Predicting Success in Product Development:
The Application of Principal Component Analysis to Categorical Data and Binomial Logistic Regression

Glauco Henrique de Sousa Mendes[1], Gilberto Miller Devós Ganga[2]

Abstract

Critical success factors in new product development (NPD) in the Brazilian small and medium enterprises (SMEs) are identified and analyzed. Critical success factors are best practices that can be used to improve NPD management and performance in a company. However, the traditional method for identifying these factors is survey methods. Subsequently, the collected data are reduced through traditional multivariate analysis. The objective of this work is to develop a logistic regression model for predicting the success or failure of the new product development. This model allows for an evaluation and prioritization of resource commitments. The results will be helpful for guiding management actions, as one way to improve NPD performance in those industries.

Keywords: new product development, applied statistics, logistic regression.

1,2 Department of Production Engineering, Federal University of São Carlos, Rod. Washington Luís, Km 235, P.O.Box 676, 13565-905, São Carlos (SP), Brazil. Phone: +55 16 3351 9540. Fax: +55 16 3351 8240. E-mail: [1]glauco@dep.ufscar.br

Introduction

In the management of NPD, the identification of critical success factors (CSFs) that contribute to increasing the probability of success of the new product is a traditional line of research with numerous reference studies (Cooper and Kleinschmidt, 1995; Song and Parry, 1996; Souder et al., 1997; Song et al., 1997; Ernst, 2002; Cooper et al., 2004a; Kahn et al., 2006; Song and Noh, 2006; Barczak et al., 2009; Kahn et al., 2012).

The traditional method for identifying these factors is the use of surveys in which data are collected with respect to a large set of potential variables. Subsequently, these data are reduced using a traditional multivariate analysis, correlating the specific factors to the success of the new product (Cooper and Kleinschmidt, 1995; Song et al., 1997; Song and Noh, 2006).

Based on the assumption that the literature concerning the NPD critical factors is thoroughly explored, the present study aims to go one step further by proposing a quantitative model capable of predicting the result of the new product. Therefore, this study seeks to fill a gap in the published literature, where there is a lack of studies utilizing more sophisticated techniques to predict the success or failure rate in NPD (Berkowitz et al., 2007).

Sophisticated techniques can be used to identify those models able to deal with problems involving a large number of variables of different types. The large number of variables reflects the need for the development of multivariate statistical models capable of dealing with problems related to a wide variety of controllable and uncontrollable factors, which may lead to the success or failure of NPD.

The present study proposes a multivariate statistical model to assist project managers in predicting the outcome of new products. The use of this model may probabilistically classify a new project as either a success or a failure. Simulations, such as "what would happen if", can be performed to identify critical areas for the efficient allocation of resources.

Classification problems are often found in many fields. In business management, the most common applications are in finance (Chen, 2011), marketing (Kaefer et al., 2005) and human resource management (Sexton and Mcmurtrey, 2005). The strong interest in classification problems has motivated many researchers (Duda and Hart, 2001; Bernadó and Garrell, 2003) to develop quantitative methods for this purpose. Linear discriminant analysis (LDA) (Johnson and Wichern, 2002) was the first method developed to solve classification problems from a multivariate perspective. In addition to LDA, other multivariate statistical tools (Flury and Riedwyl, 1990) applied in the prediction of such problems include logistic regression (LR) (Studenmund, 1992; Hosmer and Lemeshow, 2000) and multiple regression (Menard, 1993; Myers, 1990).

Logistic regression (LR) is known for assuming traditional assumptions in the application of other multivariate statistical techniques. LR, for example, is used when the independent variables do not meet the condition of multivariate normal distribution (Kleibaum, 1994; Sharma, 1996; Ohlson, 1980). LR also overcomes such assumptions as linear correlation, measures on the same scale or homogeneous error variance, which are present in the use of multiple regression (Chatterjee and Hadi, 2006).

Therefore, the objective of the present study was to develop a LR model for predicting the success or failure of new product projects. Selection of the model is mainly determined by the nature of the variables representing the success factors in their qualitative, ordinal and categorical entirety.

The paper is organized as follows. After this introduction, relevant studies from the literature are reviewed in the second and third sections. The research method is described in the fourth section followed by data analysis and the presentation of results in the fifth section. In the sixth section, the results are discussed, and their implications are explored. Finally, the limitations of this study and possible future research are presented.

Literature review

Success factors in the NPD management

A best practice can be defined as a technique, method, process or activity that is effective in generating superior results in a given business process. In the case of NPD, the use of best practices can increase the probability of the new product success (Kahn et al., 2012). The identification of best practices (critical success factors) related to NPD have shown to be notably popular in the NPD area (Ernst, 2002). Literature on the NPD management (Cooper and Kleinschmidt, 1995; Song and Parry, 1996; Souder et al., 1997; Song et al., 1997; Ernst, 2002; Cooper et al., 2004a; Kahn et al., 2006; Song and Noh, 2006; Shum and Lin, 2007; Barczak, et al. 2009; Kahn et al., 2012) was reviewed to define the variables and constructs considered in the present study, as well as to establish the dependency relationships presented in this study.

The product strategy is related to the creation and improvement of new products. Cooper et al. (2004a, 2004b) indicate that organizations with superior performance in the NPD process have specific strategies for new product projects.

The alignment among the NPD strategy, the new products projects and the main company strategies is a success factor (Cooper and Kleinschmidt, 1995; Song and Noh, 2006; Acur et al., 2012). Marketing skill is the ability of a company to detect, assess and use appropriate information about the customers, markets, competitors and external environmental forces. A market-oriented approach has been identified as a success factor (Cooper and Kleinschmidt, 1995; Song et al., 1997; Langerak et al., 2004; Haverila, 2010).

The technology strategy also must be linked to the company strategy to address technical capabilities and market opportunities during the NPD (Zapata and Cantú, 2008). A decision in the technology strategy is the selection of the technology sources. According to Scott (1999), the technology sources can contribute to the success or failure of a new product because they require different capabilities of companies with regards to the acquisition, adaptation, management and integration of the given technology with systems and people.

Company skills are defined as distinctive capabilities that assist the company in the execution of the NPD process, directly interfering in the quality of the tasks performed. The technical and marketing skills of the company have an influence on the success of the new product (Song et al., 1997; Song and Noh, 2006).

The roles of key individuals are also frequently cited as success factors. Among these factors, the project leader skill and the top managements support are highlighted (Brown and Eisenhardt, 1995; Lee et al., 2000). The project leader performs important tasks, such as facilitating communication between the project team members and top management, negotiating the resources allocation and seeking to keep the project team members motivated and focused on their responsibilities (Thieme et al., 2003). The top management also plays an important role because it is responsible for the strategic direction of NPD and the allocation of human and financial resources (March-Choda et al., 2002).

The organization of NPD is recurrent in studies on critical factors for NPD. The emphasis is on the cross functional integration and on the organization of project teams (Lee et al., 2000; Ernst, 2002). Several studies (Griffin, 1997; Souder et al., 1997; March-Chordà et al., 2002; Cooper et al., 2004a; Sherman et al., 2005; Jugend and Silva, 2012) suggest that cross functional integration has a positive influence on the new product outcome. Integration between the NPD and marketing has a positive effect on the new product success (Yap and Souder, 1994).

There are several different organizational structures for the project team. The main structures are the pure, matrix or functional project structures. The benefits of more organic structures (pure and matrix) have been reported (Larson and Gobeli, 1988). However, other studies (Lee et al., 2000; Yap and Souder, 1994) suggest that adopting a contingency perspective.

Barczak et al. (2009) suggest that companies with superior performance in NPD have formal processes that are divided into stages and activities. Many studies (Cooper and Kleinschmidt, 1995; Brown and Eisenhardt, 1995; Cooper et al., 2004c) have indicated that the quality of execution of some NPD activities has a positive influence on the new product success. Among these activities, the fuzzy front end (FFE) activities are highlighted (Griffin, 1997; Ernst, 2002; Kahn et al., 2006). The activities of the idea generation, technical and marketing studies and the new project viability (technical and economic) are the FFE activities that contribute most to the NPD success.

Multivariate data analysis

Survey data analysis requires special attention from the researchers, due to the considerable amount and types of measures used to assess the constructs used in the research conceptual model. The amount of variables reflects the multivariate aspect of the research problem and, therefore, the conceptual model proposed to measure it. The typology reflects the classification of the indicators (quantitative or qualitative) of these constructs. The treatment of the multivariate data is a relevant contribution to the construction and expansion of theories by identifying the complex causal relationships between the variables and constructs involved (Devellis, 2012).

Such techniques as cluster analysis, principal component analysis, factorial analysis, discriminant analysis, structural equation modeling and others have been used in various research fields, including biology, medicine, agronomy, engineering, business and social and behavioral sciences (Manly, 2005; Lattin et al., 2011; Brown et al., 2011; Devellis, 2012). One of the merits of these techniques is the identification of specific and precise issues of considerable complexity within a data set, transforming multidimensional information into two-or three-dimensional information (Manly, 2005; Hair et al., 2006; Lattin et al., 2011).

The multivariate analysis techniques can be classified into two groups, including: 1) dependence techniques in which one or more variables can be explained by other independent variables and 2) interdependence techniques, in which no variable or set of variables are treated as either dependent or independent. Dependence techniques include multiple regression, structural equation modeling, discriminant analysis and conjoint analysis. Interdependence techniques

may consist of factorial analysis, cluster analysis, principal component analysis, correspondence analysis and multidimensional scaling (Hair et al., 2006).

Although most multivariate techniques provide more accurate and reliable results when variables are quantitative (discrete or continuous), this property does not mean that such techniques as cluster analysis and factor analysis cannot be used as qualitative variables (categorical or nominal). However, the relevant literature (Hair et al., 2006) advises caution in the customary use of these multivariate techniques for qualitative variables.

The purpose of the present article is to predict the success or failure in NPD. Therefore, the general hypothesis was established that a multivariate statistical model is an interesting alternative to the prediction of the new product outcome (Bertrand and Fransoo, 2002). In situations that include categorical variables both in the dependent and independent variables, the use of the logistic regression model is recommended (Hair et al., 2006; Lattin et al., 2011).

Categorical Principal Component Analysis

One of the techniques applied to reduce data and enable treatment with other multivariate techniques, such as the logistic regression analysis, is principal component analysis (PCA) (Krishnakumar and Nagar, 2008). PCA seeks to construct a new set of variables, called principal components, that are less numerous than the original data, but still adequately summarize the information contained in the original variables. Each component is a linear combination of the original variables and seeks to reproduce the maximum variance of the original data. In PCA, there is no underlying explanatory model (Lattin et al., 2011).

This PCA is predominantly used in the analysis of numerical variables. However, numerous research methods in social sciences and operational management utilize qualitative variables in the proposition of constructs (Devellis, 2012; Brown et al., 2011). In these cases, multivariate statistical techniques are applied, such as multiple correspondence analysis, multidimensional scaling, factorial analysis, logistic regression and structural equation modeling (Hair et al., 2006; Brown et al., 2011; Devellis, 2012).

Many studies (Jöreskog and Moustaki, 2001; Flora and Curran, 2004; Vermunt, 2007; Wirth and Edwards, 2007; Forero et al., 2009) reveal the use of new procedures of factor analysis in categorical variables. These new procedures are primarily related to the method used to extract the factors. Frequently, in the application of factor analysis, researchers mistakenly employ the method for factor extraction by principal components in categorical variables, which is an error

caused by the assumption that the probability distribution of the categorical variables, especially the nominal variables, is known.

Therefore, methods for the extraction of factors, such as weighted least squares (WLS) or unweighted least squares (ULS), are used in exploratory factorial analysis because these approaches do not require the use of prior knowledge regarding the probability distribution of the variables (Flora and Curran, 2004; Johnson and Wichern, 2002). However, the methods for extraction of the factors for categorical variables still require further study (Jöreskog and Moustaki, 2001; Wirth and Edwards, 2007; Forero et al., 2009).

Accordingly, principal component analysis was selected as a method for reducing the variables to be considered as independent variables in a logistic regression model. However, to resolve the issue of inapplicability in the categorical principal component analysis (CATPCA), the principal component analysis was used with optimum nonlinear design for nominal and ordinal data (Meulman et al, 2004).

Generally CATPCA assigns numerical quantifiers to each of the categories of qualitative variables, thereby allowing later analysis of the main components of the transformed variables (Meulman, 1992; 1998). The numerical values assigned to each class of the original variables are defined by an interactive procedure, which is known as the alternative least squares method, such that the numerical quantifications possess metric properties (Moroco, 2003).

While traditional PCA assumes linear relationships between variables, CATPCA allows for the measurement of nonlinear relationships between variables. Moreover, CATPCA does not require that the variables be normally distributed. Thus, CATPCA also can be considered as a method to reduce the data size (Meulman, 1992; 1998; Meulman et al., 2004).

Logistic Regression

Logistic regression (LR) has been used in agriculture (Bielza et al., 2003), medicine (Martin et al., 2008), the analysis and prevention of accidents (Al-Ghamdi, 2002), education (Pavlekovic et al., 2010), the launching of software (Cerpa, 2010) and finance (Andres, 2005; Chen, 2011).

LR is a multivariate statistical technique used in situations that require the prediction of the occurrence or non-occurrence of a certain characteristic of the response variable from a set of independent variables. LR is similar to the linear regression model but is set apart by its dependent variable type, which can be categorical, binary or multinomial (Harrel, 2001; Hosmer and Lemeshow, 2000). Additionally, LR does not make certain basic assumptions

that are typically adopted in other models, as mentioned in the introduction.

The LR model for p independent variables can be written using expression 1, where P(Y = 1) is the probability of success and B0, B1,…, Bp are the coefficients of the regression model.

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p)}}$$

(1)

There is a linear regression model hidden within the logistic regression model. The natural logarithm of the ratio P(Y=1) to (1-P(Y=1)) provides a linear model in X_i, as can be observed in expression 2.

$$g(x) = \ln\left(\frac{P(Y=1)}{1 - p\,(y=1)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$$

(2)

The function g(x) has many desirable properties of a linear regression model. An advantage of the LR with respect to linear regression models is that the independent variables may be a combination of continuous and categorical variables. In the present study, the binary logistic regression was justified by the fact that all independent variables are ordinal categorical variables, and the dependent variable is binary, i.e., represents the probability of success or failure in product development.

Method

The research methods adopted in this article can be considered explicative for identifying NPD success factors, and it can be considered predictive for assuming that a multivariate statistical model is able to predict the new product outcome. The technical procedures included the research survey (Forza, 2002) and empirical modeling (Bertrand and Fransoo, 2002). The research steps were based mainly on those indicated by Forza (2002) for surveys.

Sample characteristics

The population of the companies studied is shown in Table 1.

In situations where the population is considered to be small, it is recommended that the sample constitute at least 50% of the population (Yamane, 1967). In the present survey, 62 companies were randomly selected, amounting more than 50% of the population, and this sample size is therefore considered satisfactory.

The analysis units in the survey were projects of new product developed by the companies in the past five years. Interviews were conducted with the NPD responsible (managing partner, managers or engineers). The objective was to explore two types of new product projects in each company: one success project and one failure project.

After discussions with the NPD project leaders, the target projects were defined with the specific personnel who would be interviewed. All responses should be based on records, facts and situations encountered during the product development. The interviews were conditioned according to the level of knowledge and responsibility assumed by the respondent during the project execution. In the interviews, a structured questionnaire was adopted as an instrument of data collection.

Data were collected from 112 projects, of which eight were excluded due to missing data. Thus, data were obtained for 104 projects, including 62 successful projects and 42 failures ones. This sample is consistent with the recommendations of Hair et al. (2006), who advised a minimum of 100 cases to ensure more robust results. This proportion was important to define the cutoff point in the procedures for the application of logistic regression. In this case, the cutoff value was established as 0.6.

Criteria	Description
Size	- Small enterprises with fewer than 100 employees
	- Medium enterprises between 100 and 500 employees
Industry	- Design and manufacturing of medium and high technological complexity devices
	– medical devices and automation process devices
Location	- State of São Paulo, which includes 60% of the industry
NPD	- Active
Total	99 enterprises

Table 1. Population Composition

The questionnaire and conceptual model

The questionnaire was based on the conceptual model shown in Figure 1, which was used to guide the present study. This conceptual model was based on the models of Brown and Eisenhardt (1995), Song and Parry (1997), and Souder et al. (1997). The conceptual model suggests that the product strategy, marketing skills, technology sources, company/business skills, project leader skill, functional integration, organization of the project team and quality of execution of NPD activities are constructs of NPD practices (factors) that influence the new product outcome.

The dependent variable of the model is the perceived success of the new product, for example, the result from a comparison of the company expectations with the actual performance of the new product after its launch, considering financial aspects, market share, brand strengthening and the development of new competencies for the company. The independent variables are the practices (individual variables) that form the constructs of the model.

The questionnaire was structured according to the constructs shown in Figure 1. Each construct was split into the independent variables, as related to successful practices in NPD managing. These variables were presented in the questionnaire in the form of affirmative statements. The perception of the respondents regarding the degree of compliance of the practices adopted in these projects was measured on a five-point ordinal categorical scale (1-strongly disagree to 5-strongly agree). To classify the project, a five-point Likert scale was also used. In cases where performance equaled or surpassed expectations they were classified as successful. The failures projects corresponded to products with performance below expectations.

The questionnaire was pretested in visits in four companies. Adjustments were made to facilitate the understanding of the constructs and individual variables and the scales used to measure each variable. The internal validity of the constructs was measured using Cronbach's alpha coefficient (Cronbach, 1951), which represents the reliability of the measurement scale for the constructs adopted. A coefficient greater than 0.7 represents satisfactory reliability of the measurement scale (Nunnally, 1978). The lowest of Cronbach's alpha coefficient of the constructs was 0.88, which is considered satisfactory.
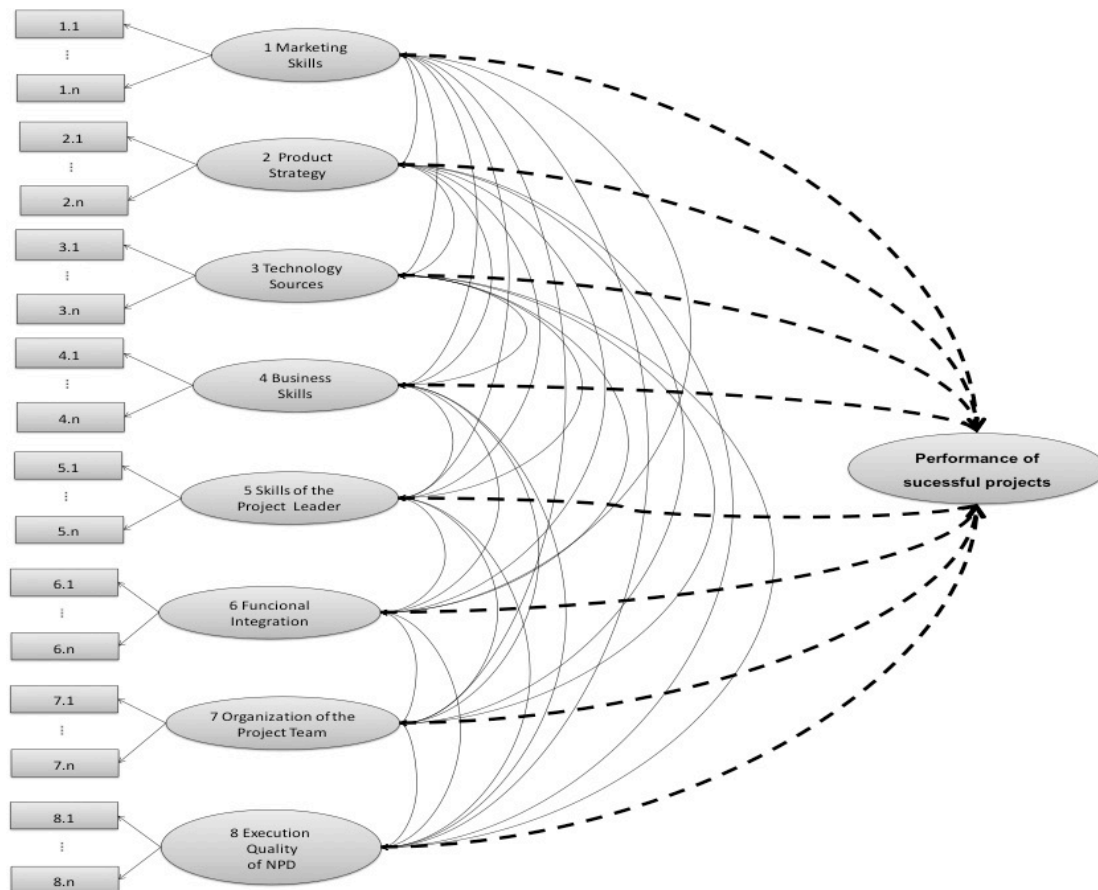

Figure 1. The conceptual model adopted in this study

Results

Data reduction

CATPCA was used to define the number of variables required to initiate the development of the logistic regression model. The same criteria used to develop the traditional PCA were used in the CATPCA. Important criteria include the database suitability (the number of cases and the correlation pattern between the variables), the definition of the number of components for each construct and the percentage of variance explained by these components. Table 2 illustrates the criteria adopted for implementation of the CATPCA.

Another indicator of the sample adequacy is the KMO test, which measures the degree of correlation between variables. The KMO test varies from 0 to 1, and higher values are considered better. Hair et al. (2006) suggest 0.50 as an acceptable level. Considering Table 2, one can observe that all constructs showed a KMO above 0.5, except for the Organization of the Project Team. In this case, all of the variables in this construct were initially included in the logistic regression model. Bartlett's test of sphericity (BTS) showed a significance value less than 0.000, rejecting the existence of an identity matrix. Based on this scenario, the data sample was considered adequate for implementation of the CATPCA.

To define the number of components, the percentage of variance explained by the components was considered. The value of the variance accumulated by the components was adopted as an acceptable index if equal to or greater than 50%. Table 2 shows that the amount of components extracted was not greater than two and the minimum accumulated

variance was 50%, which are values considered satisfactory for reducing the number of variables. Another question refers to the index of the eigenvalue. Most of the components had eigenvalues greater than 1.0, representing high explanatory power of the indicators in relation to variance. This behavior reinforces the Kaiser criterion, which recommends adopting only those components with eigenvalues greater than 1.0 (Kaiser, 1974).

Table 3 summarizes, for each construct, the variables selected by the CATPCA and the respective loads of the components. Only the components that reached at least 50% of the explained variance and components that had a load greater than 0.8 were selected.

The rules adopted in implementing the CATPCA permitted a reduction in the number of variables by 28%. These variables were included as a first step in the development of the logistic regression model and they are highlighted in the discussion section.

Logistic regression in the prediction of NPD success

The stepwise method was selected for development of the logistic regression model. Among the stepwise method options, the forward LR (likelihood ratio) was selected. The model was adjusted after four steps. The omnibus test was used to assess the significance of each step, indicating that all steps were significant ($p < 0.05$) for the model development, as can be observed in Table 4.

The explanatory power of the logistic regression model in each stage is illustrated in Table 5. As can be observed, the stage 4 model reached an overall success rate of 89.4% with a particularly high success rate for the projects classified as successful.

Construct	KMO	BTS	Cronbach's alpha	Eigenvalue		Variance explained (%)		
				1	2	1	2	Total
Marketing Skills	0,516	0,000	0,90	2,50	1,08	50	22	71
Product Strategy	0,5	0,000	0,90	1,67	1,40	42	35	77
Technology Sources	0,631	0,000	0,87	2,14	1,73	31	25	55
Company Skills	0,692	0,000	0,97	3,14	1,30	63	26	89
Project Leader Skill	0,796	0,000	0,97	4,63	2,14	58	27	85
Funcional Integration	0,674	0,000	0,92	4,56	1,25	46	12	58
Organization of the Project Team	0,267	0,000	0,98	1,54	1,34	51	45	96
Quality of execution of NPD activities	0,698	0.000	0,91	2,79	1,32	46	22	68

Where: KMO (Kaiser Meyer Olkin) and BTS: Bartelett Test of Spherecity (p<0,05)
Table 2. Criteria of the development of CATPCA

Variables	Loads of the components	
	1	2
1.1 Evaluation of the market potential for this project was well-conducted	1.0	-
1.2 The consumers/clients greatly desired this type of product	0.9	-
1.3 User requirements were understood and translated for the product specifications	0.9	-
2.1 The product offers the same solutions as the competitors, but with the advantage of a lower price	1.1	-
2.2 The product has the same characteristics of the competitors' products	1.0	-
2.3 The product has superior technical performance compared to the competitors	-	1.2
2.4 The product was well articulated with competitive strategies for the product and the company	-	0.8
3.1 Hiring of third-party staff to fill the need for skills not existent in the company	1.0	-
3.2 Alliances and partnerships with other entities	1.0	-
3.3 Use of licensing strategies	0.8	-
3.4 Alliances and partnerships with suppliers	-	1.0
3.5 Alliances and partnerships with clients	-	0.9
3.6 Self/internal development (technology completely developed by the company)	-	-1.0
4.1 The Technical Assistance team had the technical skills	1.0	-
4.2 The Commercial team had the technical skills	0.9	-
4.3 The Manufacturing team had the technical skills	0.9	-
5.1 The project leader had the technical skills needed	1.0	-
5.2 The project leader had the interpersonal skills needed	0.0	-
5.3 The leadership style adopted by the project leader was suitable for its execution, encouraging communication and conflict management	1.0	-
5.4 The leadership style allowed for participation of team members	1.0	-
5.5 The staff development team was motivated to execute the project	1.0	-
6.1 The project included participation from various areas/departments in conducting the technical development activities (product design)	0.9	-
6.2 In the project, there was an appropriate degree of integration between manufacturing and R&D	0.9	-
6.3 The project included participation from various areas/departments in conducting the activities of generation and selection of ideas	0.9	-
6.4 In the project, there was an appropriate degree of integration between Commercial and R&D	0.8	-
6.5 The project included participation of various areas/departments in conducting the feasibility analysis activities	0.8	-
6.6 The project included participation of various areas/departments in conducting the testing activities of the product/market	-	1.0
6.7 The project included participation of various areas/departments in conducting the prototype development activities	-	0.9
7.1 The project activities were performed using a functional structure.	1.0	-
7.2 The project activities were performed using a separated project structure.	-	1.0
7.3 The project activities were performed using a matrix structure	-	-0.8
8.1 Technical development activities (product design)	1.0	-
8.2 Management activities and selection of ideas	0.9	-

Variables	Loads of the components	
	1	2
8.3 Commercial activities for launching of the new product	-	1.0
8.4 Testing of the product/market	-	0.9
8.5 At the end, a general assessment was conducted to identify successes or mistakes made during the project.	0.9	-
8.6 Establishment of decision points (stage-gates) for the NPD stages	0.9	-
8.7 Meeting of necessary legal standards for the product	0.8	-
8.8 Production of documents (briefings, drawings and test results) regarding execution of the project	-	1.0

Table 3. The component loads by constructs and indicators

Steps		Chi-square	df	Sig.
1	Step	42,363	4	0,000
	Block	42,363	4	0,000
	Model	42,363	4	0,000
2	Step	24,581	4	0,000
	Block	66,943	8	0,000
	Model	66,943	8	0,000
3	Step	12,046	3	0,007
	Block	78,989	11	0,000
	Model	78,989	11	0,000
4	Step	11,964	4	0,018
	Block	90,953	15	0,000
	Model	90,953	15	0,000

Table 4. Omnibus Tests of model coefficients

Steps	Rating Project (%)		
	Success	Failure	Total
1	88,7	61,9	77,9
2	98,4	61,9	83,7
3	95,2	78,6	88,5
4	96,8	78,6	89,4

Table 5. The classification power of the models

The variables selected by the model in each step are described in Table 6.

The significant variables are those in which the equation coefficients are non-zero. The column Exp (B) is the odds ratio, which assesses the contribution of each variable in the classification; in the case of the present study, it was the successful projects.

As new variables were entered into the model, there was an improvement in its ability to adjust, given that the statistic -2LL exhibited successive reductions up to step 4. Estimates of R^2 indicated that the model of step 4 was probably the best because both statistics exhibited the highest values of R^2, with 58% being observed for the Cox and Snell estimate and 79% for the Nagelkerke estimate, as can be observed in Table 7.

Another statistic used to evaluate the goodness of fit of the model is the Hosmer-Lemeshow test, which must be greater than 0.05 for a good fit of the model to the data. In this case, the test indicated a significant statistical fit of the model in the four steps performed, as shown in Table 8.

Discussion

In analyzing the individual variables that are important to the conceptual model, four factors (individual variables) are highlighted as significant, including (i) the ability of the company to understand the customer requirements and turn them into new product specifications, (ii) articulation of the new product project to the company's strategy (product and competitive strategies), (iii) relationships with technology suppliers and (iv) the generation and selection of ideas for new products.

The first significant variable is related to the ability of a company to identify and translate customer needs into product requirements and specifications. This finding is in agreement with previous studies (Souder et al., 1997; Cooper et al. 2004c; Kahn et al., 2006) because successful projects are those in which the market assessments were

Steps	Measures	B	S.E.	Wald	df	Sig. (p<0,1)	Exp(B)
Step 1	1.3	2,2	0,73	9,58	1	0,002*	9,40
	Constant	-1,8	0,54	11,01	1	0,001*	0,17
Step 2	1.3	3,2	1,19	7,44	1	0,006*	26,00
	2.4	2,3	1,11	4,35	1	0,037*	10,00
	Constant	-3,9	1,18	10,93	1	0,001*	0,02
Step 3	1.3	2,5	1,27	4,04	1	0,045*	13,00
	2.4	2,3	1,32	3,08	1	0,079*	10,00
	8.2	2,4	1,05	5,03	1	0,025*	11,00
	Constant	-4,1	1,37	9,14	1	0,002*	0,02
Step 4	1.3	3,3	1,55	4,45	1	0,035*	26,00
	2.4	2,2	1,41	2,40	1	0,1**	8,90
	3.4	-4,2	1,58	6,96	1	0,008*	0,02
	8.2	3,0	1,21	6,20	1	0,013*	20,00
	Constant	-2,4	1,35	3,08	1	0,079*	0,09

*(p<0,05); ** (p<0,1).
Table 6. The variables included in the model

Step	*-2 Log likelihood*	*Cox & Snell R Square*	*Nagelkerke R Square*
1	97,94	0,33	0,45
2	73,36	0,47	0,64
3	61,32	0,53	0,72
4	49,35	0,58	0,79

Table 7. The statistics of each step

Step	Chi-square	df	Sig.
1	0,00	2,00	1,00
2	3,51	6,00	0,74
3	8,41	8,00	0,39
4	3,13	7,00	0,87

Table 8. Hosmer-Lemeshow Test

well-performed and the requirements of users were correctly translated into the new product specifications, which reinforces the importance of the construct marketing skills and denotes the need for greater quality of the execution of FFE activities.

The alignment between the new product strategy and the competitive strategy of the company is one a critical factor for the new product success. Four statements that address the adequacy of long-term strategies and project strategies, the average classifications attributed to successful products were statistically greater than the average classifications assigned to products not achieving success. The alignment between the products to be developed and the competitive strategies of the company was indicated as a success factor, which is in agreement with the studies of Clark and Wheelwright (1993) and Cooper et al. (2004a).

The third significant individual variable recognizes the importance of innovation by incorporating and adapting technologies obtained from external sources, particularly companies with commercial relationships (suppliers of machinery, equipment, materials, components or software). This approach is more common in the small and medium-sized technology-based companies as the companies that were considered in the present study.

The activities associated with the generation and selection of ideas must be carefully managed in NPD. These activities highlight the importance of FFE activities, as indicated by several authors (Kahn et al., 2006; Cao et al. 2011). The management of FFE has a significant impact on the performance indicators of cost, project quality and time to project development and product launch. Further, the high-quality implementation of activities for generating and selecting ideas can facilitate understanding of the characteristics desired by consumers, as well as the strategic vision of the company throughout the development cycle.

Therefore, successful projects are those in which there was a strategic alignment of the project, the market assessments were well-conducted and the user requirements were correctly translated into the new product specifications, as seen in Figure 2. This methodology requires an emphasis on the FFE activities that must be well-executed and includes the participation of external sources for technological innovation. Therefore, the significant variables identified by the forecasting model provide insights such that NPD managers and academics focus their attention on variables related to the constructs, such as the product strategy, the target market, the quality of NPD activity implementation and the technology sources.

Final considerations

The new product development process is critical for the company survive. With so many practices affecting the new product success, it is hard to predict the launch of successful new products. Given the importance devoted to NPD activities, it is critical that NPD managers focus on best practices to help ensure commercial success. This paper presented a method that NPD managers can use to predict the new product success prior to launch, allowing the efficient resource and commitment allocation.

The conceptual model incorporated a number of factors that were considered critical for NPD management. Certain factors were statistically confirmed; however, others did not significantly influence the outcome of the new product, although they are cited in the bibliography. This finding highlights practices that predict success in launching new products.
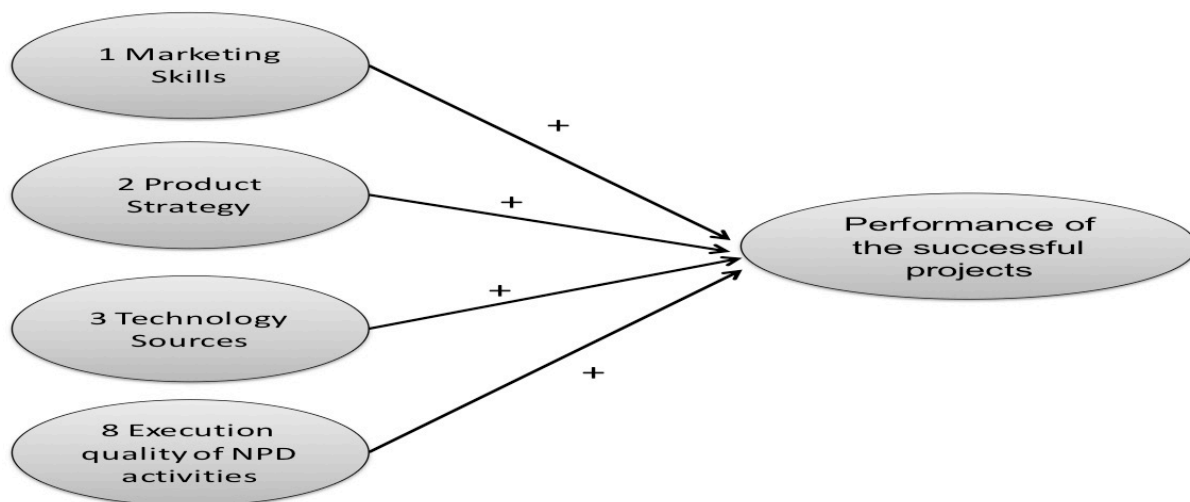

Figure 2. The constructs confirmed by logistic regression

There are several limitations of this research. It is important to note that the study was conducted in specific sectors, which, in this case, were medical devices and process automation devices and, therefore, the results cannot be generalized to all types of industries. A second limitation is the necessity of measuring the causal relationship between different NPD practices and the new product outcome.

From the perspective of mathematical models and multivariate statistics, the use of principal components analysis for categorical data is indicated as an important tool for reducing the number of variables constituting the model prescribed by logistic regression, which was adequate in predicting the success of NPD based on constructs with strictly qualitative measures of performance, obtaining a classification power of approximately 90%. The combined use of mathematical techniques and multivariate statistics was demonstrated to be an indispensable tool in the present study, and this approach should be encouraged within both the academic and professional markets.

In future studies, other multivariate techniques, such as artificial neural networks, fuzzy logic and classification trees, could be used to test and compare their respective classification powers in predicting the NPD success.

References

ACUR, N., Kandemir, D.; Boer, H. (2012). Strategic Alignment and New Product Development: Drivers and Performance Effects. Journal of Product Innovation Management, 29, 304–318.

AL-GHAMDI, A. S., (2002). Using logistic regression to estimate the influence of accident factors on accident severity. Accident Analysis and Prevention, 34, 729-741.

ANDRES, J. de, Landajo, M.; Lorca, P. (2005). Forecasting business profitability by using classification techniques: A comparative analysis based on a Spanish case. European Journal of Operational Research, 167, 518-542.

BARCZAK, G., Griffin, A.; Kahn, K. B. (2009). Perspective: trends and drivers of success in NPD practices: Results of the 2003 PDMA best practices study. Journal of Product Innovation Management, 26, 3–23.

BERKOWITZ, D; Wren, B.M.; Grant, E.S. (2007). Predicting new product success or failure: a comparison of US and UK practices. Journal of Comparative International Management, I, 50-65.

BERNADÓ, E.; Garrell, J. M. (2003). Accuracy-based learning classifier systems: Models, analysis and applications to classification tasks. Evolutionary Computation, 11, 209–238

BERTRAND, J.W.M.; Fransoo, J.C. (2002). Modelling and Simulation: Operations management research methodologies using quantitative modeling. International Journal of Operations and Production Management, 22, 241-264.

BIELZA, C.; Barreiro, P.; Rodriguez-Galiano, M. I.; Martin, J. (2003). Logistic regression for simulating damage occurrence on a fruit grading line. Computers and Electronics in Agriculture, 39, 95-113.

BROWN, B. L.; Hendrix, S. B.; Hedges, D. W; Smith, T. B. (2011). Multivariate Analysis for the Biobehavioral and Social Sciences: A Graphical Approach. Hoboken, NJ, USA: John Wiley and Sons, Inc.

BROWN, S. L.; Eisenhardt, K. M. (1995). Product development: Past research, present findings, and future-directions. Academy of Management Review, 20, 343-378.

CAO, Y.; Zhao, L; Nagahira, A. (2011). The impact of front end innovation in new product development in Japanese manufacturing companies. Nankai Business Review International, 2, 98–113.

CERPA, N.; Bardeen, M.; Kitchenham, B.; Verner, J. (2010). Evaluating logistic regression models to estimate software project outcomes. Information and Software Technology, 52, 934-944.

CHATTERJEE, S.; Hadi, A. S. (2006). Regression analysis by example. (4edition) John Willey and Sons, 375 p.

CHEN, M.Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. Expert Systems with Applications, 38, 11261–11272.

CLARK, K. B.; Wheelwright, S .C. (1993). Managing new product and process development: text and cases. New York: The Free Press.

COOPER, R. G.; Kleinschimidt, E. (1995). Benchmarking the firm's critical success factors in New Product Development. The Journal of Product Innovation Management, 12, 374-391.

COOPER, R. G., Scott, E., Kleinschimidt, E.; Elko, J. (2004a). Benchmarking best NPD practices – I. Research Technology Management, 47, 31-43.

COOPER, R. G., Scott, E., Kleinschimidt, E.; Elko, J. (2004b). Benchmarking best NPD practices – II. Research Technology Management, 47, 50-59.

COOPER, R. G., Scott, E., Kleinschimidt, E.; Elko, J. (2004c). Benchmarking best NPD practices – III. Research Technology Management, 47, 43-55.

CRONBACH, L.J. (1951). Coefficient alpha and internal structure of tests, Psychometrika, v. 16, p. 297-334.

DEVELLIS, R. F. (2012). Scale development: theory and applications. . 3 rd ed. London: Sage Publications.

DUDA, R; Hart, P. (2001). Pattern classification. John Wiley. ERNST, H. (2002). Success factors of new product development: A review of the empirical literature. International Journal of Management Review, 4, 1-40.

FLORA, D. B.; Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychological methods, 9, 466–491.

FLURY, B.; Riedwyl, H. (1990) Multivariate Statistics. A Practical Approach, New York: Chapman and Hall.

FORERO, C. G.; Maydeu-Olivares, A.; Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. Structural Equation Modeling, 16, 625–641.

FORZA, C. (2002). Survey research in operations management: a process-based perspective. International Journal of Operations and Production Management, 22, 152-194.

GRIFFIN, A. (1997). PDMA Research on new product development practices: Updating trends and bechmarking best practices. Journal of Product Innovation Management, 14, 429-458.

HAIR J.F., Black, W.C., Babin, B.J., Anderson, R.E; Tatham, R.L. (2006). Multivariate data analysis. 4th ed. Prentice Hall: New Jersey.

HARREL, F. E. Jr. (2001). Regression modelling strategies with applications to linear models, logistic regression and survival analysis. Berlin: Springer,

HAVERILA, M. (2010). The marketplace variables in successful and unsuccessful NPD projects in technology intensive companies. Journal of Technology Management & Innovation, 5(4), 121-136.

HOSMER, D. W.; Lemeshow, S. (2000). Applied logistic regression. New York: John Wiley and Sons.

JOHNSON, R.A.; Wichern, D.W. (2002). Applied multivariate statistical analysis. 5th ed. Upper Saddle Rover, NJ: Prentice Hall.

JÖRESKOG, K. G..; Moustaki, I. (2001). Factor analysis of ordinal variables: a comparison of three approaches. Multivariate Behavioral Research, 36, 347-387.

JUGEND, D; Silva, S.L. (2012): Integration in New Product Development: Case Study in a Large Brazilian High-Technology. Company. Journal of Technology Management & Innovation, 7 (1), 52-63.

KAEFER, F., Heilman, C. M.; Ramenofsky, S. D. (2005). A neural network application to consumer classification to improve the timing of direct marketing activities. Computers and Operations Research, 32, 2595–2615.

KAHN, K. B., Barczak, G..; Moss, R. (2006). Perspective: Establishing an NPD best practices Framework. Journal of Product Innovation Management, 23, 106-116.

KAHN, K. B., Barczak, G., Nicholas, J., Ledwith, A.; Perks P. J. (2012). An examination of new product development best practice. Journal of Product Innovation Management, 29, 180–192.

KAISER H.F. (1974). An index of factorial simplicity. Psychometric, 39, 31-36.

KLEINBAUM, D. G. (1994) .Logistic regression: a self-learning text. New York: Springer.

KRISHNAKUMAR, J.; Nagar, A. L. (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. Social Indicators Research, 86, 3, 481-496.

LANGERAK, F., Hultink, E. J.; Robben, H. S. J. (2004). The role of predevelopment activities in the relationship between market orientation and performance. R&D Management, 34, 295–309.

LARSON, E. W.; Gobelli, D. H. (1988). Organizing for product development projects. Journal of Product Innovation Management, 5, 180–190.

LATTIN, J.; Carroll, J. D.; Green, P. E. (2011). Analysing multivariate data. Pacific Grove, CA: Brooks/Cole, Thomson Learning.

LEE, J., Lee, J.; Souder, W. E. (2000). Differences of organizational characteristics in new product development cross-cultural comparison of Korea and US. Technovation, 20, 497-508.

MANLY, B. J. F. (2005). Multivariate Statistical Methods: a primer. 3 rd ed. Boca Raton, FL: Chapman and Hall.

MARCH-CHORDA, I.; Gunasekaran, A.; Begoña, L. (2002). .Product development process in Spanish SMEs: an empirical research. Technovation, 22, 301–312.

MARTIN, A. R. G.; Nieto, J. M. M. N.; Ruiz, J. P. N.; Jimenez, L. S. E. (2008). Overweight and obesity: The role of education, employment and income in Spanish adults. Appetite, 51, 266–272.

MENARD, S. (1993). Applied logistic regression analysis. Thousand Oaks, CA: Sage. (Series Quantitative applications in the social sciences).

MEULMAN, J. J. (1998) Optimal scaling methods for multivariate categorical data analysis. SPSS white paper. p. 1-12.

MEULMAN, J.J. (1992). The integration of multidimensional scaling and multivariate analysis with optimal transformation of variables. Spychometrica, 57, 539-565.

MEULMAN, J.J., Van der Kooij,. A.J.,; Heiser, W.J. (2004). Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data. In: D. Kaplan (ed.), Handbook of Quantitative Methods in the Social Sciences, (pp. 49-70). Newbury Park, CA: Sage Publications.

MOROCO, J. (2003). Análise estatística com utilização do SPSS. Lisboa: Sílabo.

MYERS, R. H. (1990). Classical and modern regression with applications. (2nd ed). Boston, Massachusetts: PWS-KENT Publishing Company.

NUNNALLY, J. C. (1978). Psychometric Theory. New York, NY: McGraw-Hill.

OHLSON, J. A. (1980). Financial reactions and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 18, 109-131.

PAVLEKOVIC, M., Bensic, M.; Zekic-Susac, M. (2010). Modeling children's mathematical gift by neural networks and logistic regression. Expert Systems with Applications: An International Journal, 37, 10.

SCOTT, G. M. (2000). Critical technology management issues of new product development in high tech companies. Journal of Product Innovation Management, 17, 57-77.

SEXTON, R., and Mcmurtrey, S. (2005). Employee turnover: A neural network solution. Computers and Operation Research, 32, 2635–2651

SHARMA, S. C. (1996). Applied multivariate techniques. New York: John Wiley and Sons.

SHERMAN, J. D., Berkowitz, D.; Souder, W. E. (2005). New product development performance and the interaction of cross-functional integration and knowledge management. Journal of Product Innovation Management, 22, 399-411.

SHUM, P.; Lin, G. (2007). A world class new product development best practices model. International Journal of Production Research, 45, 2007, 1609-1629.

SONG, X. M.; Noh, J. (2006). Best new product development and management practices in the Korean high-tech industry. Industrial Marketing Management, 35, 262–278.

SONG, X. M.; Parry, M. E. (1996). What separates Japanese new product winners from losers. Journal of Product Innovation Management, 13, 422□439.

SONG, X. M., Souder, W.; Dyer, B. (1997). A causal model for studying the impact of skills, synergy, and design sensitivity on new product performance. Journal of Product Innovation Management, 14, 88–101.

SOUDER, W. E., Buisson, D.; Garret, T. (1997). Success through customer-driven new product development: a comparison of US and New Zealand small entrepreneurial high technology firms. Journal of Product Innovation Management, 14, 459-472.

STUDENMUND, A. H. (1992). Using econometrics: A practical guide. New York: Harper Collins.

THIEME, R. J.; Song, X. M.; Shin, G. C. (2003). Project management characteristics and new product survival Journal of Product Innovation Management, 20, 104-119.

VERMUNT, J. K. (2007). Growth models for categorical response variables: standard, latent-class, and hybrid approaches. In: K. Van Montfort, H. Oud, and A. Satorra (eds.), Longitudinal Models in the Behavioral and Related Sciences, 139-158. Mahwah, NJ: Erlbaum.

WIRTH, R. J.; Edwards, M. C. (2007). Item Factor Analysis: Current Approaches and Future Directions. Psychological Methods, 12, 58–79.

YAMANE, T. (1967). Statistics: an introductory analysis. 2. ed. New York: HarperCollins.

YAP, C. M.; Souder, W. E. (1994). Factors influencing new product success and failure in small entrepreneurial high-technology electronic firms. Journal of Product Innovation Management, 11, 418-432.

ZAPATA, A.R.P, Cantú, S.O. (2008): Gestion Estrategica De La Tecnologia En El Predesarrollo De Nuevos Productos. Journal of Technology Management & Innovation, 3 (3), 112-122.