An empirical study of presage variables in the teaching-learning of statistics, in the light of research on competencies*

Clemente Rodríguez, José Gutiérrez, Teresa Pozo

Faculty of Educational Sciences, University of Granada

Spain

Correspondence: Clemente Rodríguez Sabiote. Departamento de Métodos de Investigación y Diagnóstico en Educación. Facultad de Ciencias de la Educación. Campus de Cartuja, s/n. 18071 Granada. Spain. E-mail: clerosa@ugr.es

* Work financed by the Spanish Ministry of Science & Innovation, under Project ED2008—03098.

Abstract

Introduction. This research seeks to determine the influence exercised by a set of presage and process variables (students' pre-existing opinion towards statistics, their dedication to mastery of statistics content, assessment of the teaching materials, and the teacher's effort in the teaching of statistics) in students' resolution of activities and acquisition of competencies relating to quantitative data analysis (product variable).

Method. We developed a correlational-predictive study with a sample of undergraduate students in Pedagogy from the University of Granada (N = 70). These students responded to a test with five problem situations, resolving the situations after having first rated the influence of the five predictive variables in their resolution process.

Results. The results obtained from analyzing a linear multiple regression indicate that the teacher's effort and students' attitude towards statistics (in this order) play an essential role in performance achieved in the resolution of statistics problems and in mastering competencies of quantitative data analysis.

Discussion or **Conclusion**: The model inferred from this study, in this particular case, gives less importance to the role of student effort or of supplementary instructional materials made available to students in their performance in statistics.

Keywords. Statistical achievement, Biggs´ 3P model, De la Fuente´ DEDEPRO model, influential variables in statistics achievement, linear multiple regression.

Received: 11/10/09 *Initial Acceptance:* 19/10/09 *Final Acceptance:* 01/11/10

Resumen

Introducción. En la presente investigación pretendemos determinar la influencia que una colección de variables de tipo pronóstico y proceso (opinión del alumnado hacia la estadística, su dedicación a los contenidos de estadística, la valoración de los materiales didácticos y la labor del docente en la enseñanza de la estadística) han jugado en la resolución de actividades y adquisición de competencias relacionadas con el análisis de datos cuantitativos (variable producto).

Método. Se ha desarrollado un estudio correlacional-predictivo con una muestra que corresponde a una cohorte de estudiantes de la Licenciatura de Pedagogía de la Universidad de Granada (N= 70) que se sometieron a un protocolo de cinco situaciones problemáticas y que, además, valoraron con anterioridad la influencia de cinco variables predictivas en la resolución de dichas situaciones.

Resultados. Los resultados alcanzados tras la implementación de un análisis de regresión lineal múltiple indican que, tanto la labor del profesorado, como la actitud hacia la estadística (por este orden) juegan un papel primordial en el rendimiento logrado en la resolución de problemas de estadística y el dominio de competencias de análisis de datos cuantitativos.

Conclusión. El modelo inferido, en este caso particular, ha restado relevancia a la influencia que puedan jugar en el rendimiento estadístico la preparación del alumnado, así como los materiales curriculares puestos a su disposición para su preparación de la asignatura.

Palabras Clave: Rendimiento estadístico, Modelo 3P de Biggs, Modelo DEDEPRO de De la Fuente, Variables determinantes del rendimiento estadístico, Regresión lineal múltiple.

Recibido: 10/11/09 *Aceptación Inicial:* 19/11/09 *Aceptación Definitiva:* 11/01/10

Introduction

The teaching of statistics has a prominent place in many university degrees, and its mastery is essential in different fields of knowledge, not only the so-called “hard” sciences, but also in social, legal and health sciences. Degrees such as Pedagogy, Psychology and Sociology include required material that enables students to acquire competencies basic to these professions. Statistics also has some presence in compulsory education, although usually at a basic level and it is addressed at the end of the syllabus, time permitting. However, as Behar and Grima (2001) argue, there has been a certain oversight, call it uninterest, in the quality of teaching-learning processes for this topic. An immediate consequence has been the stigmatization of this discipline, not only among the students, but also among a certain portion of the teaching staff that teach it.

In summary, statistics as a general rule certainly does not awaken great delight among students who study this discipline. On the contrary, it often generates a kind of phobia that Dillon (1982) named *statisticophobia,* where students perceive the subject of statistics as an obstacle to completion of their degree program (Perney & Ravid, 1990). Debate centers on their fears, anxiety and unfavorable atittude toward trying to learn and to enjoy this discipline (Mondejar, Vargas & Mondéjar 2008; Onwuegbuzie & Wilson, 2003; Justicia, Pichardo, Cano, García, De la Fuente, 2008).

With these points of reference, it is therefore important to try to establish mechanisms for regulating the quality of teaching-learning processes in statistics. One previous strategy along these lines is to consider a set of dimensions or variables that can positively or negatively affect teaching delivery of statistical content at the university level (Tempelaar, Van der Loeff & Gijselaers, 2007). With this objective in mind, we will briefly dissect some of these influencing dimensions, seeking to assess their predictive ability in developing good practices for the teaching-learning of statistics, in light of competencies required in the European Higher Education Area and of advances in contemporary research with integrative models that incorporate presage, process and product variables.

Foundational frameworks and contemporary approaches to research on teaching-learning in Higher Education

The construction of global models for understanding and improving teaching-learning processes has been a lively field for empirical research, with abundant research contributions over the 20th century and continuing today. Several lines of research have produced findings and contributions that come together into at least five groups of influence. Each of these groups begins historically with different concerns and focuses of interest, although it is not easy to establish strict boundaries between them or between the topics that describe their sequencing and evolution over time. These groups have been characterized in the literature under the founding labels: the presage-product paradigm, the process-product paradigm, the student-centered mediational paradigm, the mediational paradigm centered on teacher decision-making, and the ecological paradigm.

> *"Research studies based on these paradigms have been produced and continue to be produced in parallel. And not even today can we affirm that one of these paradigms has become definitively established as the winning paradigm. Educational science is at an earlier stage, prior to configuration of theories and consolidation of paradigms ... Nonetheless, it is possible to affirm that the presage-product paradigm is no longer in force and the process-product paradigm has received so much criticism and has had to produce so many "ad hoc" explanations and be reformulated so often that it also appears to be exhausted, or unrecognizable, as a paradigm"* (Pérez, 1983, p. 98).

More recent studies guided by foundational frameworks such as phenomenography, constructivism or studies on information processing, from both quantitative and quality perspectives, emphasize the need to address teaching-learning issues from integrating models that examine teaching-learning within its context, from the institutional demands and pressures within which it is implemented, and under the prism of the characteristics and perceptions of the players involved in the processes. This has forced researchers to reorient their studies toward the analysis of: 1) intervening cognitive, motivational and academic variables (Evans et al. 2003; Waugh, 2002; De la Fuente, & Justicia, 2007); 2) consideration of individual preferences (Hernández, Martínez, Fonseca & Rubio, 2005; Mondéjar et al., 2008; Tempelaar et al., 2007) an appreciation of and differentiation among learning styles, conditioned by contextual differences; 4) the influence of curriculum materials, virtual resources and pressures for social change and technological innovation; 5) continuous professional development and the need for ongoing preparation, brought on by changing scenarios in the work world.

This research draws from a tradition that requires new approaches in the light of contemporary advances where teaching-learning is affected by new, important context variables. The new, competency-based working models being imposed by European convergence force the teacher to establish more systematic controls over his/her teaching activity, over the materials in use, the different modalities of holding class, the time required for each activity, each activity's relevance and suitability for the level reached and for progress in content, whether task prerequisites have been met, etc. (De Miguel, 2006).

The present study seeks to line up with these new integrated approaches to research on the university classroom, while recovering some important principles from the presage-product paradigm. We emphasize some aspects such as new areas of interest that require systematic, ongoing exploration over time in order to have solid criteria for adjusting teacher decisions to the new requirements of training by competencies. On the other hand, the presence of statistics in many spheres of daily life and of professional activity justify its inclusion in new curriculum plans for the undergraduate degree in Educational Sciences, along with the need to develop programmed teaching which is well-founded in contextualized research, in the light of findings from pertinent studies in this field.

Exploration of this topic as a relevant line of research is also justified in part by students' rejection and aversion toward the subject matter, by the myths, stereotypes and phobias that prevail; while there is a simultaneous, growing demand for mastery of this discipline in different professional contexts. Within the educational dynamic of the EHEA, the process begins with describing the learning outcomes in terms of knowledge, skills, abilities, comprehension, etc., that the student will be able to show upon satisfactory completion of the educational process. These outcomes will be described based on a curriculum that responds to integrated professional profiles comprising action competencies, creating a positive climate that enables interpersonal relations and sympathy with the discipline itself as well as between the players involved (teacher-students). This is accomplished through teaching activities and tasks and assessments that allow students to use their knowledge actively and critically (Hernández et al., 2005: 28; De la Fuente, & Justicia, 2007).

Electronic Journal of Research in Educational Psychology, *8*(1), 235-262. 2010 (nº 20). ISSN: 1696-2095.

New generation models that integrate the Presage, Process and Product variables

The Biggs 3P model (2005) starts from the assumption that students use certain learning strategies as a function of the motives they have for learning, these being one of the central determinants of the teaching-learning process. Under this structure, students undertake actions and carry out tasks as a function of what they think will contribute to greater academic success and will be rewarded by the system, as a function of the motives that have prompted their learning. This requires that the teacher be aware that each student is pursuing a different path in his/her learning process, and that these paths are closely related to the reasons that prompt students to find successful solutions.

The DEDEPRO model of teaching-learning regulation, proposed by De la Fuente and Justicia (2007), shows how integrating contextual elements and teaching-learning processes contributes effectively to the students' self-regulated learning. Phenomenographic studies on modes of information processing related to different cognitive styles and student motivations, have been outlined in the literature under two large groups based on different conceptions of the use of knowledge. The first group has been characterized by a learning approach oriented toward memorization and surface reproduction of the content being studied. The second, under the label of deep learning, is the type of learning developed by students who have intrinsic goals of more reflective, active, comprehensive knowledge, and so on.

Deep and surface approaches to teaching-learning, although they have certain stability over time, are subject to ongoing modifications, the product of adaptation and adjustment processes between agents and their environments; they undergo modifications and changes determined by the context where learning and interpersonal interaction is taking place. For this reason, we start from the idea that it is possible to construct and encourage creation of a deep approach to teaching-learning through conscious processes of planning instruction. In our case, statistics and the many myths associated with its teaching-learning become the focal point of our attention in this article. A more detailed analysis of what is involved in each type of teaching-learning, for the teacher and for the students, is presented below (Tables 1 and 2):

Table 1. Surface approach in the Teaching-Learning of Statistics: implicit factors for teachers and students (Adapted from Biggs, 2005)

Implications of the Surface Approach for the Teacher of Statistics	Implications of the Surface Approach for the Student of Statistics
▪ Teaching in an unsystematic fashion. ▪ Providing lists without presenting either the intrinsic structure of the content or its relationship to the topic or the subject of statistics. ▪ Assessing independent data, as is often done when using short answer and multiple-choice tests. ▪ Showing little interest in the material being taught. ▪ Not allowing sufficient time for full dedication to the task, emphasizing getting through the program at the expense of depth. ▪ Provoking undue anxiety or frustratingly limited expectations for success: "whoever doesn't get it, that's his problem"	▪ Pursuit of a barely passing grade, resulting from an idea of the university as an imposed, compulsory task, as a meaningless trial of academic life that you must get through however you can. ▪ There are extra-curricular priorities and interests that come ahead of academics. ▪ A feeling of work overload and the attempt to bargain over assignments to bring down the time required. ▪ Frequent misinterpretation of activities and assignments, as well as mistakenly focusing one's effort on remembering anecdotes, isolated data points and disconnected facts. ▪ Skeptical view of education, constant questioning. ▪ High levels of anxiety. ▪ Incapacity to show an understanding of the discipline at complex levels, or to discuss nuances of meaning.

Table 2. Deep approach in the Teaching-Learning of Statistics: implicit factors for teachers and students (Adapted from Biggs, 2005)

Implications of the Deep Approach for the Teacher of Statistics	Implications of the Deep Approach for the Student of Statistics
▪ Teaching in such a way that clearly shows the structure of the topics and the relationships between conceptual aspects. ▪ Teaching to elicit a positive response from the students, formulating significant questions and posing important problems, rather than simply offering information. ▪ Teaching from the basis of what students already know, starting from their initial ideas. ▪ Questioning and transforming students' erroneous conceptions, considering error to be a recourse for learning. ▪ Assessment based on the deeper aspects of learning, not on format or anecdotal, disconnected data.	▪ Intent to address the task in a significant, appropriate fashion, whether due to an intrinsic curiosity or to the determination to do things well on principle. ▪ Appropriate prior knowledge, making it possible to focus at a higher conceptual level and to use high-level cognition strategies. ▪ Well-structured knowledge base, built on relevant, meaningful prior knowledge. ▪ A genuine preference for and corresponding ability to respond to unexpected, open questions, involving an intellectual challenge and requiring a connection with complex solutions.

The 3P and DEDEPRO models show learning-related factors anchored in three points of time: prognosis, before learning is produced; process, during learning; and product, or end result of learning. Proposals from Biggs (2005:38) and De la Fuente and Justicia (2007) integrate the Presage-Process-Product variables into their models, as new generation models that draw out the importance of initial variables along with the interventions of process, that is: 1) Presage variables include variables related to the teaching context such as curriculum content, teaching method, curriculum materials, new technologies, teacher performance, classroom climate, evaluation model; and student-related variables such as prior knowledge, skills, expectations and attitudes; 2) Process variables related to the effects of metacognitive processes and their corresponding approaches to learning (surface and deep); 3) The end result of learning is determined by many factors which interact among themselves, and can be described in quantitative terms (how much has been learned), qualitative terms (the meaning of what was learned, and improvements of a cognitive, attitudinal or procedural nature) and institutional terms (grades assigned and change indicators).

Studies related to the teaching-learning process in Statistics

First, we may consider a group of studies focused on the *affective dimension*, where the *student's attitude toward statistics* is crucial for adequate learning and effective use (Carmona, 2004; Estrada, Batanero & Fortuny 2004). There are many studies which reveal the close relationship between attitude towards statistics and academic performance. We note the studies by Beins (1985), Katz and Tomezik (1988), Roberts and Saxe (1982), Tempelar et al. (2007) and Wise (1985), as well as studies here in Spain, by Auzmendi (1992), Gil and Sánchez-López (1996), Gil (1999), Bayot, Mondéjar, Mondéjar, Monsalve, and Vargas (2005), and Mondéjar et al. (2008). In general, all these studies reveal that students do not face the learning of statistics with the most suitable attitude and this appears to effect learning outcomes in this discipline, where the better one's attitude, the better one's performance, and vice versa.

A second and equally important group of investigations attends to how the *teacher holds class*. In this area we refer to his or her knowledge of the material, how it is communicated, didactic resources used, and so on. Bradstreet (1996) proposes a set of strategies that can improve teaching activity, such as establishing standard guidelines for statistical good practices; ensuring that the text book is compatible with the readers' level of knowledge; connecting statistical concepts and methods with friendly software (Behar & Grima, 2001). Furthermore, Bradstreet insists that it is important for teaching to be built onto the basis of the learner's personal experiences, trying to connect abstract concepts with personal experiences, such as drawing up analogies, similes and metaphors that can be very useful instructional tools.

Burrill (1990) also proposes a set of good practices that can produce a positive effect in the teaching of statistics. She speaks of interactive, constructive learning, with students having an active role; students should first experiment with simple strategies such as counting frequencies, creating simple graphs, etc., before going on to address more complex learning; statistical topics should be dressed up in designs and presentations that make the content attractive; any assignments should emphasize not so much precision in results but rather their proper interpretation. Finding problem situations relevant to students' motivations and interests will help overcome these phobias.

A third group of studies addresses *students' dedication to learning the subject matter of statistics*. Undoubtedly, this dimension is related to one's attitude toward the discipline, one's motivation when approaching learning, and consequently, class attendance (Rodríguez & Herrera, 2009), and the importance given to these classes (Sahai, Behar & Ojeda, 1998). Where these conditioning factors take on negative connotations, the typical motive for studying statistics is to get through the degree program, where it is required to pass this class in order to earn one's degree. Unfortunately, this motive tends to generate surface learning in the terms described by Biggs (2005), memorization which has little or nothing to do with helping the student establish a deep, well-founded understanding of statistics. An appropriate strategy would be to encourage intrinsic motivation, so that the students themselves participate in the learning process and value its utility.

Likewise, Behar and Grima (2001, p.196) indicate that tasks involving a challenge that can be reached, with a moderate amount of discrepancy or apparent incongruence, stimulate students' curiosity and become an element of intrinsic motivation. In short, posing problems related to the profession, which cannot be solved without the use of statistics, is an excellent way to motivate students in their study.

Finally, we point to a fourth area of interest, regarding *technological tools* and their influence in the teaching-learning process of statistics. Along these lines, Mondéjar, Vargas and Mondéjar (2007, p.32) affirm that both new technologies and virtual education are becoming popular new options, not only because they offer methodologies that are easily adapted to a wide range of students, but also due to the importance given to auto-didactic processes, to the search for knowledge and to fostering research, as well as improving teaching-learning processes (De la Fuente, Cano, Justicia, Pichardo, García, Martínez & Sander, 2005).

Among these technology tools, Alpízar (2007:101) highlights tutorials, internet resources and microworlds, and statistical packages which, according to Balacheff and Kaput (1996), make a connection between statistics and daily life, since they permit the modeling of concrete situations and the use of real data. A recent example in this direction is the experience from Cuesta and Herrero (2008). The presence of technology in the statistics classroom became a powerful tool which made available representation systems that could be used for visualizing and experimenting with important concepts during class sessions (Alpízar, 2007, p.98).

New technologies have put in question the routine manipulation of data using paper and pencil, since they offer more effective strategies for solving problems with greater precision (Santos, 1997:5). Data organization and analysis is made easier, and calculations become simpler and more orderly, freeing the student to dedicate more time to interpretation, discussion and reflection on outcomes (Ben-Zvi, 2000; Ben-Zvi & Arcavi, 2001; Marshall, Makar & Kazak, 2002).

Objectives of the study

The general objective of the study is to determine the influence of a set of relevant variables (predictors), based on student opinion, on academic performance in the course entitled *Methodological Foundations of Educational Research*, for the portion dedicated to quantitative data analysis (statistics). This is a required course for the undergraduate degree in Pedagogy at the University of Granada.

More specifically, the objective which guides the present research is to determine the influence of the following variables: students' pre-existing attitude towards statistics, students' dedication to mastering the statistics portion of the course, teacher's effort in teaching the statistics portion of this course, help from the SPSS practice notebooks, and help from other introductory instructional materials used in solving activities related to quantitative data analysis, as carried out by students in the first year of the undergraduate Pedagogy program at the University of Granada.

Method

Participants

The population under study refers to students who, according to official enrollment data, took the subject *Methodological Foundations of Research* (group A) during the academic year 2006-07: 117 male and female students. However, for the present study we cannot speak of a sampling process, since the participants were those who sat for the final exam in June 2007, that is, 61.4% of the total. We do not delve into differential traits, such as age, gender, and so on, since there is little variability within the sample studied here, a large majority of students are female, 18- or 19-year-olds.

Instrument

As a data collection instrument, an *ad hoc* test was created using a protocol of 5 problem situations pertaining to different statistics content from the course, and presented as the June 2007 final exam for *Methodological Foundations of Educational Research*. In addition, the instrument started off with 5 additional questions requiring the student to rate 5 different variables that had acted as predictors, using a Likert-type scale. Thus, before solving the problems presented, students had to rate the five variables with their corresponding level as described above.

Procedure

The present study is clearly a correlational-predictive type. For many authors, such as Van Dalen and Meyer (1983), the correlational method is not a method is its own right. However, a greater number do consider it a method in its own right, including Arnal, Del Rincón and Latorre, (1994); Cohen and Manion (2002), Bisquerra (2004) and Etxeberría (1999), since this methodology goes beyond mere description, often generating predictive studies based on regression (Arnal et al., 1994:184). Moreover, with Arnal et al. (1994:188), we can affirm that regression consists of bringing the points of a dispersion diagram around a straight line, in order to be able to predict values based on the equation of this straight line, or regression equation. Thus we have sought to determine the influence of a set of important, predictive variables on the academic performance of pupils in the statistics portion of their course in Methodological Foundations of Educational Research. Next we will specify the variables considered, as well as the different levels that were defined for each of them.

a) *Opinion toward statistics,* with 5 response levels: very bad; bad; acceptable, good and very good.
b) *The teacher's effort in conveying the statistics content and in making it understood,* with 5 response levels: very bad; bad; acceptable, good and very good.
c) *Student's dedication to mastering* the statistics material, with 5 response levels: very little; little; moderate; much and very much.
d) *Help from the SPSS notebooks* in mastering the statistics material, with 5 response levels: very low; low; moderate; high and very high.
e) *Help from other curriculum materials providing an introduction to quantitative analysis* in mastery of the statistics portion of the subject, with 5 response levels: very low; low; moderate; high and very high.

As a *criterion* variable we have taken students' final grade on the practical portion of their exam, that is, on the statistics problems. Finally, we took the parameters of reliability and validity for quality criteria regarding the instrument of measure, which collects opinion in a Likert format. Additionally, internal consistency was assured for reliability due to the characteristics of test administration (one single administration). The most precise and most widely used coefficient for this type of situation is the Cronbach alpha, which yielded the following outcomes:

Table 3. Cronbach Alpha for the information collection instrument

Cronbach's Alpha	N of Items
.878	5

Along with Morales (2008), we can affirm that the coefficient calculated is satisfactory (α=.87), since only values less than 0.6 can be considered questionable in cases such as ours (empirical or general research).

On the other hand, for validity we considered content and criterion types. For content validity, we verified the suitability of introducing the 5 variables. For this purpose we took as our basis the Cronbach alpha obtained if each one of the variables were not included. In the following table, we observe that in every case, this coefficient diminishes when a variable is not included, with the exception of *help from other materials*, whose removal signifies a slight improvement in reliability.

As for criterion validity, we correlated each item with the total for the 5 (criterion) minus the item being tested (corrected correlation), and obtained moderate and statistically significant Pearson correlation coefficients. This denotes adequate criterion validity in each case.

Table 4. Corrected correlations for each variable and Cronbach alpha for the test when the indicated variable is deleted

Variables	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Opinion toward statistics	.770*	.842
Dedication to subject mastery	.813*	.829
Role of the teacher	.826*	.828
Help from SPSS notebooks	.698*	.861
Help from other materials	.538*	.888

*Statiscally significant correlations with bilateral $\alpha = .01$.

Statistical analyses

In order to analyze the data collected in this study, we applied a multiple linear regression that was configured in different phases, as in Cea D´Ancona (2002:65): 1) Preliminaries to multiple regression analysis: bivariate correlation analysis based on the correlations matrix (exploratory inquiry), and prior verification of the principal basic assumptions required for multiple regression analysis; 2) Estimation of the regression equation; 3) Evaluation of the resulting model.

Results

Preliminaries to multiple regression analysis

Bivariate correlation analysis based on the correlations matrix

Before undertaking the multiple linear regression procedure, we implemented the existing correlations between the criterion variable and those considered independent. The objective of this procedure is two-fold. On one hand, it serves to verify the intensity of the relationship between the criterion variable and the predictive variables, so as to determine whether or not to continue with the analysis procedure. On the other hand, this correlational intensity becomes the basis for ranking variables from greatest to least, being established as an ordering criterion for introducing them into the regression model that will be determined.

Table 5. Correlations matrix between the criterion variable and the predictive variables

Criterion variable	**Predictive variables**	Opinion toward statistics	Dedication to subject mastery	Teacher performance	Help from SPSS notebooks	Help from other materials
FINAL GRADE	Pearson correlation	.70(**)	.63(**)	.703(**)	.46(**)	.32(**)
	Sig. (bilateral)	,000	,000	,000	.000	.006
	N	70	70	70	70	70

** The correlation is significant at a level of 0.01 (bilateral).

From the results obtained in this initial exploratory approach, we can affirm that, first, all resulting correlations are statistically significant at a level of α = .01 bilateral and that some of these, in addition, are quite strong. Second, correlations obtained between the criterion variable with the rest of the predictive variables fall into a decreasing order as follows: Teacher performance (r = .703) > Pre-existing opinion towards statistics (r = .70) > Dedication to subject mastery (r=.63) > Help from SPSS notebooks (r = .46) > Help from other curriculum materials (r = .32). Accordingly, this was the order used to introduce the variables into the multiple linear regression model, using the stepwise method.

Prior verification of the principal basic assumptions required for multiple regression analysis

The correct application of multiple regression analysis requires fulfillment of a number of prior assumptions that indicate its suitability. Along with Hair, Anderson, Thatam and Black (1999); Cea D´Ancona (2002) and Etxeberría (1999), as well as others, we emphasize the following: the normal distribution of scores that make up each of the variables involved, as well as multicollinearity among them.

a) Normal distribution

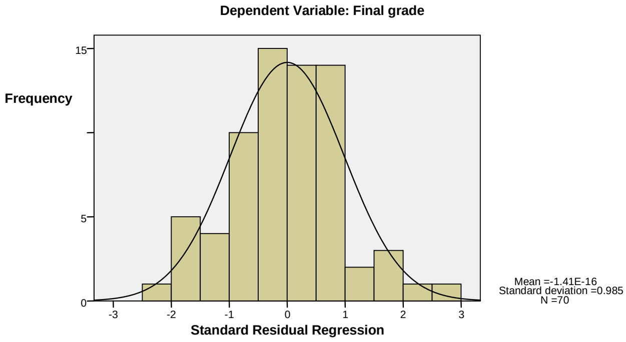

Figure 1. Bar graph of standardized residuals of the criterion variable

Table 6. Asymmetry indices of the predictive variables considered

Predictive variables	Opinion toward statistics	Teacher performance	Dedication to subject mastery	Help from SPSS notebooks	Help from other materials
Valid N	70	70	70	70	70
Asymmetry	.47	.23	.16	.11	-.26

A normal curve is appreciable from the histogram of standardized residuals, denoting a rather evident normal distribution where, moreover, we find distribution values along the symmetrical, bell-shaped curve with a mean of practically 0 and standard deviation at nearly 1 N(0,1). On the other hand, for the predictive variables the resulting asymmetry is fundamentally positive (except for one variable), with a predominance of values on the left side of the distributions. In any case, all asymmetry indices can be considered normal, as they fall within the interval of ±.5 (Gil, Rodríguez & García, 1995).

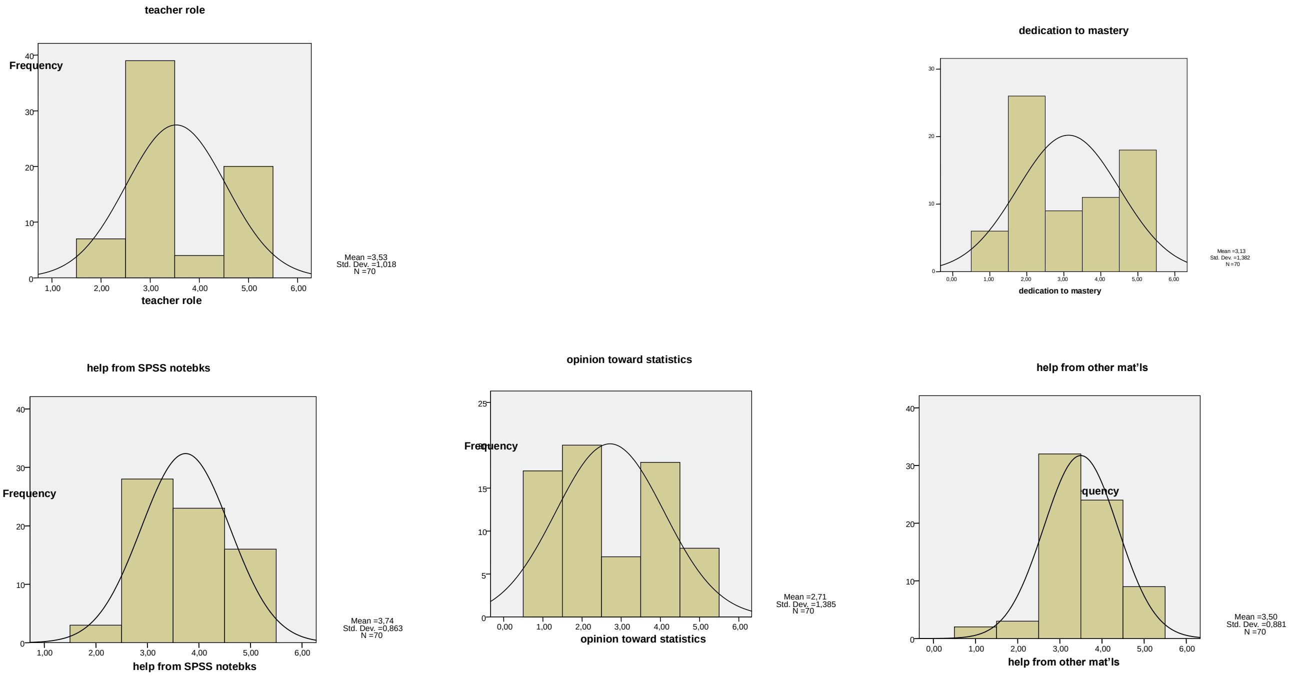

Figure 2. Set of bar graphs with normal curve, referencing the different dependent variables

b) Multicollinearity

Multicollinearity is understood to be the situation where two or more predictive variables from the model reach high bivariate correlations among themselves, possibly having negative repercussions on the results obtained from model. In order to determine the degree of multicollinearity of predictive variables, several diagnostic tests are used, most notably the Tolerance Index (TI), as well as its reciprocal, or inverse, the Variance Inflation Factor (VIF). For our study, therefore, we emphasize only the interpretive labels for both. Thus, as was also affirmed by Etxeberria (1999), Hair et al. (1999) and Cea D´Ancona (2002), values near the unit in the case of TI denote the nearly total absence of multicollinearity, such that values near 0 indicate just the opposite. Tolerance Indices (TI) < .2 begin to cause concern, and < .1 are unacceptable. For its part, the VIF is better the more that it approaches 1, so that values >5 are cause for concern and >10 are unacceptable. Next, we present results obtained for the two variables that were considered in the regression model and which we will spell out in greater detail further on.

Table 7. Tests of collinearity of predictive variables included in the regression equation

Collinearity statistics	Tolerance (TI)	VIF
Dedication to subject mastery	.496	2.016
Opinion toward statistics	.496	2.016

In light of these results, we can affirm that the variables in our model do not lack multicollinearity: they are within reasonable limits in order to be considered. Therefore, we can state finally that the predictive variables included in the regression model are suitable for the predictive purposes under consideration.

Estimation of the regression equation

Table 8. Process of introducing/deleting predictive variables in the regression model

Model	Variables introduced	Variables deleted	Method
1	Role of Teacher	.	Stepwise (criterion: F Prob. for entering <= .050, F Prob. for excluding >= .100).
2	Opinion towards statistics	.	Stepwise (criterion: F Prob. for entering <= .050, F Prob. for excluding >= .100).

** Dependent variable: final grade

Table 9. Variables excluded from the regression model finally inferred

Model		Beta within	t	Sig.	Partial correlation	Collinearity statistics
						Tolerance
1	opinion towards statistics	,382 [a]	2,953	,004	,339	,396
	dedication to mastery	,168 [a]	1,104	,273	,134	,316
	help from notebooks	,087 [a]	,834	,407	,101	,675
	help from other mat'ls	,011 [a]	,118	,906	,014	,802
2	dedication to mastery	-,014 [b]	-,089	,930	-,011	,260
	help from notebooks	,045 [b]	,444	,658	,055	,660
	help from other mat'ls	-,011 [b]	-,118	,907	-,015	,797

a. Predictive variables in the model: (Constant), Role of Teacher

b. Predictive variables in the model: (Constant), Role of Teacher, Opinion towards statistics

c. Dependent variable: Final grade

In Tables 8 and 9, we can see the variables inferred that were included in the final model after the stepwise selection process, and, in addition, their order of inclusion. Two have been included (teacher role and opinion towards statistics) and three were eliminated (dedication to subject mastery, help from the SPSS notebooks, and help from other materials).

Table 10. Model summary

Model	R	R squared	corrected R squared	Standard estimation error
1	.707*	.499	.492	1.537
2	.746**	.557	.544	1.456

* Predictive variables: (Constant), teacher role
** Predictive variables: (Constant), teacher role, opinion toward statistics
*** Dependent variable: Final grade

Table 10 presents the multiple correlation coefficients, of determination and corrected determination, as well as standard estimation errors attained in the initial and definitive models. Thus, we can affirm that a multiple correlation coefficient of nearly .75 has been attained, and a determination coefficient of .55, which amounts to saying that 55.7% of the variance of the final grade variable is explained jointly by the variables teacher role and opinion toward statistics.

Table 11. Components of the regression equation and complementary statistics

Model		Non-standardized coefficients		Standardized Coefficients	t	Sig.
		B	Typ. error	Beta		
1	(Constant)	-.364	.668		-.546	.587
	Teacher Role	1.498	.182	.707	8.236	.000
2	(Constant)	.241	.665		.362	.718
	Teacher Role	.869	.274	.410	3.172	.002
	Opinion towards statistics	.595	.201	.382	2.953	.004

**Dependent variable: final grade

As is seen, after inferring a first model with the single variable Teacher Role, a second and definitive model is proposed where the variable Opinion towards statistics is also included. Therefore, the calculated regression equation would have the following statistical notation:

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + e$$

$$Y = .241 + .869_{\text{teacher performance}} + .595_{\text{opinion towards statistics}} + 1.45 * z\alpha$$

Another important aspect of the inferred model is verifying that each of the predictive variables included therein have proved to be statistically significant on their own ($p \leq 0.05$ using t tests). This aspect is underscored in Table 9 and is one more indicator of the goodness of the inferred model.

Evaluation of the model

Different criteria can be taken into consideration in evaluating the model. Along with Etxeberria (1999) and Cea D´Ancona (2002), we emphasize certain fundamentals: the determination coefficient (R^2), the graphs showing fit to the regression line, and the significance of the model as measured through analysis of the variance (ANOVA).

We have already commented that the corrected *determination coefficient* is .544, which is equivalent to claiming that a substantial percentage of the S^2 of the criterion variable (final grade), almost 55%, is explained by the predictive variables finally included in the model (teacher role and opinion towards statistics). As for the *graphs* showing fit of the model, we offer the following:

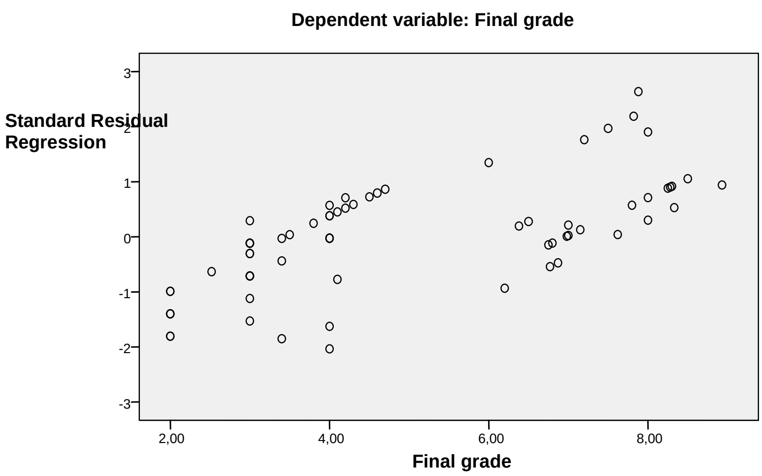

Figure 3. Dispersion graph showing model fit

We can observe a more or less moderate fit between real values and values projected by the model, since the value pairs are located to a greater or lesser degree around an imaginary line which is our regression line.

Finally, we attempt to verify whether the joint effect of the two predictive variables that have proved important in predicting the criterion variable in model 2 (definitive model) differs statistically from zero. The following ANOVA was implemented for this purpose:

Table 12. Anova associated with the regression model

Model	Sources of variation	Sum of squares	df	Quadratic mean	F	Sig.
1	Regression	160.348	1	160.348	67.834	.000*
	Residual	160.741	68	2.364		
	Total	321.088	69			
2	Regression	178.861	2	89.431	42.129	.000**
	Residual	142.227	67	2.123		
	Total	321.088	69			

* Predictive variables: (Constant), Teacher Role ** Predictive variables: (Constant), Teacher Role, Opinion towards Statistics
*** Dependent variable: Final grade

In effect, an important F ratio is associated with p = .000. This data point indicates that the inferred regression model acquires statistical significance, that is, that the explained variation is more important than the unexplained variation, and therefore becomes relevant for the predictive purposes for which it was developed.

Discussion and Conclusions

The teaching-learning of statistics constitutes a basic competency for those pursuing degrees in the Social Sciences. There is evidence that this academic subject is not one that university students find particularly motivating, nor is student performance at its best. By investigating the processes that incur in acquisition of these competencies, determining the variables that influence directly or indirectly, and by analyzing groups of factors that have the greatest input in successful learning, we hope to contribute toward optimizing the teaching of this unpopular discipline.

Over the last decades, educational research has emphasized the exploration of motivational, contextual, cognitive, and academic variables; many studies have also opened new, innovative frontiers based on problem-based learning (PBL) or a more procedural orientation focused on the acquisition of technological competence, communication in appropriate language, interpretation of results, and data analysis.

These different approaches form a global picture that makes it possible only to address disconnected arenas with isolated groups of variables; consequently, we find a need to identify integrated foundational structures and empirical analyses that lead us to comprehensive views enabling an analysis of the interaction among these groups of variables.

Returning to the results from the present study, significant correlations were obtained between the criterion variable *academic performance in the subject Methodological Foundations for Educational Research* and five groups of predictive variables oriented toward a model of deep teaching-learning. In such a model there is a sustained level of commitment, motivation and demands with regard to the subject matter, on the part of both the teacher and the student, corresponding to what the literature defines as characteristic of an active, motivational model of university work.

The five groups of variables chosen are: role of the teacher (r=0.703), pre-existing opinion towards statistics (r=0.70), dedication to subject mastery (r=0.63), help from supplementary instruction notebooks (r=0.46) and help from other teaching materials (r=0.32). The exploratory analysis of the regression equations that best predict students' academic performance, with prior verification that assumptions of normal distribution and multicollinearity are fulfilled, demonstrate that the variables that are most predictive in this model are the *role of the teacher* and *the student's pre-existing opinion about statistics;* the rest of the variables are not decisive in this model, even though another study situation may reflect a different configuration. In any case, the inferred model serves to explain nearly 55% of the variance of performance outcomes through these two predictive variables, with a significance level substantially below 0.05.

Consequently, after examining these results, and without the confirmation of possible future studies with greater complexity (causal analysis), we venture to say that there is a need to work with students on their attitudes toward statistics, and also reflect on how the teacher transmits statistical knowledge, in such aspects as: intelligible language for the students; real-life examples that they can relate to; attractive, motivational presentation of content; importance of the argumentation competency above and beyond statistical-mathematical execution; constructive learning, and so on. In short, towards those aspects which clearly belong to a deep learning style.

References

Alpízar, M. (2007). Herramientas tecnológicas en el proceso de enseñanza-aprendizaje de la estadística. [Technological tools in the process of teaching-learning statistics.] *Cuadernos de Investigación y Formación en Educación Matemática*, *2*(3), 99-118.

Arnal, J.; Del Rincón, D. & Latorre, A. (1994). *Investigación educativa. Fundamentos y metodología*. [Educational research. Foundations and methodology.] Barcelona: Labor.

Auzmendi, E. (1992). *Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitarias.* [Attitudes towards mathematics-statistics in secondary and university education.] Bilbao: Mensajero.

Balacheff, N. & Kaput, J.J. (1996). Computer-based learning environments in mathematics. In Bishop, A.J.; Clements, K.; Keitel, C; Kilpatrick, J. y Laborde, C. (Eds.) *International Hanbook of Mathematics Education* (pp. 469-501). Netherlands: Kluwer Academic Publishers.

Bayot, A., Mondéjar, J., Mondéjar, J.A., Monsalve, F. & Vargas, M. (2005). The Difficulties of Learning Concepts in the Social Sciences. In Misztal, M. & M. Trawinski (eds.) *Studies in Teacher Education: Psichopedagogy* (pp. 242-258). Wydawnictwo Naukowe Akademii Pedagogicznej: Kraków.

Behar, R. & Grima, P. (2001). Mil y una dimensiones del aprendizaje de la estadística. [1001 learning dimensions of statistics.] *Estadística Española*, *43*(148), 189-207.

Beins, B.C. (1985). Teaching the relevance of statistics through consumer-oriented research. *Teaching of Psychology, 12*, 168-169.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, *2*(1-2), 127-155.

Ben-Zvi, D. & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, *45* (1-3), 35-65.

Biggs, J. (2005). *Calidad del aprendizaje universitario*. [Quality of university learning.] Madrid: Narcea.

Bisquerra, R. (2004). *Metodología de la investigación educativa*. [Methodology of educational research.] Madrid: La Muralla.

Bradstreet, T.E. (1996). Teaching Introductory Statistics Courses so that Non-statisticians Experience Statistical Reasoning. *The American Statistician*. *50*(1), 57-71.

Burrill, G. (1990). Quantitative litaracy. Leadership training for Masters teachers. In Hawkins, A. (Ed.). *Training tearchers to teach statistics*. (pp. 219-227). The Netherlands: International Statistical Institute, UNESCO.

Carmona, J. (2004). Una revisión de las evidencias de fiabilidad y validez de los cuestionarios de actitudes y ansiedad hacia laestadística. [A review of evidences for the reliability and validity of questionnaires on attitudes and anxiety toward statistics.] *Statistics Education Research Journal, 3*(1), 5-28. *En: http://www.stat.auckland.ac.nz/serj*

Cea D´Ancona, M.A.(2002). *Análisis multivariante. Teoría y práctica en la investigación social*. [Multivariate analysis. Theory and practice in social research.] Madrid: Síntesis.

Cohen, L. & Manion, L. (2002). *Métodos de investigación educativa*. [Methods of educational research.] Madrid: La Muralla.

Cuesta, M. & Herrero, J. (2008). Analysis of Psychology Data 101: a hypermedia tutorial. *Electronic Journal of Research in Educational Psychology, 16, 6*(3), 847-862.

De la Fuente, J., Cano, F., Justicia, F., Pichardo, M.C., García, A.B., Martínez, J.M., & Sander, P. (2005). Effects of using online tools in improving regulation of the teaching-learning process. *Electronic Journal of Research in Educational Psychology, 13, 5*(3), 757-782.

De la Fuente, J. & Justicia, F. (2007). The DEDEPRO Model for Regulating Teaching and Learning: recent advances. *Electronic Journal of Research in Educational Psychology 13, 5*(3), 535-564.

Dillon, K. M. (1982). Statisticophobia. *Teaching of Psychology, 9*(2), 117-131.

De Miguel, M. (Coord.) (2006). *Metodologías de enseñanza y aprendizaje para el desarrollo de competencias. Orientaciones para el profesorado ante el Espacio Europeo de Educación Superior*. [Methodologies of teaching and learning for developing competencies. Guidance for teachers approaching the European Higher Education Area.] Madrid: Alianza Editorial.

Estrada, A., Batanero, C. & Fortuny, J. M. (2004). Un estudio comparado de las actitudes hacia la estadística en profesores en formación y en ejercicio. [A comparative study of attitudes towards statistics in teachers in training and in service.] *Enseñanza de las Ciencias*, 22(2), 263–274.

Etxeberría, J. (1999). *Regresión Múltiple*. [Multiple regression.] Madrid: La Muralla.

Evans, C.J., Kirby, J. R. & Fabrigar, L.R. (2003). Approaches to learning, need for cognition, and strategic flexibility among university students, *British Journal of Educational Psychology* , 73, 507-228.

Gil, J. (1999). Actitudes hacia la estadística. Incidencia de las variables sexo y formación previa. [Attitudes towards statistics. Effect of the variables sex and prior training.] *Revista española de pedagogía, 57*(214), 567-590.

Gil, J., Rodríguez, G., & García, E. (1995). *Estadística Básica Aplicada a las Ciencias de la Educación.* [Basic Statistics applied to Educational Sciences.] Seville: Kronos.

Gordon, S. (1995). A Theoretical Approach to Understanding Learners of Statistics. *Journal of Statistical Education*, *3*(3), 45-57.

Hair, J.F., Anderson; R.E., Thatam, R.L. & Black, W.C. (1999). *Análisis multivariante*. [Multivariate analysis.] Madrid. Prentice Hall.

Hernández, F., Martínez, P., Da Fonseca, P.S & Rubio, M. (2005). *Aprendizaje, competencias y rendimiento en Educación Superior*. [Learning, competencies and performance in Higher Education.] Madrid: La Muralla.

Justicia, F., Pichardo M.C., Cano F., García A.B., De la Fuente, J. (2008). The Revised Two-Factor Study Process Questionnaire (R-SPQ-2F): Exploratory and confirmatory factor analyses at item level. *European Journal of Psychology of Education, 23*(1), 355-372.

Katz, B. M. & Tomazic, T. Z. (1988). Changing students' attitudes toward statistics through a nonquantitative approach. *Psychological Reports, 62*, 658.

Marshall, J., Makar, K., & Kazak, S. (2002). *Young urban students' conceptions of data uses, representation and analysis.* OPME-MA XXIV, Athens, Georgia, 1292-1304.

Mondéjar, J., Vargas, M., & Mondéjar, J.A. (2007). Impacto del uso del e-learning en las actitudes hacia la estadística. [Impact of the use of e-learning in attitudes towards statistics.] *Revista Latinoamericana de Tecnología Educativa*, *6*(2), 31-47.

Mondéjar, J., Vargas, M. & Bayot, A. (2008). Measuring attitude toward statistics: the influence of study processes. *Electronic Journal of Research in Educational Psychology*, *16, 6*(3), 729-748.

Morales, P. (2008). *Estadística aplicada a las Ciencias Sociales*. [Statistics applied to the Social Sciences.] Madrid: Universidad Pontificia Comillas.

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: nature, etiology, antecedents, effects and treatments—a comprehensive review of the literature. *Teaching in Higher Education*, *8*(2), 195-209.

Pérez, A. (1983). Paradigmas contemporáneos de investigación didáctica. [Contemporary paradigms in didactic research.] In Gimeno, J. & Pérez, A.: *La enseñanza: su teoría y su práctica* (pp. 95-138). Madrid: Akal.

Perney, J., & Ravid, R. (1990, Abril). *The relationship between attitudes towards statistics, math self-concept, test anxiety and graduate students' achievement in an introductory statistics course*. Talk presented at the Annual Meeting of the American Educational Research Association, Boston, Massachusetts, USA.

Roberts, D. M. & Saxe, J. E. (1982). Validity of a statistics attitude survey: a follow-up study. *Educational and Psychological Measurement, 42*, 907-912.

Rodríguez, C. & Herrera, L. (2009). Análisis correlacional-predictivo de la influencia de la asistencia a clase en el rendimiento académico universitario. Estudio de caso de una asignatura. [Correlational-predictive analysis of the influence of class attendance on academic performance at university. A case study of one academic subject.] *Profesorado, Revista de Currículum y Formación del Profesorado*, *13*(2), 1-14.

Sahai, H., Behar, R. & Ojeda, M.M. (1998). *Un replanteamiento del problema de la educación estadística: perspectiva desde el aprendizaje*. [A new look at the problem of statistics education: perspective from the learning side.] Biometrical Education: problems, experiences and solutions. International Biometric Society Networkfor Central American and Caribean. San José. Costa Rica.

Sánchez-López, C.R. (1996). Validación y análisis ipsativo de la escala de actitudes hacia la estadística (EAE). [Validation and ipsative analysis of the scale on attitudes towards statistics, *EAE*.] *Análisis y Modificación de Conducta*, *22*(86), 799-819.

Santos, M. (1997). *Principios y métodos de la resolución de problemas en el aprendizaje de las matemáticas*. [Problem-solving principles and methods in learning mathematics.] México: Grupo Editorial Iberoamericano.

Tempelaar, D. T., Van der Loeff, S. S., & Gijselaers, W. (2007). A structural equation model analyzing the relationship of students' attitudes toward statistics, prior reasoning abilities and course performance. *Statistics Education Research Journal, 6*(2), 78-102.

Van Dalen, D.B. & Meyer, W.J. (1983). *Manual de técnicas de investigación educacional.* [Manual on educational research techniques.] Buenos Aires: Paidós.

Waugh, R. F. (2002). Measuring self-reported studying and learning for university students: linking attitudes and behaviours on the same scale. *British Journal of Educational Psychology*, 72, 573-604.

Wise, S. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, 45*, 401-405.