15. EVOLUTION OF STUDENTS' UNDERSTANDING OF STATISTICAL ASSOCIATION IN A COMPUTER-BASED TEACHING ENVIRONMENT

Carmen Batanero, University of Granada
Antonio Estepa, University of Jaén
Juan D. Godino, University of Granada

The use of computers in the teaching of statistics is receiving increasing attention from teachers and researchers (Shaughnessy, Garfield, & Greer, 1996). The introduction of computers is encouraged in different curricula, such as in the "Standards" of the National Council of Teachers of Mathematics (1989) and the "Diseño Curricular Base" in Spain (M.E.C., 1989), not only to extend what mathematics is taught, but also to affect how that mathematics is learned. However, there is still scarce research reported concerning the introduction of computers for teaching statistical concepts.

This paper presents the results of an experimental research project that investigated the effects of a computer-based teaching environment on students' understanding of statistical association. The experimental sample consisted of nineteen, 20 year-old university students who were enrolled in their first year course on exploratory data analysis and descriptive statistics. The teaching experiment included 21, 1.5 hour sessions. For seven of the sessions, the students worked in the statistical laboratory, solving problems whose solution required them to analyze different datasets provided by the teacher or collected by themselves. The planning and instruction involved the organization of an instructional sequence to meet the learning goals and contents, the selection of appropriate datasets, and a sequence of problems with increasing difficulty, including the main task variables relevant to understanding association.

The changes in the students' conceptions were assessed using two different approaches. For one approach, two equivalent versions of a questionnaire were given to the students as a pretest and posttest instrument. For the second approach, the interactions of two students with the computer were recorded and analyzed together with their written responses and their discussions during the problem-solving process. As a result, we identified some resistant statistical misconceptions as well as different *acts of understanding* concerning statistical association, which these students demonstrated during their learning process.

All of these results are presented below, starting with a summary of our previous research on students' preconceptions and strategies concerning association (Batanero, Estepa, Godino, & Green, 1996; Estepa & Batanero, 1996), which was conducted on the experimental sample as well as on an additional sample of 213 students. This additional sample was taken to compare how typical the experimental students' responses were with regard to a more typical group.

PREVIOUS RESEARCH CONCERNING STUDENTS' CONCEPTIONS AND STRATEGIES

Nineteen, 20 year-old students who had not previously studied statistical association participated in the teaching experiment described in the next section. To identify students' preconceptions, a pretest was administered to this experimental sample as well as to another sample of 213, 19 year-old students. The whole questionnaire included 10 items similar to the item shown in Figure 1.

Item 1: In a medical centre 250 people have been observed to determine whether the habit of smoking has some relationship with a bronchial disease. The following results were obtained:

	Bronchial disease	No bronchial disease	Total
Smoke	90	60	150
Don't smoke	60	40	100
Total	150	100	250

Using the information contained in this table, would you think that, for this sample of people, bronchial disease depends on smoking? Explain your answer.

Figure 1: Item 1 from questionnaire

The following task variables were used to vary the items in the questionnaire:

V1: Type of item: 2x2, 2x3, and 3x3 contingency tables; scatterplots; and comparing a numerical variable in two samples (such as studying the association between reaction time and gender).

V2: Sign of the association: Direct association, inverse association, and independence were used in 2x2 tables and in scatterplots. The sign of the association was not applicable to the rest of the items.

V3: Strength of association (measured by the correlation coefficient, phi coefficient, or contingency coefficient, depending on the item). Moderate, low, and high associations were considered.

V4: Relationship between context and prior belief: The association suggested by the context of the problem and the empirical association presented in the table sometimes coincided (theory agreeing with data), at other times did not (theory contradicting data), and in some items the context was unfamiliar to the students.

Table 1 shows the design of the questionnaire for which we tried to select a representative sample of the possible situations, properties, and representations defining the meaning of association for the teaching level and approach chosen for our research. Table 2 shows that Item 1 is a 2x2 contingency table in which there is empirical independence in the data presented, although this could contradict student's theories about a direct association between the variables. The percentage of correct answers for each item for the sample of 213 students is included in the botton row of Table 1.

Table 1: Values of the task variables and percentages of correct answers by item

	CONTINGENCY TABLES					SCATTER PLOTS			TWO SAMPLES	
	2x2 Tables		rxc Tables							
Variable	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10
V2	Indepen-dence	Inverse	Direct	Direct	Indepen-dence	Indepen-dence	Inverse	Direct	Significant	Not significant
V3	0	-0.44	0.67	0.37	0. 1	0.11	-0.77	0.55	t=3.3	t=1.5
V4	Theory contradicted	Theory agreeing with data	Unfamiliar context	Theory agreeing with data	Unfamiliar context	Theory contradicted	Theory agreeing with data	Unfamiliar context	Theory agreeing with data	Theory contradicted
% correct	39.4	50.7	90.1	87.3	60.6	83.1	85.4	21.6	83.8	73.76

For each item, we analyzed the type of association perceived by the students (direct association, inverse association, or independence). In addition, the students' written responses to the questionnaire were studied. After successive revisions, a scheme was developed for classifying students' problem-solving strategies, which were analyzed using correspondence analysis (Greenacre, 1990), as described in Batanero, Estepa, & Godino (1995). Because factor analysis of students' answers to the complete questionnaire showed a multidimensional structure, a generalizability study (Brennan, 1983) was conducted instead of computing an internal consistency index. We obtained a generalizability index of G = .86 as a measure of the possibility of extending our conclusions to the hypothetical item population and another generalizability index of G = .94 for the subjects' population.

Initial strategies

The classification of the students' strategies, with respect to a mathematical point of view, allowed us to identify some intuitively correct strategies, which pointed to correct or partially correct conceptions concerning statistical association. Below we list these strategies that have been described in further detail in Batanero et al. (1996) and Estepa & Batanero (1996).

Correct strategies

S1: Comparing either the conditional relative frequency distribution of one variable, when conditioned by the different values of the other variable, or comparing these conditional distributions and the marginal relative frequency distribution in contingency tables.

S2: Comparing either the frequencies of cases in favor of and against each value of the response variable or the ratio of these frequencies in each value of the explanatory variable in 2xc or 2xr contingency tables, where r is the number of rows and c is the number of columns in the table.

S3: Using the increasing, decreasing, or constant trend of points in the scatterplot to justify the type of association (negative, positive, or null).

S4: Using means for comparing the distribution of one variable in two different samples.

Partially correct strategies:

S5: Comparing either the conditional absolute frequency distribution of one variable, when conditioned by different values of the other variable, or comparing these absolute conditional distributions with the marginal relative frequency distribution in contingency tables.

S6: Comparing the scatterplot with a given linear function to incorrectly argue lack of association between the variables if the relationship was not linear.

S7: Computing the difference of values in paired samples and studying the sign of these differences without reaching a conclusion.

The following common incorrect strategies frequently used by students have also been established.

Incorrect strategies:

S8: Using only one cell in contingency tables, usually the cell of maximum frequency.

S9: Using only one conditional distribution in contingency tables.

S10: Incorrect interpretation of the relationship in scatterplots from isolated points.

S11: Interpreting the existence of spread in the scatterplot as no association, because for some x values two different y values occurred.

S12: Considering that there is no dependence when, in spite of having high association, there is more than one explanatory variable.

S13: Comparing only isolated values in two samples to deduce the existence of differences.

The detailed analysis of these incorrect strategies and the type of association judgments obtained from them was used to identify the following incorrect conceptions.

Determinist conception of association: Some students did not admit exceptions to the existence of a relationship between the variables. They expected a correspondence that assigned only one value in the response variable for each value of the explanatory variable. When this was not so, they considered there was no association between the variables. For example, some of these students argued that cells out of the diagonal in a 2x2 contingency table ought to have zero frequency. In scatterplots, they expected no spread and sometimes they even looked for an algebraic equation relating the two variables.

Unidirectional conception of association: Some students perceived the association only when the sign was positive, considering the inverse association as independence.

Local conception of association: Students often used only part of the data provided in the item to form their judgment. If this partial information served to confirm a given type of association, they believed that was the type of association in the complete dataset. In contingency tables, this partial information was often reduced to only one conditional distribution or even to only one cell, frequently the cell for which the frequency was maximum.

Causal conception of association: Some students only considered the association between the variables if this could be attributed to a causal relationship between them.

DESIGN OF A DIDACTIC SEQUENCE FOR TEACHING ASSOCIATION USING COMPUTERS: TASK VARIABLES IN STATISTICAL PROBLEM SOLVING

The teaching of association was part of an introductory statistics course at the university level. Different instructional methods for improving these courses have been tested, in particular computer-assisted instruction and a problem-oriented statistics course, where the emphasis is placed on having students solve real-life problems. This may result in statistics being more interesting to the students if we take advantage of their interest in the problems that statistics can solve (Willet & Singer, 1992).

The students in this study are studying to be primary teachers. Different authors (e.g., Hawkins, 1990) have discussed the difficulty of training teachers to teach statistics. Statistics is evolving into a "data science" with close relations not only to mathematics but to both computer science and its fields of application. Unfortunately, only a small minority of mathematics teachers have experienced statistics from this "practical perspective" (Biehler, 1990).

As pointed out by Steinbring (1990), in general education it is unimaginable to teach statistics as a discipline independent of school mathematics. Teachers tend to compare statistics and its feasibility for their own teaching with the methods, solutions, and patterns of reasoning of other mathematics topics. However, statistics requires more from the teacher than mathematical knowledge. These additional requirements include organizing and implementing projects, encouraging work and cooperation between students, and understanding graphical representations, computation, and so forth, not as didactic tools, but as essential statistical means of knowing. Additionally, modern computers have revolutionized the view of what statistical knowledge is, so that students also need to learn to use software for statistical problem solving. (Biehler, 1994). New representational and data analysis possibilities, together with the wider range of problems that computers allow, introduce changes in the meaning of the statistical concepts and procedures students should learn.

For this study, these considerations were taken into account when organizing the teaching, which comprised 21, 1.5 hour sessions. Seven of these sessions were held in the computer laboratory, in which the students worked on computers to solve problems. The planning of the teaching included the selection of the topics to be taught, the preparation of the software and data files, and the selection and sequencing of problems to be solved during the computer sessions. Although the content of the course was broader in scope, we centered the research on the topic of association. This topic has received scarce attention from mathematics educators, in spite of its relevance; however, some research on teaching strategies for solving specific types of correlational problems has been conducted by psychologists (Ross & Smith, 1995).

Content of the study program

The content included the fundamentals of descriptive statistics and used an exploratory data analysis approach. That is, we adopted a "multivariate perspective," even though only univariate or bivariate techniques were taught at a formal level. Therefore, students explored data files with the support of interactive computer software, and the statistical tools were combined with the possibility for selecting data subsets. The specific statistical contents were the following:

1. Random and deterministic experiments. Statistics and its applications. Use of computers in applied statistics and in teaching.
2. Population and samples. Random sampling. Measurement scales. Type of variables: nominal, discrete, continuous. Data: obtaining data, organizing data, design of a data file. Exploratory approach to data analysis.
3. Frequency, cumulative frequency, grouping data. Graphical representation: bar chart, pie charts, histograms, stem and leaf, graphical representation for cumulative frequencies.
4. Parameters and statistics. Location: Mean, mode, median; Spread: variance, standard deviation, mean deviation, variation coefficient. Order statistics: percentiles, ranges, quartiles, box plots, skewness and kurtosis.
5. Two-dimensional statistical variables: contingency tables, type of frequencies, conditional and marginal distribution. Association and independence in contingency tables.
6. Statistical association in numerical variables. Covariance and correlation. Linear regression.

Statistical software

Biehler (1994) and Shaughnessy et al. (1996) have carefully analyzed the available software as well as the ideal requirements of software to facilitate the teaching and learning of statistics. However, they also recognize that in many schools computers and modern software are still not available. When this study was conducted, there was not much didactic software for Spanish-speaking students, so the statistical package PRODEST was prepared by the research team as part of a previous project for teaching statistics with the help of computers. Although this software had rather limited capability compared to modern statistical packages, it included all the tools needed in a course of exploratory data analysis at the university level--frequency tables and graphical representation for grouped and nongrouped data, computation of statistics, stem and leaf and box and whiskers plots, cross tabulation, linear regression, and correlation. In addition, data file facilities and possibilities of selecting parts of a dataset were available.

Data files

In choosing a dataset suitable for the classroom, we wanted a domain that was familiar to the student and sufficiently rich to ensure that questions of didactic interest would arise. The first dataset was collected by the students, using a survey administered to their companions, following the model suggested by Hancock, Kaput, and Goldsmith (1992): The class of students identified a problem of interest and with the help of the teacher decided on a plan to collect and analyze the data that would help them solve the problem.

It was not possible to show the students all the different statistical procedures using only their files, because collecting their own data was very time-consuming. Thus, other datasets were provided by the teacher. Education, biology, psychology, and physical education were used as the specific contexts for the problems presented in five different datasets.

Classification of problems

In this teaching experiment, it was necessary to analyze the sequencing of traditional lessons and laboratory activities to ensure that the learning of statistics was not mainly focussed on learning to use the statistical package. The activities in the computer laboratory sessions were classified into the following three modalities:

Discovery, using experimentation and simulation, of some mathematical properties of the distributions or their statistics, such as the convergence of the frequency polygon to the density function by increasing the number of cases and subdividing the intervals.

Proposing questions from a dataset. One main difficulty posed to the statistician in the analysis of the data provided by a user is that often the person who has collected the observations does not know what can be obtained from them. We asked the pupils to propose questions concerning each file to make the students reflect on the kind of problems that statistics can help solve.

Data analysis problems. An adequate sequencing of this problem is an effective learning tool and a source of motivation for the student. Stanic and Kilpatrick (1989) indicated that the role of problem-solving in the mathematical curriculum has been characterized within three modalities--as a context to reach other didactic goals, as a skill to be learned, and as an art of discovery and invention. We intended that all these roles would be required in solving the data analysis problems.

Task variables in the data analysis problems

Once the data files were designed and the content to be taught was specified, the next step was to choose a representative sample of the population of problems concerning statistical association to provide a meaningful learning environment for our students. In the selection of such a sample (57 problems for the 7 sessions), the following task variables (Godino, Batanero, & Estepa, 1991) were considered:

TV1: Total number of statistical units: With few cases, the student may visually obtain a first idea of the characteristics of the variables. From this she/he can deduce, a priori, the most adequate type of analysis for any question considered. This is more difficult with a high number of records, where the probability that a student might modify his/her initial strategy of analysis, as a consequence of the first results obtained, grows.

TV2: Number of statistical variables in the problem and in the file. Both affect the complexity of the analysis situation, because the number of comparisons between variables (association studies) or selections of parts of the data files (study of conditional distributions) grows with the square of the number.

*TV3: Type of statistical variable*s: Qualitative (either dichotomous or not), discrete, and continuous.

TV4: Characteristics of the frequency distribution, in particular:

- *Central position values*: whether we deal with data from only one population, or whether the distribution is a mixture of several populations (multimodality).
- *The amount of dispersion* and whether this dispersion is a function of another variable or not.
- *The shape of the distribution (symmetrical or not*). Many statistical procedures based on the
- normal distribution cannot be applied to heavily skewed distributions, unless there is a
- suitable transformation of the variable.
- *Possessing outliers*: Asymmetrical distributions make the graphic representation difficult and

are best analyzed using order statistics.

All of these circumstances affect the complexity of the problem of judging association. In traditional teaching, the calculation of the regression line and the correlation coefficient, by hand or with a calculator, is usually the only activity. Due to the time necessary for training the students in computational methods, only a few examples, chosen to present a good correlation with few data, were solved. However, in practice, this procedure is not always appropriate. Quantitative complexity is inherent in statistics, and there are many statistical concepts and methods for describing the relation of two variables, thus extending the traditional concept of function. With the time saved by computation and graphical representation because of the availability of computers, the teaching was designed to allow the student to meet all these different situations, thus enriching the meaning of association that was offered to the students, where the meaning of a mathematical concept was interpreted according to Godino & Batanero (in press).

CHANGES IN STUDENTS' STRATEGIES AND CONCEPTIONS AFTER INSTRUCTION

As suggested by Shaughnessy (1992), researchers in statistics and mathematics education are natural interveners. Because our task is to improve students' knowledge, we not only observe students' difficulties with stochastic reasoning, but we wish to assess the changes after instruction. Therefore, at the end of the teaching period, a parallel version of the questionnaire was administered as a post-test to the students in the experimental sample to assess whether their misconceptions and incorrect strategies had been overcome. Tables 2-4 show the cross-tabulation of the strategies in each type of item (contingency table, scatterplot,and comparison of two samples) in the pre-test and the post-test. These tables show the improvement in the students' strategies in both contingency tables and scatterplots. The number of correct strategies changed from 17 to 31 in contingency tables (see Table 2) and from 20 to 33 in scatterplots (see Table 3); the number of incorrect strategies changed from 33 to 18 in contingeny tables and from 26 to 13 in scatterplots.

Table 2: Evolution of strategies in contingency tables

Strategies in pre-test	Strategies in post-test			
	Incorrect	Partially correct	Correct	Total
Incorrect	4	18	11	33
Partially correct	13	21	11	45
Correct	1	7	9	17
Total	18	46	31	95

Table 3: Evolution of strategies in scatterplots

Strategies in pre-test	Strategies in post-test			
	Incorrect	Partially correct	Correct	Total
Incorrect	7	7	12	26
Partially correct	0	1	10	11
Correct	6	3	11	20
Total	13	11	33	57

However, there was no improvement after the instruction in the comparison of two samples (see Table 4). Note also that although many students changed from incorrect strategies in the pre-test to partially correct or correct strategies in the post-test, other students with correct strategies in the pre-test used incorrect strategies in the post-test.

Table 4: Evolution of strategies in the comparison of two samples

Strategies in pre-test	Strategies in post-test			
	Incorrect	Partially correct	Correct	Total
Incorrect	1	1	1	3
Partially correct	4	9	7	20
Correct	3	5	7	15
Total	8	15	15	38

Moreover, the changes were not homogeneous--neither in items nor in different students. Table 5 shows this variability. In Table 5, a value of +1 was assigned to the students who, having continued in the post-test with the same type of strategy they used in the pre-test, used a more complete procedure (a value of -1 was given if they used a less complete procedure). For example, if a student changed from strategy S8 (using only one cell in contingency tables) to strategy S9 (using only a conditional distribucion) she/he changed to a more complete strategy, even though the strategy still was wrong, because she/he used two or more cells in the table, instead of just one cell. A value of ±2 was assigned if they changed from an incorrect to a partially correct strategy or from a partially correct strategy to a correct strategy (or vice versa). Finally, ±3 points were assigned when changing from incorrect to correct strategies (or vice versa). Table 5 shows that there was a general improvement from pre-test to post-test, because the total is positive. We can also observe the range of variability in the students' marks (from -5 in Student 17 to +13 in Student 10) and in different problems (from -9 in Item 9 to +37 in Item 6).

There were no improvements in the following items: (1) Item 2, an inverse association in a 2x2 table; (2) Item 3, a direct association in a 2x2 table, in which the students obtained a high percentage of success in the pre-test, and (3) Item 8, in which the correlation was due to concordance and not to causal influence. Finally, in Items 9 and 10, which concerned the comparison of two samples, there were no changes. Although in the practical sessions the students implicitly worked on the comparison of two samples, they were not introduced to the formal study of statistical procedures to assess the differences in related and independent samples, which might explain the lack of improvement. A notable improvement was noticed in the remaining items. In particular, students clearly identified independence after instruction (Items 1 and 6) and extended the use of correct strategies in 2x2 contingency tables to rxc tables (Items 4 and 5).

Regarding the judgment of association, most students overcame the *deterministic conception* of association, accepting random dependence. The *local conception* of association was also eradicated as the students noticed the importance of taking into account the complete dataset to evaluate association. Most students used all the different conditional distributions in the contingency tables, and gave up the additive procedures, using multiplicative comparison of the different frequencies in the table instead.

Table 5: Evolution of strategies from pre-test to post-test in the different items and students

Student	Contingency tables					Scatter plots			Two samples		Total
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	
1	1	0	1	2	0	2	2	-3	2	-2	5
2	2	-2	2	2	2	2	2	-2	0	-2	6
3	3	1	3	0	2	0	2	-3	-3	-3	2
4	3	2	2	-2	0	0	0	0	-2	-2	1
5	3	-2	-2	2	2	3	0	-2	0	0	4
6	2	-2	-2	3	-2	2	2	0	-2	1	2
7	-1	3	2	2	-2	2	0	3	-2	2	9
8	0	-2	0	1	2	0	3	0	-2	0	2
9	3	2	2	2	2	0	0	-3	0	0	8
10	2	0	0	2	2	2	0	3	0	2	13
11	0	-3	-2	3	3	2	0	2	2	3	10
12	1	1	1	0	2	3	-3	0	-2	0	3
13	2	0	0	2	-1	2	-2	2	0	0	5
14	3	2	1	0	2	2	0	-3	0	1	4
15	0	-2	-2	0	0	3	0	0	-2	2	-1
16	3	-2	-2	1	2	3	3	2	0	0	10
17	0	-2	-2	-2	2	3	-3	2	0	-3	-5
18	-1	2	-2	0	0	3	0	0	2	2	6
19	2	0	0	0	0	3	3	2	0	-3	7
Total	28	-8	0	18	18	37	9	0	-9	-2	91

The *unidirectional conception* of association was corrected only by some students, while others continued considering the inverse association independence. Finally, there was no improvement at all concerning the causal conception of association. Most students did not realize that a strong association between two variables is not enough to draw conclusions about cause and effect. Therefore, they argued that there was independence in Item 8, in which the correlation was due to concordance between two classifications and not to causal influence between the variables. The theoretical study of the relationships between correlation and causality during the teaching was insufficient to change the students' conceptions. As pointed out by Vergnaud, *"It is essential for teachers to be aware that they cannot solve the problem of teaching by using mere definitions, however good they may be; students' conceptions can change only if they conflict with situations they fail to handle"* (Vergnaud, 1982, p. 33). Consequently, we believe there is a need to find new practical activities that help students more deeply reflect on and accommodate their views.

ANALYSIS OF THE LEARNING PROCESS FOR A PAIR OF STUDENTS: ACTS OF UNDERSTANDING STATISTICAL ASSOCIATION

To complete our study, a pair of students was observed throughout their work in the laboratory sessions to trace their learning process. As remarked by Biehler (1994), when working with the computer, an adequate solution to statistical problems is only found through a feedback process with the specific problem and data. The students do not just choose an algorithm, they have more freedom because they have a system of options available that they can combine and select according to their strategies and partial solutions when solving the problem.

A member of the research team observed the students' work, gathering their written responses to the different problems. This observation also included the recording of their discussions and of their interactions with the teacher and the computer (Batanero & Godino, 1994). These students were also interviewed at the beginning and at the end of the experiment. When we studied in detail the observations made on these students, some recurring difficulties related to the idea of association were identified. Some of the difficulties were finally solved, either by the students themselves when discussing and looking at the results of several computer programs, or with the teacher's help, although these difficulties reappeared from time to time. At other times the difficulty was not solved, in spite of the teachers' explanations. Occasionally, the teacher did not understand the students' confusion.

In the following, we describe the learning process of these students, commenting on nine key elements of the mathematical meaning of association (Godino & Batanero, in press). We found evidence in our data that students' understanding of these elements seemed to develop at specific moments in time throughout the understanding process. These elements are essential for the student to reach a "good understanding" [in the sense described by Sierpinska (1994)] of association .

1. The comparison of two or more samples for studying the possible relationship between two variables has to be made in terms of relative frequencies.

Although the above is true, the students compared the differences using absolute frequencies of the same variables in two samples during the first session. This mistake was commented on by the teacher at the end of that session. The same incorrect procedure appeared again in sessions 2, 3, and 5, although afterwards the students seemed to overcome this difficulty.

2. The existence of differences in the same variable among two or more samples must be deduced from the comparison of the complete distribution in the different samples.

It is not sufficient to find local differences, but rather the association should be deduced from the complete dataset. In spite of this, the students started solving the problems by comparing isolated values in the two samples. For example, they only compared the values with maximum and minimum frequencies in both samples in the first session. Although these differences pointed to a possible association, they were not sufficient to quantify its intensity. This difficulty reappeared in sessions 2 and 3 and finally disappeared.

3. From the same absolute frequency in a contingency table cell two different relative conditional frequencies may be computed, depending on which is the conditioning variable. The role of condition and conditioned in the conditional relative frequency is not interchangeable.

Falk (1986) and other authors have pointed out students' difficulties in the interpretation of conditional probabilities. Students do not discriminate between the probabilities $P(A|B)$ and $P(B|A)$. Many students in our sample showed a similar confusion; that is, they referred to the conditional relative frequencies in the pre-test and throughout the experimental sessions. This confusion was noticed, during session 5, in the students who were observed; however, these students solved it with the teachers' help. They did not show this confusion during the rest of the sessions.

4. Two variables are independent if the distribution of one of these variables does not change when conditioning by values of the other variable.

Until session 5, the students did not discover that a condition for independence is the invariance of the conditional relative frequencies distribution when varying the value of the conditioning variable.

5. The decision about what size difference should be considered to admit the existence of association is, to some extent, subjective. It is difficult to obtain either perfect association or independence. The problem of association should be set in terms of intensity instead of in terms of existence.

Although the students had not studied hypothesis testing, in session 5 they discovered that judging association implies making a decision about whether to attribute small differences to sampling fluctuation or to real association between the variables. They also realized that there are different grades of association, from perfect independence to functional relationship.

6. When studying association, both variables play a symmetrical role. However, in the regression study, the role played by the variables is not symmetrical.

The fact that correlation ignores the distinction between explanatory and response variables, while in regression this difference is essential (Moore, 1995), caused much confusion for the students. When they needed to select the explanatory variable for computing the regression line (in sessions 5, 6, and 7), they did not know which variable ought to be chosen. For example, when computing the regression line between height and weight, the students were misled by the fact that there was a mutual dependence of the two variables. A great amount of discussion followed in which the students were not capable of solving this confusion. The teacher did not notice the problem and finally, the students computed the regression lines by choosing the explanatory variable at random. At the end of the teaching period, these students had not discovered that two different regression lines can be computed.

7. A positive correlation points to a direct association between the variables.

Although in session 6 the students could interpret the size of the correlation coefficient, they did not discuss the type of association (direct or inverse). At the end of the session, they noticed that when the correlation coefficient is positive and there is a linear relationship, the variables are positively associated and above-average values of one tend to accompany above-average values of the other. However, they did not explicitly use the term "direct association."

8. A negative correlation points to an inverse association between the variables.

When (in session 6) the students first encountered a negative correlation coefficient, they were so surprised that they asked their teacher if this was possible. They also had trouble when comparing two negative correlation coefficients. The students knew that a negative number with a high absolute value is smaller than a negative number with a low absolute value. However, a negative correlation coefficient with a high absolute value points to a higher degree of dependence than a negative correlation coefficient with a lower absolute value. This fact caused much misinterpretation for the problems in which a negative correlation occurred. Therefore, the knowledge of the properties of negative number ordering acted as an obstacle to dealing with negative correlation.

Although with the teachers' assistance the students observed that a negative correlation coefficient corresponded to a negative slope of the regression line and that this meant that the *y* value decreased when increasing the *x* value, they did not explicitly use the term "inverse association." They never did differentiate between the two types of association, even at the end of the sessions.

9. The absolute value of the correlation coefficient shows the intensity of association.

Although the students related the high absolute value of the correlation coefficient with the intensity of association, they did not relate this idea to the spread of the scatterplot around the regression line.

THEORETICAL AND METHODOLOGICAL CONSEQUENCES OF THIS RESEARCH

We believe that the demand for statistics education, linked to the availability of computers, is going to increase the interest in research on these topics in the years to come. The possibilities for multiple linked representations in dynamic interactive media is a major potential for new ways of learning and doing mathematics (Kaput, 1992). However, there is a danger that students will no longer do any significant statistical thinking but will spend most of their time learning to use the software. It is also necessary to take into account the relevant influence of the teacher and the context in the students' learning. As this research has shown, it is not enough for the student to solve realistic problems of data analysis using statistical software in order to acquire a good understanding of conceptual mathematical objects.

We have reported the resistance of some misconceptions concerning statistical association for some students, based on the results from a teaching process based on the intensive use of data analysis packages. The observation of the learning process of a pair of students also revealed the complexity of the meaning of association, which should be conceived as a composed entity (Godino & Batanero, in press) and whose elements have to be studied. Each of these elements of meaning needs to be contextualized in adequate problematic situations, helping students to develop conceptual tools for solving statistical problems and to grasp a more complete meaning of association.

We have also discussed some specific task variables of the descriptive study of association problems, showing the diversity of association problems. The systematic study of the specific variables of the data analysis problems is another point to be undertaken by researchers in statistical education, to enable teachers to design didactic situations that ease the acquisition of statistical concepts and procedures and the development of students' problem solving capacity.

Acknowledgments

This report has been funded by Projects PR95-064 and PS93-196 (Dirección General de Investigación Científica y Técnica, M.E.C., Madrid).

REFERENCES

Batanero, C., Estepa, A., & Godino, J. D. (1995). Correspondence analysis as a tool to analyze the relational structure of students' intuitive strategies. In R. Gras (Ed.), *Méthodes d'analyses statistiques multidimensionnelles en Didactique des Mathématiques* [Multivariate methods in mathematics education] (pp. 245-256). Caen: A.R.D.M.

Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27*(2), 151-169.

Batanero, C., & Godino, J. D. (1994). A methodology to study the students' interaction with the computer. *Hiroshima Journal of Mathematics Education, 2*, 15-25.

Biehler, R. (1990). Changing conceptions of statistics: A problem area for teacher education. In A. Hawkins (Ed.), *Training teachers to teach statistics* (pp. 20-29). Voorburg: International Statistical Institute.

Biehler, R. (1994). Software tools and mathematics education: The case of statistics. In C. Keitel & K. Ruthnen (Eds.), *Learning from computers: Mathematics education & technology* (pp. 68-100). Berlin: Springer Verlag.

Brennan, R. L. (1983). *Elements of generalizability theory*. Iowa City, IA: ACT Publications.

Falk, R. (1986). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the second conference on teaching statistics* (pp. 292-297). Voorburg: International Statistical Institute.

Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatter plots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education, 4*, 21-41.

Godino, J. D., & Batanero, C. (in press). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. In J. Kilpatrick & A. Sierpinska (Eds.), *Mathematics education as a research domain: A search for identity* (pp. 177-193). Dordrecht: Kluwer.

Godino, J. D., Batanero, C., & Estepa, A. (1991). Task variables in statistical problem solving using computers. In J. P. Ponte, J. P. Matos, & J. M. Matos (Eds.), *Mathematical problem solving and new information technologies. Research in contexts of practice* (pp. 193-203). Berlin: Springer Verlag.

Greenacre, M. J. (1990). *Correspondence analysis in practice*. London: Academic Press.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27,* 337-364.

Hawkins, A. (Ed.). (1990). *Training teachers to teach statistics.* Voorburg: International Statistical Institute.

Kaput, J. J. (1992). Technology and mathematics education. In D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515-555). New York: McMillan.

Ministerio de Educación y Ciencia. (1989). *Diseño curricular base para la educacion secundaria* [Basic curricular design for secondary education]. Madrid: Author.

Moore, D. S. (1995). *The basic practice of statistics.* New York: Freeman.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Ross, J. A., & Smith, E. (1995). Thinking skills for gifted students: The case for correlational reasoning. *Roeper Review*, *17*(4), 239-243.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 205-237). Dordrecht: Kluwer.

Sierpinska, A. (1994). *Understanding in mathematics.* London: The Falmer Press.

Stanic, G. M. A., & Kilpatrick, J. (1989). Historical perspectives on problem solving in the mathematics curriculum. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 1-21). Reston, VA: Erlbaum & National Council of Teachers of Mathematics.

Steinbring, H. (1990). The nature of stochastical knowledge and the traditional mathematics curriculum. Some experience with in-service training and developing materials. In A. Hawkins (Ed.), *Training teachers to teach statistics* (pp. 2-19). Voorburg: International Statistical Institute.

Vergnaud, G. (1982). Cognitive and developmental psychology and research in mathematic education: Some theoretical and methodological issues. *For the Learning of Mathematics*, *3*(2), 31-41.

Willet, J. B., & Singer, J. D. (1992). Providing a statistical model: Teaching applied statistics using real-world. In F. Gordon & S. Gordon (Eds.), *Statistics for the twenty-first century* (pp. 83-98). Washington, DC: The Mathematical Association of America.