Measuring and Monitoring Movements in Course Interest at the University Computing Centre: A Statistical Platform for Quality Improvement and New Course Deployment Decisions

Vesna Lužar-Stiffler[1,2], Miroslav Milinović[1] and Charles Stiffler[2]

[1]University Computing Centre, Zagreb, Croatia
[2]CAIR Research Centre, Zagreb, Croatia

A sample of 7000+ students was analyzed by the University Computing Centre (SRCE) staff to evaluate course satisfaction, teacher performance, and changing course demands. Findings indicated adequate overall performance, and changing course content needs over a four year time period. The major conclusion of the Project (SOCRATES) is that the software product developed during the research can become an economical and useful tool in helping to shape administrator efforts and decisions regarding the allocation of scarce organizational resources and in the development of new course offers at SRCE.

Keywords: statistical analysis, ANOM, course/teacher evaluation, SAS application, CRM

1. Introduction

Crafty men contemn studies;
Simple men admire them; and
Wise men use them.
(Sir Francis Bacon, 'Of Studies', 1625)

This academic research project (Project SOCRATES) sprang out of a desire (by one of the authors) to build an easily managed software-based system for measuring, monitoring, and improving the quality of course offerings, and for enhancing and reallocating scarce teaching resources (via graphics-oriented, SAS®[1] Institute decision maker software) at the University Computing Centre, Zagreb (SRCE). The "non-credit" courses being offered (to students, faculty and staff) at SRCE, have been developed over the years by both SRCE and CARNet (Croatian Academic and Research Network).

A. **Background and Motivation:** limited teaching resources coupled with increasing demand for computer, software, and internet literacy training at SRCE demanded the creation of a more economical approach to the forecasting and deployment of both current and new course offerings. Achievement of this goal on a sustainable basis required an improved internal management reporting system. Prior approaches to the evaluation and analysis of SRCE courses have been described in the papers by Bekić, et. al. (1998) and Milinović, et. al. (2000).

B. **Aim:** long term measurement, and monitoring of the course activity with properly "modeled", sensibly designed, easily maintained, and user friendly application. The application would allow managers and administrators to establish an unbiased, quantitative (i.e., statistical) foundation for better decision making on the course content, course development, and teacher training and assignment.

---

[1] SAS is a registered trademark of SAS Institute Inc., NC, USA.

C. **Preview:** this paper includes a brief presentation and overview of the research process used to complete the "proof-of-concept" phase of Project SOCRATES as follows: Problem definition, questionnaire design, data collection (continuously for four years), data analyses (univariate, bi-variate, and multivariate), interim results evaluation, preliminary analytical reporting results, and decision-oriented, software application development.

D. **Intended contribution:** the goal of this research is to help shape management decision making behavior in an economical and relatively unbiased fashion... so as to better allocate scarce public resources in accordance with student-community needs, and to identify and either remedy, or else learn from "statistically unusual" course/presenter satisfaction ratings.

Methodology

Problem definition: The research problem was defined as "an attempt to test whether courses are evaluated as essentially equal, professors rated as essentially equal, and demand for various courses was essentially static" over the study period (4 years: 1996–2000). During this period, there was a total of 936 courses given, evaluated by 9,249 students, representing a total of 54,617 student-hours. Ultimately, the objective was to identify potential causes for any differences in course/teacher evaluation, and when changes were found, to establish "probable cause" for movements/ trends in demand for various courses as an aid to planning.

The research objectives were to test the following three major hypotheses:

H_{01}: All courses are being rated as essentially equal in terms of student's overall satisfaction;

H_{02}: All professors are being rated as essentially equal in the minds of students;

H_{03}: The demand for courses and the course ratings are essentially static over time.

Again, the business objective of the research (as was established in conjunction with SRCE administrators and staff) was to use this information to help shape management resource and scheduling decisions related to the course content, faculty assignment/training, and new course development.

The instrument/questionnaire was designed after accessing information available on the Internet (primarily from the University of Colorado, Boulder and the University of Maryland, European Division) regarding methods for evaluating both course content and teacher/presenter attributes. Attributes, mostly satisfaction ratings, (please contact SRCE for a copy of the instrument) were measured on both five (5) anchor point scales and binary type ("yes"/"no") scales. Various basic demographic information on the students and on SRCE facility was also gathered using categories selected based on Central Bureau of Statistics guidelines. This allowed for a comparison between SOCRATES demographic results with Croatian students, (excluding staff and faculty) in general... an indirect method sometimes used for inferring the likelihood of having achieved "representativeness" in the sample. In fact, SOCRATES respondents were found to be not significantly different in age (by number of years of education) from the general population of students at large. That is, SRCE's target users seem to reflect a representative cross section of students by AGE. However, additional research is required to determine potential differences that may exist across "areas of concentration" (e.g., Arts, Mathematics, Electrotechnical/Computing, Economics, etc.).

For purposes of this research, the **universe** was defined as all students, staff and faculties that were authorized to take SRCE courses. The sample frame would then logically consist of a list of names of everyone that took a course at SRCE during the time period under consideration.

However, because all students did not complete all courses, because of the difficulty in obtaining timely lists, and because of administrative costs, the actual respondents **sampled** for this research represent a "convenience sample". That is, all students that attended complete course were asked, by the teacher/presenter at the end of the final class session, to fill in the course evaluation form and return it to the instructor before leaving. Since the students were not tak-

ing courses for grades, and since the teachers expressed interest in improving the course content and teaching skills, it was assumed that the lack of total student anonymity did not seriously influence the students' willingness to provide unbiased feedback on the course and on the teacher/presenter.

However, because of the nature of the sampling design, the reader is cautioned regarding the wholesale adoption of the results presented below in that they may be subject to "non-observation", non-sampling errors of unknown magnitude. For example, students who either refused to cooperate in the research, or who abandoned the course prior to completion...may have given the course/professor much lower or a much higher overall satisfaction ratings. Based on this "non-probability design", the chance of this having happened is unknown... and unknowable. It means that the results are limited to those who did respond to the survey ($n = 7000+$) and cannot be generalized to the students who did not[2].

Future changes in research design could include provisions for sealed envelopes, absence of teachers from the room during survey completion, preprinted questionnaire completion instructions, and a centralized envelope drop/collection point so as to better guarantee total anonymity of the students.

Statistical Analysis: Besides some essential descriptive statistics on key productivity and satisfaction measures, bivariate associations and correlations among the variables were also produced and tested for significance. The **first two hypotheses** above (H_{01}, H_{02}) were tested using the analysis of means (ANOM, Ott (1967), Schilling (1973)) and analysis of variance (ANOVA). As compared to ANOVA, ANOM is much less known and less widely used, although more appealing to statistics/quality practitioners due to its inherent graphical nature and similarity to both ANOVA and control charting methodologies. In this study, it was decided to apply it for three main reasons:

1. It is more appropriate than ANOVA because of the large number of groups (i.e., courses, groups of courses, or teachers) being compared;

2. It can be used to identify the groups (i.e., courses, groups of courses, or teachers) that are rated significantly higher than average, and those that are rated significantly lower than average (with ANOM one compares group means against the overall mean.);

3. The graphical output is relatively self-explanatory.

Canonical discriminant analysis was applied in order to identify the key drivers of discrimination among the course (or instructor) ratings and to position (graphically) individual courses (or instructors) on the first few (statistically significant and most important) canonical dimensions. Alternatively, logistic regression was used to identify what drove key "satisfaction measures" (as defined below).

The third hypothesis was tested using the tests for proportions and Cochran-Mantel-Haenszel (CMH) correlation statistics (for testing for the change between the last two years) and the Cochran-Armitage test for trend (over the four-year period). The analysis of structural changes was also performed, but in this paper the results are only shown in the graphical form (using comparative Pareto charts and the growth-share matrix bubble plots).

These simple plots allow administrators/ budgeters to quickly "see" which courses are consuming significant enterprise resources and whether the rates of "resource spend" are going into areas of relative growth or decay. Also, key productivity and effectiveness measures can be generated.

Evaluation and modification: Project SOCRATES has thus far exceeded management/administrator expectations in terms of hoped-for analytic performance, simplicity in use, time of prototype completion, and software development budgeting milestones. As currently developed, the SOCRATES application for course/ trainer evaluation can readily be "scaled up" to handle much higher course/teacher/attribute loading requirements. Data mining efforts are currently being instigated to study a number of interesting trends and potentially useful (multivariate) insights that have been gleaned from this "gold

---

[2] Total number of questionnaires available was 9249, of which approximately 2000 were excluded due to incompatibilities resulting from a change in the original questionnaire design. Additionally, Croatian Academic and Research Network (CARNet) courses offered at SRCE are also included as part of the above analyses (and can be accessed and analyzed separately, if so desired).

mine" of easily obtained... academic... "primary research data".

Application: As a key result of this project, a customized Analytic Reporting Application was developed. The general aim of the application is to provide support for the process of continuous improvement of course offerings at SRCE. The specific purpose of the application is to provide statistical evidence (in the form of reports and graphs) as inputs to management decisions on such things as: which courses should be dropped, which should be adjusted, what new courses to add, which instructors need more training, how to optimally allocate training resources, etc. The key features of the application are presented in the Application section below.

Results

"TQM comprises more than statistics; but, without statistics, it is often 'smoke and mirrors'." (Harry V. Roberts, University of Chicago)

For the purpose of the analyses presented in this study (and because of the assumptions underlying the statistical methods applied), data have been restricted to only those courses (or instructors) that have been rated by at least 20 students within the last two school years ($n = 7442$). Because of space limitations, only a few of the most important results can be presented here. For reasons of privacy, more emphasis will be given to the discussion of course differences than will be given to teacher differences.

The following six key response variables (satisfaction measures) have been identified:

- Overall course rating,

- Overall teacher rating,

- Course ratings of 2 or less (i.e., very low),

- Teacher ratings of 2 or less (i.e., very low),

- Students' intention to take another course with the same teacher,

- Students' intention to take another course at SRCE.

Altogether, there were 20 questions ("satisfaction variables"), which form three groups: variables that measure satisfaction with the course (content, quantity, etc.), variables that measure

satisfaction with the instructor, and variables that measure future intentions (e.g., to take another course) or previous experience with the courses ("first timers" vs. "repeats"). Values of the Spearman correlation coefficients among pairs of variables (not shown) ranged from 0.01 to 0.84, most of them being greater in absolute value than 0.50 and highly significant. The highest correlation coefficient is between overall course rating and "the fulfillment of expectations" ($r = 0.84$), while overall teacher rating is most highly correlated with "the quality of presentation" ($r = 0.83$). The best (i.e., most highly significant) predictor of both "intention to take another course at SRCE" and "intention to take another course with the same teacher" is "fulfillment of expectations".

The first hypothesis (H_{01}) regarding equality of courses (as measured by overall course rating) has been rejected using F test ($F = 14.17$, $ndf_1 = 21$, $ndf_2 = 7420$, $P < 0.0001$). A sub-issue, concerning the question which courses are actually different, is handled by ANOM (instead of e.g., multiple comparison techniques) and graphically presented in Figure 1. The ANOM plot indicates that there are 4 underperforming (i.e., courses with means significantly lower than the overall mean) and 6 overperforming courses (i.e., courses with means significantly greater than the overall mean) out of a total of 22 courses evaluated below.

The canonical discriminant analysis of course/teacher ratings found discrimination among courses resulting in 8 statistically significant canonical correlations, with the first two accounting for 58% of total variation. The first canonical variable can be defined as "course content" and the second as "course quantity" (i.e., difficulty and amount of the material presented). Therefore, a plot of "course means" positioned on the first 2 canonical dimensions (please see Figure 2), for example, revealed which courses are perceived by students as too difficult (up on the CAN2 axis), too easy (down on the CAN2 axis), or just right (around zero on the CAN2 axis).

Another important discriminator among the courses is whether the student is a "first-timer" or a "repeat" (i.e., has already taken at least one course at SRCE). The "repeats" are rating courses higher than the "first-timers" on course
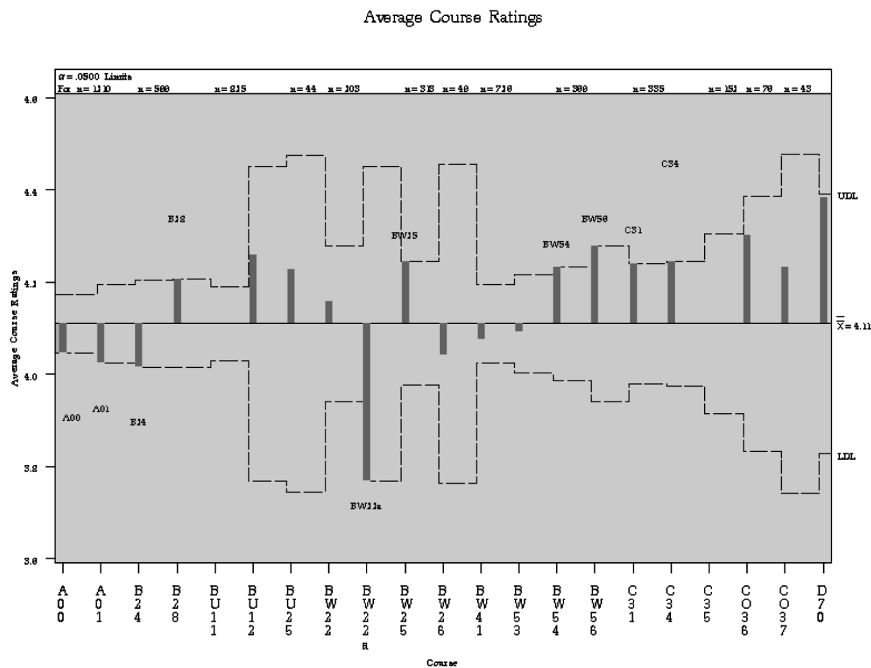
Fig. 1. ANOM plot showing under- and over-performing courses (as measured by average student satisfaction rating).
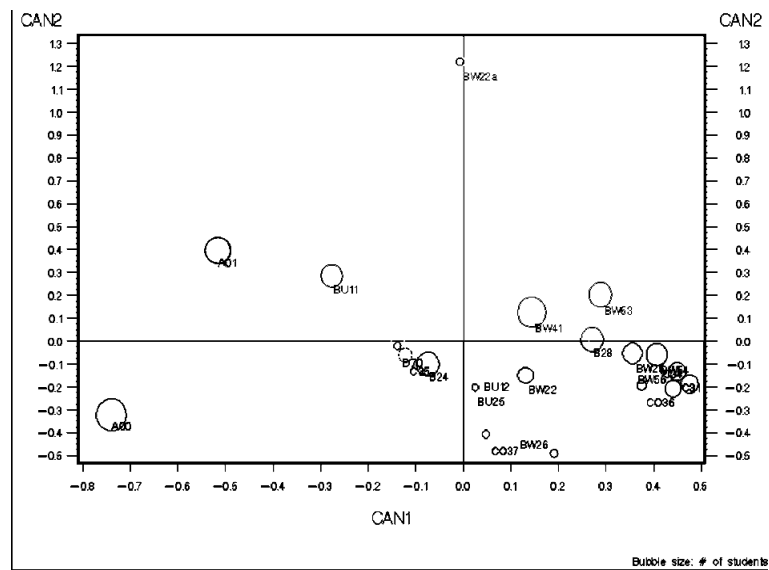

Fig. 2. Course means in the space of the first two canonical dimensions.

content (usability, quality of examples, etc.), while the "first-timers" are more apt to complain about inadequacy of the level of their prior knowledge. Therefore, additional "introductory level" ("orientation") courses may be required at SRCE. In general, there seems to be more dissatisfaction with "too much" (34%) than with "not enough" (11%), with slightly more of the "too much" complaints coming from the "repeat" students.

The second hypothesis (H_{02}) (equality of teacher rating) has been rejected as well. Again, the ANOM was used to identify the above and below average rated teachers/instructors. Since the teacher and course ratings are highly correlated, ANOM was also applied to the teacher ratings adjusted for course ratings. After the adjustment, only 7 teachers (out of 41) were significantly higher or lower rated than average, as compared to 13 before the adjustment.

	Key Productivity Measures			Key Satisfaction Measures			
School Year	No. of courses	No. of students	No. of student hours	Avg. course rating	Avg. teacher rating	% intend repeat course	% intend repeat teacher
1998/1999	215	1,981	11,493	4.10	4.42	98.8%	93.4%
1999/2000	222	2,253	14,098	4.04	4.44	99.2%	93.7%
change	3.3%	13.7%	22.7%	−0.06	0.02	−0.4	−0.3
P	0.74	< 0.0001	< 0.0001	0.02	0.43	0.23	0.67

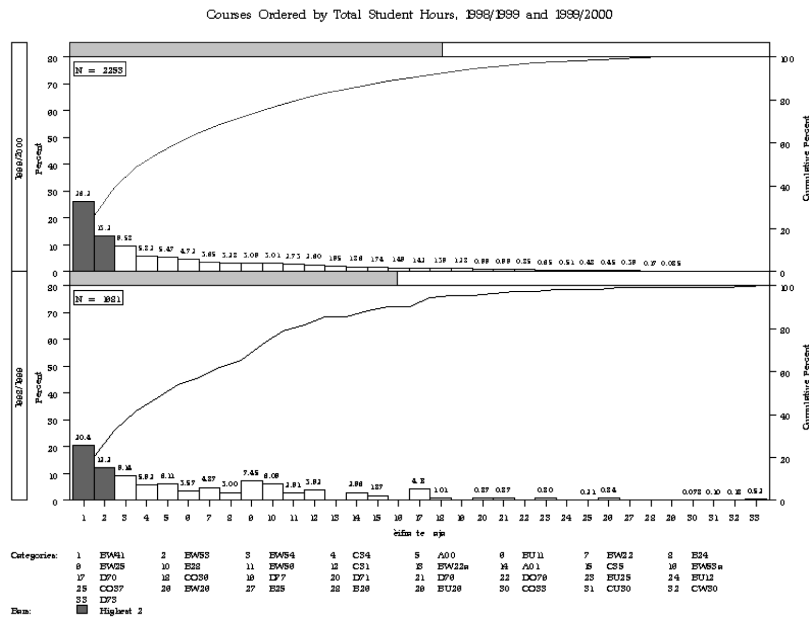*Table 1.* Summary of change in demand and satisfaction rating statistics.

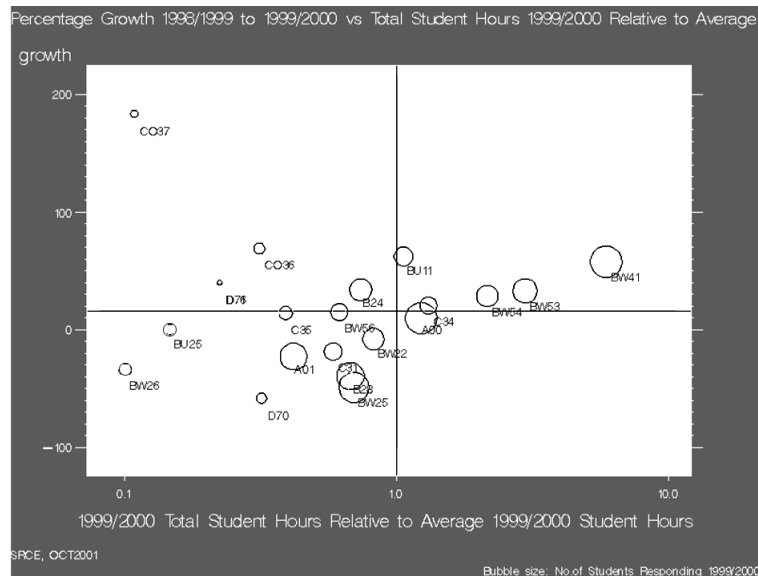*Fig. 3.* Comparative Pareto chart (demand for courses in 1998/1999 vs. 1999/2000).

Fig. 4. Growth-share matrix (demand for courses in 1998/1999 vs. 1999/2000).

Actually, two teachers' average rating moved from the "average" zone to "above average" after being adjusted for course ratings. This would indicate that these 2 teachers were probably rated lower because of dissatisfaction with the course content. Canonical correlation analysis confirmed the presence of both common "teacher/course" dimension and "teacher only" dimension.

The third hypothesis (H_{03}) (no change over time) was tested separately for the demand variables (number of courses, students, student hours) and "satisfaction measures" (i.e. ratings). As shown in Table 1, there was a significant increase in the number of students (13.7%) and an even larger increase in the number of student hours (22.7%) between the last two school years. There seems to have been a slight, but significant drop in the overall course ratings. Preliminary analysis indicates a need for newer courses.

The structure of the change between the last two years is graphically presented in Figures 3 (comparative Pareto charts) and 4 (Boston Consulting Group (BCG) "growth-share" matrix, Porter (1985)). Figure 4 can be used for identifying four groups of courses: "fast growing and high use courses", "fast growing, but low use courses", "low growth, high use courses", and "low growth, low use courses". The first group of courses requires more resource allocation; the second requires either more promotion or competitive analysis, or both; the third group represents stable core course offerings in need of careful management attention; the final group may need evaluation for possible phase out.

Application

The application "SOCRATES" that has been developed (using SAS modules: base, STAT, GRAPH, QC, INSIGHT, AF, and EIS) for the purpose of ad hoc "point-and-click" analytical reporting is composed of three main parts: 1) Productivity measures reporting, 2) Satisfaction measures reporting, and 3) multidimensional database (MDDB) "drill-down" charts and reports. The main screens for 1) and 2) are (partially) shown in Figures 5 and 6, respectively. On the screens there are a number of
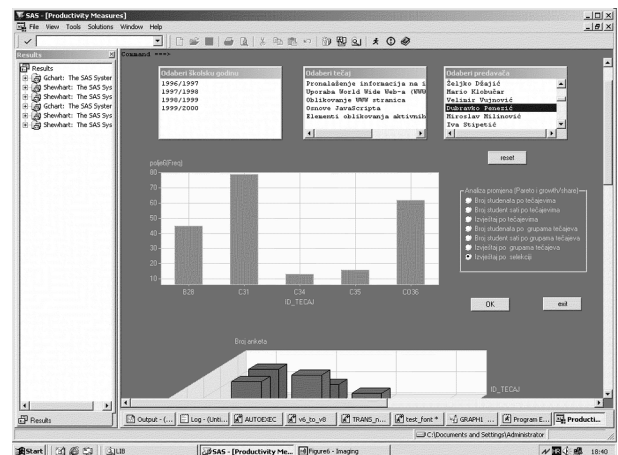

Fig. 5. Application SOCRATES Productivity Measures screen (partial view).
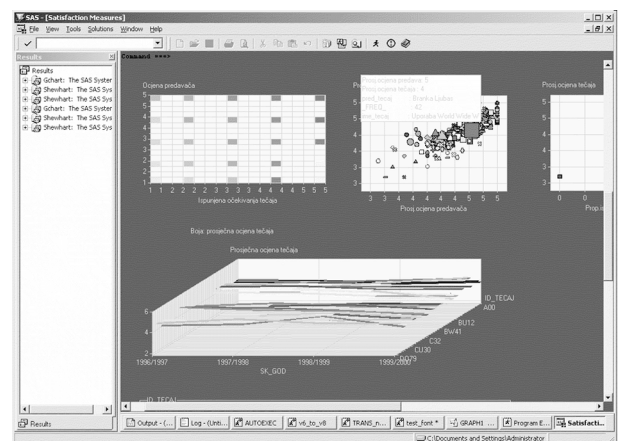

Fig. 6. Application SOCRATES Satisfaction Measures screen (partial view).

"dynamic" graphs and charts that are automatically updated upon selection of a year and/or a course and/or a teacher from the corresponding lists. Additionally, a user can select from a list of predefined analytical graphs and reports (e.g., various ANOM charts for means and proportions, Pareto charts, growth-share plots, etc.) which are, upon selection, instantly created in "printer-ready" and (optionally) HTML (or any standard graphics) format.

Summary

This research project has clearly shown that, indeed, courses and teachers differ in terms of various student satisfaction measures, and that demand for different courses at SRCE is quite dynamic.

The project also indicates that (with a few exceptions) course/teacher satisfaction measures seem to be "in control... at a constant and reasonably acceptable level of performance"... with ample room for continuous (and specifically measurable) quality improvements.

The teaching staff at the University represents precious intellectual assets that, in many cases, have taken decades to create.

The importance of encouraging practical teaching experience cannot be denied.

On the other hand, it is said, "There is nothing so practical as a good theory" (Kurt Luwin, circa 1957). Theory is a path for learning.

This research has just begun a continuous examination of the relationships that seem to exist among 1) course attributes, 2) selected teacher/presenter attributes, and 3) various measures of student characteristics and satisfaction with SRCE. Learning will continue as measures and methods improve and as new theories are tested.

The key findings can be used to help shape the important decisions made by students, administrators and teachers in a huge number of specific ways. The findings naturally lead to a number of relatively unbiased conclusions (thus relying much less on "management opinion"... no matter how "experienced" and "well intentioned" it may be) regarding the need for changes in course content or structure, and regarding teacher performance, teacher fitness for particular course loads, course assignments, and their possible need for, or use in quality improvement efforts.

SOCRATES provides useful evidence in support of argument. But rational decision making requires a sometimes delicate balance between the desire for a change and the need to maintain the status quo. SOCRATES attempts to provide a decision making platform and environment that respect the needs and desires, as well as the capabilities and aspirations of teachers, students, and administrators... in a right and simple way.

Learning moves on the slow path of uncertainty, by trial and error.

But there will always be a dynamic tension between the need for stability as opposed to the need for change,... practice versus theory.

This odd tension is captured clearly by the following quotations:

"Experience without theory teaches nothing." (W. Edwards Deming, circa 1986)

"You must learn by doing it, for though you think you know it, you have no certainty until you try." (Socrates)

One theory being tested at SRCE holds that, for the benefit of our students (i.e., better Customer Relation/Retention Management (CRM)), and in the name of organizational learning, new ideas are worth a try.

References

[1] AGRESTI, A. (1990), *Categorical Data Analysis,* NY: John Wiley & Sons, Inc.

[2] BEKIĆ, Z., L. FILIPČIĆ, M. MILINOVIĆ (1998), "Development of a modular course system in Internet training", *ITI '98*, Pula, Croatia.

[3] KUHFELD, W. F. (1994), Graphical Methods for Marketing Research, In *Marketing Research Methods in the SAS System.* SAS Institute.

[4] MILINOVIĆ, M., J. TINGLE, V. VRGA (2000), "CAR-Net: Educated users for effective utilization of information technology", *SoftCOM 2000*, Split, Croatia.

[5] OTT, E. R. (1967), "Analysis of means — a graphical procedure", *Industrial Quality Control* 24(2): 202–209.

[6] PORTER, M. E. (1985), *Competitive Advantage: Creating and Sustaining Superior Performance,* Freepress publications.

[7] SAS Institute User's Guides and References (Version 8) (1998), (1999), (2000).

[8] SCHILLING, E. G. (1973), A systematic approach to the analysis of means, Part I, Analysis of treatment effects, *J. of Quality Technology* 5 (3): 93–108.

Contact address:
Vesna Lužar-Stiffler
University of Zagreb
University Computing Centre
J. Marohnića bb
10000 Zagreb
Croatia
Phone: +385 1 616 55 97
Fax: +385 1 616 55 59
e-mail: vesna.luzar.stiffler@srce.hr

Miroslav Milinović
University of Zagreb
University Computing Centre
J. Marohnića bb
10000 Zagreb
Croatia
e-mail: `miroslav.milinovic@srce.hr`

Charles Stiffler
CAIR Research Centre
Vlaška 26
10000 Zagreb
Croatia
e-mail: `c|harles.stiffler@cair-center.hr`

VESNA LUŽAR-STIFFLER obtained her B.Sc. in Mathematics at the University of Zagreb, her M.Sc. and Ph.D. in Computer Science from the Faculty of Electrical Engineering and Computer Science in Zagreb, specializing in Computational Statistics. She was awarded Fulbright Postdoctoral Grant for research at the Department of Statistics, Stanford University, specializing in computational statistics and multivariate analysis (data mining foundations). Her work experience includes research, development, and teaching at the University Computing Centre, teaching of statistics at the University of Zagreb, Stanford University, University of Maryland, University of Naples, SAS Institute, Ljubljana. She has consulted in the area of applied statistics and SAS software with various companies (including pharmaceutical, aeronautic, automotive, semiconductor manufacturing technology, insurance, food retail, beverage) and government organizations in the US, Italy, Croatia, Slovenia, Macedonia.

MIROSLAV MILINOVIĆ has B.Sc. degree in mathematics from the Faculty of Natural Sciences and Mathematics – Department of Mathematica, University of Zagreb. For more than 14 years he has been working in University Computing Centre (SRCE), University of Zagreb, where he is currently Assistant Director and Head of Information Services and Applications Department. He is also acting as CEO Advisor in Croatian Academic and Research Network – CARNet. Since 1994 Miroslav Milinović has been participating in establishment and development of information and user support services in CARNet. He is lecturer (since 1997) and track leader (since 1999) for Network Knowledge Systems track in Workshops on Network Technology annually organised by Central and Eastern European Networking Association – CEENET. He has been a member of evalution panels for two international projects in information services and user support area (ETINU, GNRT) carried out by Trans-European Research and Education Networking Association – TERENA. He is author of several scientific and technical papers, publications and courses in the field of information technologies, mainly Internet. He also serves as Program Committee Chair of CARNet Users Conference since 1999 and is a member of the Program Committee the TERENA Networking Conference 2002.

CHARLES STIFFLER obtained his MBA in 1980 at San Diego State University, specializing in Consumer Psychology and Market Research, and his doctorate (Ph.D.) in Business and Administration from the University of Colorado, Boulder in 1985, specializing in Business Strategy and Market Research. He taught at the University of Colorado, Boulder until 1987, where over a 7 year period he completed 120+ market research projects. He was consulting in industry, taught in the area of enterpreneurship and was owner/director of a research firm (Environmental Marketing Inc.) during the late 80's. From 1990 until 1995 he taught with the University of Maryland at various locations in Europe. For the past several years he has worked as a private consultant in Europe and the US, and has periodically taught at the University of Zagreb, SAS Institute, Ljubljana, and the University Computing Centre, Zagreb. Recent areas of research activity include beverage industry, food retailing, insurance marketing, automotive sales, customer relationship management, total quality management, international consumer behavior, and excellence in teaching.