This paper describes the conceptual base for the development of a computer-based expert system. After reviewing developments in computer-based learning and experiments with computer-assisted learning in statistics, the paper describes the nature of expert systems and desired attributes of expert systems in statistics. An overview of proposed research projects to develop a computer-based expert system research outline/statistical tutor is presented. Current progress, anticipated timelines and methodological concerns are provided. Two figures—The Changing Focus of Attention in Technology for Computer-Assisted Learning and System Delivery Tools—are included. (Contains 37 references.) (Author/ALF)
New Directions for Teaching Research Methods and Statistics:
The Development of a Computer-Based Expert System
Lauran H. Sandals
Michael C. Pyryt
Department of Educational Psychology
The University of Calgary

Paper Presented at the meeting of the
American Educational Research Association
San Francisco, California
April, 1992

Running head: New Directions...Expert Systems
Abstract

This paper describes the conceptual base for the development of a computer-based expert system. After reviewing developments in computer-based learning and experiments with computer-assisted learning in statistics, the paper describes the nature of expert systems and desired attributes of expert systems in statistics. An overview of proposed research projects to develop a computer-based expert system research outliner/statistical tutor is presented. Current progress, anticipated timelines, and methodological concerns are provided.

INTRODUCTION

Although significant methodological and statistical advances such as structural equation modeling, meta-analysis, item response theory, generalizability theory, exploratory data analysis and new methodologies for longitudinal research and quasi-experimentation have occurred during the past 20 years, these developments have not been incorporated in the research methods curriculum. Aiken, West, Sechrest, and Reno (1990) recently surveyed 186 of the 222 psychology departments in the United States and Canada identified by the American Psychological Association as offering the Ph.D. degree. Their results suggest a crisis in the teaching of research methodology at the graduate level. Across all programs, a median of 2.6 full-time faculty teach statistics, measurement, or both at the graduate level. One-third of the programs have no faculty trained to teach these courses. The median number of faculty trained to teach statistics and research methods across all programs was 0.9. Results also indicate that topic coverage in current courses resembles that of 20 years ago. ANOVA, contrasts and comparisons, and regression are emphasized; while multivariate analysis, power analysis, and causal modeling are rarely incorporated into the curriculum. Rated measurement competencies of program graduates indicated that only one quarter of recent graduates were judged competent at applying classical test theory
concepts such as reliability, validity, and item analysis. Less than 10% were judged competent at applying item response theory, generalizability theory, and item bias analysis. Aiken et al. (1990) recommended substantial revamping of the methodological curriculum, greater utilization of course offerings in other disciplines such as sociology and economics, informal seminars, faculty retraining, and appropriate statistical review of journal articles. In their discussion, Aiken et al. (1990) also noted that the median age of quantitative psychologists was 51 for members and 65 for fellows of the Evaluation, Measurement, and Statistics Division of APA.

This paper describes the conceptual basis for the development of a computer-based expert system software package that provides consultation in research methods and statistical analysis. It can be used as an instructional aid in graduate courses dealing with research methodology and statistics.

**Developments in Computer-assisted Learning**

Skinner (1986) provided a historical review of the role of the computer as an instructional support tool. From their antecedents in the early work of Pressey (1926) and Skinner (1954), teaching machines evolved into the initial computer-based learning systems on DEC PDP-1 and PDP-11 systems and IBM 1500 systems in the late 1960's and early 1970's. This was followed by the PLATO and TICCIT projects of the mid 1970's to the early 1980's. Major progress occurred with the development of the commercial and affordable microcomputer in the late 1970's early
1980's with the original Apple II, Commodore Pet, and the Tandy Models I and II. Computer-Assisted Learning (CAL) was now available to the masses at various levels of our educational and training institutions. The power of the microcomputer was enhanced by the introduction of the IBM XT computer in 1981 and its generations the AT, PS/1, PS/2, and the respective MS-DOS clones as well as the Apple Macintosh Operating Environment introduced in 1985. Presently these tremendously sophisticated microcomputer based systems, with vast amounts of Random Access Memory (RAM) and hard disk capacity, along with their Graphical-User Interface (GUI)-based Operating Systems such as Windows and Apple System 7 have allowed the development of powerful expert system software shells that operate under these GUI's. These shells allow knowledge engineers and subject experts to create beta test versions and even final expert system products that were not possible even two years ago.

Gaines (1987) has also examined the historical development of computer-assisted learning. His integration of parallel developments in the areas Knowledge Science and Computer Science incorporated in a timeline is shown in Figure 1. The model demonstrates the many forces that have come together both from a behavioral and technical perspective so that we can now jump from lab-based experimental work in expert systems to field-based research and finally commercial-based projects. Gaines (1987) predicted that major innovations would occur in the 1990's.
Sandals (1989) provided a critical evaluation of research and meta-analysis in the field of computer-assisted learning. This review article examined issues in evaluating the efficacy of CAL. A six-step research continuum ranging from concern with hardware to multi-site field testing and cross-validation was introduced as a structure for examining research and development activities in CAL. Too often, CAL products are commercialized without undergoing empirical evaluation. Sandals (1989) challenged researchers to demonstrate that Computer-assisted learning is more effective than well-designed traditional instructional activities.

**Computer-Assisted Learning and Statistics**

One of the earliest major long term CAL projects in the area of research methods and statistics was developed and carried out at the University of Alberta in the Faculty of Education beginning in 1974 on their Educational Psychology 502 course named STAT1 for a IBM 1500 system (Hunka, Romaniuk & Maguire, 1976; Kearsley 1976). The course was later transferred to a DEC VAX 11-780 system in 1980 and renamed STAT2. This system took over 3000 hours of time to design, programme, and to revise over a period of four years in its STAT1 version. Student terminal time to took between 29 to 160 hours for completion with an average course completion time of 69 hours. This course ran as
mainline CAL with additional labs and discussion groups supplementing the CAL course instruction. In order to go through all routes in the course approximately 3000 responses were needed on part of the students. There were 11 formative tests spread out through the course. The time needed to write seven of these exams which were computer administered ranged from 2.9 to 21.9 hours with an average completion time being 8.3 hours. In the completion of this course through CAL, students accomplished their goals as demonstrated by their scores on tests and affective responses to surveys. Students also saved both themselves and their instructors an average of 24 hours of lecture and 84 hours in traditional laboratory sessions. Since the professors and lab instructors were freed up from traditional teaching activities, they had more time for one-on-one sessions with students when the need arose (Sheridan, 1980). This system was also used at The University of Calgary from 1988-1990 by the first author to help students review basic skills before pursuing other graduate courses in research design and statistics, to prepare for candidacy examinations, and to encourage preliminary work on their theses. Unfortunately the system is now unavailable due the termination of the VAX computer systems at the Universities of Alberta and Calgary.

Other systems have been developed by individuals across North America to examine the use of computers especially micros as adjunct support in the area of behavioral statistics (Gale, 1986; Slawinski, Jamieson, Ells & Wasko, 1987). These systems
usually emulate the computer as lab tool and are used to support such topic areas as sampling, central tendency, measures of dispersion, charting and graphing data, and univariate tests of inference.

**Expert Systems**

The British Computer Society Specialist Group defined an expert system as "the embodiment within a computer of a knowledge-based component, from an expert skill, in such a form that the system can offer intelligent advice or take an intelligent decision about a processing function. A desirable additional characteristics, which many would consider fundamental, is the capability of the system, on demand, to justify its own line of reasoning in a manner directly intelligible to the enquirer. The style adopted to attain these characteristics is rule-based programming" (Forsyth, 1984, p. 10). The four essential components for an expert system are: (a) the knowledge base, (b) the inference engine, (c) the knowledge acquisition module, and (d) the explanatory interface. The most important component of the expert system is the knowledge base, which consists of facts and rules about a knowledge domain and heuristics for searching through the knowledge base (Bunderson & Inouye, 1987). The inference engine is the system for the series of steps and queries of the knowledge base that lead to conclusions (Naylor, 1984). The knowledge acquisition module is the presentation of the knowledge in a symbolic code as to be useable by the computer. The explanatory interface provides the
reasons why a program reached the given conclusion or asked a particular question. Doukidis and Whitley (1988) summarized Humphreys' (1984) description of the five level of analysis needed in order to make a decision about how much is known about a specific knowledge domain. The five levels are as follows:

- Level 5 --Verbal Description of Knowledge
- Level 4 --Selection of Knowledge Representation Technique
- Level 3 --Selection of Variables/Values within the Representation
- Level 2 --Specification of Knowledge according to different Values
- Level 1 --Development of Rules

**Expert Systems and Statistics**

Discussions of the issues regarding the development of expert systems in statistics are provided by Gale (1986), Hand (1984, 1985a, 1985b, 1986a, 1986b, 1987, 1990a, 1990b, 1991), and Smith and Hand (1983). Hand has developed his own expert type system in the area of Non-Parametric Statistics titled the "Knowledge Enhancement System" or (KENS). Hand (1985) has provided recommendations for expert systems in statistics. The system needs to be adaptable to incorporate new knowledge and new techniques. The system needs to explain why it is asking a particular question or recommending a particular technique. The system needs to explain technical terms so as to be user-friendly for the statistically naive. The system needs to recommend multiple techniques. The system needs to permit both exploratory
data analysis and formal hypothesis testing. The system needs to adapt and make changes as the analysis occurs. For example, a different analysis may be required if the assumptions for a statistical technique are not met. The system must be flexible to allow experienced users to override recommendations. The system should be adaptable to backtrack when necessary. The system needs storage capabilities. An ordering of techniques is necessary. Questions should be sensible and build on previous questions. A teaching function would be helpful. Ideally, the program would be self-contained.

The most comprehensive attempt to use expert system technology in research methodology and statistics is the METHODOLOGIST'S TOOLCHEST PROFESSIONAL (Brent, 1991). This package consists of seven different expert systems programs. DESIGNER RESEARCH assists the investigator in choosing efficient and valid research design for a given problem. This program explores strategies for providing comparable groups, strategies to reduce expectancy effects, assignment strategies, external validity strategies, monitoring and assessment strategies, timing strategies, and strategies to ensure construct validity and statistical conclusion validity. STATISTICAL NAVIGATOR PROFESSIONAL assists the investigator in choosing from over 200 possible statistical analyses that range from non-parametric techniques such as the Kolmogrov-Smirnov test to causal modelling techniques such as LISREL. EX-SAMPLE performs to power analysis to determine necessary sample size. MEASUREMENT & SCALING
STRATEGIES guides the researcher through all phases of the tests construction process. It also helps identify existing measurement scales in the social sciences. DATA COLLECTION SELECTION guides the researcher in determining appropriate data collection strategies. WHICHGRAPH provides consultation in the selection of over 100 types of graphs and advice on graph construction to minimize bias and misunderstanding. HYPERSTAT is an on-line dictionary of methodological, statistical, and graphical terms.

Using Hand's criteria, Brent's system gets mixed reviews. The system provides a comprehensive list of techniques from median tests to covariance structure analysis. The latest version provides the opportunity to browse definitions of statistical terms. The system recommends multiple techniques and ranks order the ratings. The system provides some insight into the explanations for the questions asked and recommendation (i.e. the inference-engine). Backtracking is possible. The system does not appear to be user-friendly. A great deal of sophistication is needed to interact with any of the programs. The phrasing of questions seems idiosyncratic. Only one class of techniques is considered at a time. There is no direct tie in to data analysis results. Although this system provides an explanatory interface and description of the programming language used in the knowledge acquisition modules, it is difficult to determine if any of the recommended procedures have a sufficient knowledge base beyond the recommended practices in selected
textbooks. Since the knowledge base is the key determinant of the construct validity of an expert system, it is essential that rigorous validation procedures be applied in developing the knowledge base.

**METHODOLOGY FOR THE ESTABLISHMENT OF THE KNOWLEDGE BASE**

In order to even start the development of an expert system research/thesis proposal consultant, it is extremely important that as wide a scan as possible should be made of other Social Science experts who are involved in teaching and advising students and faculty on research methods and descriptive and inferential statistics. In addition to collecting results from other experts in the area of research methodology and statistics initially in university Psychology and Educational Psychology Departments, we have been contacting and surveying as many institutions and individuals as possible regarding their research projects, surveys and courses such as those listed in other parts of this proposal. For example, the Social Statistics Research Unit at the City University, London, England has recently completed a full masters program in Social Research Methods and Statistics which is offered on a core module basis along with 16 other modules which may be taken on an individual basis. The program caters to students who have prior degrees and career expertise in such fields as Education, Psychology, Sociology, Nursing, Social Work and Management. The Polytechnic of Central London and Surrey University also have an Interdisciplinary Program in Social Research Methods that looks at unique ways of
defining the knowledge bases covered in such programs. Because of the need to collaborate among academics across disciplines in this area conferences have been held in England on the "Teaching Methods of Social Research". The report of the most recent conference was edited by Gubbay (1991). It is hoped the knowledge databases from these programs and courses, in addition to others, can be used in the development of our three expert system projects. On completion of the development of the knowledge base and system prototype, the proposers intend on field testing the system by releasing the package to interested university departments and centres which would be willing to beta test the system and its related knowledge base. The feedback from these test sites will be used in developing the final system product. A similar process for field testing and feedback will also be used on the final expert system packages and the CD-ROM research database.

OVERVIEW OF THE PROPOSED RESEARCH PROJECTS

1. A research/thesis proposal outliner and writer which will include a checklist and guide for the development of a research or thesis proposal that will make use of expert system help routines. This research project/product will include both basic and advanced research methodological concepts and issues.

2. A computer-based guide and tutorial for the proposed Statistical Techniques which will cover concepts similar to the University of Michigan "Guide for Selecting Statistical Techniques for Analyzing Social Science Data" (Andrews, Klem,
Davidson, O'Malley, & Rodgers, 1981) and other recently developed software products such as DESIGNER RESEARCH and STATISTICAL NAVIGATOR (Brent & Mirielli, 1989) and STATISTICAL NAVIGATOR PROFESSIONAL (Brent, Mirielli, Detring & Ramos, 1991). Our system will have a more logical interlink among its phases. It will utilize state-of-the-art Graphical User Interfaces (GUI'S). This product as well as the Research Thesis outliner will probably utilize the advantages of running under the WINDOWS (3.1 and above) Operating System for the IBM/PC and clone family of computer users, or X-Windows for the Unix user and the Mac System 7 (and above) Operating Environments for the Macintosh user.

3. A CD Rom with cross-referencing to the Research/Thesis Outliner and Statistical Tutor. This CD Rom will include a comprehensive data bank of research studies in the last 10 years broken down by social science discipline that have used research designs and statistical techniques which were recommended by through the use of the two products. The studies will be cross-referenced by descriptors to identify sample size, sampling techniques, type of instruments used, and additional information still to be determined. All three research projects and final products listed above will run as stand alone packages but they will be more powerful if the end user utilizes them as a three-step integrated package which will allow the user to move around the system until a final solution and/or proposal is developed. The order of development will be described as Project 1, 2, and then 3. The target dates for completion of the research projects
are to be: Phase (1) Establishment of the knowledge base - Sept. 1992; Phase (2) Project 1 and 2 - Sept. 1993; and Phase (3) Project 3 - Sept. 1994. Our projections take into account the present financial and human resource limitations.

At the present time this project is still at its early stages. Our work on the knowledge base is between Level 5 and Level 4 as described by Humphrey (1984). In choosing a language to author our prototype, several prior expert research system projects and related languages have been reviewed. Many different system delivery tools have been developed over the last 30 years with one of the first being LISP which was developed by McCarthy in the late 1950's. A partial breakdown of some of the more popular Expert delivery tools appears in Figure 2 as developed by Doukidis and Whitley (1988). The two categories for graphing these tools are the "Data and Knowledge Representation" and the "Control and Inference Strategies" dimensions with each running on a scale from Rapid to Flexible. For the most part Rapid means that a system is easy to use but with a shell that can not be modified to any great extent. A system on the flexible scale requires a more sophisticated programmer or knowledge engineer. A flexible system is more powerful since the developer is usually only restricted by his or her programming skills. We will be prototyping our initial pilot versions in two formats probably PESYS (Pascal Expert System) developed by Edgar Whitley (1990) at the London School of Economics and a WINDOWS expert system shell called KNOWLEDGE PRO from Knowledge Garden
Inc. The final version will probably be rewritten in C or a similar language that can be ported from system to system with more ease than the dedicated expert shells.

SOME FURTHER ISSUES

The quest for consensus on the identification of skills and competencies regarding personnel involved in research in education has been an ongoing issue for debate. Worthen (1975) described the research activities of the AERA Task Force on Research Training. As a first step, the AERA Task Force on Research Training listed the skills needed to conduct research and evaluation in education and discussed these skills with 60 potential employers (Worthen and Gagné, 1969). As a second step, the competencies were reviewed and refined based on further feedback from the field (Glass and Worthen, 1970). As a third step, Anderson, Soptick, Rogers, & Worthen (1971) carried out a task analysis of research and evaluation activities in 13 agencies with 109 professionals in the area of research and evaluation. A factor analysis was carried out on the tasks in order to come up with a cluster of competencies in research and evaluation. Worthen (1975) summarized the work of the preceding reports into a list of "Twenty-five General Research and Evaluation Tasks and Related Competencies" along with 82 subcompetencies identified as necessary to perform the 25 general
competencies. Worthen commented, "With the increasing sophistication and specialization in the various fields on which educational evaluation and research depend, it seems increasingly unlikely that any single individual will be such a paragon as to be highly skilled in all of the important areas. The more reasonable stance seems to be to assert only that the full range of relevant competencies be either possessed by or readily available to whatever entity is carrying out the work - whether that entity be an individual, a team, or an organization." (p. 14). When defining the "entity" Worthen (1975) did not realize that it might be a computer-based expert system. The competencies as defined are still relevant today. These competencies and others that have been subsequently identified will be incorporated in the development of this proposed expert system knowledge base.

Another issue that has arisen in the discussion of the proposed knowledge base relates to the issues of the qualitative and quantitative research paradigms. There are questions about the feasibility of having the expert system capable of conducting a meaningful dialogue with the researcher if the proposed study is determined to be qualitative in nature. After discussion with such experts as Martin Hammersley and others at the conference in London on "Teaching Methods of Social Research" it appears that at least from the qualitative researchers perspective this may be very difficult because of the hierarchical nature of quantitative methods and the almost complete lack of it in qualitative
research. Thus, it would be difficult to build a rule-based system to analyze research issues in the qualitative domain. Hammersley (1992) suggests that the quantitative vs qualitative dichotomy is counterproductive. He argues that what is typically classified as qualitative or quantitative represent a continuum of research tools and skills instead of two very different methodological approaches. We are hopeful that we will be able to address many of the research issues and related analysis techniques in our proposed system if we have a continual dialogue and support from traditional researchers in the two methodological domains as well as leaders in this debate such as Hammersley.
References


Hand, D. J. (1990b). Emergent themes in statistical expert systems. In M. Schader & W. Gaul (Eds.), *Knowledge, data,


New Directions...Expert Systems

Socially Organized Systems
Autonomous Activity Systems
Inductive Inference Systems
Knowledge-Based Systems
Human-Computer Interaction
Problem-Oriented Languages
Virtual Machine Architecture
Electronic Device Technology

Generation
0 1 2 3 4 5 6 7

Mechanical "teaching machines"
Electronic "teaching machines"
CAL languages, coursewriters
Interactive CAL, personal computers
Knowledge-based CAL, expert systems
Adaptive CAL, self-optimizing
Goal-driven CAL, self-generating

Figure 4 The changing focus of attention in technology for computer-assisted learning

Figure 1. The Changing Focus of Attention in Technology for Computer-Assisted Learning
Figure 2. System Delivery Tools