® UNIVERSITY-LEVEL DATA ANALYSIS COURSES WITH THE EMPHASIS ON UNDERSTANDING AND COMMUNICATION OF STATISTICS – A TEN YEARS ACTION RESEARCH PROJECT

NIGLAS Katrin and OSULA Kairi
Tallinn Pedagogical University
Estonia

In 1994, an action research project was initiated in Tallinn Pedagogical University (Estonia) with the aim to work out and implement a new course in statistics for students of social sciences and education. It was considered important to develop the students' ability to understand the practical, real life meaning of statistical concepts as well as the ability to communicate statistics in two different (Estonian) languages: in so called "statistical language" and "everyday language understandable also to a layman." During the year 2004 about 300 students who had completed data analysis course were given semi-structured survey questionnaires focussing especially on the communication aspects of the course. This paper will give an overview of the main results which tend to support strongly the approach that assumes the active involvement of students and emphasises the development of communication skills.

INTRODUCTION

In the beginning of 1990s, in line with the development of the Department of Social Sciences and increasing availability of computers in Tallinn Pedagogical University (Estonia), the limitations of traditional statistics education were perceived and the need to reform the statistics courses, especially for psychology majors, but also for other students, was declared. For that reason, an action research project was initiated with the aim to develop and implement a new course in statistics for students of social sciences and education, which would be based on the use of a real data, non-mathematical (conceptual) approach, with active learning and using the computer as a tool for study and actual data analysis (see Niglas, 1996a; 1996b). The need for a critical-pedagogical approach that evaluates the meaningful interpretations of studied problems and the enhancement of motivation for students to work and study independently as to better manage in information society was also stressed. In addition, it was considered important to develop the students' ability to understand the practical, real life meaning of statistical concepts as well as the ability to communicate statistics and the results of the analysis.

To meet the aims of the project there was a need for a methodological approach where the role of creativity is emphasised and the experience gained during the research process utilised. Therefore the action research model, where the inquiry is seen as a spiral of cycles including planning or design, action, evaluation, replanning, action, etc., was chosen. The principles of the action research were first introduced by Kurt Lewin in 1940's, but only in 1970's the approach was starting to gain wider acceptance (Syrjälä et al., 1994). By now the relevance of the action research approach in the context of educational research is strongly argued by several authoritative methodologists (see, for example Carr & Kemmis, 1986; Elliott, 1991). Thus, the action research project was initiated in 1994 with the following main goals in mind:

- To elaborate the aims for the course on data analysis;
- To develop the structure and the supporting materials for the course;
- To devise the teaching method which would help to achieve the aims set for the course.

It was anticipated that the project should include at least four cycles of (re)design, teaching and evaluation to reach the satisfactory results and give the answers to all more concrete research questions arising from the main aims described above. During the stages of evaluation several types of data were collected and utilised: memos of the teacher-researcher, students' assessment works, semi-structured feedback surveys and focus-group discussions with students.

Ten years ago the formulation of the basic principles for the new course was based on an extensive literature review, and also on the author's experience as a lecturer in the University. On one hand, it is surprising that most of the novel ideas advocated by the leading authors in the field of statistical education more than ten years ago (see Haack, 1979; Hawkins, et al., 1992; Råde & Speed, 1985) are still topical and discussed as new ways to advance statistical education today. For example, Taffe has as early as 1986 convincingly argued the virtues of the practical model of teaching statistics compared to the traditional mathematical model, but Boland (2002) and many other authors on ICOTS6 recognise that statistical education is "often of a mechanical and tedious nature with little or no emphasis on data analysis and practical examples." The use of projects, real-world data, short stories, problem-based approach and active learning are still the keywords often introduced as the ways to improve statistical education (see, for example, Binnie, 2002; D'Andrea & Waters, 2002; Jolliffe, 2002; Rossmann & Chance, 2002).

On the other hand, the fact that the ideas concerning statistical education which were posited as central within this project 10 years ago, are still on the agenda today, encourages sharing the positive results of our action research project to improve the quality and the relevance of statistical education at the university level. Due to the limitations set to the length of this paper and the specific focus of the IASE satellite conference, hereinafter we will concentrate only on one aspect of our action research project – the teaching method which emphasises the understanding of and communication skills within statistics.

SETTING

As a result of the first two cycles of the action research project the structure of the introductory course of statistics was proposed and the preliminary results of the project reported (see Niglas, 1996a; 1996b). The content of the course is rather classical: we begin from an overview of data types and the ways the data-table should be composed, then continue with an overview of descriptive statistics, and the second half of the course introduces basics of inferential statistics. During the following cycles of the action research project the structure of the basic course has not changed considerably (for some groups it has been divided between two semesters as according to the regular feedback survey many students asked for more time to work on their projects needed for an assessment). However, during the project the scope of the courses was extended so that in addition to the basic course there is also an advanced course in data analysis available for students covering some most popular multivariate techniques.

While during the first stages of the project our new data analysis course was taught only to psychology majors, the interest has grown rapidly whereby today we teach data analysis courses to the students of various fields: social sciences (sociology, social work, psychology, information sciences, public administration), educational sciences and teacher training, informatics and mathematics. Although the course can take different labels depending on the curriculum, statistics is always taught by us as a tool for research (in most cases it is called "Data Analysis" to emphasise the practical nature and the wider scope than traditional statistics courses have). This means that, during an introductory lecture, basic aspects and steps of empirical research are summarized and the main emphasis throughout the course is on the selection of suitable methods, the interpretation of the results and the ways of presenting the results. Computations are left mainly to the computer (we use SPSS and for distance education MS Excel). Instruction is usually organised so that the proportion of lectures and workshops is 1:2, but for some groups, where we have to cope with a smaller number of instruction hours, it is 1:1.

According to the results of the regular feedback surveys, the overall attitudes towards the course over the years of the project have mainly been positive (average evaluation scores by study groups between 4,0 to 4,8 on the 5-point scale). At the same time the assessment tests have indicated that after completion of the course some students have not fully grasped the concepts or cannot communicate their knowledge. In the feedback surveys many students have reported that this is one of the most difficult subjects within their curriculum, and that they experience difficulty in comprehension of the concepts studied. Feedback from the colleagues from various fields also states that although they have taken one or even several courses in statistics, while reading research reports, they encounter difficulty in understanding what was said as "every second word" is a statistical term which they have heard about, but can not really interpret in the practical context. These were the reasons why special attention was given, within the project, to the development of a teaching style that would facilitate a meaningful interpretation of studied

problems and develop the students' ability to understand the practical, real-world meaning of statistical concepts as well as the ability to communicate statistics and the results of the analysis.

One of the main distinguishing features of the new approach to teaching statistics, compared to the formal approach which is still prevalent in the statistics courses in Estonia, and according to the aforementioned literature, also in many other countries, is the demand for communicating every statistical concept introduced in the course in two different (Estonian) languages: in so called "statistical language" and "everyday language understandable also to a layman". This demand concerns both the teacher and the students and permeates all the stages of study (e.g. introduction and discussion of the concepts in the lectures, posing questions about data and interpreting the results of workshop exercises, writing a report and assessment test). To practice the communication skills the students are expected to be actively involved both in the lectures and workshops. In the lectures every now and then guided discussion about the relevant material among the participants is encouraged, which gives them the opportunity to "discover" the train of thought and the logic of the statistical method presented, but also to assess whether they have understood the concept and can put it into practice. The answers are usually expected to be given by volunteers as to make it as close as possible to the "natural" conversation, but sometimes the teacher has to interrupt to stop some overly active student to prevail the communication, or sometimes, if in spite of the encouragement students are too shy to volunteer, the teacher has to turn personally to somebody. In the workshops the teacher supplies students with database from a real research project and the general guidelines, but thereafter takes the role of a "critical and informed friend" leaving the role of the "inquirer" - who asks the questions about data, chooses appropriate methods and interprets the results - to the students.

On the other hand, to facilitate the effective and appropriate use of statistical methods in practice, which is obvious precondition for fluent communication of statistical results, it is considered important in our courses to give the students examples of good practice and to tell them short "stories" (real or imaginary) about what has happened or could happen in real research projects, what can go wrong and how to learn from errors other people have made. The "stories" are illustrating practically all aspects of data handling process: issues of data quality (preparation of data collection instrument, process of data collection, data entry), asking questions about data, data analysis, interpretation and presentation of results, etc.

Thus, in order to help students to understand the essence and the practical meaning of statistical methods and parameters, but also to help them to realize that behind complicated statistical terms and formulae are simple and logical ideas which we actually use in our everyday reasoning, we use examples from (real or imaginary) situations and try to explain the meaning of statistical concepts using the words which we use in our everyday conversation. Students are asked to do the same – to interpret the results of statistical analysis using words "which people who have never taken a course in statistics and do not know special terminology would also understand".

To give a simple example: some students who come to the course have previously learned that *median is a middle value of an ordered sample*. As there are also students who have never heard about median it is possible to set up an experiment and ask whether they, given this definition and the result from the computer analysis stating that for salary the median equals 7000, have understood what exactly is the practical meaning of the statistical figure 7000? Most of them say "no" (in Estonian the definition sounds even more technical as the term for *ordered sample* is *variatsioonirida* which by word for word translation is *variation series*). Teacher can now ask students, who were familiar with the definition, to explain the meaning of the concept in a way that everybody would understand what the figure 7000 says about data. Usually with the help of the teacher students reach to the explanation which says that *half of studied people got salary less than 7000 crowns and half of them got more than 7000 crowns* (now everybody understands what was said). After that a new example will be given, where students are asked to explain what it means, if median for a number of children is 1? Well, somebody will rather quickly answer that *half of people had less than one child and half of people had more than one child*. Now the question of correctness of this interpretation can be discussed and the dependence of the interpretation on the nature or type of data illustrated and stressed.

Mathematics, in its traditional form, plays a relatively small role in our teaching of data analysis. However, the students are expected to understand the concepts in order to be able to realise whether the statistical result calculated with the aid of computer is realistic or not in a

given situation. If the result does not "make sense", then the students should be able to reflect whether the problem arises because they have chosen a wrong statistical method, or they have not used the computer program correctly (e.g. they have chosen wrong options), or was it so that the question asked wasn't reasonable for given data in the first place. Therefore the students are expected to be able to read basic formulae and to remember the algorithms for the most important calculations. This is the component which we vary the most depending on the background of students – mathematics, informatics and psychology majors are stronger in mathematics than others and therefore a more thorough explanation of the influence of different components of the formula to the final result is provided.

SURVEY ON COMMUNICATION ASPECTS OF THE COURSE

Reflective memos taken by the teacher in the course of the action research project indicate that it has not always been easy to get students to actively participate in discussions and that students often seem to experience difficulties when they are asked to answer the questions or interpret the results of data analysis in the lesson. However, on the other hand the regular feedback surveys have been indicating students' support to the chosen teaching style. To get a better overview of the students' attitudes towards didactical approach developed in the course of the action research project, semi-structured survey questionnaires with open-ended questions focussing especially on the communication aspects of the course were given to more than 300 students who had completed data analysis course in spring and fall terms in 2004. The participation in the survey was voluntary for the students, but as the participants were allowed to skip two questions in their course assignment test, almost all students (96%) who completed an assignment and were given the feedback questionnaire chose to fill it in. The anonymity of responses was warranted by the fact that questionnaires were collected separately from assignment tests. The final sample size is 297 students.

In the following, we will give an overview of the main results of the survey which surprisingly strongly support the approach to teaching that assumes an active involvement of students and emphasises the development of communication skills. In the questionnaire focussed but open-ended questions were used with the request to also give reasons for the attitudes reported. This, we hope, gave us more reliable descriptions of students' intrinsic feelings about the course, but as expected, generated quite a lot of missing values. For this overview, all the answers were categorised using open-coding. Although only a few students gave for any particular question answers consisting of more than one aspect or idea, several categories were used for one answer where applicable. For the sake of compactness, the questions given as titles and labels in the tables below are shorter than in the original questionnaire, but the most important questions are fully translated where necessary.

Although the actual attendance in the lectures wasn't very high (70% of students attended more than 75% of lectures and 86% of students more than 75% of workshops), 95% of respondents considered it necessary to participate both in the lectures and workshops. This result is remarkable in the context that the lectures are not compulsory in our university and that there is a growing mentality both by the student body and the leaders of the university that students should be able to learn theoretical parts of the courses exclusively with the aid of printed materials. The reasons for non attendance were mainly external: some students have jobs and could therefore not attend at all times (9%), some were sick for a longer period, some had trouble waking up early, and there was also some overlap with other lectures reported. Only 7% of students said that it was sufficient to work with the printed materials and textbooks, although there are several readily available textbooks of basic statistics in Estonian. From Table 1 it can be seen that the main reason why students considered it important for them to participate is that they value the thorough and understandable explanations and the practical examples which were used throughout the course.

Table 1 *Reasons Why Students Consider Necessary to Attend Lectures and Workshops on Data Analysis*

Why did you consider necessary to attend?	*Count*	*Percent of respondents*
Good examples and understandable explanations, which help you to learn more	124	41.8%
Useful for an assessment	45	15.2%
Possibility to master SPSS	39	13.1%
Difficult subject	34	11.4%
New knowledge, useful, interesting	18	6.1%
Possibility to ask questions	11	3.7%
Useful for a thesis	9	3.0%
No reasons given	63	21.2%
Number of respondents	*297*	

The developed teaching style assumes an active involvement of students which can be a problem for several reasons: first, students in Estonia are still used to the traditional lectures where only a teacher is expected to talk; secondly, occasionally the number of students participating in the lectures are rather big (about 60-70 students) and thirdly, a situation where only very few students are interacting with a teacher and other students are still passive listeners may develop. According to the results of our survey we have managed to get more than a half of the students actively involved in discussions – about 53% of students reported that they had at least sometimes answered the questions the teacher asked during lessons, 6% mentioned that they did it, but only quietly in their thoughts and about 40% had not actively been communicating with the teacher. The main reason given for silence was simply that a student did not know the answer or was not sure about it (this is related to the attendance frequency), but also the hesitation to speak loudly in front of the fellow students. Also the language problems were mentioned (some students are Russian).

However, while asked *"Do you prefer the teaching style where the students are expected to think along and be actively involved in discussion, or the teaching style where the students are given the role of the listener and the follower of the instructions?"* most of the students (89%) considered it important and useful that the students are given an active role instead of a passive listener and follower. The students were also asked to give reasons for their opinion whatever the opinion was for a particular respondent (see Table 2).

Table 2 *Reasons Given To Support the Preference of Students' Active Involvement*

Reasons to support an active involvement of students	*Count*	*Percent of respondents*
Makes you to think along	100	33.9%
Easier and better to learn	69	23.4%
Makes it interesting	29	9.8%
Possibility to ask questions and get deeper explanations	23	7.8%
Gives activity and possibility to have a say	16	5.4%
Comfortable, cosy	10	3.4%
Lecturer gets feedback	5	1.7%
No reasons given	74	25.1%
Number of respondents	*297*	

Table 2 indicates that according to the students' remarks, the demand for activity makes them "think along" which helps them to acquire knowledge and skills, makes the teacher to explain concepts more deeply, makes the subject more interesting and the atmosphere more

comfortable. The few students (7%) who clearly preferred the teaching style where the students are not expected to take an active role, based their opinion on the argument that they personally just do not like to express ideas in the lectures, or that there are many students who do not like to participate in public discussions and will therefore feel uncomfortable if they are asked to do so.

The second central question asked *"Whether the approach intentionally used by the teachers where the (new) concepts were discussed in two "languages" - "statistical language" and "everyday language understandable also to a layman" – helped to understand the subject or was it sooner a cause for a confusion?"* According to the results the demand to communicate in two different (Estonian) languages was considered helpful by 87% of respondents, about 7% said that it was helpful but also confusing and only 3% were reporting that it was disruptive. There was also a small number of students (3%) who had not noticed that this kind of "language games" had occurred or did not answer the question. According to the given comments (see Table 3), students liked to communicate in two "languages" because it helped them to understand the concepts and to see the relevance and the use of statistics in real life. This fact has, obviously, a clear positive influence on students' attitudes towards the usefulness of the acquired knowledge: 80% of them considered it useful for their further studies or work, additional 13% considered it useful at least to a certain extent and only 3% said that it is of no use to them. The remaining 4% of students gave no answer to the question about the usefulness of the acquired knowledge.

Table 3 *Reasons Given To Support the Preference of Discussing Statistical Concepts in Two "Languages"*

Reasons to support the use of two "languages" in the discussion	*Count*	*Percent of respondents*
Easier to understand, can master better	140	47.1%
Wouldn't have understood only statistical language, but only everyday language is not enough	24	8.1%
Can see the relevance for a real life	18	6.1%
Because the subject is difficult	4	1.2%
No reason given	111	37.5%
Number of respondents	*297*	

At the end of the questionnaire, the students were asked to add any comments about positive and negative aspects of the course and give suggestions which would help to design a course that is of more significance to their needs and expectations. Sixty percent of students had some critical comments, most of which requested more time for lectures, workshops or for their project work. There was also some criticism about the workshops organised by a junior teacher and the request to link lectures and workshops better. Students wished there could be smaller groups in the lectures and complained that sometimes they could not listen and work as their fellow students were chatting. On the other hand there were a lot of positive comments; the most relevant in this context is the students' opinion that regardless of the complexity of the material teachers could present it in a sensible and understandable manner, at the same time giving good examples and demonstrating the significance of statistics for their activities.

Thus, although there were some critical comments about the organisation of the course, the results suggest that the chosen teaching style is favoured by the students and helps them to understand statistics better. This in turn lessens the anxiety towards statistics and makes students to see the relevance of statistics in their studies, future work and everyday life.

ACKNOWLEDGEMENTS
We are grateful to the Estonian Science Foundation for their grant no 6148.

REFERENCES

Binnie, N. (2002). Using projects to encourage statistical thinking. ICOTS6, South Africa. See *http://www.stat.auckland.ac.nz/~iase/publications.php?show=1*

Boland, P.J. (2002). Promoting statistics thinking amongst secondary school students in the national context. ICOTS6, South Africa. See *http://www.stat.auckland.ac.nz/~iase/publications.php?show=1*

Carr, W., & Kemmis, S. (1986). Becoming critical: Education, knowledge and action research. London: Falmer.

D'Andrea, L., & Waters, C. (2002). Teaching statistics using short stories: reducing anxiety and changing attitudes. ICOTS6, South Africa. See *http://www.stat.auckland.ac.nz/~iase/publications.php?show=1*

Elliott, J. (1991). Action research for educational change. Philadelphia: Open University Press.

Haack, D.G. (1979). Teaching statistical literacy. *Teaching Statistics, 1*(3), 74-76.

Hawkins, A., Jolliffe, F., & Glickman, L. (1992). *Teaching statistical concepts*. Longman Publishing, New York.

Jolliffe, F. (2002). Statistical investigations – drawing it all together. ICOTS6, South Africa. See *http://www.stat.auckland.ac.nz/~iase/publications.php?show=1*

Niglas, K. (1996a). *Basic Course in Statistics for Non-Math Students. An Action Research Project.* Unpublished master thesis in Estonian.

Niglas, K. (1996b). Basic course in statistics for students of social sciences and education with spss for windows. *Tartu Conference on Computational Statistics and Statistical Education* (p.38). *Tartu, Estonia.* (Abstract).

Råde, L., & Speed, T. (Eds.) (1985). *Teaching statistics in the computer age. Proceedings of the sixth ISI Round Table Conference on Teaching Statistics in Canberra, Australia, August 20-2.* Lund: Studentlitteratur.

Rossman, A.J., & Chance, B.L. (2002). A data-oriented, active learning, post-calculus introduction to statistical concepts, methods, and theory. ICOTS6, South Africa. See *http://www.stat.auckland.ac.nz/~iase/publications.php?show=1*

Syrjälä, L., Numminen, M., Syrjäläinen, E., & Saari, S. (1994). Laadullisen tutkimuksen työtapoja. Helsinki.

Taffe, J. (1986). Teaching statistics – mathematical or practical model. In R. Davidson and J. Swift (Eds.), *The Second International Conference on Teaching Statistics, 11-16 August* (pp. 332-336), *University of Victoria.*