encouraged to participate in the forum discussions.
Q14	Os cursistas são convidados a compartilhar suas ideias e conhecimentos.	The students are invited to share their ideas and knowledge.
Q15	Os cursistas são encorajados a responder a questão central do fórum.	The students are encouraged to answer the central question of the forum.
Q16	Os cursistas são estimulados pelo tutor a distância a propor ideias/questionamentos transversais ao tema central do fórum.	The students are encouraged by the online instructor to propose ideas/questions that are transversal to the central theme of the forum.
Q17	O tutor presencial foi amigável na relação com os cursistas.	The presential instructor was friendly in relation to the students.
Q18	O tutor presencial fez com que os cursistas se sintam confortáveis com sua ajuda no polo de atendimento.	The presential instructor made the students feel comfortable with their help at the call center.
Q19	O tutor presencial tem interesse genuíno em relação ao aprendizado do cursista.	The presential instructor has a genuine interest in the student's learning.
Q20	O tutor presencial se mostra disponível no horário de atendimento no polo.	The presential instructor is available to help the students in the place the courses have been taken.
Q21	O tutor à distância relaciona as implicações do conteúdo com várias teorias.	The online instructor relates the implications of content to various theories.
Q22	O tutor à distância apresenta fundamentos preliminares de ideias/concepções que são desenvolvidas nas atividades virtuais.	The online instructor presents preliminary foundations of ideas/concepts that are developed in the virtual activities.
Q23	O tutor à distância apresenta seu ponto de vista quando julga adequado.	The online instructor presents his point of view when he/she deems it appropriate.
Q24	O tutor à distância comenta adequadamente as pesquisas atuais desenvolvidas na área de estudo.	The online instructor adequately comments on the current research developed in the area of study.

Continued on next page

Table 5 – continued from previous page

	Original version (Portuguese)	English version
Q25	Há a disponibilidade das correções das avaliações/trabalhos de forma adequada.	The corrections of the tests and questions are available in a adequate time.
Q26	Os métodos de avaliação do cursista são justos e apropriados ao curso.	The students' assessment methods are fair and appropriate to the course.
Q27	As avaliações/materiais para os testes são trabalhados pelo tutor à distância.	The assessments/materials for the tests are considered by the online instructor.
Q28	O curso requer a leitura de textos que estão disponíveis.	The course requires reading of texts that are available in the VLE.
Q29	Leituras complementares, chat, fóruns, portfólios contribuem para apreciação e compreensão dos conteúdos.	Complementary reading materials, chat, forums, portfolios contribute to the appreciation and understanding of the contents.
*Q30	A sua capacidade de organização contribuiu para uma melhor absorção das informações trabalhadas ao longo do curso.	Its organizational capacity contributed to a better comprehention of the information processed during the course.
*Q31	Você foi capaz de desenvolver atividades além das propostas pelo curso.	You were able to develop activities beyond those proposed by the course.
*Q32	Você foi em busca de informações que pudessem complementar o conhecimento adquirido no curso.	You have searched information that could complement the knowledge acquired in the course.
*Q33	Você disponibilizou material complementar (textos, imagens, vídeos) ao longo das atividades do curso.	You have made available complementary material (texts, images, videos) throughout the course activities.
*Q34	Você acessou as informações extras (textos, imagens, vídeos) trazidas pelo tutor e pelos demais cursistas.	You have accessed the extra information (texts, images, videos) brought by the tutor and the other students.

*Included statements for the tested factor.

4.2.4 Data collection

In the application of the QEOn, the coordinators of each one of the three courses were asked to send an e-mail to the concludents, inviting them to participate in this research. In the e-mail, a brief text of awareness was placed, emphasizing the importance of the serious participation of the students in this evaluation. The QEOn was made available online using Google Docs. It was enough to access the link available in the email, and the students would have access to the QEOn.

The QEOn was available for 30 days, due to the end of the activities of the three courses occuring at different times. For each course, the form remained online during the 10 days prior to its conclusion and another five days after the ending of the class in the VLE, in order that the online instructors reinforced the request to fill it.

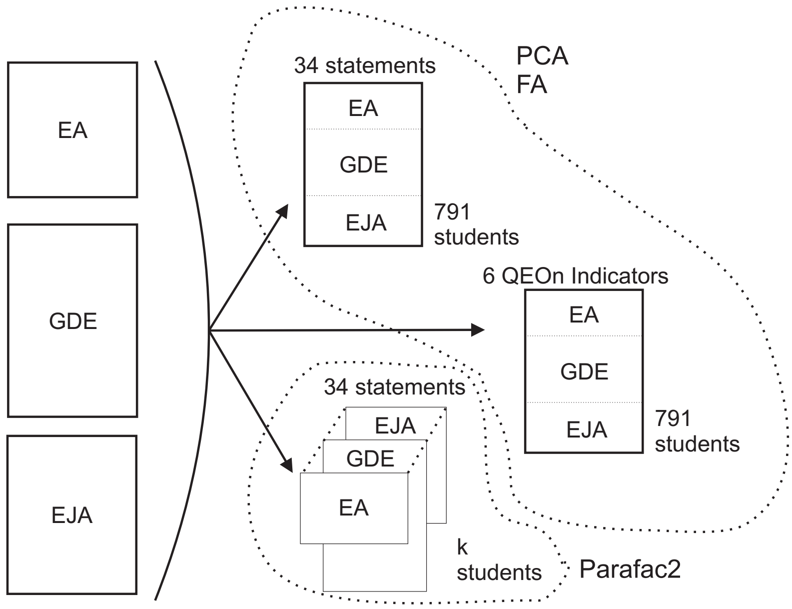


Figure 6: Data organization by the PCA and FA models and the Parafac2 model.

In addition to the awareness process, a database treatment procedure was performed, excluding those students who did not respond to the QEOn completely.

Figure 6 shows the structural layout of the data and which techniques will be used for each one of the presented structures. The bilinear models, FA and PCA, will be applied to the matrices in the upper part of the figure. The multilinear model, Parafac2, will be applied in the tensorial structure shown at the bottom of the figure. For the Parafac2 model, we complete by zeros the matrices that have lower sizes in the second mode and then we apply the constrains presented in Section 3.4.2.4.

The procedure mentioned above does not have any similarity to matrix or tensor data completion, since we are not dealing with sparse data.

4.3 Summary

In this chapter, the methodological issues, such as: context of application, data collection, sampling, QEOn questionnaire, mathematical models used and the dataset structures, were shown. In the following, the results obtained by the computation of Factor Analysis, Principal Component Analysis and the Parafac2 will be presented.

5 RESULTS

In this chapter, the results obtained with the application of the adequacy tests of the sample collected by the QEOn, the analysis of the representativeness of each statement in the FA, and finally the analysis of the load matrix from the application of the EFA will be presented. After validating the updated QEOn questionnaire, the PCA will be applied to the two data structures presented in Figure 6. Furthermore, a multidimensional analysis of the data will be done using the Parafac2 model to extract intrinsic structures associated with the analyzed variables.

5.1 Factor Analysis

This section aims to validate the application of the updated QEOn questionnaire in blended learning courses for the extraction of quality indicators. This validation will be done using a classical data mining technique called Factor Analysis (FA) (Romero and Ventura, 2010). FA aims to extract the largest amount of information from the observed variables, which are analyzed and grouped according to their specificities, manifesting common attributions to the mentioned items. These variables, called originals, can be organized to form a new set of variables, of lesser information than the original set, which are identified as latent variables of the environment (Andriola, 2002). The development of indicators, based on the application of FA, is common in the field of psychometrics, and is therefore a widespread technique for validation of a data collection instrument (Gorsuch, 1997; Fabrigar *et al.*, 1999). This model, and others used in research in the context of engineering and computation, has proved useful for the analysis of educational contexts (Silva *et al.*, 2012; Nunes *et al.*, 2015; da Silva *et al.*, 2017).

5.1.1 KMO and Bartllet's tests

For validation of data collected, the Kaiser-Meyer-Olkin (KMO) and Bartllet' tests were performed to verify the applicability of the FA (Gorsuch, 1997; Majors and Sedlacek, 2001). The KMO test measures the suitability of the sample, indicating the representation of the variances of the variables. Regarding the Bartllet's sphericity test, it indicates whether the data is correlated, rejecting the null hypothesis that the correlation matrix is an identity matrix.

Table 6: KMO and Bartllet's tests.

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy		0,956
Bartllet's test of Sphericity	Approx. Chi-square	18051,04
	df	561
	Sig.	0,000

As can be seen in Table 6, the KMO (0,956) and Bartllet (Approx. Chi-square = 18051,04; Sig. = 0,000) tests validated the application of the FA in the collected database, indicating an excellent scenario for EFA application (Majors and Sedlacek, 2001).

5.1.2 Selecting the Number of Components

As presented in Section 3.4.1.1, prior to the application of the FA, an important step is the selection of the number of components of the model. For this, the energy associated with each component was calculated by means of its respective eigenvalue. The total number of components is equal to the number of variables measured, than we can have 34 principal components. Thus, as presented in Table 7, six components were selected for the FA. The Kaiser criterion ($\lambda > 1$) was used for the selection of these components (Gorsuch, 1997; Nunes *et al.*, 2015; Fabrigar *et al.*, 1999; Majors and Sedlacek, 2001).

Table 7: Eigenvalues and Variance Explained.

Components	Variance Explained		
	Eigenvalues (λ)	% of Variance	% of Cumulative Variance
1	14,082	41,418	41,418
2	2,437	7,168	48,586
3	2,083	6,127	54,713
4	1,461	4,297	59,010
5	1,117	3,286	62,296
6	**1,003**	**2,951**	**65,247**
7	0,863	2,539	67,785
8	0,823	2,421	70,207
⋮	⋮	⋮	⋮
34	0,141	0,415	100,0

Also according to Table 7, the cumulative variance of the components is highlighted, whose total value of the six components explains slightly more than 65% of the total information collected by the QEOn. It should be noted that the first component explains 41,418% of the total information contained in the data collection.

5.1.3 Commonalities

The commonality of a variable indicates how well that variable is adequately represented in the FA. In Table 8, among the 34 assertions that make up the QEOn questionnaire, all of them have commonalities above 41%, indicating that the variables are significantly represented in the FA (Fabrigar *et al.*, 1999; Nunes *et al.*, 2015).

5.1.4 Reliability Analysis

In this research, Cronbach's α was used to verify the reliability of the data (Cronbach, 1951). This index represents the distribution of the variables of the same factor. That is, it is a test that represents the stability of the results provided by the measured questionnaire.

Cronbach's α was computed in each of the factors in order to guarantee a greater reliability of the data. The results varied between 0.747 and 0.949 in the factors that have lower and higher loading, respectively (see Table 8). Based on these results, it is possible to guarantee an excellent condition for the analysis of the arithmetic averages originated by these factors.

5.1.5 Validating the Updated QEOn Questionnaire

For the validation using Factorial Analysis (FA), all the loading factors extracted are satisfactory according to psychometric criteria (> 0.4) (Gorsuch, 1997; Fabrigar *et al.*, 1999; Majors and Sedlacek, 2001; Nunes *et al.*, 2015), guaranteeing the stability of the instrument, as well as confirming the internal consistency of each of the factors evaluated, as was presented in Section 5.1.4 with the respective Cronbach's α.

Although the SEEQ instrument, used as basis for QEOn development, has nine factors, the updated QEOn presented a different structure, assessing the peculiarity of the relations in the distance education, and covering important characteristics of this teaching and learning modality.

Another important feature presented in Table 8 is the fact that the component matrix needed to be rotated to improve the visualization of loading factors. The rotation used was the orthogonal Equamax rotation, which allowed to changing of the direction of vectors contained in the matrix loadings, guaranteeing a better visual representation of the factorial loadings according to the six analyzed components. This method change the coordinate axes keeping the orthogonality among them, in such way that the components are still uncorrelated.

It is important to emphasize that some statements have high loadings in other components besides that component whose statement was assigned, as can be observed in the statements Q5, Q8, Q16, Q25 and Q27. In this way, we can attest the existence of a hybrid statement, and it can be classified as belonging to more than one component. The classification of each statement is based on the information of each evaluator, when he/she analyzes it they intrinsically choose to allocate it in a cluster of greater similarity of content.

Then, as a final result of this section we have a factorial structure of the updated QEOn questionnaire with six factors.

Table 8: Factor Analysis with Equamax Rotation. Commonalities and Cronbach's α.

Factors		Components						Commonality	Cronbach's α
		1	2	3	4	5	6		
Online Instructor Profile	Q5	**0.590**	0.525	0.245	0.099	0.223	0.052	0.746	0.948
	Q6	**0.602**	0.484	0.232	0.144	0.227	0.136	0.743	
	Q7	**0.608**	0.394	0.270	0.207	0.283	0.087	0.730	
	Q8	**0.574**	0.509	0.301	0.160	0.230	0.080	0.767	
	Q9	**0.557**	0.406	0.261	0.164	0.317	0.083	0.678	
	Q12	**0.545**	0.314	0.201	0.312	0.242	0.123	0.606	
	Q21	**0.569**	0.362	0.295	0.145	0.282	0.114	0.658	
	Q22	**0.623**	0.347	0.280	0.263	0.197	0.171	0.725	
	Q23	**0.524**	0.293	0.203	0.300	0.134	0.127	0.526	
	Q24	**0.673**	0.297	0.22	0.274	0.215	0.138	0.730	
Online Instructor and Students	Q13	0.201	**0.759**	0.208	0.258	0.116	0.128	0.756	0.904
	Q14	0.202	**0.725**	0.171	0.315	0.197	0.072	0.740	
	Q15	0.221	**0.770**	0.192	0.279	0.171	0.101	0.796	
	Q16	0.488	**0.549**	0.23	0.23	0.205	0.183	0.721	
Interaction Learning	Q1	0.266	0.002	0.127	0.228	**0.632**	0.143	0.559	0.788
	Q2	0.109	0.117	0.072	0.049	**0.782**	0.085	0.652	
	Q3	0.188	0.083	0.1	0.133	**0.711**	0.104	0.586	
	Q4	0.133	0.22	0.003	0.197	**0.581**	0.202	0.484	
	Q11	0.2	0.215	0.18	0.355	**0.556**	0.090	0.562	
Presential Instructor and Students	Q17	0.071	0.203	**0.840**	0.133	0.048	0.098	0.784	0.921
	Q18	0.155	0.146	**0.88**	0.112	0.134	0.105	0.861	
	Q19	0.152	0.11	**0.858**	0.112	0.13	0.097	0.812	
	Q20	0.132	0.092	**0.854**	0.169	0.052	0.057	0.792	
Interaction Assessment and Evaluation	Q10	0.113	0.246	0.072	**0.496**	0.263	0.165	0.420	0.821
	Q25	0.508	0.117	0.189	**0.551**	0.148	0.088	0.642	
	Q26	0.352	0.057	0.217	**0.628**	0.197	0.169	0.637	
	Q27	0.506	0.176	0.241	**0.549**	0.134	0.196	0.703	
	Q28	0.010	0.246	0.126	**0.596**	0.194	0.174	0.502	
	Q29	0.025	0.296	0.171	**0.633**	0.17	0.195	0.584	
Self-Assessment	Q30	0.051	0.042	0.057	0.37	0.156	**0.496**	0.414	0.747
	Q31	0.109	0.048	0.087	0.095	0.068	**0.748**	0.593	
	Q32	0.01	0.379	0.330	0.090	0.218	**0.705**	0.557	
	Q33	0.095	0.369	0.057	0.290	0.104	**0.745**	0.580	
	Q34	0.539	0.218	0.086	0.121	0.048	**0.669**	0.522	

5.2 Principal Component Analysis on Statements

The Principal Component Analysis (PCA) was applied to the database that converts the statements as variables. The idea is to perform an exploratory analysis between the variables and objects that are being used by this research. For all subsequent analyzis the data was preprocessed through autoscaling.

5.2.1 Scree Plot - On Statements

To analyze the PCA in this application, we need to select the number of principal components that will be part of the analysis. In Figure 7, it can be observed that the first two components present a large part of the variance of the whole sample, however, when using the Kaiser criterion (Eigenvalue > 1), we select the first 6 components that represent 65,27% of the variance cumulative and the Root Mean Square Error of Cross-Validation (RMSECV) is 1.682.

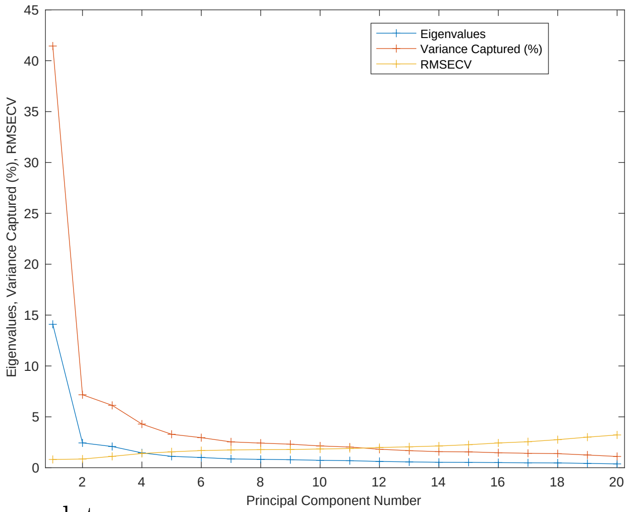


Figure 7: Scree plot

5.2.2 Score Plots

Analyzing the score plots generated by the relations of the principal components (PC) taken in pairs, we can associate each of the graphics to each of the features raised in the characterization of the sample (Section 4.2.2). An important fact to note is that as we know that the PC1 has the largest explained variance (41.44%), it was chosen to become the standard x-axis, producing an analysis of PC2 (7.19%), PC3 (6.04%), PC4 (4.30%), PC5 (3.29%), PC6 (2.95%) and Q Residuals (34.74%) comparing to PC1. Just to clarify, the Q Residuals is the sum of squares of each sample in the error matrix and it helps us to explain if the model is describing well a given sample well.

The first group of labels analyzed, presented in Figure 8, is related to the age range. As can be seen in the dispersion of the data in all the graphs presented, the points that do not make up the great mass of respondents are mostly students with an age range between 40 - 50 and 20 - 30 years old.

The second group of labels analyzed was related to the evaluated course. Among the three evaluated courses, it can be observed in Figure 9 that the PC1 $\times$ PC2 graph shows that the GDE course has few respondents outside the cluster of points in the

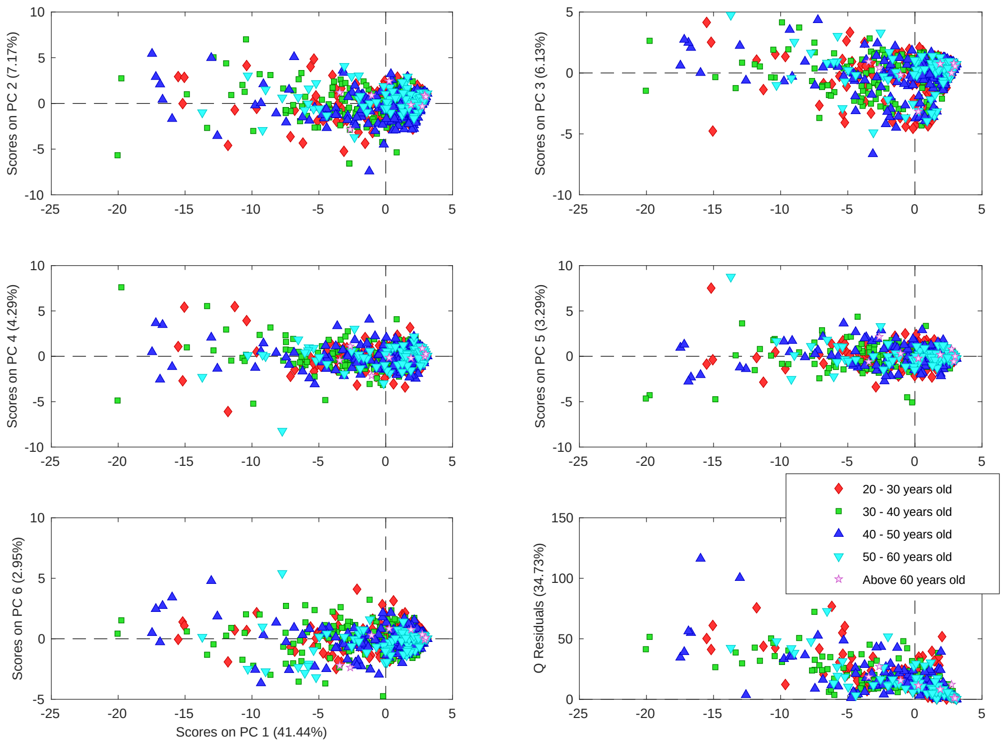


Figure 8: Score plot - Age Range - 6 components

center of the graph and this variability can be observed with greater clarity throughout PC1.

The students who were taking the courses for the first time in blended learning had the information related to their answers grouped around the central axis of the graphs presented in Figure 10. It is worth mentioning that in all the graphs, there is a visible dispersion of the students who carried out between 1 - 2 and 3 - 4 courses.

The gender discussion is constant in research involving educational evaluation (Dweck and Bush, 1976). In this way, we analyze the difference in the response of male and female. In all the graphs in Figure 11, it can be seen that male's responses are much more concentrated than female's. This point deserves a more detailed investigation into what is causing this difference.

As most students scored between 8.1 - 10 on the final grade (see Table 4), there is a dominance of the points on that label. However, we can observe a great variability of the data of the students that obtained a mean of 6,1 - 8 throughout the PC1, as can be seen in Figure 12.

No further information could be extracted from the processed data. Next, the loading plots will be analyzed and the latent relationship among the statements will be verified.

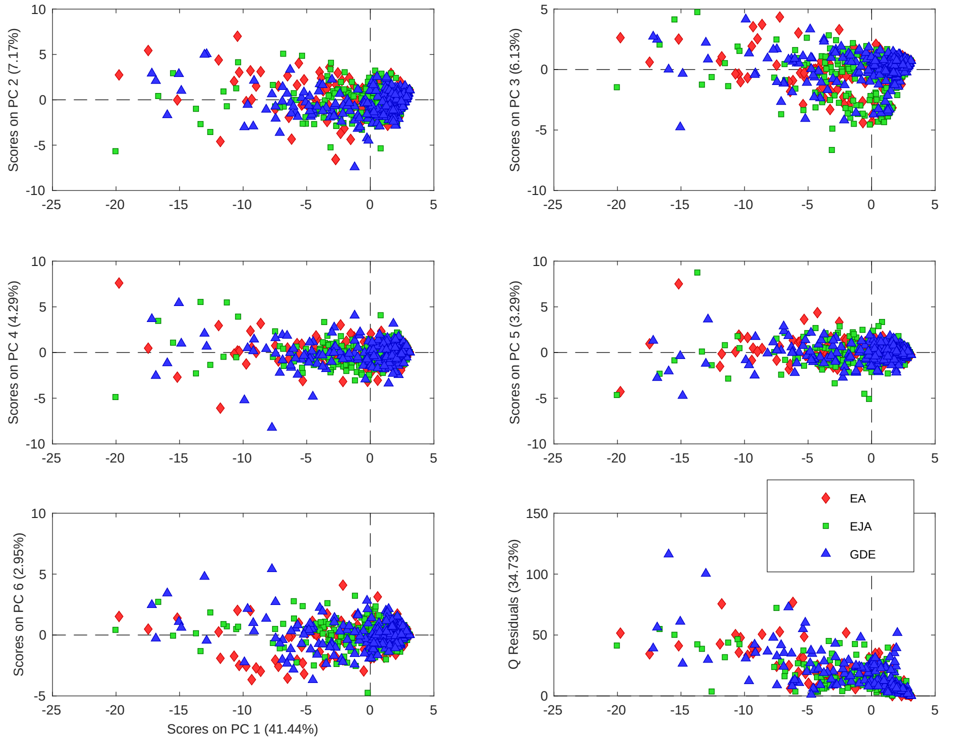


Figure 9: Score plot - Courses - 6 components.

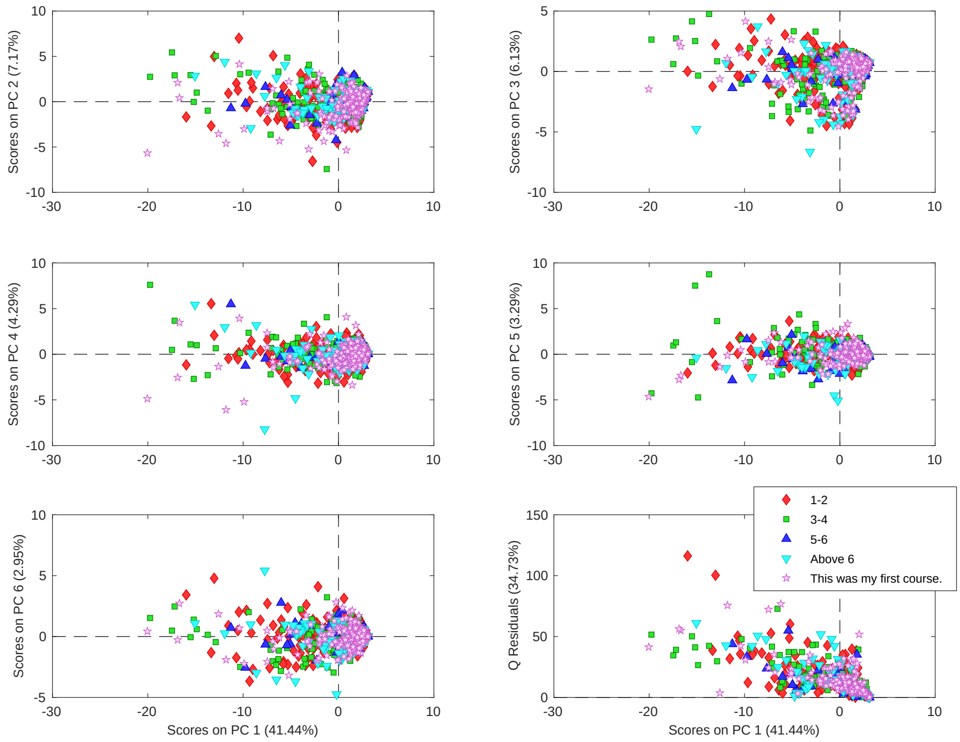


Figure 10: Score plot - Blended courses that have been taken by the students so far - 6 components.

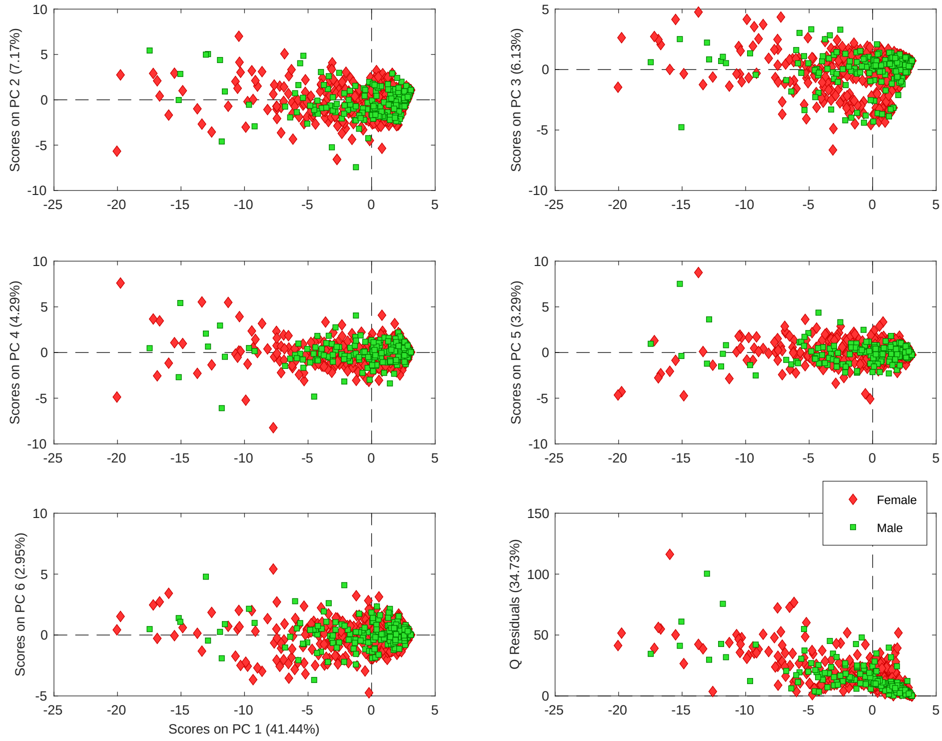


Figure 11: Score plot - Gender - 6 components.

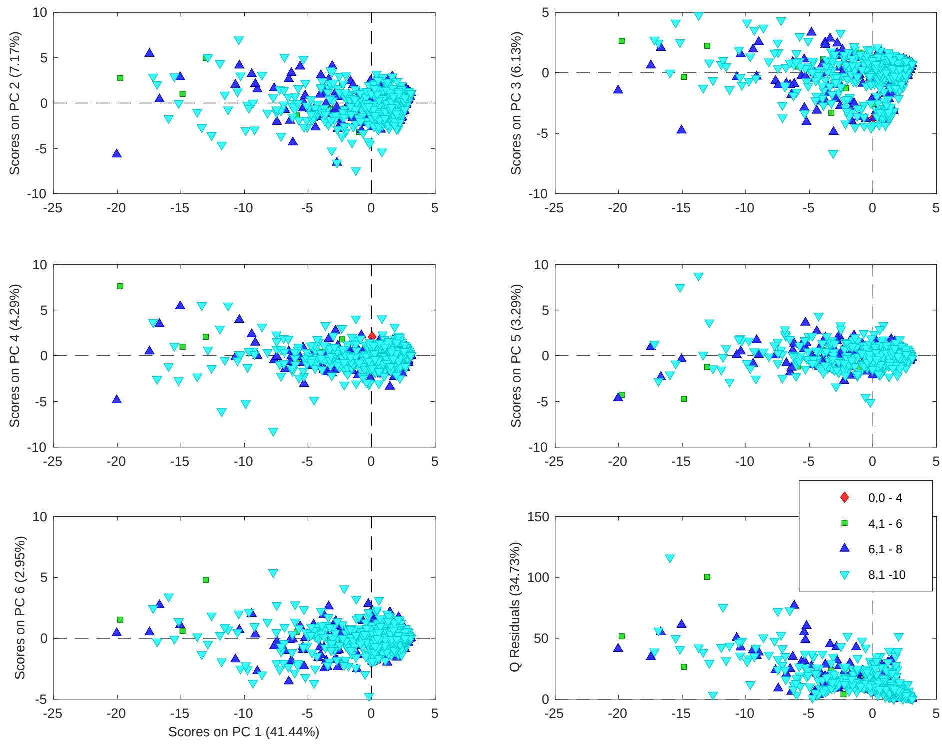


Figure 12: Score plot - Range of the final grade obtained by the students - 6 components.

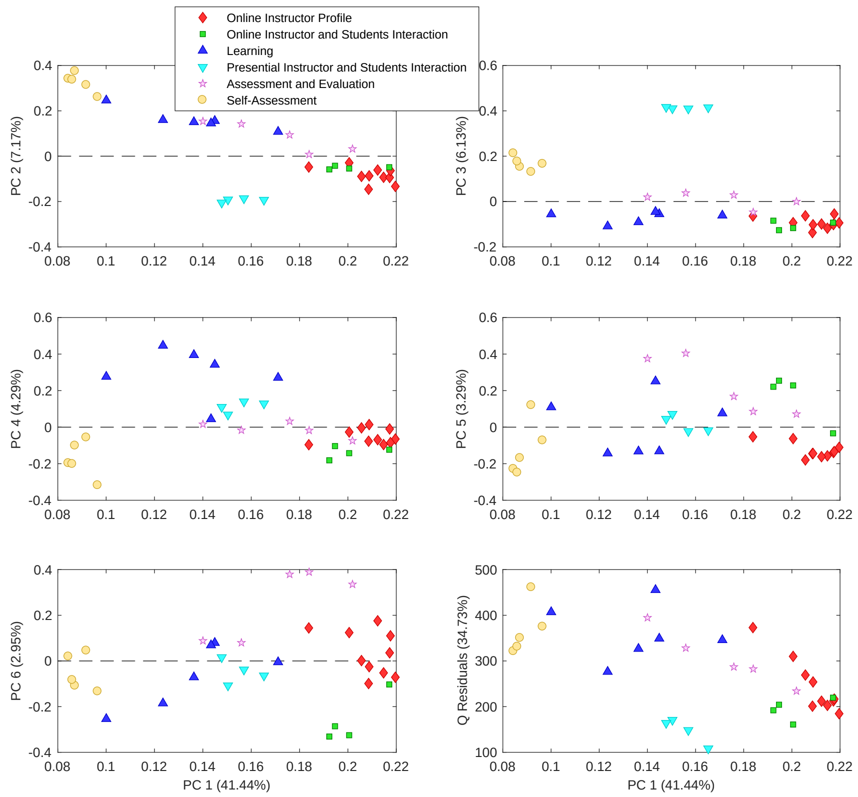


Figure 13: Loading plots - 6 components

5.2.3 Loading Plots

The loading plots verify the relationships among the analyzed variables from the collected database. In this analysis, the 34 statements of the updated QEOn questionnaire are considered.

As presented in Section 3.4.1.2, PCA has the ability to separate observed variables into latent structures, and can create groupings of the observed variables and classify them from a linear combination. In Figure 13, it can be observed that PC1 (41.50%) can retain much of the variability of the statements. However, in the PC1 $\times$ PC2 graph, the creation of the Presential Instructor and Student Interactions and the Self-Assessment clusters are very or clearly visible. This classification was based on the results originated by the application of the factorial analysis (see Section 5.1.5).

Still analyzing the Figure 13, we emphasize that the Learning and Online Instructor Profile indicators can be easily detected in the PC1 versus PC4 chart. The other two indicators, Online Instructor and Students Interaction and Assessment and Evaluation, could not be clearly visualized in the biplot representation.

In order to try to obtain a better representation that could improve the view of the clusters formed with the PCA application, in Figure 14 we have a three-dimensional chart of the variables PC1, PC2 and PC3, and it is possible to clearly establish the clustering of the Assessment and Evaluation indicator, in addition to the others already

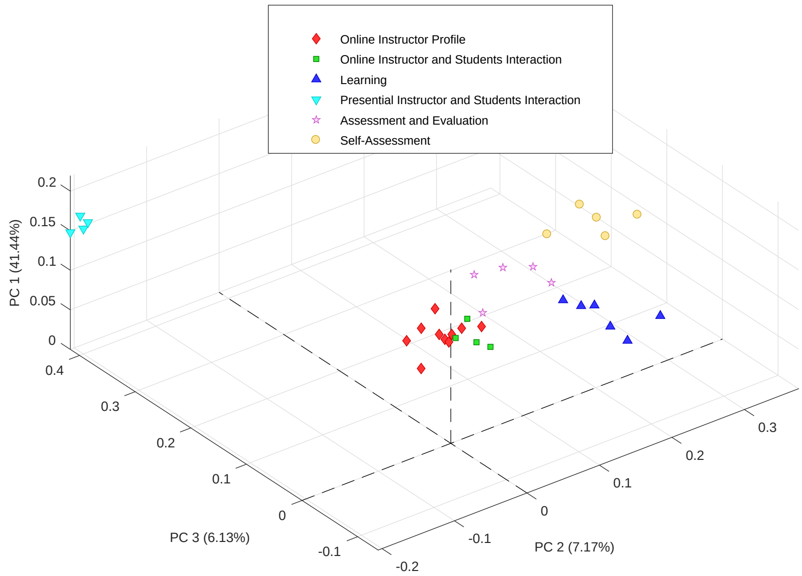


Figure 14: Loading triplot - PC1 × PC2 × PC3

identified in Figure 13.

In the following section, the PCA will be applied taking into account the variables as the QEOn indicators.

5.3 Principal Component Analysis on QEOn Indicators

IAs done in Section 5.2, PCA will now be applied to the database that takes into account the QEOn indicators as variables. In this way, we try to identify how these indicators are intrinsically related and what is the educational relevance of this result.

For the extraction of the QEOn indicators, the mean value was computed among the statements that make up each indicator. This approach is valid only because of the accurate results presented by the degree of reliability of the indicators, Cronbach's α (see Section 5.1.4).

5.3.1 Scree Plot - QEOn Indicators

In order to select the number of principal components to be analyzed in this section, we observe Figure 15.The first two principal components present a cumulative variance of 73.87% of the original data (RMSECV = 0.9782). Although the second component does not satisfy the Kaiser criterion, we include it in order to enrich the analysis, since it holds 13.81% of the information contained in the original variables.

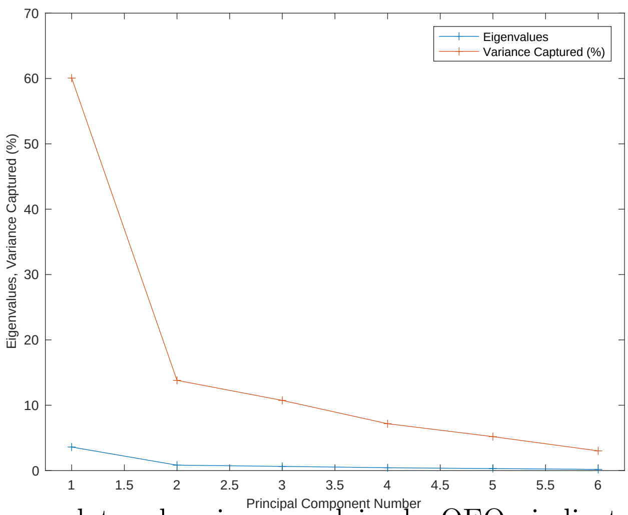


Figure 15: Scree plot and variance explained - QEOn indicators.

5.3.2 Score Plot

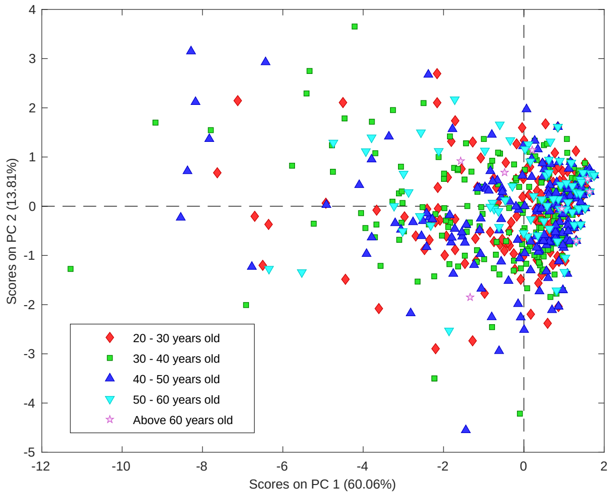
Since we have only two principal components selected, PC1 and PC2, with explained variance of 60.03% and 13.81%, respectively, the labels Age Range, Courses and Gender were used to try to identify hidden patterns in the score plot (see Figure 16). Most of the information was compressed in order to obtain the QEOn indicator and this compression excluded some internal variance by making the analysis even more difficult, i.e., the options in Gender label cannot be easily seen in Figure 16c.

5.3.3 Loading Plot

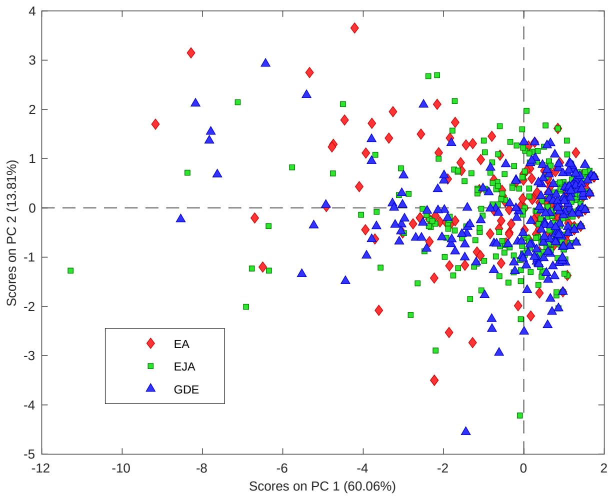
The 6 QEOn indicators are considered as observed variables in this PCA. The Figure 17 shows the latent relationship among the variables according to the biplot PC1 $\times$ PC2.

Based on the previous knowledge we have, some indicators have, conceptually, a high degree of similarity since they evaluate different aspects of the same object, such as the online instructor case that has its profile and interaction with the students analyzed. This similarity is observed in Figure 17 and it also brings the Self-Assessment as the most different indicator when compared to the others. Somehow it attests that the Self-Assessment indicator does not highly influence the other ones, and vice versa. It is a personal characteristic of each student and cannot be developed during this modality of learning.

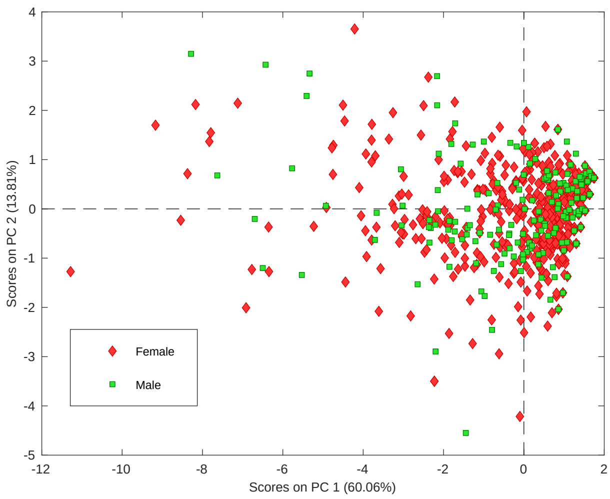
In the next section, the three-way Parafac2 model will be applied in order to understand the intrinsic relationships of each of the analyzed courses.



(a) Age Range



(b) Courses



(c) Gender

Figure 16: Score plot with previous labels obtained.

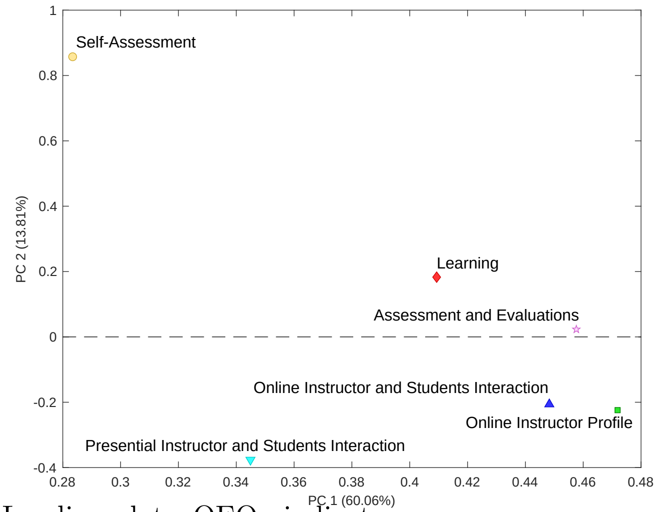


Figure 17: Loading plot - QEOn indicators.

5.4 Parafac2

A new way of analyzing the data collected is based on a three-way structure. As the response matrices, associated with each analyzed courses, have a different number of objects (respondents), we have chosen to use the tensor decomposition Parafac2 to extract patterns associated with the analyzed variables, according to each tensor-mode.

5.4.1 Model Order Selection

As in previous models, one of the main points of the Parafac2 model is the selection of the number of components, also called as model order.

For our application, we used the Core Consistency Diagnostic (Corcondia) test (Bro and Kiers, 2003), which, from the general Tucker3 model, establish the superdiagonal constraint of the $\underline{\mathbf{G}}$-core tensor and verified if the estimated core tensor, by the Parafac2 model, after adding components, is a superdiagonal tensor. If it is not then the order of the model is incorrect. Then, the correct number of components is the maximum number reached till the core tensor changed its superdiagonal structure.

The core consistency plot (see Figure 18) represents a the core tensor $3 \times 3 \times 3$ superdiagonal with ones and all the remaining elements are close to zero. The circles are ideally non-zero target, representing the number of selected components and the stars are components with close to zero representativeness. As can be seen in Figure 18,the first three components achieve an internal core consistency of 100%, which is the ideal number of components for our dataset. Iterations terminated based on fit error relative change.

The variation per component can be seen in Figure 19. The components in Parafac2 are non-orthogonal meaning that the variances are not additive. The total variation can be seen in the blue bar and the unique variation is in the yellow bar, the least one means the variation which is not correlated to any of the other components.

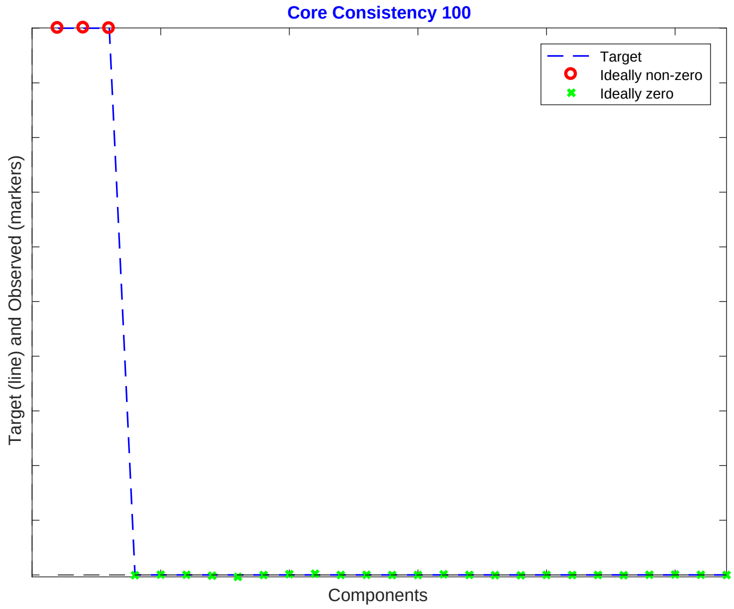


Figure 18: Corcondia 3 components.

For our dataset, the diference between the two bars is very small. When you get a model which is not correct we can usually see a large diference between the two bars. Then according to what we see, we may confirm that the number of components is correct.

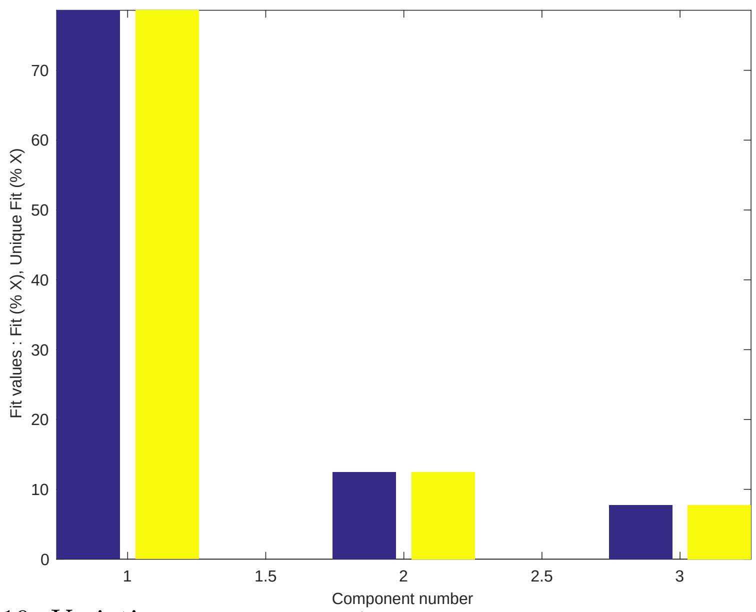
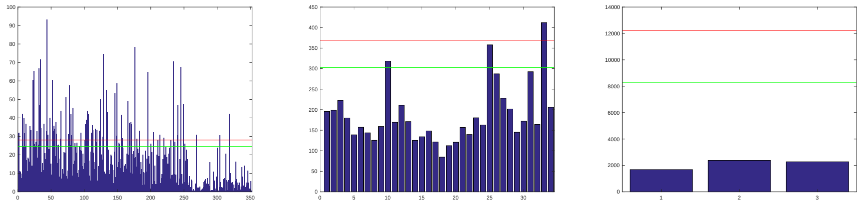


Figure 19: Variation per component.

Figure 20 represents the residual sum of squares of tensor $\underline{\mathbf{E}}$ by each mode (see Figure 6). If one of the samples has an unusually high sum squared residual variation then that sample would be an outlier. In mode 1, we have some unusual residuals related to some of the respondents. If we consider the data as personal information of each student, this variation may be associated with completely different opinions regarding the



(a) Mode 1. (b) Mode 2. (c) Mode 3.

Figure 20: Residual sum square by each mode

statements provided.

As a final result of this section, we guarantee a good model with 3 components.

5.4.2 Loadings

As we are using the Parafac2 decomposition, for each of the 3 components selected, we will have a different sample for each $k-$slabs of the tensor shown in Figure 6. Based on this, Figure 21 represents the signature in students domain, which means how is the pattern variation of students' opinion across each blended course.

We can observe that in component 1, the loadings are concentrated on the positive part of the graphs. It is important to note that we have different sizes for each slab, so the top and bottom graphs keep up where there is still information to be treated. The first component for each sample presents, on its majority, a non-zero loadings, meaning that the weight of this component is more representativeness when comparing to the others.

Figure 22 represents the signature in QEOn statements domain. Four areas of the figure were highlighted because they presented a different behavior among the analyzed components. Parts I and III show a very marked variation at component 3, most of which are below the other two components in Part I and well above the other components in Part III. This variation indicates that the information contained in component 3 is related to the classification of the Online Instructor Profile (Q5, Q6, Q7, Q8, Q9, Q12, Q21, Q22, Q23 and Q24), Learning (Q1, Q2, Q3, Q4 , Q10 and Q11) and Self-Assessment (Q30, Q31, Q32, Q33 and Q34).

Part II, highlighted in component 2, is associated with the classification of the indicator Presential Instructor and Students Interaction (Q17, Q18, Q19 and Q20). Finally, Part IV is highlighted because it presents a very similar behavior between components 1 and 2 with high correlation, and another component is necessary to explain the difference between the assertions of the Self-Assessment indicator.

Loading and Score plots are the same in the 3rd mode. Figure 23 represents the signature in the courses domain. We have three scores for the first course (EA), which has a high value in component 1, an average value for component 3 and a negative value for component 2. Since each course presents a non-zero value of the components, we can

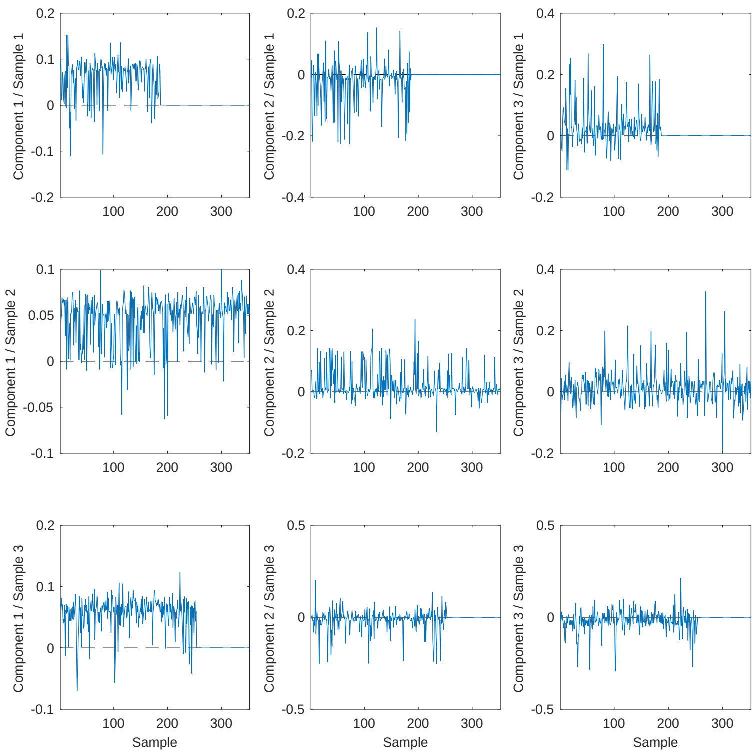


Figure 21: Loading Mode1

consider that we have components related to uncorrelated signature, that is, the analysis of the components need to be done together. Such analysis makes sense given that data related to educational contexts are complex and intrinsically interrelated, some with a high and others with a low correlation.

The following section presents the discussion of the results obtained as well as the comparison with the current literature on the subject of evaluation on blended learning courses.

5.5 Are these results reliable according to the literature?

Also according to Table 8, the six components generated by the FA can be considered as intrinsic indicators of the QEOn, which are related to the quality of online education. In this way, each indicator can be characterized as follows:

- **Online Instructor Profile -** This indicator is directly related to the quality of tutoring by the online instructor. If this online instructor demonstrates dynamism, enthusiasm, not being limited to the content available on the platform, attentiveness to the student's learning, presenting the content in a satisfactory way, relate theory and practice, they are guaranteed to obtain a good indicator score (Sarmel and Abrahão, 2007; Attwell, 2006).
- **Online Instructor and Students Interaction -** This indicator measures how

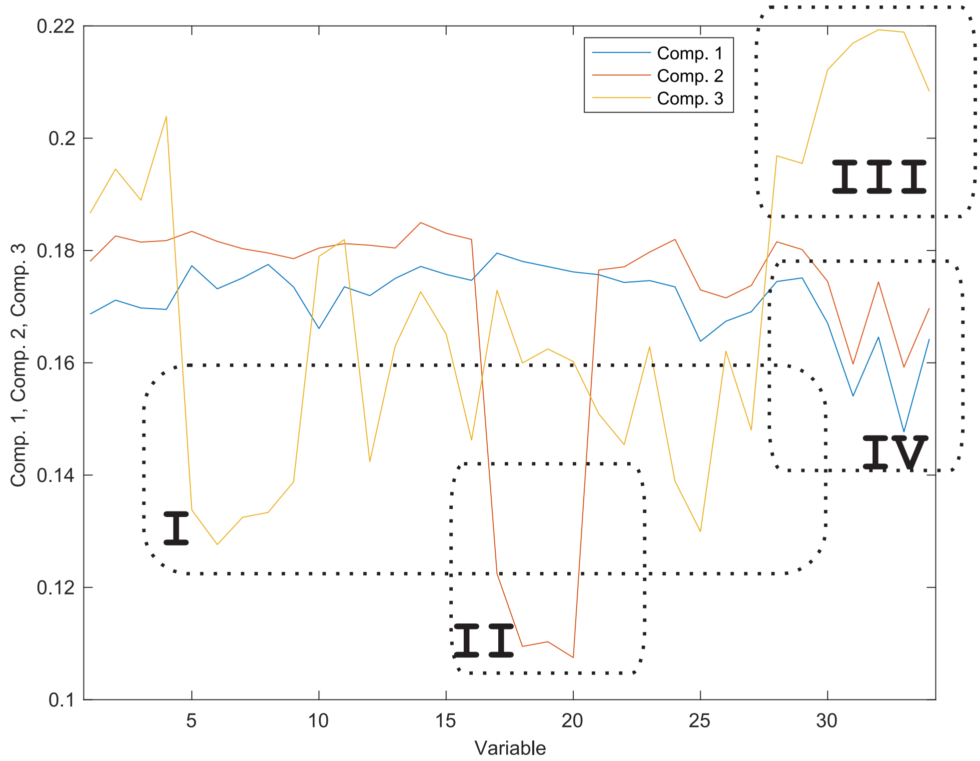


Figure 22: Loading Mode2

stimulating the online instructor is in the forum discussions. Some researchers advocate online instructor - courser interaction as being one of the key factors for online education success (Sarmel and Abrahão, 2007; Attwell, 2006).

- **Learning -** The level of learning acquired by the student is measured by this indicator. The students themselves punctuate their learning along the course, and this indicator can be considered as a student self-assessment regarding the knowledge retained (Ribeiro *et al.*, 2013).
- **Presential Instructor and Students Interaction -** In some structures of online education there is the figure of the presential instructor, who works as a support professional present in the poles of face-to-face support. This indicator measures the relationship between the presential instructor and the student (Barni, 2011).
- **Assessment and Evaluation -** The evaluative processes in online education are still motives for recent research, since it is necessary to have a systematic evaluation according to the modality of teaching, taking into account the different tools present in the VLE's (Laguardia, Portela, and Vasconcelos, 2007; Sales *et al.*, 2011; Andriola and Loureiro, 2005). This indicator is related to the adequacy of the evaluation processes to the online course.
- **Self-Assessment -** reflects the autonomy of the student throughout the formative process. Students in distance learning need this autonomy to play their central roles as apprentices, making them a partner of the teacher (Jara and Mellar, 2010).

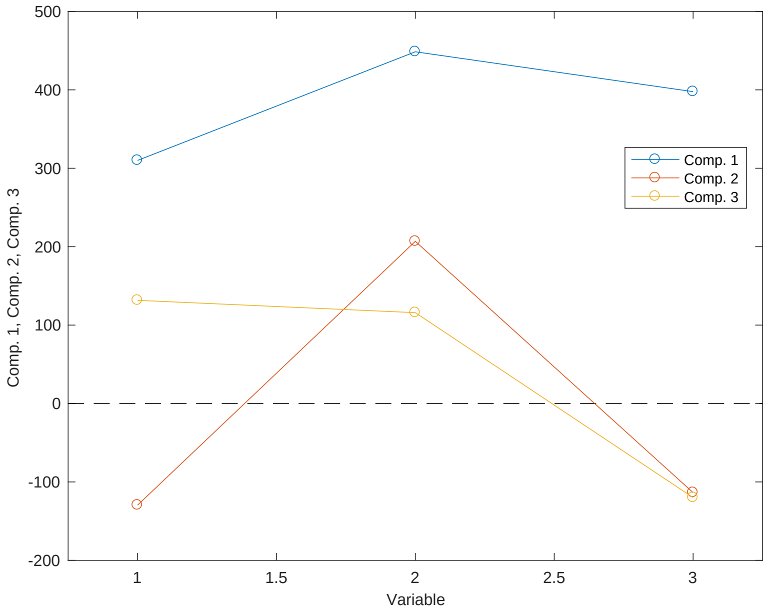


Figure 23: Score plot - Courses

The research cited in the above topics (Sarmel and Abrahão, 2007; Attwell, 2006; Ribeiro *et al.*, 2013; Barni, 2011; Laguardia, Portela, and Vasconcelos, 2007; Sales *et al.*, 2011; Andriola and Loureiro, 2005) present results related to the quality indicators obtained by this research. However, in these surveys there is no mathematical model that describes how these factors can be constituted. Carle (2009) evaluates college students' evaluations of teaching effectiveness across time and instruction mode, which could be online and face-to-face, using a multilevel growth modelling approach. In fact this is quite what we want to do here and this research could be classified in the educometrics field.

The PCA and Parafac2 presented interesting results from the relationship among the statements and among the quality of education indicators. In PCA, the two indicators associated with the online instructor appeared with a high proximity, considering that the two indicators evaluated the comparison of the same professional. In Parafac2, with only 3 components, the behavior of the statements' clustering can be detected along the third component, which in with Figure 22 is divided into 4 parts.

These indicators become fundamental for the various spheres of teaching management to work on improving online courses, including and providing feedback to the instructors involved in the course (Jara and Mellar, 2010).

5.6 Summary

In this chapter, three mathematical models were used to analyze the data collected. The factorial analysis was able to validate the latent structure of 6 QEOn questionnaire indicators. The principal component analysis revealed proximity among the obtained indicators, keeping the latent structure when crossed through a triplot graph. The Parafac2 model presented the signature for each of the modes analyzed, based on a 3-component model.

6 FINAL REMARKS AND PERSPECTIVES

This work has strong adherence to the current organizational situation in Brazil, with regards to the area of teacher training, initial or continuing formation, with the use of the blended learning modality or online education. Thus, verification of the quality of the courses offered in the blended mode becomes essential for this training to be successful, corroborated with the training network of public education professionals.

The development of the updated QEOn is aimed to identify and compare the students' perceptions regarding the course offered in the blended learning courses. These perceptions can be identified through FA, which enables the extraction of 6 quality indicators from blended learning courses. The analysis of these indicators makes it possible to make decisions, based on more consistent criteria, which enables pedagogical interventions, and can consequently reduce dropout rates.

The factors extracted from the QEOn were: Online Instructor Profile, Online Instructor and Students' Interaction, Learning, Presential Instructor and Students' Interaction, Assessment and Evaluation, and Self-Assessment. These factors can be adopted as indicators of the quality of teaching, in order to contribute to the pedagogical-administrative decision making that can optimize resources and efforts in the search for better learning.

This research took into account the overall evaluation of three extension courses, where no specific course was analyzed separately. Facing this fact, we perceive the need to specify the evaluation process, both investigating the course individually, as well as each of the classes. This interest is justified by the need to compare the performance of instructors in order to understand good and bad teaching practices in each class.

Because the QEOn is adaptable, new contributions can be added, enhancing its investigative power, especially with regard to management support. Such concerns stem from the belief that looking at different angles of student feedback analysis can improve the quality of teaching, as advocated by SET.

Some researchers in the field of education are against the use of statistics and mathematical models at the educational process, arguing that the educational phenomenon is complex and any numerical measure attributed to it would not be enough to represent it as a whole. We understand that researchers in that field have difficulty in describing educational processes without any mathematical tool. Large scale assessments prove that it is not only useful, but it also presents how to take the analysis provided by the data collected into account, directing public policies in order to improve educational contexts around the globe.

Although the term Educometrics is not yet very common, researchers in some areas have already looked at the need to analyze educational data. These researchers are usually economists, computer scientists, and for the most part, educational psychologists, among others. It is important to emphasize that there is an area in computer science that deals with educational data mining, which would make a lot of sense in educometric terms, but a great part of the work related to this subject comes from computer systems used in education, especially in distance learning (Romero and Ventura, 2010).

6.1 Future perspectives

We can point to the use of the factors extracted from the QEOn in reducing the dropout rate of blended learning courses and the development of statements that seek to measure the degree of expertise of the respondent. The inclusion of these factors in new research will aim to improve issues related to decision making by managers, to improve the quality of teacher training in the blended mode. Additionally, some other tools must be taken into account in the educational context, such as: coclustering (Papalexakis, Sidiropoulos, and Bro, 2012) and multilevel analysis (Carle, 2009).

It is also necessary to have better development on the performance of educometricians in the world today, so that the recognition of these professionals is obtained. For this, the faculties of education need professors specialized in statistics and education to form generations of professionals capable of dealing with and acting on this new paradigm. It is clear that intersections will be found with other metrics, especially psychometry and sociometry, but it is necessary for researchers, teachers and professors at the educational field to take further steps regarding the comprehension and analysis of well structured educational data since they are the professionals whose are capable to better understand the real applicability of the results in a classroom.

As future perspectives, we may highlight: anticipate the adaptation of this questionnaire for application among teacher as a whole; seek correlations between the students and students' perceptions of the course under analysis; and investigate to what extent the experience of the students in other online courses has in the consistency of the answers (Silva *et al.*, 2012; da Silva *et al.*, 2017).

REFERENCES

Abrami, P. C. *Improving judgments about teachinf effectiveness using teacher rating forms*, New Directions for Institutional Research, v. 109, p. 59–87. 2001.

Acar, Evrim; Yener, Bulent. Unsupervised Multiway Data Analysis: A Literature Survey. *IEEE Transactions on Knowledge and Data Engineering*, v. 21, n. 1, p. 6–20, 2009.

Aleamoni, L. M. Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, v. 13, n. 2, p. 153–166, 1999.

Algozzine, B.; Beattie, J.; Bray, M.; Flowers, C.; Gretes, J.; Howley, L.; Mohanty, G.; Spooner, F. Student evaluation of college teaching: A practice in search of principles. *College Teaching*, v. 52, n. 4, p. 134–141, 2004.

Almeida, A. L. F.; Favier, G.; Mota, J. C. M. Constrained Tensor Modeling Approach to Blind Multiple-Antenna CDMA Schemes. *IEEE Transactions on Signal Processing*, v. 56, n. 6, p. 2417–2428, 2008.

Anderson, Lorin W.; Krathwohl, David R.; Airasian, Peter W.; Cruikshank, Kathleen A.; Mayer, Richard E.; Pintrich, Paul R.; Raths, James; Wittrock, Merlin C. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Abridged Edition.* Pearson, 2. ed., 2000.

Andriola, W. B.; Loureiro, R. C. Sistematização da avaliação da aprendizagem em comunidades organizadas no ciberespaço. *Revista Iberoamericana de Educación*, v. 37, n. 1, p. 1–6, 2005.

Andriola, Wagner B. Evaluación: La vía para la calidad educativa. *Ensaio. Avaliação e Políticas Públicas em Educação*, v. 7, n. 25, p. 335–368, 1999.

Andriola, Wagner B. Factores caracterizadores de centros educativos eficaces. *Bordón: Revista de Pedagogía*, v. 53, n. 2, p. 175–183, 2001.

Andriola, Wagner B. Detección del funcionamiento diferencial del ítem (DIF) em tests de rendimiento. Aportaciones teóricas e metodológicas. *Doctoral Thesis - Universidad Complutense de Madrid*, 2002.

Attwell, Graham. *Evaluating e-learning: a guide to the evaluation of the e-learning*, v. 2 of *Evaluate Europe Handbook Series.* European Commission, 2006.

Baldwin, T.; Blattner, N. Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching*, v. 51, n. 1, p. 27–32, 2003.

Barni, E. M. O papel do tutor presencial na Educação a Distância. *Procedings of X Congresso Nacional de Educação*. Curitiba, 2011, p. 10671–10683.

Beran, T.; Violato, C.; Kline, D. What's the 'use' of student ratings of instruction for administrators? One university's experience. *Canadian Journal of Higher Education*, v. 17, n. 1, p. 27–43, 2007.

Beran, T.; Violato, C.; Kline, D.; Frideres, J. The utility of student ratings of instruction for students, faculty, and administrators: A "consequential validity study". *Canadian Journal of Higher Education*, v. 35, n. 2, p. 49–70, 2005.

Bloom, B. S. Some major problems in educational measurement. *Journal or Educational Research*, v. 38, n. 1, p. 139–142, 1944.

Bloom, B. S. What we are learning about teaching and learning: a summary of recent research. *Principal*, v. 66, n. 2, p. 6–10, 1986.

Bloom, B. S.; Hastings, J. T.; Madaus, G. F. *Handbook on formative and sommative evaluation of student learning*. New York: McGraw-Hill, 1971.

Bloom, Benjamin S.; Engelhart, Max D.; Furst, Edward J.; Hill, Walker H.; Krathwohl, David R. *Taxonomy Of Educational Objectives: Handbook 1, The Cognitive Domain*. Boston: Allyn and Bacon, 1956.

Braskamp, L. A.; Ory, J. C. *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass, 1994.

Bro, Rasmus; Kiers, Henk A. L. A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics*, v. 17, n. 5, p. 274–286, 2003.

Campbell, J. P.; Bozeman, W. C. The value of student ratings: Perceptions of students, teachers and administrators. *Community College Journal of Research and Practice*, v. 32, n. 1, p. 13–24, 2008.

Carle, A. C. Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach. *Computers & Education*, v. 53, p. 429–435, 2009.

Carroll, J. D.; Green, P. E.; Chaturvedi, A. *Mathematical Tools for Applied Multivariate Analysis*. United States of America: Elsevier Science, 2. ed., 1997.

Carroll, J. Douglas; Chang, Jih Jie. Analysis of individual differences in multidimensional scaling via an Nway generalization of "Eckart Young" decomposition. *Psychometrika*, v. 35, n. 3, p. 283–319, 1970.

Cattell, R. B. Parallel proportional profiles and other principles for determining the choice of factors by rotation. *Psychometrika*, v. 9, p. 267–283, 1944.

Centra, J. A. *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. Jossey-Bass, 1993.

Cichocki, Andrzej; Zdunek, Rafal; Phan, Anh Huy; Amari, Shunichi. *Nonnegative Matrix and Tensor Factorizations*. Wiley, 2009.

Clayson, D. E. Student Evaluations of Teaching: Are They Related to What Students Learn? *Journal of Marketing Education*, v. 31, n. 1, p. 16–30, 2009.

Cronbach, L. J. Coeficient alpha and the internal structure of tests. *Psychometrika*, v. 16, p. 297–334, 1951.

da Silva, T. E. V.; Ribeiro, G. O.; Nunes, A. O.; Vasconcelos, F. H. L.; Andriola, W. B.; Mota, J. C. M. QEOn Questionnaire for Assessing Experiences in Virtual Learning Environments. *IEEE Latin America Transactions*, v. 15, n. 6, p. 1197–1204, 2017.

de Leeuw, Jan. R in Psychometrics and Psychometrics in R. Presentation, 2006.

Dweck, C. S.; Bush, E. S. Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. *Developmental Psychology*, , n. 12, p. 147–156, 1976.

Eiszler, C. F. College students' evaluations of teaching and grade inflation. *Research in Higher Education*, v. 43, n. 4, p. 483–501, 2002.

Fabrigar, L. R.; Wegener, D. T.; Maccallum, R. C.; Strahan, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, v. 4, n. 3, 1999.

Feldman, K. A. Grades and college student's evaluations of their courses and teachers. *Research in Higher Education*, v. 4, p. 69–111, 1976.

Franco, A. O. R.; Silva, T. E. V.; Vasconcelos, F. H. L.; Paiva, I. R. Multilinear Decomposition Application into Students' Evaluation of Teaching Effectiveness. *Procedings of IEEE Global Engineering Education Conference*. Berlin, 2013, p. 917–923.

Franklin, J. *Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately*, New Directions for Teaching and Learning, p. 85–100. 2001.

Franklin, J.; Theall, M. Who reads ratings: Knowledge, attitude and practice of users of student ratings of instruction. *Procedings of Annual meeting of the American*

Educational Research Association. San Francisco, 1989.

Gibbs, G.; Habeshaw, S.; Habeshaw, T. *53 interesting ways to appraise your teaching.* Technical and Educational Services, 1988.

Gorsuch, R. L. Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, v. 68, n. 3, p. 532–560, 1997.

Hervas, A.; García, F. B.; Penalvo, F. J. G. A Method of Assessing Academic Learning Experiences in Virtual Learning Environments. *IEEE Latin America Transactions*, v. 12, n. 2, 2005.

Hodges, L. C.; Stanton, K. Translating comments on student evaluations into the language of learning. *Innovative Higher Education*, v. 31, p. 279–286, 2007.

Jara, Magdalena; Mellar, Harvey. Quality enhancement for e-learning courses: The role of student feedback. *Computers & Education*, v. 54, p. 709–714, 2010.

Kiers, Henk A. L.; ten Berge, Jos M. F.; Bro, Rasmus. PARAFAC2 - Part I. A direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics*, v. 13, n. 3-4, p. 275–294, 1999.

Kolda, T. G.; Bader, B. W. Tensor Decompositions and Applications. *SIAM REVIEW*, v. 51, n. 3, p. 455–500, 2009.

Kroonenberg, Pieter M. *Applied Multiway Data Analysis.* Wiley, 2008.

Laguardia, J.; Portela, M. C.; Vasconcelos, M. M. Avaliação em ambientes virtuais de aprendizagemEvaluación: La vía para la calidad educativa. *Educação e Pesquisa*, v. 33, n. 3, p. 513–530, 2007.

Lathauwer, L. De; Moor, B. De; Vandewalle, J. A multilinear singular value decomposition. *SIAM. J. Matrix Anal. Appl.*, v. 21, p. 1253–1278, 2000.

Liaw, S. S. Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers and Education*, p. 864–873, 2008.

Linke, R. D. Performance indicators in higher education: report of a Trial evaluation study commissioned by the Commonwealth Department of Employment, Education and Training. *Australian Government Publishing Service*, 1991.

Lorge, Irving. Thorndike's Publications from 1940 to 1949: A Bibliography. *Teachers College Record*, v. 51, n. 1, p. 42–45, 1949.

Lulla, B.P. *Concept of Educometrics*. B.I.E. Publications Unit, Bombay Teachers' Training College, 1980. URL https://books.google.com.br/books?id=UV9SGwAACAAJ.

Majors, M. S.; Sedlacek, W. E. Using factor analysis to organize student services. *Journal of College Student Development*, v. 42, n. 3, 2001.

Marsh, H. W. Students' Evaluations of Universisty Teaching: Research Findings, Methodological Issues and Directions for Future Research. *International Journal of Educational Research*, v. 11, p. 253–388, 1987.

Marsh, H. W.; Roche, L. A. Making students' evaluations of teaching effectiveness: The critical issues of validity, bias and utility. *American Psychologist*, v. 52, p. 1187–1197, 1997.

Marsh, Herbert W.; Bailey, Michael. Multidimensional Students' Evaluations of Teaching Effectiveness: A Profile Analysis. *The Journal of Higher Education*, v. 64, n. 1, p. 1–18, 1993. URL http://www.jstor.org/stable/2959975.

MEC. Ministério da Educação. Portaria No 1.087, de 10 de agosto de 2011. *Portaria Ministerial*, 2011a.

MEC. Ministério da Educação. Resolução No 1, de 17 de agosto de 2011. *Resolução Ministerial*, 2011b.

MEC. Ministério da Educação. Portaria No 1.105, de 08 de Novembro de 2013. *Portaria Ministerial*, 2013.

MEC. Ministério da Educação. Decreto No 8.752, de 6 de Maio de 2016. *Decreto Ministerial*, 2016.

Meyer, J. H. F.; Muller, M. W. Evaluating the quality of student learning: I. An unfolding analysis of the association between perceptions of learning context and approaches to studying at an individual level. *Studies in Higher Education*, v. 15, p. 131–154, 1990.

Morup, M. Applications of Tensor (Multiway Array) Factorizations and Decompositions in Data Mining. *WIREs Data Mining and Knowledge Discovery*, v. 1, p. 24–40, 2011.

Nasser, F.; Fresko, B. Faculty view of student evaluation of college teaching. *Assessment and Evaluation in Higher Education*, v. 27, n. 2, p. 187–198, 2002.

Nolen, S. B.; Haladyna, T. The Encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, v. 15, n. 2, p. 116–130, 1990.

Nunes, A. O.; Silva, T. E. V.; Mota, J. C. M.; Almeida, A. L. F.; Andriola, W. B. Developing an Instrument for Assessment of Academic Management in Engineering Courses. *IEEE Latin America Transactions*, v. 13, n. 1, 2015.

Nunes, Albano Oliveira. Matrix and Tensorial Methods as Management Support in Engineering and Systems and Digital Media Courses at the Federal University of Ceará (in portuguese). *Doctoral Thesis - Federal University of Ceará*, 2016.

Ory, J. C. *Faculty thoughts and concerns about student ratings*, New Directions for Teaching and Learning, p. 3–15. 2001.

Papalexakis, E.; Sidiropoulos, N. D.; Bro, R. From K-means to higuer-way co-clustering: Multilinear Decomposition with Sparse Latent Factors. *IEEE Transactions on Signal Processing*, v. 61, n. 2, p. 493–506, 2012.

Parsons, P. G. The Lancaster approaches to studying inventory and course perceptions questionnaire: a replicated study at the Cape Technikon. *South African Journal of Higher Education*, v. 2, p. 103–111, 1988.

Ramsden, P. A Study of the relationship between student learning and its academic context. *PhD Thesis*, 1981.

Ramsden, P. A performance indicator of teaching quality in higher education: the Course Experience Questionnaire. *Studies in Higher Education*, v. 16, p. 129–150, 1991.

Ribeiro, G. O.; Silva, T. E. V.; Nunes, A. O.; Pinto, F. A. P.; Vasconcelos, F. H. L. Avaliação da efetividade do ensino em um curso de formação continuada semipresencial. *Procedings of XIX Workshop de Informática na Escola*. Campinas, 2013.

Richardson, J. T. E. Instruments for obtaining student feedback: a review of the literature. *Assessment and Evaluation in Higher Education*, v. 30, n. 4, p. 387–415, 2005.

Rickards, J. P.; Friedman, F. The Encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, v. 3, n. 3, p. 136–143, 1978.

Romero, C.; Ventura, S. *Data Mining in e-Learning*. WIT Press, 2006.

Romero, C.; Ventura, Sebastián. Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, v. 40, n. 6, p. 601–618, 2010.

Ryan, J. J.; Anderson, J. A.; Birchler, A. B. Student evaluation: The faculty responds. *Research in Higher Education*, v. 12, n. 4, p. 317–333, 1980.

Sales, G. L.; Prestes, P. A. N.; Soares, J. M.; Barroso, G. C. Aplicação do Modelo de

Avaliação Formativa Learning Vector (LV) à Ferramenta Wiki do LMS Moodle. *Revista Novas Tecnologias na Educação*, v. 9, p. 1–10, 2011.

Sarmel, M. M.; Abrahão, J. I. O tutor em Educação a Distância: análise ergonômica das interfaces mediadoras. *Educação em Revista*, , n. 46, p. 109–141, 2007.

Silva, T. E. V.; Vasconcelos, F. H. L.; Almeida, A. L. F.; Mota, J. C. M. Multivariate analysis for the students' evaluation of teaching effectiveness in teleinformatics engineering. *Procedings of IEEE International Conference on Teaching, Assessment And Learning For Engineering (TALE)*. Hong Kong, 2012, p. H1A–1 –H1A–6.

Smilde, A.; Bro, R.; Geladi, P. *Multi-way Analysis with Applications in the Chemical Sciences*. Wiley, 2004.

Spooren, Pieter; Vandermoere, Frederic; Vanderstraeten, Raf; Pepermans, Koen. Exploring high impact scholarship in research on student's evaluation of teaching (SET). *Educational Research Review*, v. 22, n. Supplement C, p. 129 – 141, 2017. URL `http://www.sciencedirect.com/science/article/pii/S1747938X17300337`.

Sproule, R. Student evaluation of teaching: A methodological critique of evaluation practices. *Education Policy Analysis*, v. 8, n. 50, 2000.

Sternberg, R.; Zhang, L. Styles of thinking as a basis of diferentiated instruction. *Theory into Practice*, v. 44, n. 3, p. 245–253, 2005.

Struyven, K.; Dochy, F.; Janssens, S. Students' perceptions about evaluation and assessment in higher education: a review. *Theory into Practice*, v. 30, n. 4, p. 325–341, 2005.

Tepper, Jonathan A. Measuring Constructive Alignment: An Alignment Metric to Guide Good Practice. *Procedings of 1st UK Workshop on Constructive Alignment*. United Kingdom: Higher Education Academy Information and Computer Sciences (ICS) Subject Centre and Nottingham Trent University, 2006.

Theall, M.; Franklin, J. Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research*, v. 27, n. 5, p. 45–56, 2001.

Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika*, v. 31, p. 279–311, 1966.

Tyler, Ralph W. *A Rationale for Program Evaluation*, Springer Netherlands, p. 87–96. 2000.

Vasconcelos, F. H. L.; Silva, T E. V.; Almeida, A. L. F.; Andriola, W. B.; Mota, J. C. M. Evaluating the Students' Perception of the Teaching Effectiveness in a

Teleinformatics Engineering. *Procedings of IEEE Global Engineering Education Conference*. Berlin, 2013, p. 910–916.

Vasconcelos, Francisco Herbert Lima. Analysis of Context and Learning Outcomes at an Educational Assessment in an Engineering Undergraduate Course (in portuguese). *Doctoral Thesis - Federal University of Ceará*, 2016.

Wachtel, H. K. Student evaluation of coolege teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, v. 29, n. 2, p. 191–121, 1998.

Wagenaar, T. C. Student evaluation of teaching: Some cautions and suggestions. *Teaching Sociology*, v. 23, n. 1, p. 64–68, 1995.