Batanero, C. (Ed.), Training Researchers in the Use of Statistics, 339-353.

Printed in Granada, Spain.

JUAN D. GODINO, CARMEN BATANERO, & RAMÓN GUTIÉRREZ JAIMEZ

THE STATISTICAL CONSULTANCY WORKSHOP AS A PEDAGOGICAL TOOL

In this article we present and analyse the results of three related experimental studies: (1) the use of statistics in a sample of mathematics education doctoral theses in Spain; (2) the attitudes towards data analysis and statistical consultancy by doctoral students in education; (3) the future statistics consultants perception of their competence for consultancy work. We also describe a project aimed to implement two didactical devices, which would improve the researchers' attitudes and use of statistics and the future consultants' competence. This project would serve to link together prospective consultants and clients within a Statistical Consultancy Unit at the Faculty of Education.

1. INTRODUCTION

In this work we present information about the statistical training of researchers in mathematics education and about the use of statistics in educational research in Spain. This information, as well as our experience in both teaching research methods courses, and doing co-operative consultancy in educational research projects, led us to realise how difficult it is to carry out their own data analysis for those who have not specialised in statistics, and their need for the co-operation of statisticians.

In Spain, moreover, there is neither a culture that favours statistical consultancy, nor a deserved scientific or economical recognition of this work. It is also necessary to make researchers aware of the resources and possibilities that statistics can offer them, and of the specific technical skills and knowledge required to carry out good statistical analyses. On the other hand, students training in statistics do not have the possibility to do practices in companies, and, therefore, their training is mainly theoretical.

In synthesis, in this paper we analyse the following points:

1. The use of statistics in mathematics education research and the impossibility to provide doctoral students in education with a statistical training that allows them to solve all their data analysis problems. Given the complexity and constant evolution of advanced statistics techniques, it is unrealistic to expect that potential users are able to make an autonomous and appropriate use of statistics.
2. The perception by doctoral students in the field of education of their data analysis needs, knowledge and capabilities.
3. The perception by students training as professional statisticians (in Spain) of their capacity for consultancy work, that is, for solving real data analysis problems in co-operation with researchers in other disciplines.

These points show that the statistics curricula at both graduate and postgraduate levels should be revised to develop a culture favourable to statistical consultancy among statistics users and suppliers of statistical services. We analyse the educational problems implied by statistical consultancy, in relation to both the practical training of future statistical consultants, and to the creation of positive attitudes towards consultancy on the part of future researchers.

We finally describe plans for implementing two interrelated didactical tools that might contribute to solving these problems: (1) the Statistical Consultancy Workshop (directed to trainee statistics students), which is conceived as an optional subject where students would contact real applied research problems; (2) the Statistical Consultancy Units where these students would practice, guided by a Tutor. These units are conceived as meeting points for future clients (applied researchers) and consultants (statistics students), where applied researchers' positive attitudes towards statistical consultancy and trainee consultants' practical experience would be developed. We analyse the basis and characteristics of both didactical tools, and the human and material resources needed to implement this project. We also present the results of a survey carried out to the different agents implied in this project, as regards their attitudes and preconceptions about statistical consultancy.

2. THE USE OF STATISTICS IN MATHEMATICS EDUCATION RESEARCH: AN EMPIRICAL STUDY

2.1. THE STUDY

In this section we analyse the use of statistics in a sample of doctoral dissertations in mathematics education carried out in Spain, to show that educational researchers need the statistician's support in the analysis of their data and in the interpretation of their results. We chose this area because of our knowledge of the specific research problems and our own experience in teaching research methods and in doing statistical consultancy in this area in the University of Granada and other Universities in Spain. On the other hand, we think that the problems related to the use of statistics found among these researchers, who have an advanced training in mathematics, will be higher among researchers in other areas of Education, where the training in mathematics and statistics is weaker.

In Spain, doctoral programs in mathematics education were started in 1988 at the Universities of Granada and Valencia. In addition to these specific programs, doctoral dissertations in this area of knowledge have been carried out in mixed doctoral programs (sciences and mathematics education), pedagogy, general didactics or mathematics. The production of theses has concentrated in the period 1991-2000, since mathematics education was not recognised as an official area of knowledge until 1983, and the doctors who started the first doctoral programs had carried out their doctoral theses in other branches of mathematics, like statistics or analysis. The majority of students entering the doctoral programs have graduated in mathematics, and have taken one or more courses of mathematical statistics during their training. Many of them have also taken a course of statistics applied to education.

In the doctoral program at the University of Granada, we offered four methodological courses in the period 1988 to 2000, one of which (course 1) was compulsory and the remaining were optional. The content of these courses, strongly centred in the specific problems of the area of knowledge, and which lasted 30 hours each, is summarised below:

1. *Research methods in education.* The research process and its main stages. Research paradigms and agendas in mathematics education. Different approaches to research. Bibliographical sources in mathematics education. Methods and techniques of data collection.
2. *Educational research design.* Research variables and hypotheses. Types of variables and their control. Basic types of experimental and quasi-experimental design. Their application to educational research. The process of causal inference. Validity and reliability.
3. *Analysis of educational data.* Coding and preparing data for analysis. Exploratory data analysis. Introduction to inference. Correlation and regression. Analysis of variance. Introduction to the use of SPSS.
4. *Applications of multivariate data analysis to educational research.* Geometrical representation of multivariate data. Classification methods. Factorial methods. Correspondence analysis. Implicative analysis.

In practice, most students took some of the courses 2, 3 and 4, besides the compulsory course 1, although few students took the four courses, due to the fact that there were a limit of 32 credits in the Program, including 9 credits of research work. The situation in the new Doctoral Program is getting worse, since the maximum number of credits that a student can take is 20, courses 3 and 4 have been suppressed, and both courses 1 and 2 are optional. In other doctoral programs in education, the number of methodology and data analysis courses is still more reduced.

2.2. METHODOLOGY USED TO ANALYSE THE DOCTORAL THESES

The analysis has been carried out in a total of 25 doctoral thesis in the area, from the Universities of Barcelona, Cadiz, Granada, Huelva, Basque Country, Santiago de Compostela and Zaragoza. For each of them, we analysed the methodological approach, type and size of samples used, research design, data analysis methods, their adequacy for the problem being researched and the interpretation of their results.

Most researchers took intentional samples of students or teachers, and in 13 theses one or several pilot samples were also analysed. One work of research used a multi-stage-stratified sampling of schools where a proportional number of schools were selected among all the schools in each stratum of a Spanish region. The number of children in each school was also considered in the second stage of sampling, so that the final sample was random. In another thesis a sample of 60 schools was taken, and a very complete statistical study of the sociological and economical variables that define the sample was done. In our opinion, the sample was a good representation of the population studied.

Two theses only performed a study of documents and four of them used a qualitative approach, with scarce use of statistical methods (just a descriptive study). The sizes of the samples were very variable, ranging from the study of 3 cases up to a sample of 1904 students. In general, the number of variables analysed was quite extensive, and thus, even with a sample not too big, the data set was quite large.

The software used was, generally SPSS, although some more specific software were also used, such as Bilog, Lisrel, CHIC, Iteman and Statworks. Auxiliary software were built for collecting data (for example, to record the students' interaction with the computer when solving a task), and we even found a case were a scalogram analysis was carried out with the help of a calculator. Many theses combined more than one statistical software, since the variety of methods used was not available in just one package.

The frequency of use of different statistical methods is presented in Table 1, where we can appreciate the difficulty of preparing the doctoral students in such a wide range of methods within the scarce time available for data analysis courses. Sometimes, the statistical analysis was carried out by statisticians who participate in the research teams (this is our case for some of the theses analysed).

Table 1: Frequency of Statistical Methods Used in 25 Doctoral Theses

Statistics methods used	(1)	(2)
Descriptive study	22	2
Item analysis: difficulty and discrimination indexes	17	
Validity, reliability	11	6
Item response theory, generalizability theory	7	2
Association: contingency tables, association coefficients; correlation	5	4
Hypothesis tests, confidence intervals, estimation of effects	3	3
Non parametric inference: Kruskal - Wallis, range correlation	5	4
Lineal models: anova, ancova, regression; multiple regression; repeated measures	5	3
Lineal logarithmic models	1	2
Lisrel models	1	
Analysis cluster	5	1
Descriptive analysis of variables deduced from content analysis of texts	7	
Factor analysis, principal components	5	5
Correspondence analysis	7	1
Scaling (Guttman scalogram)		1
Multivariate inference: T2 Hotelling, manova, mancova	4	2
Experimental design	6	3
Total	111	39

(1) Correct use; (2) Errors in analysis or interpretation

In other cases the analysis was performed by external statisticians, specialists in educational research methods with no specific statistical studies, or by the doctoral students, who learned to operate the statistical software and carry out the analysis of their own data.

In relation to these last two cases we observed an excess of use of statistical methods, consisting of testing a given hypothesis or analysing a variable by several logically equivalent methods that produce the same result. For example, in some theses the reliability was computed, by the two-halves method, Alpha and Theta coefficients, and by generalizability coefficients deduced from analysis of variance with repeated measures. Of course the four values obtained only varied in the last decimal figures. Another example is using correspondence analysis to study the association between rows and columns in a contingency table with only 3 columns, where the associations are observed by simple visual inspection. This suggests the researchers' lack of confidence in the method chosen or recurring to statistics in order to assure the "scientific status" in a work of research that, otherwise, would scarcely be relevant.

Another fact observed is that, statistical methods were sometimes adapted to the specific field of education and quite a thorough knowledge of this field was required on the part of the data analyst. For example, some works used correspondence analysis to relate students' strategies or errors with task variables in the items of a questionnaire.

This questionnaire was elaborated with the help of experimental design in such a way that projecting certain supplementary variables in the correspondence analysis could serve to carry out informal hypotheses tests about qualitative data (Batanero, Estepa, & Godino, 1995). Another example was using implicative analysis to show the evolution of students' conceptions as a consequence of a teaching experiment (Batanero, Godino, & Navarro-Pelayo, 1995). Experimental designs such as Latin- square, grecolatin-square, or factorial fractional, were used in several works to control task variables in the questionnaires or to divide an over-sized questionnaire into equivalent parts that were distributed to diverse groups of students. The analysis of variance with repeated measures was used to compute different components of generalizability.

2.3. SOME DIFFICULTIES OBSERVED

The difficulty that the correct use and interpretation of statistics implies for researchers is shown in the current controversy around significance tests (Harlow, Mulaik, & Steiger, 1997; Batanero, 2000), as well as in the students' difficulties in understanding even elementary statistical methods and concepts (Batanero et al.., 1994; Vallecillos, 1999). In spite of the small number of works analysed, we found a variety of incorrect uses and interpretations of statistical methods and results, some of which had been previously described in research literature, particularly in mathematical education (White, 1980; Menon, 1993). Below we describe some of these errors and their implications for the quality of research work.

Association

Computing several association coefficients between two variables, without discussing their relevance and taking the coefficient with the highest value to argue the existence of association.

Analysis cluster

Using unclear reasons to determine the number of clusters, when neither the dendogram nor the distances serve to clearly differentiate a given number of them. Arguing, based on the said clusters, that a given number of subject typologies were determined, even when the association coefficients between the dependent variables and the groups identified are near to zero.

Using a contingency table crossing each task with a list of strategies to carry out a cluster analysis of students' strategies in a series of tasks, instead of using the original data set. That is, significantly reducing the data set and using the percentage of each strategy in each problem, instead of each student's strategy in each problem

Contingency tables and logarithmic lineal models

Not taking into account the minimum advisable expected frequency in the table cells. For example, we found a Chi-square test of a large contingency table where only 10% of the cells had frequency greater than 5, and 20% of the cells were empty.

Using logarithmic lineal models in contingency tables where the hypothesis of independence among diverse dimensions in the table did not apply. The data were obtained from several samples, each of which provided data for a certain combination of categories for each dimension. Deducing from the analysis that there was an interaction among the different dimensions.

Factor analysis

Applying factor analysis to data sets with too few cases in relation to the number of variables analysed (less than 2 cases per variable in a thesis), without noticing that correlation coefficients have very wide confidence intervals in small samples.

Oblique rotation was used to get a number of factors bigger or smaller than the number of factors obtained in the initial extraction by principal components, as a way to show the validity of a questionnaire. For example obtaining only one factor by principal components and using the oblique rotation to justify that there were two differentiated factors, even when the two factors were correlated and the variance explained by the second factor was very small.

Confusion among the total variance explained by a factor and the variance explained in the reduced factorial space. In this way a researcher interpreted that a given group of factors explaining 70% of the variance before rotation, could explain 100% of the variance after rotation. It is symptomatic that these errors appear in doctoral students with a high mathematical preparation, who previously studied analytical geometry. The relevance of the context in the understanding of concepts is shown in these examples. None of these researchers would doubt that a rotation of a solid in the space preserves the solid form (number of factors) and relative dimension of each axis (contribution to the explained variance).

Variance analysis

Confusing random effects (for example the effect of different schools) with fixed effects. In particular not keeping in mind the lack of robustness of random effects models when the distribution are no normal or the variances are too heterogeneous.

Finding a significant effect of the school on the dependent variable and considering all the schools equivalent when carrying out inferences in the remaining analysis.

Carrying out a series of one-way variance analysis to study the effect of a series of factors, instead of a factorial analysis of variance or a multivariate analysis of variance.

Scaling techniques

To show that a given competence increases in levels, one author applied scalogram analysis to a questionnaire with 16 tasks that was designed to assure the validity and to control the relevant variables. Since the complete questionnaire did not fit a scale pattern, he recursively applied scalogram analysis and suppressed tasks, until he only kept 4 of the initial tasks. Even then two different factors with a very close proportion of explained variance appeared in factor analysis, yet the author thought he had proved the unidimensionality of the construct studied.

Experimental design

Few studies used experimental design to explicitly control concomitant variables and to assure a better possibility of generalising the results. Since the samples are, generally, intentional, it is important to study the instrument validity and to analyse the type of tasks to which the results could be generalised in students similar to those who participate in the study. Out of 20 researchers who prepared their own questionnaires, only 9 of them controlled the task variables with experimental design and only 17 of them analysed the instrument validity and or reliability.

The methodology of quasi - experimental design suggested by Cook and Campbell (1979) should also be considered to make inferences to other students or to evaluate the effectiveness of instruction. This is based on controlling the threats to validity and on analysing the changes' patterns. Ten out of the 25 doctoral theses implemented teaching experiments; seven of them did not use a control group; only in one thesis the study suggested by Cook and Campbell was carried out and in two of them the initial knowledge of the participant sample was compared with a bigger sample. Finally we also found a comparison between two groups, in one of which there was a year of instruction before the pre-test was applied.

Inference

One author interpreted significance level in a hypothesis test as "probability of error." Another researcher thought he proved the null hypothesis, because he had found a non-significant result. We have also found confusion between probability and percentage.

In the later theses the study of effects size in statistical tests was included to substitute or supplement the classical T or F tests. We think that this change is due to the current controversy regarding hypothesis tests and to the recommendations to improve its use (Harlow, Mulaik, & Steiger, 1997; Levin, 1998). However, after computing the effects size, some of these authors only analysed whether the effect was significant or not, and only reported the effect's p-value, not doing any analysis of its practical significance. In this way, a method criticised (only reporting the T or F value and its significance) is changed by another equivalent method, that in fact, presents the same philosophical problems.

The problem of multiple comparisons was also found in these theses, where the significance level that would be necessary to carry out a great number of tests on the same sample was usually not taken into account. Finally we found a wrong formula for the sampling error of the mean.

Sampling

*C*onsidering random an intentional sample of children from the same school, from which a random subsample was selected, after eliminating the students that showed learning problems. Considering that a sample size of only 12 cases is enough to apply methods, where the basic assumptions do not hold, but that are robust enough for big samples; confusion between conglomerate and stratified sampling; using stratified sampling, with a very variable size of sample in heterogeneous stratum (regarding the variable measured, for example, different educational levels) and not keeping in mind the subsamples or strata sizes in computing global estimates.

All these results show the difficulty that researchers who have not specialised in statistics find in carrying out their own data analysis and the consequences that this might imply for the quality of their research work. The difficulties found in the theses analysed increase in other related areas, such as language education, psychopedagogy, psychology, science education and physical activity education, where we have sometimes carried out consultancy work. Therefore, it is necessary to recognise that mastering advanced statistical concepts and methods to guarantee their appropriate and pertinent use in solving real data analysis problems is the competence of professional statisticians.

3. ATTITUDES AND NEEDS IN RELATION TO STATISTICS BY STUDENTS IN THE DOCTORAL PROGRAMS IN EDUCATION

Besides analysing the use of the statistics by educational researchers, we aimed to evaluate the extent to which future researchers were aware of their limitations and to which they value the relevance of the consultant's work in relation to the final quality of their research work. To assess these attitudes we gave a questionnaire to 50 doctoral students in mathematics education (10), sciences education (8), pedagogy / psychology (17), and other educational field (15). 20 of them had good previous training in statistics (2-3 statistics courses in their undergraduate studies), 11 had some training and 19 had no previous training in statistics. 28 students were still defining their research project and the remaining were in different stages of collecting or analysing data.

Analysis and discussion of results

A first set of questions were intended to evaluate the researchers' attitudes regarding data analysis and how they valued the consultant's work (Table 2) .

Table 2: Frequency of Answers to Questions about Researchers' Attitudes Towards Data Analysis and Consultancy

Researchers' attitudes in relation to data analysis	Yes	No	No answer
Have you carried out some statistical analyses in your previous research?	28	21	1
Do you plan to carry out some statistical analyses in your current research?	46	3	1
Do you consider that your statistical knowledge is enough to carry out your data analysis and to interpret your results yourself?	2	46	2
Would you be willing to request collaboration from a statistical consultant?	48	2	
Would you invite a statistician to co-supervise your thesis ?	36	10	4
Would you invite the person who did the data analysis to co-author some derived publications?	26	20	3
Do you consider that your training allows you to properly set your problems to a statistician?	25	19	6

Table 3: Researchers' Subjective Perception[a] as Regards the Need of Statistical Help

Research stages and methods of data analysis where help is needed	Median	Interquartile range
Research design, identifying variables, selecting samples	1	1
Identifying the statistical techniques appropriate to the problem	2	1
Coding and recording data	2	0
Producing descriptive univariate tables and graphs	2	1
Producing descriptive bivariate tables and graphs	2	1
Computing statistical summaries (central position, spread, shape)	2	1
Studying association in simple or multiple contingency tables	2	1
Correlation analysis and simple or multiple regression	2	1
Variance and covariance analysis	2	1
Fitting of distributions of probability	2	1
Estimation and hypothesis testing	2	1
Time series and longitudinal studies	2	1
Multivariate methods (cluster analysis, factor analysis, etc.)	3	1
Analysis of questionnaires and studies of reliability	2	1
Using statistical software	2	1
Interpreting the results of the statistical programs	2	1
Writing the report	2	1

[a] Responses based on a 4 point scale from 0=not needed to 3= very important

Most subjects were planning to carry out some statistical analyses in their current research and many of them had also done some data analysis before. We note the high percentage of students (40%) that do not consider necessary to include the statistician as an author of the publications derived from the consultant's contribution, which suggests a scarce recognition of the knowledge contributed by data analysis experts. We believe that, even when consultants are paid for their work, they do not lose the responsibility and property of the scientific knowledge produced by their collaboration. We also gave the doctoral students a series of questions to assess (in a scale 0-3) their subjective perception about their need of help for doing specific data analyses. In Table 3 we present the average values for each question in the whole sample.

The median score was 2 or greater in all the questions, except in "research design, identifying variables, selecting samples". This suggest that doctoral students are aware of their needs for collaboration from a statistical consultant, especially in the most advanced methods.

4. A DIDACTICAL PROPOSAL: THE STATISTICAL CONSULTANCY WORKSHOP AND UNITS

Belli (1998) analyses the consultants' work features, stressing their educational role: Consultants have to make users understand their data limitations and possibilities and the requirements of research designs to reach the intended conclusions. A process of collaboration among educational researchers and statistical consultants incorporated in the research teams would significantly reduce part of the problems and needs described.

We conceive statistical consultancy as a device to co-operatively study data analysis problems, and therefore as a didactical system. Consultants need the client's contribution, as much as clients need consultant's knowledge, and both of them require some adequate structure supporting this work (material resources, knowledge, attitudes, etc.). It is clear that the circumstance that leads to statistical consultancy is the lack of enough statistical knowledge on the part of the client. However, a minimum knowledge is required to value and be aware of the necessity of the expert's collaboration.

In the same way, the consultant should know enough of the application area to guarantee the mutual communication and understanding. Barnett (1988) indicates that the consultant should be a solver of problems posed by another person, and therefore he/she should be a translator and a communicator: he/she should understand enough of other disciplines to appreciate the problems, must express them in statistical terms and what is more important, to communicate the results in an understandable way.

We think that the practical training of professional statisticians should include specific courses where the diverse contents and necessary skills for the application of statistics are approached in a systematic way. Even when the student was finally directed to teaching, the practice of consultancy carried out would have a positive impact in his teaching, as he could use real cases of application of statistics to make more attractive and less abstract his/her lectures (Wisenbaker & Scott, 1998). Consultancy courses should have an essentially practical orientation and be based on the philosophy of workshops and seminars, with the support of a network of centres where practical work would be carried out.

A bibliographical survey in statistical education reveals that the various aspects of statistical consultancy has been analysed by different authors. A first and important issue analysed is the consultants' training. For example, Rangecroft and Wallace (1998) describe the process of consultants' preparation at Sheffield Hallam University, which is based on an intense year of practice in companies, after the student has carried out some team works in the second year of University studies. Ruberg (1998) describes the training of statistical experts within a multinational pharmaceutical company, where graduate statisticians form part of multidisciplinary research teams.

Not all these contents are currently considered in the curricula for training statisticians in Spain. The main difficulty in organising practice for future statisticians in some regions like Andalusia is in fact the shortage of suitable companies in the University geographical area. Since small companies do not invest many resources in research, students of statistics (and other university specialities, such as economy) are prevented from having periods of practice as an main component of their preparation.

In cities like Granada, the University is in fact the biggest company and the place where there are more opportunities to use statistics. The University is also the place where the professionals who will need the consultant's collaboration are trained. Therefore, we consider that the conditions that facilitate the meeting and mutual exchange of these potential data analysis clients and consultants should be created within the University. We notice that in other university specialities like, teacher training, engineering, interpretation or computer science, students can practice in companies or other institutions.

Based on these contextual factors, we describe a project for implementing two interrelated didactical tools aimed to improve the application of statistics by researchers: (1) A course offered to statistics students, which is conceived as a Workshop of Statistical Consultancy (WSE); (2) A series of Units of Statistical Consultancy (USC), which will be started up at the different University Faculties where experimental research is developed; these units would be assisted by the students entering the WSE, who will be supervised by an expert statistician in the role of Tutor.

This WSE workshop could be offered as a free configuration subject, since in the Spanish University curricula students can optionally take a percentage of credits from other Faculties or degrees. Therefore, the WSE could be offered in specific faculties like the Faculty of Education or Medicine to provide future consultants with a specialisation in the use of statistics in education, medicine, psychology, etc. Students would receive a practical training, and would receive some payment from Projects, Research Groups, or different types of scholarships (collaboration grants, teaching assistantships, etc.). Another possibility is that the work carried out in the statistical workshop could be valued as a part of the future statistician's graduation projects or Master's thesis.

This workshop would be supplemented by the creation of Statistical Consultancy Units (SCU) within some University centres for graduate degree students who would in fact be assisted by the statistics students entering the WSC. We think that future researchers and future statistical consultants should meet at the University centres and departments, where collaboration between them should be started.

One problem hindering the use of consultancy and the incorporation of statistical experts in research teams at Universities is the reduced size of such teams. At the University, where the attitudes and work habits are modelled, research teams are frequently formed by only a doctoral student and his/her thesis supervisor; or at best the research group is constituted by a few researchers. Consultancy has to be faced at a higher level, either the Department or Faculty, or even at the whole University. Consultancy work at the SCU would be supervised and assessed by professors or lecturers, who would be selected among professionals with experience in statistical consultancy. The SCU would in addition be a good resource to favour the development of positive attitudes toward statistical consultancy, and an opportunity for statistics students to practice their theoretical knowledge.

Below we describe the objectives, contents, procedures and resources needed for starting up these two didactical devices.

4.1. OBJECTIVES

With this project we intend to reach the following purposes:

1. Creating a network of centres where trainee statisticians could practice statistical consultancy.
2. Favouring the development of a positive attitude towards statistical consultancy among future applied researchers.
3. Reinforcing the future researchers' culture of working with statistical experts, through consultancy and/or by including expert statisticians into research teams.

4.2. CONTENTS

The course content would be essentially practical and would be organised in two components: a) student's practice of co-operative data analysis of real problems, inside research teams; b) Seminars guided by the workshop supervisor, where students would present their work in progress to their classmates and the difficulties found would collectively be discussed by the students and the lecturer.

It would also be interesting to organise specific sessions directed at increasing the students' knowledge of the following non-statistical knowledge, which is basic for doing consultancy and for practising statistics: Defining the data analysis problems; interviewing the client; identifying questions of interest, presenting and discussing results; communicative strategies; psychological and educational aspects of consultancy work.

4.3. PROCEDURE AND RESOURCES NEEDED

Besides the SCW course, we intend to start up a *statistical consultancy unit* (SCU) at the University Faculties where experimental research requiring data analysis is being carried out by diverse departments and groups. These units would support, in particular, the data analysis of research carried out by doctorate students.

Human resources

At least two students training as statisticians would take over the responsibility of each SCU, during their period of practice or when holding a scholarship. These students will be supported by the lecturer responsible for the SCW and the lecturers of the remaining subjects, who will answer any doubts during tutoring time. The users requiring the services of the SCU will also be supported by their respective tutors and theses' supervisors.

The lecturer responsible for the SCW would co-ordinate the different SCU within the University, and this work (teaching, advising and evaluating the consultancy work carried out) would be recognised as teaching credits. This lecturer should have specialised in statistics and have a wide experience in consultancy within the Faculty where the SCW is functioning.

Material resources

In each Faculty a physical, identified space would be prepared to host the SCU. The University computing resources would be used, either at the computer labs or in the research support units at each Faculty. The bibliographical resources at the different Departments and University libraries would also be available.

Academic frame

The consultancy work carried out by a student training as statistician will be valued with credits in his/her academic record, as the workshop can take part in the curriculum as an optional subject. If enough resources were available the course might also be offered to students majoring in other university careers, whenever they have an appropriate previous statistical formation (e.g., Economy and Psychology).

5. EVALUATION OF THE PROJECT BY TRAINEE STATISTICIANS

It would not be possible to carry out the project presented without an interest on the part of students training in statistics, who should take the responsibility for an important part of the consultancy work inside the SCW, guided by the tutor. To evaluate the interest in the project on the part of these students, we carried out a study of their attitudes towards the consultancy work. In particular we were interested in their interest in doing practice and we wanted to know if they considered themselves well enough trained to perform consultancy work in their future professional life.

With this aim, we gave a questionnaire to 43 students in their 4º year of statistics studies at the University of Granada. All of them were interested in enrolling in the "Statistical Consultancy Workshop", in case the project was carried out, assigning between 4.5 and 10 credits (x = 6.48, s = 1.16) to the course.

Table 4: Future Statisticians' Subjective Perception of their Capacity for Consultancy Work[a]

Student's preparation in diverse aspects of consultancy work	Median	Interquartile range
Knowledge of the fields where statistics is applied	1	1
Main stages in experimental research and the role of statistics in the same	1	0
Research types and approaches (exploratory/confirmatory; transversal/longitudinal; experimental / quasi-experimental	1	1
Philosophical and ethical problems in the application of statistics	1	1
Working in teams to solve data analysis problems	1	1
Dealing with clients and understanding their problems	1	1
Research design, identifying variables, selecting samples	2	1
Identifying the appropriate statistical techniques	2	1
Using the software needed to perform the statistical methods	2	0
Interpreting the software results	2	0
Writing the report	1	1
Presentation of reports to clients or in an audience	0	1
Oral or written communication of reports	1	1

[a] Responses based on a 4 point scale from 0=not needed to 3= very important

Only 12 of the students had had the opportunity to carry out practical work of data analysis for companies and only 3 of them considered that their studies qualified them well enough for their future functions of statistical consultants. In Table 4 we present the median and interquartile range of the scores given by the students to the different items in the questionnaire, valued in a scale 0 (nothing), 1 (little), 2 (enough) to 3 (much).

The students training in statistics considered themselves to have “little” preparation to confront real problems, except by using the software, interpreting the results, identifying statistical techniques and working in teams. These students receive a high technical preparation, that should be reinforced with additional practical knowledge about ethical and philosophical problems, dealing with clients and understanding their problems, producing and communicating reports, points that received a punctuation equal to 1 or even 0 and that are needed for their future work

6. DISCUSSION AND IMPLICATIONS FOR TRAINING RESEARCHERS AND FUTURE STATISTICIANS

A realistic vision of researchers' training should recognise the complex character of statistical knowledge, even when increasing the teaching time and improving the didactical resources. It is difficult for researchers who are not specialised in statistics to acquire a complete mastering of statistical concepts and methods, beyond the most basic content, or that which becomes familiar due to its frequent use. The time assigned to study statistics in undergraduate and graduate courses in human, social, scientific and technological careers is too limited to prepare these investigators to be self-sufficient in solving their data analysis problems. As Belli wrote:

> *"graduate students in applied areas had some introductory background in statistics, but lack many of the skills to actually analyse the data they collect"* (1998, p. 344).

We also observe the need to encourage a more positive attitude among researchers in human, social, scientific and technical disciplines towards statistical consultancy. Many of these potential users are unaware of the necessity to consult statisticians or to incorporate statistical experts in the research teams, and they do not value the difficulty of this work. As a consequence, the statistical training curricula should take into account this problem, and try to make the users aware of their own limitations. We consider it necessary to increase the team work culture among researchers, by integrating the applied statistician's role in the design and data analysis phases of experimental investigations and by fostering statistical consultancy.

At the same time it is necessary to revise the statistical experts' training program (in mathematical or applied disciplines) to qualify them for developing co-operative work and consultancy. In the case of many Spanish universities, we observe that future statisticians do not carry out adequate practical work to apply their theoretical knowledge.

> *"Lack of consulting experience, which is a chance to apply the tools of statistics, leaves them without an appreciation for the artistic side of statistical reasoning. They learn the formulas and theories of analyses, but have no opportunity to become involved in planning a study, gathering real experimental data, nor having to present the fruits of their analyses to professional audiences".* (Bentley, Schneider, & Bentley, 1998, p. 347-48).

The project of Statistical Consultancy Workshops and Units can serve as a meeting point for potential clients and suppliers of statistical consultancy and can follow the example of this type of consulting services in other countries (as in the examples discussed in Belli, 2001; Jolliffe, 2001, Ospina, 2001). This will allow students of experimental disciplines (humanities, social, sciences, technology) to appreciate the benefits of using statistics in a reasonable way. For statistics students, this project can serve to learn the bases of consultancy and to appreciate that

> "*The problems faced by statisticians acting as consultants are varied, not only by the origin of data and research questions posed, but also by the type of personal abilities required to solve them with success. Communication with researchers with insufficient statistical and mathematical preparation is an arduous task, and sometimes, frustrating, and it will require in good measure a dose of patience and tolerance. Nevertheless, working within a team to solve practical problems can be very exciting, and contribute a great intellectual recompense.* (Hand & Everitt, 1987, p. 9).

ACKNOWLEDGEMENT

This research has been supported by the grant BSO2000-1507 (Ministry of Science and Technology, Madrid).

REFERENCES

Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Journal of Mathematical Teaching and Learning,* 2(1& 2), 75-97.

Batanero, C., Godino, J. D. Green, D. Holmes, P., & Vallecillos, A. (1994). Errors and difficulties in understanding statistical concepts. *International Journal of Mathematics Education in Science and Technology,* 25(4), 527-547.

Batanero, C., Estepa, A., & Godino, J. D. (1995). Correspondence analysis as a tool to analyse the relational structure of students' intuitive strategies. In R. Gras, & M. Artigue (Eds), *Colloque "Méthodes d'analyses statistiques multidimmensionnelles en Didactique des Mathematiques."* (pp. 155-166). Caen: A.R.D.M.

Batanero, C., Godino, J., & Navarro-Pelayo, V. (1995). The use of implicative and correspondence analysis for assessing pupil's combinatorial reasoning. In R. Gras, & M. Artigue (Eds.), *Colloque "Méthodes d'analyses statistiques multidimmensionnelles en Didactique des Mathematiques."* (pp. 245-256). Caen: A.R.D.M.

Barnett, V. (1988). Statistical consultancy. A basis for teaching and research. In R. Davidson y J. F. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 303-307). University of Victoria.

Belli, G. M. (1998). The teaching aspects of consultancy. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Vol. 1, pp. 341-346). Singapore: International Association for Statistical Education and International Statistics Institute.

Belli, G. M. (2001). The teaching/learning process in University statistical consulting labs in the United States. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 325-338). Granada: International Association for Statistical Education and International Statistical Institute.

Cook, T. D., & Cambell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field setting.* Chicago: Rand Mc.Nelly.

Hand, D. J., & Everitt, B. S. (1987). Statistical consultancy. In D. J. Hand, & B. S. Everitt (Eds.), *The Statistical consultant in action* (pp. 1-10). Cambridge: Cambridge University Press.

Harlow, L.L., Mulaik, S. A. & Steiger, J. H. (Eds.), *What if there were no significance tests?*

Mahwah, N.J.: Erlbaum.
Jolliffe, F. (2001) Learning by experience. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 355-370). Granada: International Association for Statistical Education and International Statistical Institute.
Levin, J. R. (1998). What if there were no more bickering about statistical significance tests? *Research in the Schools,* 2, 45-53.
Menon, R. (1993). Statistical significance testing should be discontinued in mathematics education research. *Mathematics Education Research Journal,* 5(1), 4–18.
Ospina, D. (2001). Statistical research and consulting in Universities from developing countries: The Colombian case. In C. Batanero (Ed.), *Training Researchers in the Use of Statistics* (pp. 295-310). Granada: International Association for Statistical Education and International Statistical Institute.
Rangecroft, M., & Wallace, W. (1998). Group consultancy, as easy as falling off a bicycle?. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics*. (Vol. 1, pp. 359-364). Singapore: International Association for Statistical Education and International Statistics Institute.
Ruberg, S. J. (1998). Consulting: The two way street. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics*. (Vol. 1, pp. 365-370). Singapore: International Association for Statistical Education and International Statistics Institute.
Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher,* 25(2), 26–30.
Vallecillos, A. (1999). Some empirical evidencies of students difficulties in understanding statistical tests. *Bulletin of the International Statistical Institute*, Tome 52, Book 2, p. 201-204. Helsinki: International Statistical Institute.
White, A. L. (1980). Avoiding errors in educational research. In R. J. Shumway (Ed.), *Research in mathematics education* (pp. 47-65). Reston, Va: National Council of Teachers of Mathematics.
Wisenbaker, J. M., & Scott, J. S. (1998). Examining the consultancy -teaching relationship for university based statisticians. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.) *Proceedings of the Fifth International Conference on Teaching Statistics*. (Vol. 1, pp. 371-376). Singapore: International Association for Statistical Education and International Statistics Institute.

Juan D. Godino and Carmen Batanero,
Department of Mathematics Education, Faculty of Education
University of Granada,
Campus de Cartuja, 18071 Granada, Spain
E-mail: godino@ugr.es; batanero@ugr.es

Ramón Gutiérrez Jáimez
Department of Statistics, Faculty of Sciences
University of Granada,
Campus de Fuentenueva, 18071 Granada, Spain
E-mail: rgjaimez@ugr.es