Validation of the academic management evaluation instrument based on principal component analysis for engineering and technological courses

Validación de un instrumento de evaluación de la gestión académica basado en el análisis de componentes principales para cursos tecnológicos y de ingeniería

A. O. Nunes¹, T. E. V. Silva², J. C. M. Mota³, A. L. F. Almeida⁴, and W. B. Andriola⁵

ABSTRACT

In recent years, the expansion of higher education in Brazil has led to a series of demands related to aspects concerning training at the college level. These processes relate to: academics, professionals, entering in the labor market, among others. In this context, an important aspect is the quality of the courses. Thus, the evaluation becomes a critical diagnostic process of reality and starting point for possible interventions to be put in practice by the coordinators of the programs. This article presents the results of a questionnaire administered at the Federal University of Ceará (UFC), especially to Systems & Digital Media and Engineering Programs professors. This research aims to identify how the professors from each department see the administrative procedures developed by the departments and also investigate the possibility of using Principal Components Analysis (PCA) as a support for management of the higher education training. The methodology included the implementation of Likert scale questionnaire and subsequent mathematical treatment with PCA. The results indicate the potential application of PCA to support the management of higher education; it was possible to extract preliminaries inferences related to management methods and their characteristics. This suggests the possibility of developing the Educametrics field.

Keywords: Courses management, principal components analysis, quality in higher education.

RESUMEN

En los últimos años, la expansión de la educación superior en Brasil ha dado lugar a una serie de demandas relacionadas con la calidad de la formación en el nivel universitario. Estas demandas están vinculadas a los siguientes aspectos: elementos de gestión académica, calidad de los profesionales, mercado de trabajo, entre otros. En este contexto, un aspecto importante es la calidad de los cursos y de la formación de los alumnos. Por lo tanto, la evaluación se convierte en una importante herramienta de diagnóstico de la realidad y también en punto de partida para posibles intervenciones que pueden ser llevadas a cabo por los coordinadores de los cursos. Este artículo presenta los resultados de un cuestionario realizado en la Universidad Federal de Ceará (UFC), especialmente a profesores de los cursos de Sistemas y Medios Digitales e Ingenierías, cuyos objetivos fueron, por un lado, identificar la forma en que los profesores de cada departamento ven los procedimientos administrativos desarrollados en sus departamentos, y por otro lado, investigar la posibilidad de utilizar el Análisis de Componentes Principales (PCA) como apoyo a la gestión de la formación en educación superior. La metodología utilizada incluyó la aplicación del cuestionario tipo Likert y el tratamiento matemático posterior con PCA. Los resultados indican la posible aplicación de las herramientas de análisis para apoyar la gestión educacional y también la posibilidad de extraer inferencias relacionadas con los métodos de gestión y sus características particulares. Esto sugiere la posibilidad de desarrollar el campo de la Educametría.

Palabras clave: Gestión académica, análisis de componentes principales, calidad en la educación superior.

Received: November 18th 2014
Accepted: July 6th 2015

Introduction

Issues related to high dropout and failure rates in higher education are historical in Brazil. However, the advent of shares increased supply of higher education, these issues become even more evident, especially in courses of engineering and related areas (Schuwartman, 1992), (Rompelman, 1999) and (Silva Filho, et al., 2007). In this context, educational...
diagnostic evaluation is emerging as a starting point, becoming prominent for understanding these phenomena and suggesting possible interventions. Throughout this study, we will discuss the process of administration in higher education, specifically the one carried out by the managers of the courses and their role to solve this problem.

From the systemic planning perspective, the academic administration must present a fundamental role for the correct objective to solve problems related to the educational quality provided to the undergraduate students, and their consequent formation for the Labor Market (UNESCO, 1998). This is due to the fact that academic environment takes a broader position than just its instructional services, becoming prominent with regard to the management of development processes (Vanghese, 2007). It is noteworthy that the proposed actions must follow the current public policies (BRASIL, 1996), as well as those recommended by the institution (UFC, 1982). It is noticed that the courses’ managers are responsible for: regulations, policies, supporting and guiding students throughout their university education within the limits of the academic and administrative aspects (Lagadão, 2013).

These limits confront us with the need to evaluate the academic and managerial skills of coordinations at Brazilian higher education. Generally, education management has been studied subjectively, mainly by evaluating its qualitative aspects. However, this study seeks to contribute to the analysis of the educational context under a quantitative point of view regarding qualitative inferences, according to the researches that have been developed recently in a field called “Educametrics” (Silva, et al., 2012), (Vasconcelos et al., 2013), (Nunes, et al., 2013) and (Nunes, et al., 2015).

This study aims to validate an assessment tool that provides the analysis of academic higher education management skills, from the perspective of the respondents of the courses, who filled a form based in a set of statements (\(A_i, i = 1, 2, \ldots, 14, 15\)) developed for this purpose. These statements were based on the Federal University of Ceará’s internal statute (UFC 1982). The UFC’s statute point out some directions that must be followed by the managers of the institution regarding the improvement of the quality in higher education. These statements had been previously presented in Nunes, et al. (2013) and Nunes, et al. (2015). Thus, we sought to investigate the applicability of Principal Components Analysis to achieve educational indicators for management courses and vocational formation in engineering, technology and related areas at Federal University of Ceará, based on the results of these indicators the coordination can manage the processes regarding better practices for the student formation.

### Evaluation and Management in Higher Education

Currently, the Higher Education Institutions in Brazil had been undergoing a process of renewal and changes in their ways of working, in order to improve their service to society (Schwartzman, 1992) and (Sguissardi, 2009), which may include the higher education management and its practices in order to improve its function.

It is necessary to understand that educational management has its basis in several areas, highlighting the university administration and education. Thus, the main responsible for conducting the process is the coordination of higher course (Marquesin, et al., 2008). This feature assigns a higher complexity for the activity performed by these professionals, due to the need to manage the processes of teaching and learning at different levels and with different institutional responsibilities, from the planning (educational, structural, financial), execution and evaluation (Valoi, 2007).

Therefore, the management of higher education has multimodal characteristics, allied to these evaluative processes, which may contribute to the improvement of educational quality with respect to the formation of higher education level (Gómez, et al., 2014). Among the knowledge that should give management support the evaluation of higher education (Bogoya, et al., 2013), can be used to check various aspects, going through the simplicity of institutional process of teaching and learning, such as: mission, institutional development plan, research and extension policies, graduate, social responsibility, human resources policies, physical infrastructure, and educational materials (Stufflebeam, et al., 2007).

According to Tyler (1942) and (1993), the evaluation provides grants for critical analysis from the perspective of the institution, enabling the investigation of ineffective elements, rearranging the various stages of the process. Still on this model, some aspects should be highlighted: professor/student interaction, education as a process of establishing standards of conduct, the diversity of instruments in the evaluation process, among others.

Thus, it can be seen that the evaluation becomes more significant when we take into account the analysis of programs, curricula and educational systems, crossing the boundaries of the specific measuring student performance. This gives to evaluation some parametric aspects for intervention in reality/context of the program/course. This with the intention to further develop the institutional function (Tyler, 1942).

According to the information above, the qualitative and quantitative assessment of the management procedures of undergraduate degrees in engineering, technology and similar fields at UFC, may provide a basis for the quality improvement of the courses (Gómez, et al., 2014) and their graduates; and also enable the adoption of strategies that aimed the reduction in dropout and failure rates, providing an improvement in the academic outcomes of undergraduate students. Thereby, we were looking for a mathematical mo-
del that provides the link between the elements of qualitative and quantitative aspects of the assessment proposed.

Mathematical Model

In the present study the multivariate statistics method called Principal Components Analysis – PCA was used, aiming the evaluation and validation of the instrument developed, regarding the extraction of information related to the management processes of Engineering, Technology and similar courses at UFC. The PCA is a method that reduces the space of representation of measures of the information brought by the data, whose most significant information is preserved. The process starts by discarding linear combinations with small representation of the data characteristics, thus the overlap information is not considered.

The PCA is used to analyze integrated data obtained from statistical measures relationships in order to identify patterns, possible inferences and classifications in the data set for analysis.

In this work, we started from a database in matrix form, \( \mathbf{X} = [x_1, \ldots, x_i, \ldots, x_P]_{M \times P} \) in which \( x_i, i \in [1, P] \) is a column vector which components originated in data set. The PCA seeks to relate the entries of \( \mathbf{X} \) through measures of covariance or correlation coefficients of the variables. This new set of information has some important algebraic characteristics (Green, 2011) and (Tabachnick, et al., 2001).

The PCA-based analysis in this work, part of the correlation coefficients matrix, which highlights the relationships of the data in a domain of observation, is represented by \( P \) linearly independent columns of the matrix \( \mathbf{X} \), assuming \( P \leq M \). This indicates that the matrix \( \mathbf{X} \) has full rank \( P \). It should be noted that the eigenvector associated with the largest eigenvalue, of the covariance or correlation matrix of the data, represents the direction of the projection with the greatest intensity information from the original dataset.

The PCA method is calculated from the original database, represented by the matrix \( \mathbf{X} \), following the sequence below.

Algorithm PCA

Consider the Matrix \( \mathbf{X} = [x_1, \ldots, x_i, \ldots, x_P]_{M \times P} \), formed by the vectors of observed data \( x_i, i \in [1, P] \), where \( P \) is the number of statements.

(i). calculate the \( \mu_p \), sample mean and the \( s^2_p \) sample variance of the variable \( p \)-th:

\[
\mu_p = \frac{1}{M} \sum_{i=1}^{M} x_{ip} \\
\sigma^2_p = \frac{1}{(M-1)} \sum_{i=1}^{M} (x_{ip} - \mu_p)^2
\]

Where the value of \( x_{ip} \) is the value of \( x_i \) in the \( p \)-th observation, which comprises the statements evaluated by respondents.

(ii). Based on the eq. (1), compute the entries from covariance matrix \( \mathbf{c} = [c_{ij}]_{P \times P} \), between the vectors \( x_i \) and \( x_k \) in the statements \( A_j \) e \( A_k \):

\[
c_{jk} = \frac{1}{(M-1)} \sum_{i=1}^{M} (x_{ij} - \mu_j)(x_{ik} - \mu_k)
\]

(iii). Calculating the correlation coefficient between the statements \( A_j \) and \( A_k \), considering the Equations (2) and (3):

\[
r_{jk} = \frac{c_{jk}}{s_j s_k}
\]

(iv). To obtain the matrix of correlation coefficients, hereinafter called correlation matrix.

\[
\mathbf{R}_{P \times P} = \begin{bmatrix}
1 & r_{12} & \cdots & r_{1P} \\
 r_{21} & 1 & \cdots & r_{2P} \\
 \vdots & \vdots & \ddots & \vdots \\
 r_{P1} & r_{P2} & \cdots & 1
\end{bmatrix}
\]

Where \( \mathbf{R}_{P \times P} \) is obtained from Equation (4)

(v). From the Eigenvalue Decomposition - EVD of \( \mathbf{R} \), obtaining matrices of eigenvectors and eigenvalues \( \Lambda \) and \( \mathbf{Q} \), respectively:

\[
\Lambda = \begin{bmatrix}
\lambda_1 & & \\
& \ddots & \\
& & \lambda_P
\end{bmatrix}
\]

\[
\mathbf{Q} = [q_1, \ldots, q_i, \ldots, q_P]
\]

and wherein \( \lambda_1 \) and \( q_i \), \( i \in [1, P] \), are associated eigenvectors and eigenvalues, respectively.

For the largest eigenvalues and its eigenvectors associated are chosen as main components.

Methodology and Research Context

The study was conducted at the Federal University of Ceará (UFC), with professors from six engineering courses from the Centre of Technology and Agricultural Sciences (Teleinformatics Engineering, Electrical Engineering, Chemistry Engineering, Mechanical Engineering, Civil Engineering, Fishing Engineering and Mechanical Producer Engineering) and the Systems and Digital Media (SDM) course at UFC. These courses had been chosen by the fact that the institutional affiliation of the researchers, who has conducted this research, has free access to the professors. It is important to notice that the evaluated courses had starting their curricula reformulation with this study could be useful to help the managers’ decision making.

The research sought to identify the professors’ point of view about procedures, methods and actions taken by the coordination of each course with regard to administrative processes. The instrument developed was based on the Institution rules, where the skills and competencies of Course’s Coordinator are pointed out (UFC, 1982). About 20% of professors answered the form.

Instrument design of the research evaluation

As noted, it was developed an instrument in the form of a poll, that containing 15 statements to obtain the informa-
tion that were quantified (numerical data) (Nunes, et al., 2013) and (Nunes, et al., 2015).

Regarding the statements, the participants suggested their level of agreement through a 6-point Likert-type Scale, which, at the time of processing data, received the conversion values of: Totally Disagree (1); Strongly Disagree (2); Disagree (3); Agree (4); Strongly Agree (5); and Totally Agree (6). The purpose of this instrument is to identify possible inconsistencies about management processes under the professors’ point of view, especially focused on academic and administrative management. Purposely, the statements were randomly rearranged, not allowing any kind of classification based on its structural features, such as: teaching and learning relationship, academic and vocational formation, and management procedures.

The data collection was performed using an electronic type, available at this storage service and sync of the Google Drive service, as well as in printed form. In the data analysis, the software Matrix Laboratory - MATLAB and the Statistical Package for Social Sciences-SPSS were used. The plots were obtained by the software OriginPro8®.

Results and Discussion

Validation of results

To ensure that the database was suitable for PCA, some tests of reliability and adequacy were performed as described below

We emphasize that all the data variables were used to analyze the eigenvalues obtained a priori (PCA-1), and after this, only the most important variables were selected to the final analysis (PCA-2).

1. α de Cronbach is an index that seeks to ensure the reliability of the data from the measure of its internal consistency (Hörnquist, et al., 2002). For the collected data, the Cronbach’s α was 0.9402 (Table 1 – PCA1), which ensures a good statistical distribution of the data around the mean, i.e., respondents has following the same thinking to answer the poll. Therefore, the data is reliable for analysis, according to this parameter.

Table 1. Comparative validation of the indicators in PCA-1 end PCA-2

<table>
<thead>
<tr>
<th>Number of Item/stage</th>
<th>KMO</th>
<th>A</th>
<th>Scale/ cumulative variance %</th>
<th>Scale/ cumulative variance %</th>
<th>Scale/ cumulative variance %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1 eigenvalue)</td>
<td>(2 eigenvalues)</td>
<td>(3 eigenvalues)</td>
</tr>
<tr>
<td>15 items/PCA-1</td>
<td>0.824</td>
<td>0.9402</td>
<td>82.7 / 55.16 %</td>
<td>82.7 / 55.16 %</td>
<td>1,24 / 75.20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2 eigenvalues)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.824</td>
<td>0.9402</td>
<td>82.7 / 55.16 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 items (excluding A4, A10 e A13)/ PCA-2</td>
<td>0.892</td>
<td>0.9406</td>
<td>7.31 / 60.93 %</td>
<td>0.88 / 68.26 %</td>
<td>0.82 / 75.09%</td>
</tr>
</tbody>
</table>

2. Kaiser-Meyer-Olkin (KMO) and Bartlett’s Tests: Calculate the correlation among all variables analyzed and also verify the hypothesis of correlation matrix of variables is an identity matrix (Tabachnick, et al., 2001).

The value obtained for KMO test was 0.824 (Table 1 – PCA-1), showing that the data are significant, thus guarantees a good adaptation of the sample for the application to use the PCA, since this must be equal or upper than 0.6 (Green, 2011). We find that the Bartlett’s test reject the null hypothesis that the correlation matrix would be an identity matrix (Green, 2011).

Application of PCA

1. Communalities: the variances of the data extracted are compared with the variances of the original variables, checking how much variance in common h² (communalities) exists among observed variables and those obtained from the PCA model, being communalities h² related to the specific variance s² (specificity) and the measurement error variance e², according to the expression h² = 1–(s² + e²) (Hörnquist, et al., 2002) and (Pasquali, 2012). The values of the communalities are set in Table 2 (PCA-1), and all the 15 statements presented values higher than 0.6, indicating high representativeness of the statements in the first case (PCA-1).

2. Selecting the Number of Components: A Figure 1 (PCA-1) shows the magnitude of the eigenvalues (λ) extracted from the correlation matrix. To select the number of principal components to be considered later, we used the Kaiser criterion, that says that the components to be considered must submit eigenvalues greater than 1 (λ>1) (Tabachnick, et al., 2001).

Table 2. Values Communalities of the Statements.

<table>
<thead>
<tr>
<th>Item</th>
<th>Inicial</th>
<th>h² (15 items)</th>
<th>h² (12 items)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PCA-1</td>
<td>PCA-2</td>
</tr>
<tr>
<td>A1</td>
<td>1,000</td>
<td>0,712</td>
<td>0,765</td>
</tr>
<tr>
<td>A2</td>
<td>1,000</td>
<td>0,693</td>
<td>0,719</td>
</tr>
<tr>
<td>A3</td>
<td>1,000</td>
<td>0,720</td>
<td>0,738</td>
</tr>
<tr>
<td>A4</td>
<td>1,000</td>
<td>0,770</td>
<td>-</td>
</tr>
<tr>
<td>A5</td>
<td>1,000</td>
<td>0,614</td>
<td>0,668</td>
</tr>
<tr>
<td>A6</td>
<td>1,000</td>
<td>0,633</td>
<td>0,640</td>
</tr>
<tr>
<td>A7</td>
<td>1,000</td>
<td>0,810</td>
<td>0,879</td>
</tr>
<tr>
<td>A8</td>
<td>1,000</td>
<td>0,735</td>
<td>0,701</td>
</tr>
<tr>
<td>A9</td>
<td>1,000</td>
<td>0,773</td>
<td>0,875</td>
</tr>
<tr>
<td>A10</td>
<td>1,000</td>
<td>0,910</td>
<td>-</td>
</tr>
<tr>
<td>A11</td>
<td>1,000</td>
<td>0,659</td>
<td>0,606</td>
</tr>
<tr>
<td>A12</td>
<td>1,000</td>
<td>0,800</td>
<td>0,855</td>
</tr>
<tr>
<td>A13</td>
<td>1,000</td>
<td>0,716</td>
<td>-</td>
</tr>
<tr>
<td>A14</td>
<td>1,000</td>
<td>0,737</td>
<td>0,766</td>
</tr>
<tr>
<td>A15</td>
<td>1,000</td>
<td>0,700</td>
<td>0,799</td>
</tr>
</tbody>
</table>

In Figure 1 (PCA-1) we can see an abrupt decay curve and apparently stabilizes at values between 0 e 1 after the first 3 components, which represent an accumulated variance that exceeds 7% as seen in Table 1 (PCA-1). This result shows sufficiency on the representation of the original data (Pasquali, 2012).
3. **Cluster Analysis By PCA**: For the PCA-1, the three principal components obtained, by the Kaiser criterion, do not make any sense in the analysis. Then, the first two Principal Components (PC1 and PC2) were selected presenting great results according with the preliminaries expectations (Figure 2).

The occurrence of outliers showed the need to remove their items from the analysis, since this could negatively influence the clustering formation. The removal of items demanded a new validation process where the data in were obtained Table 1 (PCA-2).

It can be observed that the new validation was consistent with higher values than the previous validation step. It is possible to observe that the model selects only one component equal to or greater than 1 (Kaiser Criterion), and this alone represents approximately 61% of the variance (Table 1. PCA-2), where we can still view the stability in the decay of magnitudes of the eigenvalues in Figure 1 (PCA-2).

In PCA-2, the Kaiser criterion selected only one principal component, however, as showed in the first analyze (PCA-1), this criterion was revised to make reasonable inferences about the results obtained and, in this case (PCA-2), we have considered the second principal component in the analysis, as can be seen in Figure 3. A new consistent pattern can be observed in Figure 3, with the formation of three clusters.

The cluster II presents statements related to the institutional training environment and its external visibility, more broadly in society, with regard to aspects related to meaningful learning and ENADE, or more specifically on their performance in the labor market through his first contact through internships. The cluster was named Management of Professional Training x Labour Market (PM).

The cluster III presents statements related to the institutional training environment and its external visibility, more broadly in society, with regard to aspects related to meaningful learning and ENADE, or more specifically on their performance in the labor market through his first contact through internships. The cluster was named Management of Professional Training x Labour Market (PM).

Analyzing the cluster III, the statement points to an affinity in the matters relating to Academic Training of students, es-
pecially with the regard to the enviroment to support learning as library and extracurricular activities. This cluster was named the Academic Management Training x Environments to support learning (AE).

The Figure 3 can still provide other information such as: statement of the cluster I present higher loading values on PC1, followed by the other two clusters with almost the same order of magnitude; towards higher values of PC2 loading are presented by the statements of cluster II, followed by the other two, being these of similar values.

We also observed that PC1 and PC2 in the cluster III was identified as the second most important in relation to these two components. We can infer that professors showed an awareness of the important role that the management of higher education has in conducting academic training processes in engineering courses, and related technology.

Conclusions and Perspectives

The aim of this study was the validation of an assessment instrument constructed to identify the skills of academic management, using educational indicators, in the context of professors’ point of view of the Systems & Digital Media and Engineering courses at Federal University of Ceará. In this sense, this instrument can be used to provide support for the curriculum design and management process regarding the improvement of the quality in higher education.

The results show the feasibility of using PCA for the analysis and validation of the data collected related to academic higher education management. Results showed the need to remove some items that were corrupting the analysis. This was generated by the professors, understanding the role of coordinators who have no autonomy in the management of financial resources. This result does not correspond to what is practiced in the institution, thus indicating a clear shift of management processes to improve the quality of its courses, this being based on the professors’ point of view.

The remaining statements were classified into three groups: Planning x Assessment; Management of Vocational Training and Labour Market; and Management of Academic Training and their support learning environments, and these can be seen as educational quality indicators. We can also highlight the important role of coordination presented in the meaning given to management in major components extracted. Thus, it was established that the management in higher education is more focused on academic training than in professional training. This corroborates that this instrument allows us to make evaluative inferences about actions and management methods. Then, it can support possible interventions proposed by the course coordination, regarding the improvement of teaching and learning and thereby improving the quality of training in technology and engineering areas.

The analysis of results points to the existence of a relationship between the methods of management in higher education, where it was found that this provides important elements for academic and vocational training. This corroborates the possibility to make evaluative inferences about the actions and methods of management. Thus enables the design of interventions aimed at improving teaching and learning in undergraduate engineering education. We could even say that the vision of professors points to the need for decentralization of financial management within the courses.

Although the results obtained are presented fairly consistent and, therefore, promising for the use in evaluations of management processes in undergraduate courses, the validation study showed a limitation with regard to the application context: only 7 courses were evaluated, and the results obtained can not be extended for the whole university.

These studies point to possible perspectives, such as: the possibility of application of the instrument in other courses and areas; can make use of inferences as subsidy for reformulation of policies that aims to decrease dropout and failure rates, contributing to the intervention with regard to the academic success of the learner; investigation of the relationship between the results found in this research with the indices calculated by external evaluations; adaptation of the instrument to implement with students of the courses, that aimed to identify the vision of the graduate student on the level of training received; the application of other mathematical models which exploit the inherent information data, such as multilinear decomposition and tensor analysis (Vasconcelos et al., 2013).

Acknowledgments

This paper was supported by CAPES, CNPq, FUNCAP and GPed research group – Federal University of Ceará (UFC) – Brazil.

References


