The Design and Validation of a Hybrid Information System for the Auditor's Going Concern Decision

MARY JANE LENARD, GREGORY R. MADEY, AND PERVAIZ ALAM

MARY JANE LENARD is an Associate Professor of Accounting at Barton College in Wilson, NC. She received her Ph.D. at Kent State University, a B.S. in Economics from Carnegie Mellon University and an M.B.A. in Finance at the University of Akron. She is a Certified Management Accountant and a member of the Association for Information Systems, Decision Sciences Institute, and the American Accounting Association. Dr. Lenard has recently published in *Decision Sciences*. Her current research interests include hybrid systems and neural networks applied to auditing and other business decisions.

GREGORY R. MADEY is an Associate Professor of Information Systems at Kent State University. He received his Ph.D. from Case Western Reserve University. He has published articles in many journals, including *Communications of the ACM* and *IEEE Transactions on Engineering Management*. His research interests include electronic commerce, artificial intelligence, neural networks, and hypermedia/multimedia.

PERVAIZ ALAM is an Associate Professor of Accounting at Kent State University. He received his Ph.D. from the University of Houston. His areas of research interests are auditing, financial economics, and international accounting. Dr. Alam's recent publications have appeared in *Decision Sciences*, *Financial Review*, and *Journal of International Business Studies*.

ABSTRACT: Decision making in a semistructured environment often involves the use of quantitative, structured analysis along with the qualitative judgment of an expert. Decision support systems and expert systems are often developed to assist in this judgment process. The hybrid information system described in this paper combines a statistical model with a rule-based expert system in order to integrate the quantitative and qualitative aspects of decision making. The GC Advisor hybrid system is designed for use by auditors to assess the ability of the client firm to continue as a going concern. The guidelines for expert system validation given in previous literature are then applied to the validation of GC Advisor.

KEY WORDS AND PHRASES: auditor's going concern assessment, expert systems, hybrid information systems, system validation.

A HYBRID SYSTEM INTEGRATES TWO OR MORE TECHNOLOGIES in order to utilize their strengths and minimize their weaknesses. Often one or more intelligent technologies

are integrated [31, 38]. The hybrid model developed here combines a statistical model with a rule-based expert system to assist in the audit opinion decision and is imple-mented as the GC Advisor hybrid system. The going concern uncertainty opinion is issued by the auditor to a client company when that company is at risk of failure or exhibits other signs of distress that threaten its ability to continue as a going concern [27]. This decision facing the auditor involves the use of both quantitative and qualitative information. Statistical models have been used with varying degrees of success in predicting whether a firm should receive a going concern modification from publicly available information. Studies by Bell and Tabor [7], Chen and Church [15], and Lenard et al. [27] use logit models or neural networks. These models represent the use of quantitative information to predict whether a firm should receive an audit report with a going concern modification, or whether a standard audit report indicating a healthy firm should be issued. In contrast to statistical models, expert systems provide the auditor with a qualitative analysis of the going concern decision [10, 21].

The GC Advisor is programmed using an object-oriented system. The object-ori-ented method is combined with rule-based procedures in order to apply GC Advisor's logic. Incorporated into the logic of the auditor's going concern judgment is the process of generating as well as confirming or disconfirming hypotheses. Libby [28] describes many of the ill-structured accounting decision situations as diagnostic problems. As such, when a diagnostician must search for further information in order to investigate the symptoms of a problem, he or she must formulate a hypothesis to direct the search. A recent expert system designed for diagnostic problems and presented by Benjamins and Jansweijer [8] describes the processes of symptom detection, hypothesis generation, and hypothesis discrimination. The feature included in their expert system for hypothesis discrimination allows for additional observations to be incorporated into the analysis before the diagnosis is complete. The GC Advisor includes a similar feature that provides hypothesis discrimination.

The next step in evaluating the models that predict the going concern decision is to combine the quantitative and qualitative decision making models to produce a hybrid model. The focus of this paper is the development and testing of the hybrid model that predicts the going concern decision by combining the quantitative and qualitative decision making models. An expert system that predicts the going concern decision is combined with a statistical model that predicts bankruptcy. One factor that contrib-utes to the going concern decision is an analysis of financial variables that could indicate bankruptcy. However, previous studies show only a weak association be-tween the issuance of an audit report modified with a going concern uncertainty and the occurrence of bankruptcy [15, 32]. Therefore, if a bankruptcy prediction model is included as a component of the analysis that the auditor performs during the going concern assessment, it can potentially provide a good deal of support for the auditor's decision. In our hybrid model, the bankruptcy prediction is included in the logic of the expert system and represents a component of the going concern decision. In other words, the decision reached by the hybrid includes the consideration that, if a firm is predicted to go bankrupt, it should receive a going concern opinion.

Once the hybrid system is designed, it must be validated. Validation is an important

part in the process of designing, developing, and implementing an expert system because it is included in the decision-making success of the system [36, 37]. Validation methods can be either qualitative or quantitative, or a combination of the two [5]. The validation of the system must address a number of areas, including construct validity, content validity, and criterion validity [37]. The validation can be accomplished through subsystem validation, face validation, and predictive validation [5, 36].

Construct validity emphasizes the existence of a theory on which the system is based [37]. In this study, the theory is supported by the principles supplied by the auditing standards and their consistent interpretation by the experts involved in the development of the model. Content validity can be assessed through direct examination of the system by the expert, by a system test against human experts, and by a system test against other models, while criterion validity involves a definition of the level of expertise of the system [37]. Back [5] uses subsystem validation "interlaced" with face validation to determine content validity. Back's method uses examples from real companies to test specific sets of inputs and outputs. Then face validation involves demonstrating the system to selected experts in order to elicit comments about the system's performance in terms of functionality. Back accomplishes criterion validity through predictive validation involving the use of test cases. The objective of the validation in this study is to expand upon Back's work and to apply her validation methods to the GC Advisor hybrid system.

This study first presents a summary of prior research in hybrid models, followed by the hypotheses that we propose to test the expert system and hybrid models. Next, we summarize the design of the statistical model that performs the bankruptcy prediction. Input to the model is a training set of financial variables, as well as a binary response variable, that are used to indicate the financial health of a firm. The next section describes the development and testing of the expert system that uses rule-based logic to recommend whether or not a company should receive an audit report modified with a going concern uncertainty. Finally, we describe the hybrid system that combines the quantitative, statistical model and the qualitative, expert system model into a rule-based system that uses the statistical bankruptcy prediction model to assist in the determination of the audit opinion for a company. Once the statistical model and the expert system have been combined to form the hybrid model, we test the performance of the hybrid model using a series of test cases of bankrupt and nonbankrupt firms. The conclusions to be reached are that the bankrupt firms in the test set should receive an audit report modified with a going concern uncertainty, and the nonbankrupt firms in the test set should receive a standard, or unmodified, audit report. The results of the testing show that the hybrid system performs better than either the statistical model or the expert system alone.

Previous Research in Hybrid Models

THE AUDITING LITERATURE HAS NOT YET MODELED THE AUDITOR'S GOING CONCERN decision using hybrid systems. However, hybrid systems have been developed for manufacturing and related applications. Rabelo, Alptekin, and Kiran [38] have combined

neural networks and expert systems in the design of flexible manufacturing systems. Madey, Weinroth, and Shah [30] describe a neural network that is embedded into a factory simulation for modeling continuous improvement policies. Maren [31] describes a series of neural network hybrids that report improved results compared with traditional methods. Two of these are neural networks operating in parallel and hierarchies of neural networks. Specifically, Gevins and Morgan [19] have developed systems of interacting neural networks to analyze multiple EEG signals recorded from the brain. Their network has replaced one of two human scorers who previously hand-marked the data they collected. Anikst et al. [3] used two binary-tree networks for encoding speech information at two different levels. Another example of different networks that are fused together is reported by Hering, Khosia, and Kumar [23], who separately trained and combined a system of networks for tactile sensing. Rossen and Anderson [41] have used two modules of hidden layer networks within a larger network system for speech recognition. Rabelo and Avula [39] use a hierarchical neural network system for intelligent control of a robotic arm.

The advantages of hybrid systems over traditional expert systems are that there are more choices of artificial intelligence (AI) techniques. As such, the hybrid system can perform the task with greater efficiency. Hybrid models also provide a way to combine the solutions or solution procedures of analytic and simulation models [42]. In addition, hybrid systems have the potential of addressing more complex tasks because of their combination of techniques. Potential disadvantages, however, are that a greater demand is placed on system resources and that the developer may be faced with programming a task that is best evaluated using several different types of models.

Hypotheses

THE FOLLOWING HYPOTHESES ARE FORMULATED TO MEASURE the significance of various factors that are proposed to assist a decision maker in the evaluation of a company as would be performed in an annual audit by an external auditor.

Hypotheses for Expert System Validation

Once the system has been programmed using the experts' logic, development continues with the validation and evaluation of the system. According to Hansen and Messier [20], validation is an iterative process that takes place throughout the development of the system, and the evaluations become more formal as the system matures. Evaluation in the early stages consists of a demonstration that the prototype can be used on simple cases. Then, once the knowledge base is expanded, the evaluation process continues to include more complex cases and feedback from experts and potential users.

The iterative process of development of the GC Advisor hybrid system can be described as follows: We begin with the development of the expert system and demonstrate it to our experts. We perform separate validations of the expert system and the statistical model that are combined to form the hybrid. We then perform a validation of the hybrid system using a series of test cases.

The hypotheses that we test for the expert system reflect the expert system's use by two small groups of auditors at each of the three Big Six public accounting firms. For purposes of this study, the auditors are called novices because they have a maximum work experience of six years, and none of them is at the level of manager, which is the level of expertise usually required to recommend a conclusion to the audit report. The first hypothesis is based on a test similar to the one used by Hansen and Messier [20]. The test refers to the responses to a questionnaire that is completed by the novices before and after they work with the expert system. The questionnaire seeks their opinions about whether they believe the use of the expert system will have a positive impact on the performance of their job. Thus, the following alternate hypothesis is tested:

> *HA1: Novices' scores on the belief questionnaire will be above average and show positive improvement after the use of the expert system as compared with their scores before the use of the expert system.*

The second test assesses the level of expertise provided by the system. Novice auditors complete, or "solve," two test cases, one with and one without the expert system. The alternate hypothesis considers whether a novice using the system lists a similar number of responses to explain the logic of the audit report decision, as compared with the novice evaluating the case without the system.

> *HA2: The number of responses provided by novices using the expert system will be consistently within an acceptable range, as opposed to the number of responses provided by novices not using the expert system.*

Hypothesis for Comparison of Models

Once the statistical model and the expert system are combined to form the hybrid model, preliminary testing is performed by comparing models. The comparison uses a series of test cases of bankrupt and nonbankrupt firms. We expect that the bankrupt firms in the test group would receive an audit report modified with a going concern uncertainty, and the nonbankrupt firms in the test group would receive a standard, or unqualified, audit report. Hence, the following alternate hypothesis is formulated:

> *HA3: The hybrid model performs better than the statistical model or the expert system in predicting the type of audit report that a group of bankrupt or nonbankrupt firms should receive.*

Design of the Statistical Model

THE STATISTICAL MODEL HAS BEEN DEVELOPED TO INCLUDE a series of financial and qualitative variables that identify firms likely to go bankrupt based on stress signals exhibited by those firms. This information about stressed firms can then help the auditor decide whether or not to modify the audit report with a going concern uncertainty.

Financial Variables Used

Studies that analyze bankruptcy and studies that predict the going concern decision have demonstrated a great deal of accuracy in using financial ratios as variables in the statistical modeling process. Typically, these ratios are used to assess liquidity, solvency, and profitability [34, 44]. We have relied on previous research by Mutchler [35] and Chen and Church [15] to identify a set of financial ratios to use as variables in our model. Mutchler [35] lists the following ratios identified by auditors as those used in the audit opinion decision: (1) Net Worth/Total Liabilities (NWTL); (2) Cash Flow from Operations/Total Liabilities (CFTL); (3) Current Assets/Current Liabilities (CACL); (4) Total Long-Term Liabilities/Total Assets (LTDTA); (5) Total Liabilities/Total Assets (TLTA); and (6) Net Income before Taxes/Net Sales (IBTS). Of these, CFTL and CACL are liquidity ratios, NWTL, LTDTA, and TLTA are solvency ratios, and IBTS is a profitability ratio. In addition, the study by Mutchler [35] analyzes how good news or bad news variables might affect the auditor's going concern judgment. She compiles a list of good news versus bad news variables. *Good news about debt* includes information such as having a line of credit available, issuance of new debt, forgiveness of debt, restructuring of debt, and waivers obtained for violation of debt covenants. *Bad news about debt* (BND) includes default on debt or significant uncertainty about whether future debt payments can be made, accounts receivable factoring, and preferred dividend arrearages. The variables that test as significant in our study are: (1) CACL, (2) NWTL, (3) LTDTA, (4) TLTA, and (5) BND. Although the profitability measure IBTS is not significant in the statistical model, profitability issues are applied as part of the expert system logic in the next stage of the model.

M-Estimator Discriminant Analysis

We performed preliminary testing to determine the best statistical model based on the highest prediction accuracy. Models considered for evaluation were logit, a neural network that used the GRG2 optimizer algorithm, discriminant analysis, and the *M*-estimator discriminant model described below. The *M*-estimator model proved to have the highest prediction accuracy for the test data set used, so it was chosen as the best statistical model to be included in the hybrid system.

   M-estimator discriminant analysis has been proposed by Randles et al. [40] and employed successfully by Booth et al. [11], Hu et al. [24], and Booth and Montasser [12]. The *M*-estimator method is a procedure that calculates the Mahalanobis distances and their associated weights. It is a robust procedure that diminishes the deleterious effect of outlying data observations without removing them from the data set [11]. The procedure employed here consists of the *M*-estimator modification and a zero cutoff for classification into the two groups of bankrupt or nonbankrupt firms (a solution value above zero indicates a bankrupt firm, a solution value below zero indicates a nonbankrupt firm). The method follows the description by Booth and Montasser [12], and Lenard et al. [26].

Expert System Development

THIS SECTION DESCRIBES THE DEVELOPMENT OF AN OBJECT-ORIENTED expert system that is applied to the auditor decision of whether a firm is likely to continue as a going concern. First, we describe the application of object-oriented technology to the problem and explain the contribution of this type of system; then we present the expert system design.

Application of Object-Oriented Technology

The objective in developing an expert system to assist in the auditor's going concern decision is to propose a system that offers ease of use, flexibility, accuracy in audit judgment, and adaptability. Ease of use refers to a system that can be learned in a relatively short time frame. Flexibility refers to the ability of the developer to make changes in the programming of the model at the implementation stage. Accuracy refers to the ability of a system to function well in a problem area where there is uncertainty and ambiguity involved in the decision-making process. The choice about what type of audit report to issue is one type of problem that requires the use of expertise and heuristics in order to deal with ambiguity and reach a decision. Uncertainty is also involved in the form of the risk of consequences if an incorrect decision is reached. Adaptability refers to a system that can adjust responses or variables for a particular problem situation, yet still provide the user with "expert" support that can assist in the judgment necessary to make a decision.

The object-oriented approach (OOA) fits the objectives of our study because development can be accomplished in a reasonable time frame while offering advantages of flexibility and expansion. Henderson-Sellers and Edwards [22] have found that "a system based on object representation can remain more flexible since changes at the implementation level are more easily accomplished without requiring changes to the systems design itself." Furthermore, Muhanna [33] believes that object-oriented approaches encourage the use of software engineering principles such as decomposition and stepwise refinement and, as such, promote and facilitate reusability.

Level 5 Object [25] is an object-oriented expert system shell that runs under Windows. It combines the use of rules and procedures with object-oriented programming. The user's application is a domain that can be divided into classes. In addition, any button or value box added to the system's interface can also be programmed with a function. The expert system is designed to achieve one or more goals. These goals may be reached through the context of the application, through a programmed set of rules, or through the use of the instances of a class that are solved when needed or when changed. The Level 5 software can also incorporate hypertext features and forward-chaining logic. Objects that are part of the display can also activate procedures.

Expert System Design

A group of partners from three of the Big Six public accounting firms, for a total of four different experts (experts 1, 2, 3, and 4), participated in the design of the rules for

the expert system. These individuals were interviewed to determine the approach each of them uses to decide whether a company should receive the going concern modification. They were also asked to examine at least one case study so that their approach could be observed in practice.

The development of the expert system logic is divided into three phases. The phases represent the development of the system and the procedures for working with the experts. Phase I consists of the development of the prototype and the initial consultation with the first expert. The second expert is brought in during phase II, and system development proceeds with the two experts until phase III. During this final phase, the remaining two experts are included, so that all four experts are consulted for the final phase of the development. Any changes made by experts 3 and 4 were presented to the first two experts and accordingly revised. Development of the system was not considered complete until there was consensus among all the experts. This procedure for developing the system in phases provided a framework for successfully incorporating the knowledge of multiple experts into the system. Throughout the development process, the experts expressed a concern that the user could not reconsider any of the responses entered during a session. We therefore created a "summary screen," which presents the current state of the system before the final conclusion is made. This summary screen incorporates the application of hypothesis discrimination to the system logic. This is the same concept of hypothesis generation that has been proposed by Benjamins and Jansweijer [8] in their application of a medical diagnosis. Use of the summary screen enables the user to access "edit" screens through a series of push-buttons that activate links to these screens. If changes are made, the system invokes forward-chaining logic that adjusts the system's decision.

Expert System Validation

THERE ARE TWO OBJECTIVES OF THE VALIDATION IN THIS STUDY: (1) to conduct a preliminary evaluation of the system through consultation with both experts and a limited group of novice users, and (2) to develop a quantitative procedure to test the outcomes of the system. The first objective establishes content validity by eliciting an evaluation from experts, as in Back's [5] study. The first objective also establishes criterion validity in terms of the level of expertise provided by the system, because novice users complete one test case without the system, and one test case with the system. The two cases are different in order to determine whether a novice evaluating a test case using the system elicits a similar number of responses to explain the audit report decision, compared with a novice evaluating the case without the expert system. Novices also answer a questionnaire designed to elicit their beliefs about expert systems, both before and after they solve a test case on the computer. The number of users testing the system is relatively small (ten individuals) at this stage of development. The second objective uses predictive validation, which also helps assess criterion validity. A prediction sample based on 1990 financial statement data is used to compare the results of the expert system with the results of the statistical model and

the results of the hybrid system. Hypotheses 1 and 2 are examined for expert system validation and hypothesis 3 for the comparison of models.

Procedures for Testing with Novices

Novices were selected from each of the three public accounting firms to participate in the testing of the system. At this stage of the study, the system was not completely developed, so the testing was a preliminary investigation of the system's knowledge and explanation capabilities, as opposed to a full-scale field test. The testing was performed on just the expert system to determine subjects' acceptance of an "expert" decision model. The results of this testing were to assist us in the development of the hybrid system. Each of the individuals voluntarily participated in the testing. Four individuals were selected from firm 1, three from firm 2, and three from firm 3, for a total of ten test subjects. The objective of the testing is to determine whether the novices believe the system will have a positive impact on the performance of their job (hypothesis 1). The usefulness of the system is addressed by the ability of the novices to complete the task using the system, and also by several groups of questions on the belief questionnaire. The questions address several areas: (1) ease of use, (2) system flexibility, (3) the system's effect on audit judgment, and (4) the issues of the cost and time to complete an audit when assisted by a system. These areas are important because one question that measures the impact of an expert system is whether interaction with the system leads to changes in the user's perception of the domain [43]. The question-naire was administered both before and after the novices completed the task. This procedure is similar to the one used by Hansen and Messier [20] in testing their system, EDP-XPERT. The procedure for testing with the novices is reviewed below.

The task itself consisted of reviewing and solving two cases, each of which contained a description and financial information about a company. The information was taken from actual corporate annual reports. Included in each case discussion was a descrip-tion of the company, highlights from management's discussion and analysis, relevant financial reports consisting of the balance sheet, income statement, and cash flow analysis, and any footnotes that were appropriate, such as those that explain long-term debt, liquidity analysis, and litigation, if any. In order to complete the task, the individual had to decide the type of audit report that the firm was to receive: standard audit report or audit report modified for going concern uncertainty. Each individual received one case of each type, although this fact was not revealed to the individual before the test. Case 1 was a company whose correct solution is a standard audit report, while case 2 was a company whose solution is an audit report modified for going concern uncertainty.

The ten test subjects were randomly divided into two groups. Group A consisted of four of the test subjects, who completed case 1 manually and case 2 with the expert system. Group B consisted of the remaining six individuals, who completed case 2 manually and case 1 with the computer. In each situation, the manual case was completed first.

All subjects completed a pretest questionnaire concerning their "beliefs" about

expert systems. The subjects evaluated a series of statements about expert systems and indicated their agreement or disagreement with the statements. Next, each subject read the first case study about a company and was asked to indicate the type of audit report the company should receive. As they completed the case, they were asked to write down all relevant information they used to make their decision. They then were given a brief introduction and demonstration of the GC Advisor expert system, which they then used to solve the other case, again to determine the type of audit report the firm should receive. The system's summary screens were used to identify the number of areas considered for the solution. The expert system considered anywhere from nine to eleven areas during the solution process, depending on how the subject responded to the questions posed by the system.

Subjects received one point for the correct solution and one point for each statement that matched an area indicated by the experts (and hence the expert system). Since it was possible that a company receiving a standard audit report (case 1) might draw less information because there were fewer problems, we did not want that case to be the one that was always done manually, which is why the test subjects were divided into two groups. Because the test procedure administered to each group was slightly different, the results for the two groups have been reported separately. After completing the case on the expert system, the individuals evaluated the "beliefs" statements as a posttest questionnaire.

Comparison of Novices' Belief Scores

The beliefs questionnaire was scored on a five-point scale, depending on the degree of agreement with the statement. The first three questions, although only having three responses each, were scored as 1 point for the least agreement, 3 points for the neutral response, and 5 points for the statement with the most agreement.

For each group (A or B) of test subjects that completed the testing, the belief scores were compared using paired t-tests. Table 1 shows the results of these comparisons, using the scoring system described above. Question 1 asks about the amount of effort that is expected in order to learn the system. This question reflects the system's ease of use. The paired difference t-test shows that both groups A and B rated the system as requiring very little effort to learn once they had used it to solve a test case. Questions 2 and 3 refer to the effort required to update the software and the expected reliability of the system, respectively. These two questions reflect the flexibility of the system. Questions 4, 5, and 6 refer to the anticipated effect of an expert system on auditor judgment, the client's image of the auditor given that the auditor makes use of an expert system, and the effect on the auditor's self-image. Questions 7, 8, 9, and 10 as a group are auditing cost issues. These questions assess how well the system can be adapted to the auditor's use on the job. Question 7 asks a general question about the effect on the cost of auditing (increase or decrease), and question 8 refers to the anticipated effect on the need for specialists given that the auditor is using an expert system. Question 9 refers to the anticipated effect of the use of an expert system on the auditor's use of time (less efficient or

Table 1.   Test Subjects' Belief Scores Before and After the Use of the Expert System (Paired t-tests)

Question(s)	Group A average score			Group B average score		
	Before	After	t-value	Before	After	t-value
1. Ease of effort required to learn the system	3.5	5.0	3.0**	3.0	4.67	5.0***
2. Ease of effort needed to update the system's knowledge	3.5	3.5	0.0	3.0	3.67	1.6
3. Expected reliability of computer hardware and software	5.0	5.0	0.0	3.67	4.0	1.0
4. Effect on auditor judgment	3.5	3.5	0.0	4.17	4.33	1.6
5. Effect on client's image of the auditor	4.0	3.75	−0.4	4.0	4.0	0.0
6. Effect on auditor's self-image	4.25	4.25	0.0	4.0	4.33	1.6
7. Effect on cost efficiency of auditing	3.25	3.5	0.4	2.83	3.67	2.1**
8. Effect on job improvement without need for specialists	3.5	3.75	1.0	3.0	3.67	1.6
9. Effect on auditor's ability to manage time	4.25	4.5	1.0	4.0	4.17	1.0
10. Protection of auditor's personal and professional liability	3.0	2.75	−1.0	3.17	3.33	1.0

*** $p < 0.01$; ** $p < 0.05$.
Response scale: Questions 1–3: 1 = hard/low reliability; 3 = medium/moderate; 5 = easy/high reliability; Questions 4–10: 1 = reduce/diminish greatly; 2 = reduce/diminish somewhat; 3 = no effect; 4 = increase/improve somewhat; 5 = increase/improve greatly.

more efficient), and question 10 refers to the effect on personal and professional liability if an expert system is used.

Table 1 indicates that test group A scores are higher after the novices use the system, for four of the ten questions, than before they use it. However, the change in score is significant only for question 1. Test group B scores show positive improvement on all questions except question 5, where there is no change. The changes are significant for questions 1 and 7. These results suggest that the expert system is relatively easy to learn and the subject auditors believe that its use could actually reduce the cost of auditing. Based on the positive differences, hypothesis 1 is supported for four of the questions for test group A and for nine of the questions for test group B. For question 5, the scores show either diminished effect (group A) or no effect (group B) on the client's image of the auditor if the auditor uses an expert system. Overall, only one of

the belief scores taken after the novices use the system is below average. Consistent with hypothesis 1, the auditors agree to some extent that the system will have a positive effect on their job performance.

Comparison of Novices' Response Scores

The novices' response scores are shown in Table 2. The purpose of the analysis is to determine if the expert system is *at least* comparable to what a novice, or less experienced auditor, does in rendering the audit opinion. Since experts have validated the solutions provided by the system, a range of nine to eleven responses is considered acceptable. In order to test hypothesis 2, we use a paired *t*-test to compare the differences between the observed results [36]. This comparison shows that the number of responses were significantly different when a novice used the expert system compared with the novice working without the expert system.

Table 2 indicates that, on average, the novices working without the expert system use fewer responses in support of their decision than novices using the system (group A: average of 5.75 versus 10.75; group B: average of 5.67 versus 9.00). The responses of the novices using the expert system are within the acceptable range, thereby supporting hypothesis 2. The results also show that novices who do not use the expert system have fewer responses than the range of responses considered acceptable by experts. Furthermore, the novices used slightly different terminology than the expert system in phrasing their statements and responses. This is important in gauging the effectiveness and level of expertise of the expert system because, although the expert system uses slightly different terms than the novices do, these are the terms that the experts use in making their decision. In terms of accuracy, all of the test subjects responded with the "correct" type of audit report when they used the expert system. The "correct" response was a standard audit report for case 1 and an audit report modified with a going concern uncertainty for case 2.

Hybrid System Development

A HYBRID SYSTEM INTEGRATES TWO OR MORE TECHNOLOGIES IN ORDER TO OPTIMIZE each technology. The hybrid model developed here combines the expert system with the statistical model to assist in the audit opinion decision and is implemented as the GC Advisor hybrid system.

Placement of the Statistical Model in the Expert System Logic

Statistical models that predict bankruptcy arrive at a conclusion using historical information. For the auditor to make an informed decision about a company's ability to continue as a going concern, he or she evaluates how the company will obtain funds for operations. For this reason, a model that predicts bankruptcy is most useful for the evaluation of operating risk, when financial ratios are normally assessed. The statis-

Table 2.   Number of Responses Used to Render the Audit Opinion by Novices: Manual versus Expert System

	Group A			Group B		
	Manual (case 1)	Expert system (case 2)	t-value	Manual (case 1)	Expert system (case 2)	t-value
Average number	5.75	10.75	4.5***	5.67	9.0	3.4***
Standard deviation	0.96	0.87		1.21	0.0	

*** $p < 0.01$

tical model replaces the analytical procedures previously performed by the rule-based analysis of the expert system in order to evaluate operating risk.

The Hybrid Model—How DDE Works

Once we determined the best statistical model for predicting bankruptcy, it was incorporated into an Excel spreadsheet. The hybrid system accesses Excel using a feature supported by the Level–5 Expert System called Dynamic Data Exchange (DDE). Level–5 facilitates DDE by assigning a separate class for DDE actions. The first screen to appear contains a push-button to access the Excel spreadsheet. The user activates Excel and is automatically placed into the spreadsheet, which contains the prediction equation that is derived from the statistical model. The values of the variables are automatically calculated by the spreadsheet after it requests the basic information, such as Current Assets, Current Liabilities, Total Assets, and Total Liabilities. The user then must enter a "1" or a "0" for the BND variable (Bad News about Debt). The definition of items that are included as variables appear as a pop-up note in the Excel spreadsheet. The calculation formula, which is entered in the "Prediction" column of the spreadsheet, automatically recalculates a value for the prediction. The link to transfer the data is activated when the user closes the spreadsheet and is automatically transferred back to the expert system.

If the prediction value of the statistical model is greater than zero, operating results are not acceptable and the system responds with a question about whether the firm has other methods of obtaining operating funds. If no methods are available, the operating risk is considered high. If other methods are available, or if the prediction value is less than or equal to zero, then operating risk is considered low and the system continues operation as described for the expert system, progressing to assess other business conditions. The GC Advisor hybrid model applies the same type of logic to risk scoring as the expert system model. The complete program logic for the GC Advisor hybrid system is shown in figure 1.

Hybrid System Validation

THE USEFULNESS OF THE GC ADVISOR HYBRID SYSTEM is tested by hypothesis 3. The prediction accuracy, as a measure of performance, of GC Advisor is compared with the statistical model and the expert system model for the set of firms in the 1990
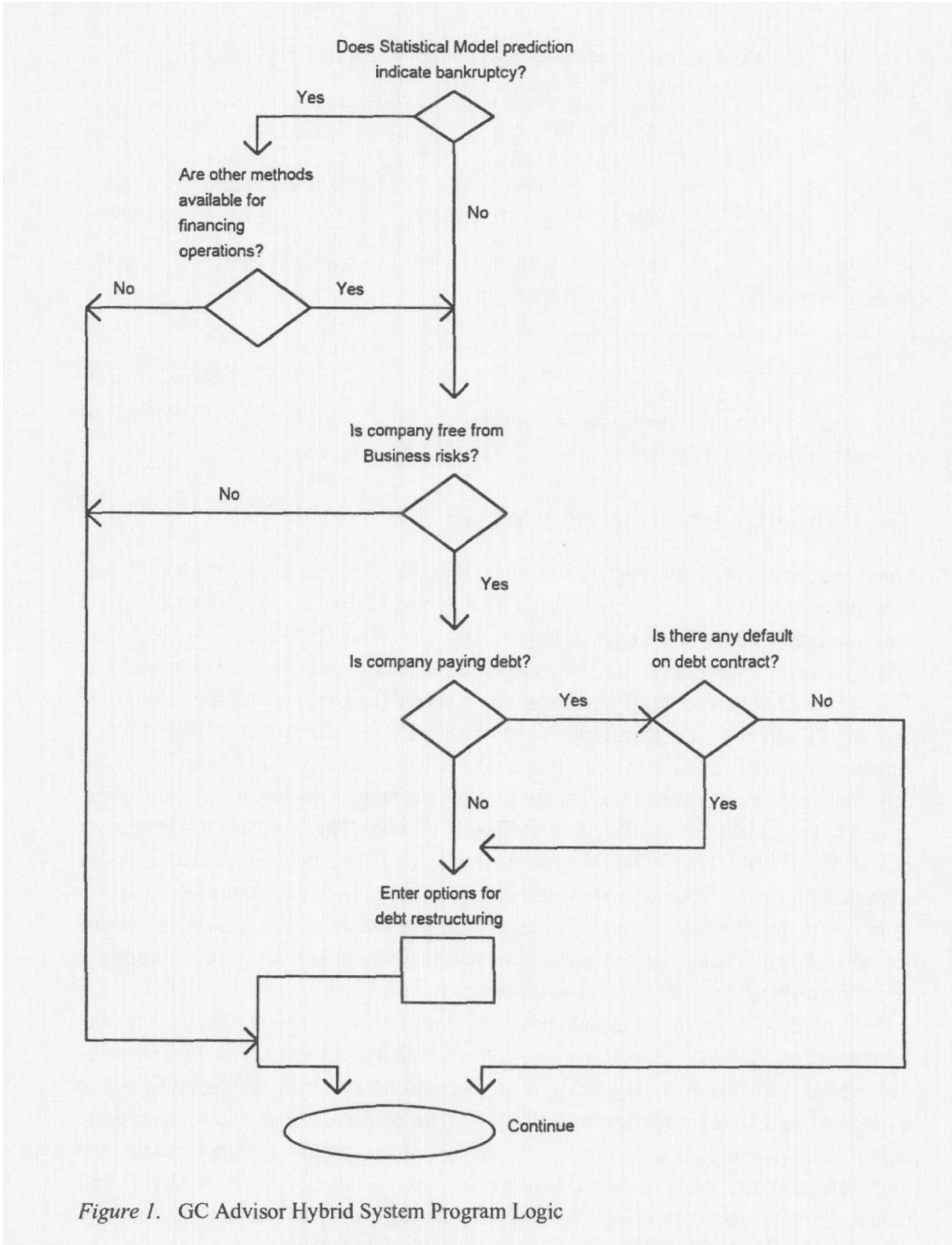
Figure 1. GC Advisor Hybrid System Program Logic

prediction sample. The reasoning of the hybrid model is the same as the reasoning of the expert system in that the model is expected to recommend an audit report modified with a going concern uncertainty for the bankrupt firms, and a standard audit report for the nonbankrupt firms.

Data for the full sample consist of firms randomly selected from the Compustat Industrial Tapes and the Compustat Industrial Research Tapes. We randomly selected
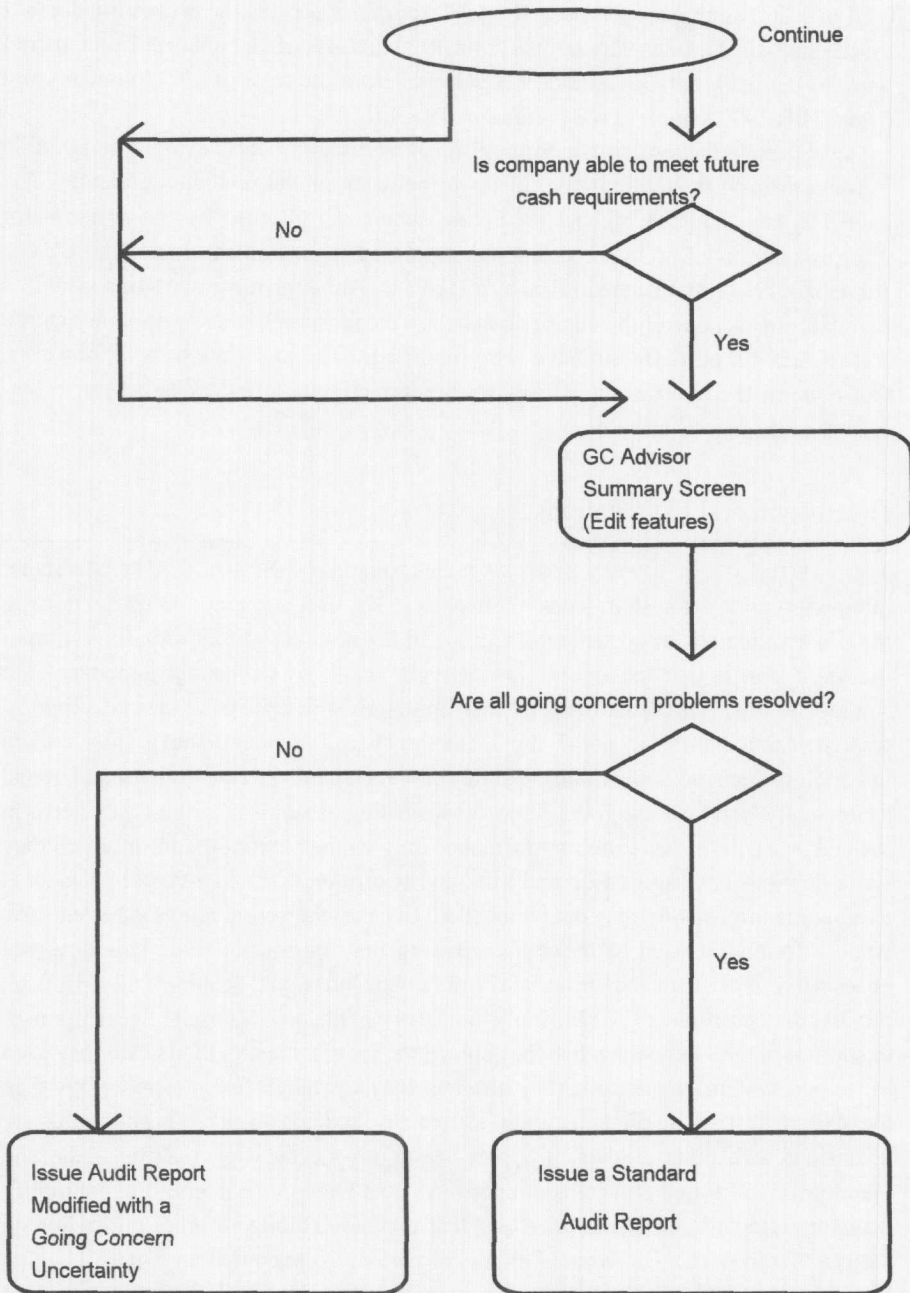
Figure 1. Continued

firms from the years 1989 and 1990, representing manufacturing companies and retail firms. The firms for each year are grouped as bankrupt (BR) or nonbankrupt (NBR). Financial information obtained for the firms described as bankrupt represents the last year of data before the firms filed for bankruptcy. The sample from 1989 consists of the 32 bankrupt firms, paired with 32 nonbankrupt firms randomly selected from the

total of 1,562 nonbankrupt firms. The 1989 sample is used as the estimation data set for the statistical model. The sample from 1990 consists of 26 bankrupt firms, paired with 26 nonbankrupt firms randomly selected from the total of 1,602 nonbankrupt firms. This 1990 sample is the prediction data set.

Table 3 shows the result of the comparison of models. In order to evaluate hypothesis 3, we compared the total number of firms that each model predicted correctly. The hybrid system had the highest prediction accuracy, followed by the M-estimator discriminant model, so hypothesis 3 was supported. The two firms missed by the GC Advisor hybrid system were missed by all models. Although these two firms exhibited financial stress, none of the other risk factors were present, so there was no indication that the audit reports should have been modified. It is interesting to note, however, that none of the nonbankrupt firms were incorrectly classified by either the expert system or the hybrid system.

Discussion and Conclusions

THE VALIDATION OF EXPERT SYSTEMS IS PERFORMED IN ORDER TO ASSESS whether interaction with the system leads to changes in the user's perception of the domain [45]. We validated our system with a group of ten novice auditors. One limitation of the validation is that the sample is relatively small, so we cannot generalize our findings to other test situations. Another limitation is that we used the test subjects' written statements to evaluate the factors they use to make the going concern uncertainty decision. The limitation of this process is similar to the limitation of verbal protocol analysis. Although the subjects are writing or summarizing as they perform the task, we can measure their performance only on the basis of the information they write down or recall. Ericsson and Simon [18] caution that when verbal protocol is used, some information may not be recalled. Because the verbal report is incomplete, some information used to make the decision may be unavailable. This does not necessarily invalidate the information obtained, but some details of the decision process may be missing [9]. Our study indicates that the novices used fewer responses in support of their decision when they solved the cases manually than when they used the expert system. In addition, they often used somewhat different terminology than the expert system in phrasing their statements and responses. These results are consistent with other studies of expert versus novice decision making. Alba and Hutchinson [2] found that inexperienced subjects were more likely to oversimplify decisions. Bedard and Mock [6] found that experts exhibited a more global search pattern than novices. Bouwman [13, 14] provided evidence that an expert relies on hypotheses and rules of thumb to guide an information search. Weber [44] noted that experts tend to cluster wide-ranging incoming information into appropriate categories, a behavior that is lacking among novices. The GC Advisor expert system provides the novice with the ability to engage in this direct, goal-driven behavior. According to Abdolmohammadi and Wright [1], it is important to recognize the difference between expert and novice behavior because it is "extremely valuable in designing decision aids and in developing staff training programs."

Table 3.   Comparison of Models: Prediction of Audit Opinion
Using 1990 Data Sample

	Total correct	Percentage correct
M-estimator discriminant model	49/52	94.2
Expert system	45/52	86.5
Hybrid system	50/52	96.2

Another purpose of the validation testing is to determine whether the subjects believe that use of the expert system will have a positive effect on the performance of their job. This outcome helps the developer decide whether a prototype should be developed into a full-scale system [5]. One finding of the beliefs questionnaire we administered was that both groups of subjects rated the system as requiring little effort to learn once they had used it to solve a test case. This finding is similar to the findings of Hansen and Messier [20], whose questionnaire formed the basis of our study. One reason for our positive findings could be that both groups tested had a fair amount of computer experience (8.75 years for group A and 9.16 years for group B). Previous research by Mackay and Elam [29] has shown that a lack of computer expertise can affect the application of the functional area of knowledge, even if the person is an expert in the functional area. Our results indicate that the novices had appropriate computer experience to learn the system quickly and were able to focus on the content of the model.

The GC Advisor hybrid system contributes to the literature in the development of expert and hybrid systems because the design using object-oriented methods allows the audit task environment to be divided into classes for more efficient programming and operation of the system. Our study also contributes to the validation literature because it extends Back's methods of assessing content validity and criterion validity to hybrid systems. We have performed preliminary evaluations to assess ease of use, system flexibility, model accuracy, and adaptability to the audit environment. In a subsequent study, a more formal evaluation will consist of field testing with a larger group of test subjects, and testing with a new or larger set of cases.

In comparison with other hybrid models, the GC Advisor reflects a similar result. Rabelo et al. [39] found, when using their model for scheduling, that the system had a higher probability of success than traditional approaches. Madey et al. [30] report that, when using their simulation model, running under the neural network's recommendations, simulated results were over 12 percent better than when running without them. Hering et al. [23] report that their combination of a series of networks resulted in a system that could make distinctions that were difficult to encode into a single equivalently sized network. Our study demonstrates that a hybrid system could be designed as a tool to assist auditor decision making.

the study. We would like to express our appreciation to J. Winchell, L. Allen, and R. Lenard for their computer assistance. Finally, we wish to thank the Editor-in-Chief and two anonymous reviewers for their helpful comments and suggestions.

REFERENCES

1. Abdolmohammadi, M.J., and Wright, A. An examination of the effects of experience and task complexity on audit judgments. *Accounting Review, 62,* 1 (January 1987), 1–14.

2. Alba, J., and Hutchinson, J. Dimensions of consumer expertise. *Journal of Consumer Research, 13,* 3 (March 1987), 411–454.

3. Anikst, M.T.; Meisel, W.S.; Newstadt, R.E.; Pirzadeh, S.S.; Schumacher, J.E.; Shinn, P.; Soares, M.C.; and Trawick, D.J. A continuous speech recognizer using two-stage encoder neural networks. *Proceedings of the Second International Joint Conference on Neural Networks,* vol. 2, Washington, DC, 1990.

4. Auditing Standards Board. *Statement on Auditing Standards No. 59: The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern.* New York: AICPA, 1988.

5. Back, B. Validating an expert system for financial statement planning. *Journal of Management Information Systems, 10,* 3 (Winter 1993–94), 157–177.

6. Bedard, J., and Mock, T.J. Expert and novice problem-solving behavior in audit planning. *Auditing: A Journal of Practice & Theory, 11,* supplement (1992), 1–20.

7. Bell, T., and Tabor, R. Empirical analysis of audit uncertainty qualifications. *Journal of Accounting Research, 29,* 2 (1991), 350–370.

8. Benjamins, R., and Jansweijer, W. Toward a competence theory of diagnosis. *IEEE Expert, 9,* 10 (1994), 43–51.

9. Biggs, S.F., and Mock, T.J. An investigation of auditor decision processes in the evaluation of internal controls and audit scope decisions. *Journal of Accounting Research, 21,* 1 (Spring 1983), 234–255.

10. Biggs, S.F., and Selfridge, M. GC-X: a prototype expert system for the auditor's going concern judgment. Working paper, University of Connecticut, 1986.

11. Booth, D.E.; Alam, P.; Ahkam, S.N.; and Osyk, B. A robust multivariate procedure for the identification of problem savings and loan institutions. *Decision Sciences, 20,* 2 (Spring 1989), 320–333.

12. Booth, D.E., and Montasser, S.H. Robust discriminant analysis and the periods of modern Egyptian economic development. *Industrial Mathematics, 35,* 1 (1985), 81–91.

13. Bouwman, M.J. The use of accounting information: expert vs. novice behavior. In G.R. Ungson and D.N. Brannstein (eds.), *Decision Making: An Interdisciplinary Inquiry.* Boston: Kent Publishing, 1982.

14. Bouwman, M.J. Expert vs. novice decision making in accounting: a summary. *Accounting, Organizations and Society, 9,* 3/4 (1984), 325–327.

15. Chen, K.C.W., and Church, B.K. Default on debt obligations and the issuance of going-concern opinions. *Auditing: A Journal of Practice & Theory, 11,* 2 (1992), 30–49.

16. Dopuch, N.; Holthausen, R.W.; and Leftwich, R.W. Predicting audit qualifications with financial and market variables. *Accounting Review, 62,* 3 (July 1987), 431–454.

17. Edwards, J.M., and Henderson-Sellers, B. Application of an object-oriented analysis and design methodology to engineering cost management. *Journal of Systems Software, 23,* 1 (1993), 123–138.

18. Ericsson, K., and Simon, H. Thinking-aloud protocols as data. *Psychological Review, 87,* 3 (May 1980), 215–251.

19. Gevins, A.S., and Morgan, N.J. Application of neural network signal processing in brain research. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 36,* 7 (July 1988), 1152–1166.

20. Hansen, J.V., and Messier, W.F., Jr. A preliminary investigation of EDP-EXPERT. *Auditing: A Journal of Practice & Theory, 6,* 1 (Fall 1986), 109–123.

21. Harris, C.R. An expert decision support system for auditor going concern evaluation. Ph.D. dissertation, University of Texas at Arlington, 1989.

22. Henderson-Sellers, B., and Edwards, J.M. The object-oriented systems life cycle. *Communications of the ACM, 33,* 9 (September 1990), 143–159.

23. Hering, D.; Khosla, P.; and Kumar, B.V.K.V. The use of modular neural networks in tactile sensing. *Proceedings of the International Joint Conference on Neural Networks*, vol. 2, Washington, DC, 1990, pp. 355–358.

24. Hu, M.; Fisher, D.M.; Booth, D.E.; and Hung, M.S. Robust discriminant analysis in marketing research: methods and applications. *Industrial Mathematics, 38*, 2 (1988), 181–196.

25. Information Builders, Inc. *LEVEL5 OBJECT for Microsoft Windows*. New York: Information Builders, 1993.

26. Lenard, M.J.; Alam, P.; Booth, D.E.; and Madey, G.R. A hybrid statistical expert system that assists the auditor with the going concern decision. Working paper, Kent State University, 1995.

27. Lenard, M.J.; Alam, P.; and Madey, G.R. The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision. *Decision Sciences, 26*, 2 (March–April 1995), 209–227.

28. Libby, R. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research, 23*, 2 (Autumn 1985), 648–667.

29. Mackay, J.M., and Elam, J.J. A comparative study of how experts and novices use a decision aid to solve problems in complex knowledge domains. *Journal of Information Systems Research, 3*, 2 (June 1992), 150–172.

30. Madey, G.R.; Weinroth, J.; and Shah, V. Hybrid intelligent systems: tools for decision making in intelligent manufacturing. In C.H. Dagli (ed.), *Artificial Neural Networks for Intelligent Manufacturing*. New York: Chapman & Hall, 1993.

31. Maren, A.J. Hybrid and complex networks. In A.J. Maren (ed.), *Handbook of Neural Computing Applications*. San Diego: Academic Press, 1990.

32. McKeown, J.C.; Mutchler, J.F.; and Hopwood, W. Towards an explanation of auditor failure to modify the audit opinions of bankrupt companies. *Auditing: A Journal of Practice & Theory, 10*, supplement (1991), 1–13.

33. Muhanna, W.A. An object-oriented framework for model management and DSS development. *Decision Support Systems, 9*, 1 (1993), 217–229.

34. Mutchler, J.F. A multivariate analysis of the auditor's going-concern opinion decision. *Journal of Accounting Research, 23*, 2 (Autumn 1985), 668–682.

35. Mutchler, J.F. Empirical evidence regarding the auditor's going-concern opinion decision. *Auditing: A Journal of Practice & Theory, 6*, 1 (Fall 1986), 148–163.

36. O'Keefe, R.M.; Balci, O.; and Smith, E.P. Validating expert system performance. *IEEE Expert, 24* (1987), 81–89.

37. O'Leary, D.E. Validation of expert systems—with applications to auditing and accounting expert systems. *Decision Sciences, 18*, 3 (Summer 1987), 468–486.

38. Rabelo, L.C.; Alptekin, S.; and Kiran, A.S. Synergy of artificial neural networks and knowledge-based expert systems for intelligent FMS scheduling. *Proceedings of the International Conference on Neural Networks*, vol. 1. New York: IEEE, 1990, pp. 359–366.

39. Rabelo, L.C., and Avula, X.J.R. Intelligent control of a robotic arm using hierarchical neural network systems. *Proceedings of the International Conference on Neural Networks*, vol. 2. New York: IEEE, 1991, pp. 747–751.

40. Randles, R.; Broffit, J.D.; Ramberg, J.S.; and Hogg, R.V. Generalized linear and quadratic discriminant functions using robust estimates. *Journal of the American Statistical Association, 73* (1978), 564–568.

41. Rossen, M.L., and Anderson, J.A. Representational issues in a neural network model of syllable recognition. *Proceedings of the International Joint Conference on Neural Networks*, vol. 1, Washington, DC, 1989, pp. 19–25.

42. Shantikumar, J.G., and Sargent, R.G. A unifying view of hybrid simulation/analytic models and modeling. *Operations Research, 31*, 6 (November–December 1983), 1030–1052.

43. Teach, R.L., and Shortliffe, E.H. An analysis of physicians' attitudes regarding computer-based clinical consultation systems. *Computers and Biomedical Research, 14* (1981), 542–558.

44. Weber, R. Some characteristics of the free recall of computer controls by EDP auditors. *Journal of Accounting Research, 18*, 1 (Spring 1980), 214–240.

45. Wensley, A. Expert systems and audit planning: evolving research issues. *Expert Systems with Applications, 5*, 1 (1992), 351–357.

46. Zavgren, C.V. Assessing the vulnerability to failure of American industrial firms: a logistic analysis. *Journal of Business Finance & Accounting, 12*, 1 (Spring 1985), 19–45.