Pergamon

Expert Systems With Applications, Vol. 7, No. 2, pp. 305–321, 1994

0957-4174/94 $6.00 + .00

The Development and Validation of a Campus Recruiting Expert System Using Expert Opinions and Historical Data

MADJID TAVANA, PATRICK LEE, AND PRAFULLA JOGLEKAR

La Salle University, Philadelphia, PA

Abstract—*This article describes the development and validation of an expert system used for screening entry level candidates for employment in an accounting firm. The system is designed to assist a recruiter conducting campus interviews in deciding which candidates to call for a second interview at company premises. Such expert systems have the potential for improving consistency and efficiency of the many human recruiters' decisions, and for minimizing personal biases or the use of unlawful criteria (e.g., race, gender, or age) in these decisions. Whereas a hypothetical expert system for this recruiting problem has been previously described in the literature, we actually built one. We describe how we built the necessary knowledge base and how and why our inference engine calculates the probabilities of acceptance, hold, and rejection based on a combination of expert opinions and historical data base. The system is validated by comparing it against recent case-by-case decisions of human experts. We recognize the limitations of our study and discuss the lessons learned from this development effort.*

1. INTRODUCTION

INFORMATION SYSTEMS for the analysis and statistical reporting of data in such areas as hiring, performance appraisal, and salary administration have been widely used for several decades (DeSanctis, 1986; Linder, 1985; Perry, More, & Parkinson, 1987; Simon, 1983; Walker, 1986). It has been suggested that one reason for this wide use is the many government regulations requiring documentation of compliance with federal and state laws, including The Civil Rights Act, The Employee Retirement Income Security Act, Occupational Health and Safety Act, and The Toxic Substances Act [Staff, 1989]. In any case, today most companies possess rich and comprehensive computerized personnel data bases. Yet, these data bases are rarely used for actually monitoring and improving the decision-making processes themselves in the human resources area.

In recent years, with the availability of an increasing number of expert system "shells," applications of expert systems are being proposed to help improve management decision making in a variety of areas. Potential benefits of these systems include: more consistent decisions, efficiency, operational cost savings, better utilization of human resources, and easier access to rare or dispersed knowledge (Liebowitz, 1990; Lin, 1986). In the field of human resource management, it has been proposed that expert systems would be useful in such decisions as performance appraisal, hiring, and training (Briggs & Doney, 1989; Extejt & Lynn, 1988; Krebs 1988; Whaley, 1989). However, at present, most of these proposals are reported only on a conceptual level, and actual implementation of expert systems in this field remains largely undocumented. In this article we describe the development and validation of an expert system used for screening entry level candidates for employment in an accounting firm.

Our system builds on Extejt and Lynn's (1988) description of a hypothetical expert system for the campus recruiting problem of a bank. Extejt and Lynn (1988) explain the many potential advantages of such a system, including improving the consistency and efficiency of the many human recruiters' decisions, and minimizing their personal biases or the use of unlawful criteria (e.g., race, gender, or age) in these decisions.

Extejt and Lynn (1988) emphasize that such an expert system should not be used as a decision maker. Instead, it must be used as a decision aid, and campus recruiters must be allowed to make their own final decisions using certain information not captured by the system. This is important because:

1. No expert system can incorporate all relevant information.
2. In a dynamic environment, a company's decision criteria undergo a change over a period of time.
3. The need to update an expert system may become evident only when there is a pattern of conflict be-

Requests for reprints should be sent to Madjid Tavana, Management Department, La Salle University, Philadelphia, PA 19141.

tween the system's recommendations and the human recruiters' decisions.

With that philosophy, Extejt and Lynn (1988) visualize only a probabilistic recommendation from the system, with the final decision being left to the recruiter's discretion. Thus, the expert system is not to provide a recommendation on the "best decision," but it is simply to indicate the likelihood of each of the choices (e.g., accept, hold, or reject) being "right" in the sense that a human expert would have made the same choice. We believe that such probabilistic estimates encourage the recruiters to be responsible for their decisions, rather than feeling like powerless clerks who simply follow the computer's directives.

Finally, Extejt and Lynn (1988) provide a scenario for developing such a system by using expert opinions. As they visualize it, the personnel manager of their bank could identify some three "experts" in recruiting and query them regarding their decision processes to arrive at, albeit after much discussion, the "decision rules" the expert system could use. Extejt and Lynn (1988) speculate that based on these rules, for a sample candidate, the expert system may report the probabilities of alternative decisions being correct as:

A. Schedule second interview . . . 60% chance of correct decision
B. Place resume in "hold file" . . . 70% chance of correct decision
C. Reject candidate . . . 20% chance of correct decision (p. 14).

In Extejt and Lynn's (1988) scenario, the campus recruiter is to take these results, supplement them with his or her additional information about the candidate from the campus interview, and arrive at the final decision.

We did have some misapprehensions about Extejt and Lynn's (1988) framework. First, Extejt and Lynn (1988) never quite define what they mean by a "correct decision." One could interpret it many different ways. For example, one definition could be that an accept decision is correct only if the candidate is eventually hired and actually turns out to be a highly productive member. Another possible definition is that an accept decision is correct if the candidate is recommended for hire by the team of interviewers on company premises. However, such definitions would require data and analysis far beyond the resources we had. Therefore, we settled for a definition whereby a screening decision is "correct" if a recognized human expert would have arrived precisely at that decision. With this definition, it does make sense to talk about a probability of a decision being correct because different human experts (each, equally recognized) may arrive at different decisions given the same set of candidate qualifications.

Second, we could not logically accept Extejt and Lynn's (1988) suggestion, in the above quotation, that in a set of mutually exclusive and collectively exhaustive choices such as accept, hold, and reject, the chances of the various decisions being correct could add up to greater than unity. However, the rest of Extejt and Lynn's (1988) proposal for an expert system for the college recruiting problem seemed reasonable. We also had a client who was willing to help us develop such an expert system. Thus, we embarked on this effort to explore the feasibility and usefulness of the expert system development scenario proposed by Extejt and Lynn (1988).

We discovered that although the experts in our study agreed with Extejt and Lynn's (1988) notion of probabilities of certain decisions being correct, they found it difficult to provide such probabilistic decision rules. Our experts were more comfortable in indicating their "most likely decision" about a candidate with given attributes. Our experts also questioned, as we did, Extejt and Lynn's (1988) suggestion that the probabilities of mutually exclusive and collectively exhaustive decision alternatives being right need not add up to 1.00. Finally, in our efforts at validating our expert system, we found that general decision rules described by a group of experts did not match with the frequency of actual case-by-case decisions by another group of equally competent experts. Consequently, we could not rely on expert opinions alone in the development of our knowledge base, and we had to integrate the expert opinions with the historical data base. In short, we believe that our study brings many important insights to the process of developing expert systems.

In what follows, first we describe our problem setting. We then describe how we used expert opinions to build the knowledge base and the inference tree for our expert system. Next, our efforts in validating the system, and the resultant "improvised" system that combined expert opinions with the historical data base are explained. We conclude with a recognition of some of the limitations of our study and a discussion of the lessons of our development efforts.

2. THE PROBLEM SETTING

As Extejt and Lynn (1988) point out:

> Campus recruiting is an expensive, labor intensive activity conducted by numerous corporations. Often the persons who conduct the on-campus interviews are line personnel (engineers, accountants, sales managers, etc.) who are given minimal training in making selection decisions. Within one firm the numerous persons who conduct on-campus interviews may use some common decision rules, but each may introduce his/her own preferences, biases and rules of thumb. The result might be a set of candidates whose qualifications, aptitudes and interests are somewhat inconsistent. Using an expert system should help standardize this decision process (p. 12).

Armed with Extejt and Lynn's (1988) article, we approached a Philadelphia-based accounting firm and offered to build an expert system for their campus recruiting problem. Because our purpose was to learn (and help our students learn), there was no consulting fee. All we asked for was the firm's cooperation. Under those terms and with a promise of confidentiality, we had no trouble obtaining the requisite cooperation—particularly because one of our MBA students was an Assistant Director of recruiting at the home office of the firm.

Our client had a branch office in New York, and together the Philadelphia and New York offices conducted some 450–550 campus interviews per year, asking some 120–170 of those for interviews on company premises (at times, calling some that were initially put in the "Hold" category), ultimately offering jobs to 40–60 of the applicants, of which 25–35 accepted and were actually hired each year. Note that the company considered all historical personnel data to be confidential and provided us with only the data on the on-campus interviews after deleting the names of individual candidates. For all other pieces of information, the Director of Recruiting provided us with his best-guess estimates. Thus, we did not have the actual year-by-year statistics on how many candidates came for interviews on company premises, how many were actually hired, and so forth. Company growth, economy, employee turnover, and many factors specific to the individuals involved influenced these statistics from year to year.

In his decision to cooperate with us, one of the stated objectives of the company's Director of Recruiting was to reduce the number of candidates called for an interview on company premises, without reducing the "yield," that is, the number actually hired in a year. We warned that, given the lack of accurate historical statistics, this result may be difficult to prove. However, we agreed that this was a reasonable expectation for our expert system, and the project was commissioned.

Our first step was to understand the existing campus interview process and its inputs and outputs. Figure 1 documents this understanding. As can be seen, in the current system, a recruiter obtains many different pieces of information on the candidate at the time of a campus interview. Some of these pieces of information (e.g., degree type, degree level, GPA, etc., represented by bold circles in Fig. 1) are obtained by every recruiter for each candidate he or she is considering. However, some pieces of information sought (e.g., the reputation of the candidate's school, the candidate's appearance, age, sex, sexual preference, etc., represented by dotted circles) vary from recruiter to recruiter. It may be worth noting here that some of these factors are illegal to consider in the recruiting process, and others may be against company policy. On the other hand, skillful recruiters can pick up insightful pieces of information about a candidate's criminal record, mental stability, ethics, ability to communicate, and so forth that can help find the best employees and avoid future liabilities to the company.[1] Based on these pieces of information, for each candidate, a recruiter decides to "accept" (i.e., call the candidate for an interview on company premises), "hold" the application (for consideration at a future time), or "reject" (i.e., not to give any further consideration to the candidate). Overall, in the opinion of the Director of Recruiting, the current process and criteria were often arbitrary and inconsistent. Furthermore, these criteria were different from year to year because company policies had changed considerably over the past few years. Thus, the Director of Recruiting indicated that an expert system that simply imitates past decisions would not suffice.

Of course, some recruiters are more experienced and are considered "recruiting experts," and their decisions are almost always final. They also tend to put a very small percentage of the applicants into the "hold" category. Other not-so-experienced recruiters often consult these recruiting experts when in doubt. Consequently, almost all of the initial "hold" decisions of the recruiters are reviewed by the recruiting experts before a recruiter makes a final decision, which may still be either "accept," "hold," or "reject." For all practical purposes (98% of cases), a final "hold" is equivalent to a "reject" decision. Only in approximately 2% of the cases may a candidate on hold actually be called for an on-site interview. In Figure 1, this is indicated by the dotted loop from hold back to the screening decision.

It is this campus recruiting process that our expert system was intended to assist. In addition to the many advantages of such an expert system outlined by Extejt and Lynn (1988), it was hoped that with an expert system, the recruiters would be able to make a definitive final decision in a larger number of cases without consulting with the recruiting experts as often as they did under the existing system. Once again, we warned that this result may be difficult to document because the number of consultations were not tracked in the existing system.

Although the expert system is not intended to alter any portion of the process following the campus screening decision described above, it is important to understand what that process is. As Figure 1 depicts, when a candidate is "accepted" for an interview on the company premises, if the candidate chooses to come, he or she is interviewed by a team consisting of a recruiting expert and several relevant managers and potential colleagues. All the members of this team have

[1] As Extejt and Lynn (1988) suggest, one benefit of an expert system is to help minimize the use of illegal pieces of information while encouraging continued use of insightful pieces of information that may not be captured by the expert system.

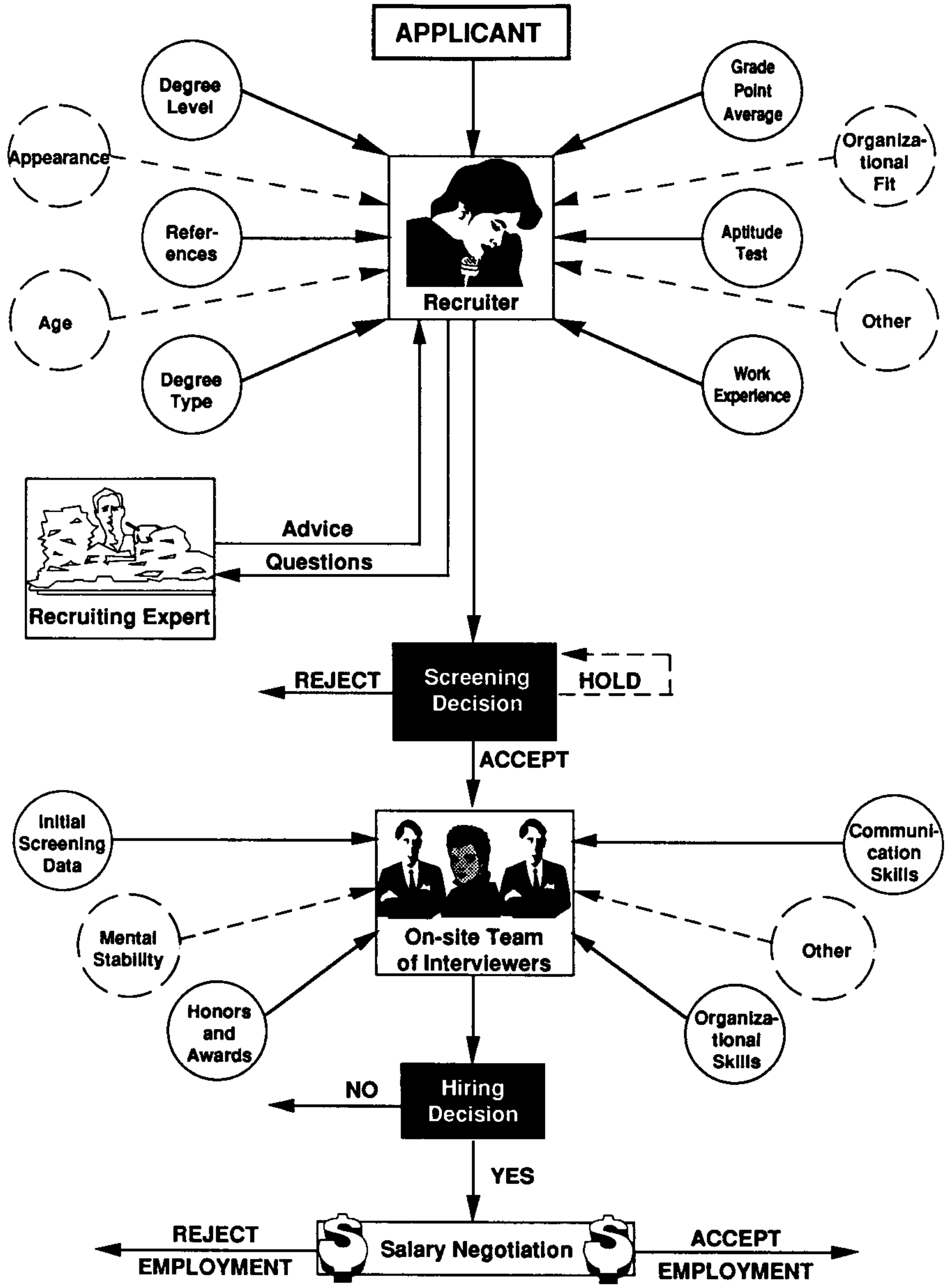

FIGURE 1. The existing recruiting process.

full access to the data and recommendation from the campus screening process, and an opportunity to talk to the candidate and obtain whatever pertinent information they want (e.g., appearance, ethics, mental stability, age, sex, etc.). The team then makes a judgment on such issues as the candidate's "fit" with the company and makes a recommendation on whether the candidate should be hired. One of the partners of the firm then negotiates the position and salary with the candidate. Of course, sometimes the company's offer is not acceptable to the candidate, and the process of negotiation is terminated. Most of the time, the candidate with a positive recommendation is hired. The "yield" that our recruiting director wanted to maximize is the finally hired candidates as a percent of candidates accepted for on-site interview as a result of the campus screening. The company agreed, however, that we would not be able to prove this, given the fact that we

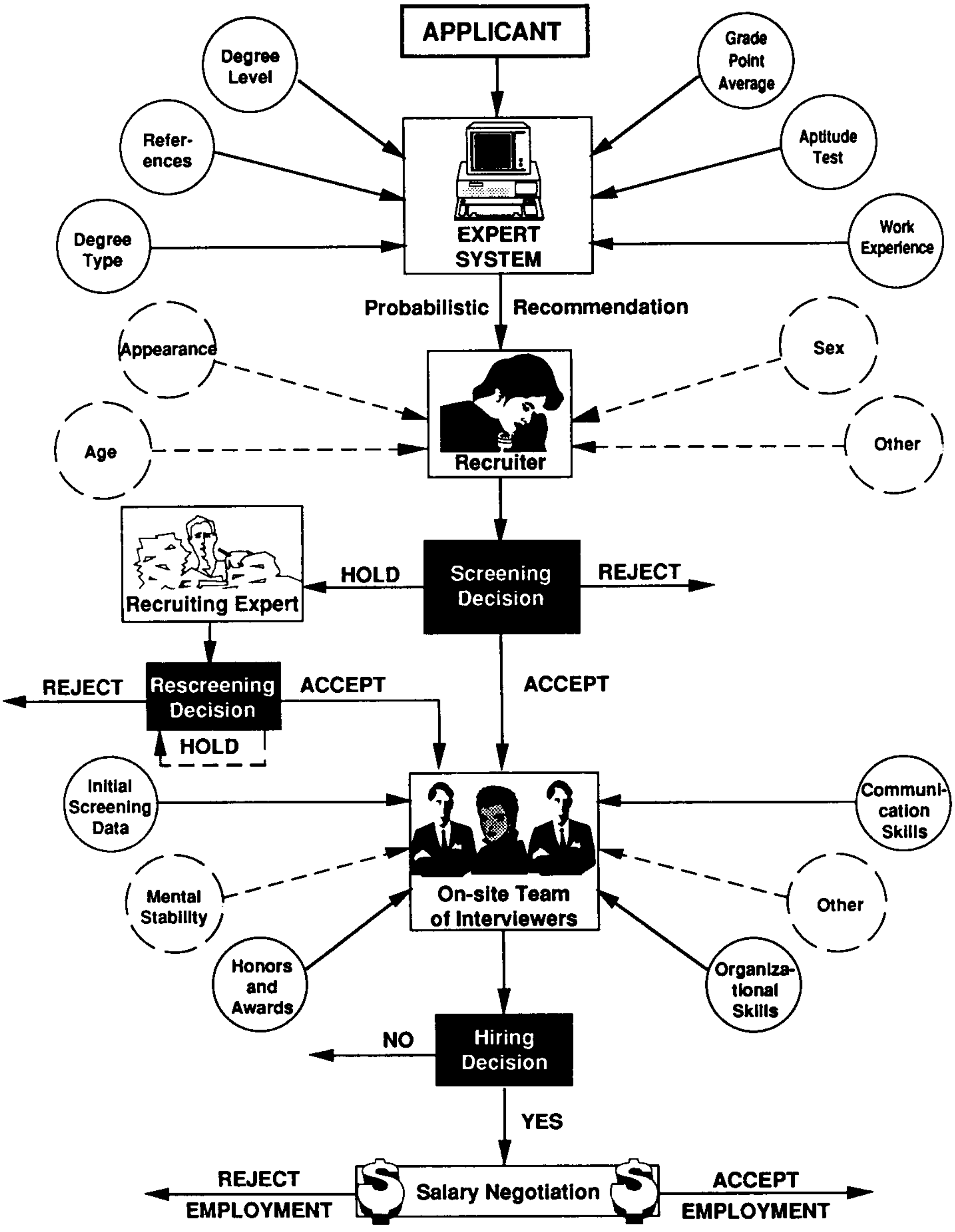

FIGURE 2. Computer-aided recruiting process.

had only a limited access to the historical data. The company was to make an internal determination of the effectiveness of our expert system.

Figure 2 depicts our visualization of where the expert system fits in the screening process. Based on the most commonly used pieces of information, and rules defined by a team of recruiting experts, the system is to provide an initial estimation of the probability of "accept," "hold," and "reject" to the recruiter. For a discussion of how these commonly used pieces of information and rules (called the knowledge base) were identified, and what they are, see the next section. Given our philosophy that an expert system is not to be a substitute, but only an aid, for human decision making, the recruiter is still free to use other pieces of information to make his or her final decision about a candidate. Only when the recruiter's final decision about a candidate is "hold" is a recruiting expert to review the case. The rest of the recruiting process remains exactly identical to the process in Figure 1.

An expert system consists of three major components:

1. its Dialog Structure, which serves as the language interface with the user
2. its Knowledge Base, which is domain-specific and incorporates the experts' knowledge about what information to obtain and how to use it
3. its Inference Engine, which allows hypotheses to be generated and tested so that recommendations based on the knowledge base can be made in specific cases of application of the system (Liebowitz, 1988).

In this article, it will suffice to say that we used the expert system shell named VP-Expert to create a user-friendly dialog structure[2] and VP-Expert's inference engine to derive the recommendations based on our knowledge base. Here, our focus is on describing how we arrived at our knowledge base, what it is, and how we validated it.

3. THE KNOWLEDGE BASE

Having agreed with the Director of Recruiting on the scope and the purposes of the expert system, we initiated the knowledge acquisition process. We requested the company to identify for us several of their "recruiting experts." Accordingly, a committee of five partners of the firm identified a total of nine recruiting experts whose experience and judgment they trusted. Five experts were from the Philadelphia office and four were from the New York office. Because the Director of Recruiting had indicated that an expert system must not merely imitate historical decisions in the company, we decided to use the opinions of the experts from the Philadelphia office to develop our knowledge base (i.e., the rules defining the average probabilities of acceptance, hold, and rejection) and the case-by-case judgments of the experts from the New York office to validate that knowledge base. We wanted these groups to be separate because those defining the rules may be biased toward adhering to the rules even when individual cases call for exceptions. (For a clearer understanding of this division of experts, see the validation section.)

[2] The user interface of our system controls the screen display used to communicate with the recruiter. The screen layout consists of two active horizontal windows. The top window displays the questions and answers, and the lower window presents the system's probabilistic recommendations. At all times, a recruiter has access to two pull-up windows "WHY" and "HOW." Using the WHY window, a recruiter can find out why the system is asking a particular question. The HOW window shows how the system is arriving at certain recommendations. Printed reports generated by the system provide summary statistics about individual recruiters' decisions, number of applicants and acceptance by universities, and year-to-date totals of these types of statistics for the entire company.

Verbal, unstructured interviews constitute the most frequently reported knowledge acquisition approach (Olson & Rueter, 1987; Wright & Ayton, 1987). Accordingly, we interviewed the Philadelphia-based experts on the current, as well as the ideal, process and criteria for screening candidates for entry level accounting positions. The experts' primary complaint about the current process was that during the busy recruiting season they lacked the time. This often led to hasty decisions. They all believed that their time could be better utilized if routine tasks could be assigned to or shared with nonexperts, and if the recruiters did not have to consult with the experts so often. Finally, the experts also noted that, in campus screening, often the criteria used by the nonexpert recruiters were arbitrary, inconsistent, and at times, illegal. On the other hand, the experts themselves did not seem to fully agree on the criteria that should be used, or on the relative rankings of those criteria. As a result of our unstructured interviews, we identified 14 criteria as the most frequently mentioned criteria (i.e., mentioned by at least two of the Philadelphia experts).

Next, we held several Delphi rounds in order to narrow down this list of the ideal criteria to be used in campus screening. Delphi Technique was devised by a research group at the Rand Corporation to obtain the most reliable consensus of opinion from a group of experts about an issue or a problem. A basic premise of the Delphi method is that if the opinion of one expert on a certain point is good, then the combined opinion of several experts will be even better. According to Dalkey (1969), there are three features of Delphi:

1. anonymity (effected by the use of formal communication channels such as questionnaires), which helps reduce the effect of dominant individuals
2. statistical analysis of the group response, which helps assure the representation of all the members' opinions
3. repeated trials (following the anonymous, statistical feedback), which help increase objectivity and promote consensus

Using the list of the 14 most frequently mentioned criteria, we asked the five experts to rank order them in terms of their importance, and to write brief statements of rationale for their individual rankings. The results of the first round of the Delphi rankings are presented in Table 1.

As can be seen from Table 1, the rankings were fairly close to one another for a few of the 14 criteria, but quite divergent for others. By summarizing the ranking statistics (i.e., the range, mean, and mode for each of the 14 criteria) as well as the anonymous responses about the underlying rationale in each round of the Delphi survey, and asking the experts to revise their rankings and rationales in the next round, we were able to see a reasonable consensus emerge about the top 7 criteria after five rounds (see Table 2). In fact,

TABLE 1
Initial Ranking of 14 Most Frequently Mentioned Attributes

	Ranking Assigned By						
Attribute	Director	Assistant Director	Recruiting Specialist A	Recruiting Specialist B	Recruiting Specialist C	Mean Ranking	Overall Rank
Appearance	12	7	11	12	6	9.6	10
Aptitude test	4	4	5	1	3	3.4	4
Communication skills	11	9	8	9	8	9.0	8
Degree level	3	3	4	3	2	3.0	3
Degree type	2	1	1	4	1	1.8	1
Grade point average	7	5	2	6	5	5.0	5
Honors and awards	9	8	10	10	10	9.4	9
Organizational fit	5	13	14	8	13	10.6	11
Organizational skills	13	11	12	14	11	12.2	13
Problem-solving skills	14	12	9	11	12	11.6	12
References	1	2	3	2	4	2.4	2
Secondary field of study	8	10	6	7	9	8.0	7
Technical skills	10	14	13	13	14	12.8	14
Work experience	6	6	7	5	7	6.2	6

there were no changes in the individual rankings of the top 7 criteria from Round #4 to Round #5. Note that although Tables 1 and 2 identify each of the experts with his or her rank orders, during the Delphi rounds, complete anonymity was maintained, and the experts were provided with only a statistical summary of the group's responses.

To make sure that our very first attempt at the development of an expert system was a manageable one, we wanted to include no more than six or seven of the most important criteria in that system. Based on Table 2, we concluded that Degree Type, References, Degree Level, Grade Point Average, Aptitude Test, and Work Experience were the six most important criteria to be examined during the campus screening. However, when we asked the experts exactly how to categorize the various possibilities of a candidate's secondary field of study, there was little agreement. Some experts valued science as an important secondary field; others thought English and communication was more important. Still other experts felt that a graduate degree made the secondary field of study irrelevant. Finally, the Recruiting Director suggested that what the company needs is a team of accountants with a balanced variety of secondary fields of specialization. Thus, inclusion of this criterion seemed problematic for the expert system, and it was deemed to be a criterion best handled by the team of on-site interviewers.

TABLE 2
Final Ranking of 14 Most Frequently Mentioned Attributes

	Ranking Assigned By						
Attribute	Director	Assistant Director	Recruiting Specialist A	Recruiting Specialist B	Recruiting Specialist C	Mean Ranking	Overall Rank
Appearance	12	7	11	10	7	9.4	9
Aptitude test	4	4	5	4	4	4.2	4
Communication skills	10	11	8	11	9	9.8	10
Degree level	3	3	3	3	2	2.8	3
Degree type	2	1	1	2	1	1.4	1
Grade point average	6	5	4	6	5	5.2	5
Honors and awards	9	8	9	8	10	8.8	8
Organizational fit	8	12	12	12	13	11.4	12
Organizational skills	13	14	13	14	14	13.6	14
Problem-solving skills	14	10	10	9	11	10.8	11
References	1	2	2	1	3	1.8	2
Secondary field of study	7	9	7	7	8	7.6	7
Technical skills	11	13	14	13	12	12.6	13
Work experience	5	6	6	5	6	5.6	6

In a meeting following the Delphi rounds, the experts also agreed that a candidate's qualifications in terms of the remaining criteria, including those that did not make our list (e.g., the reputation of the candidate's school, the candidate's mental stability, ethics, etc.), could be better evaluated by the team of interviewers on company premises, although the campus recruiters were also free to use any of them in making their screening decisions. Thus, it was agreed that the six criteria: Degree Type, References, Degree Level, Grade Point Average, Aptitude Test, and Work Experience, were the only ones to be included in the knowledge base of our expert system (see Table 3). During this meeting, we were also able to identify the inference tree (see Fig. 3) to be used in our expert system.

The first criterion is whether a candidate is an accounting major or not (Degree Type). An entry level employee must pass the CPA and/or the CMA exam within a designated period of time after employment. These exams have a prerequisite of specific undergraduate or graduate course work in accounting. Practically speaking, a degree in accounting is a must for the CPA/CMA certification. Hence, it is necessary to eliminate any candidates without an accounting degree.

The second criterion is the candidate's references. In our client's system, those submitting references are asked to provide some insights about their interaction with the candidate, along with an overall rating of the candidate on the scale of 1 (Poor) to 4 (Excellent). These ratings are then used to calculate the cumulative score. Three references are required from each applicant. The cumulative score, the sum of the three ratings, may range between 3 and 12. The firm's policy is to reject candidates with an unfavorable cumulative score (7 or less). There are no exceptions to this policy. Hence that is the second criterion on our inference tree (Fig. 3).

TABLE 3
Classification of Screening and Hiring Attributes

Screening attributes used by the expert system
Degree type
References
Degree level
Aptitude test
Grade point average
Work experience
Screening attributes used by the recruiter
All of the above and others
Screening attributes used by the on-site team of interviewers
All of the above, plus
Secondary field of study
Honors and awards
Appearance
Communication skills
Problem-solving skills
Organizational fit
Technical skills
Organizational skills
And others

The third step in our inference tree is to classify candidates according to their degrees (Degree Level). Candidates with graduate degrees are more likely to meet certification requirements than those with undergraduate degrees only. Consequently, our expert system considers these two groups of candidates separately in deciding about their probabilities of acceptance, hold, and rejection.

Within each of these two groups, according to our experts, three criteria should determine whether a candidate is put in the "accept," "hold," or "reject" category:

1. grade point average (GPA)
2. aptitude test score
3. work experience

In our client firm, the GPA classification ranges for a candidate with undergraduate degrees are: A (3.50–4.00), B (3.00–3.49), C (2.50–2.99), and D (2.00–2.49).[3] For candidates with graduate degrees, the GPA classification ranges are: A (3.75–4.00), B (3.50–3.74), C (3.25–3.49), and D (3.00–3.24).

The second classification pertains to the candidate's score on an aptitude test conducted by the firm prior to the campus interview. These test scores are reported as E (excellent), G (good), A (average), and F (fail).

The last classification deals with a candidate's work experience. Due to the high cost of training, prior experience of a candidate for an entry level position is a benefit to an accounting firm. Because candidates with 3 or more years of experience seem to dramatically reduce the training required, they would be considered more favorably. Our experts agreed to classify experience into four categories of M (3 or more years of related experience), R (less than 3 years of related experience), U (unrelated work experience), and N (no work experience).

Thus, for a candidate with an undergraduate degree in accounting, and favorable references, there can be 64 possible combinations of qualifications in terms of his or her grade point average, aptitude test, and work experience. Given the Extejt and Lynn (1988) work, we planned to ask each one of our experts to indicate his or her probability of acceptance, hold, and rejection of a candidate with each possible combination of qualifications. Accordingly, we shared a summary of related paragraphs from Extejt and Lynn (1988) with our experts and asked their reactions to this scheme.

As we have already indicated in the introduction section, our experts agreed with Extejt and Lynn's

[3] Because most schools do not graduate individuals with a grade point average less than 2.00, the default of rejecting applicants with a less than 2.00 average is redundant.

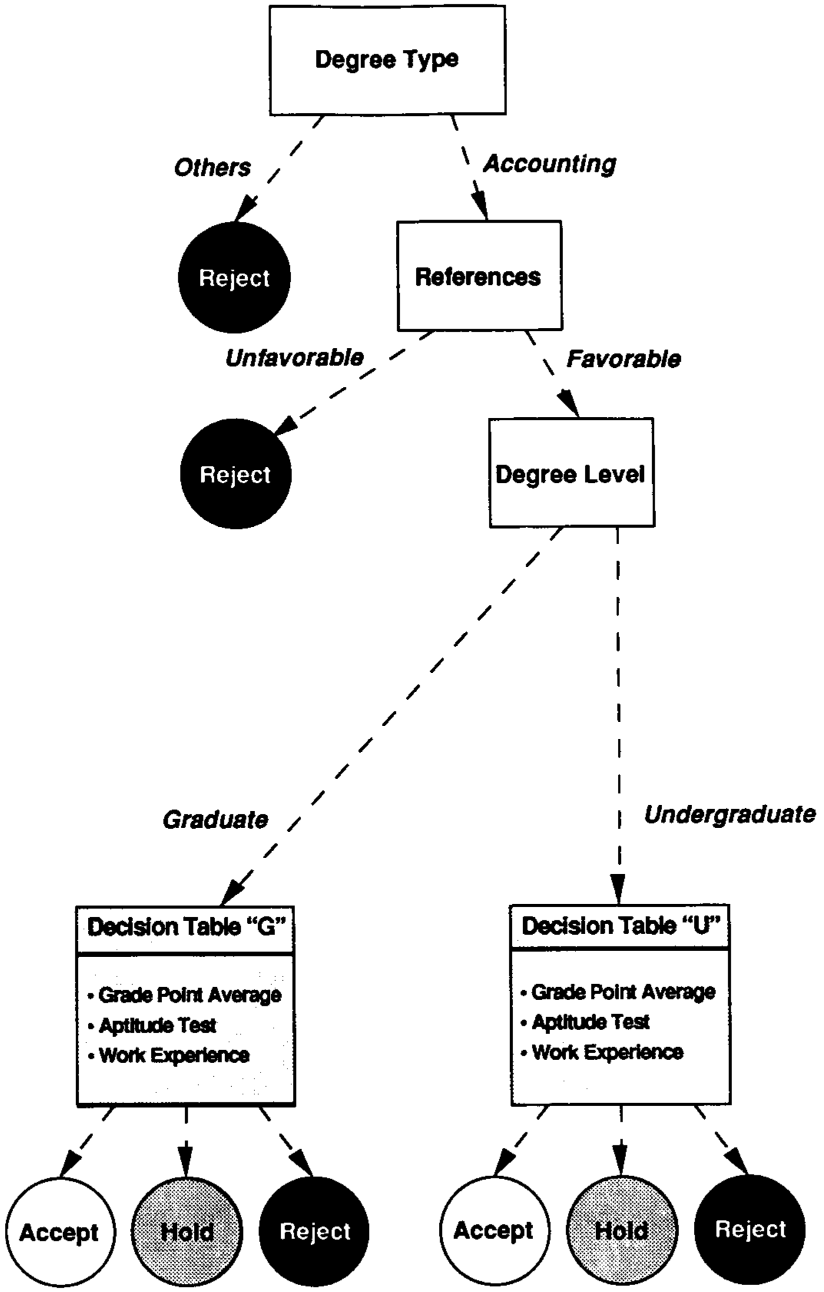

FIGURE 3. Inference tree.

(1988) notion that an expert system is to be an aid rather than a substitute for human decision making, and saw the need for providing the recruiters with only probabilities of certain decisions being the choice of an expert, and not some definitive decisions. However, our experts questioned Extejt and Lynn's (1988) suggestion that the probabilities of mutually exclusive and collectively exhaustive decision alternatives being right need not add up to 1.00. Instead they required that in our system, such probabilities must add up to 1.00.

Individually, our experts also found it impossible to estimate their own probabilities of acceptance, hold, and rejection of candidates with each one of the possible combinations of qualifications. Instead, they were willing to indicate only what their most likely decision would be in each case. Therefore, using a questionnaire, we asked each one of our five Philadelphia experts to check mark whether he or she would accept, put on hold, or reject a candidate with each possible combination of qualifications. The probabilities were calculated using the frequencies of the experts' most likely decisions about particular combinations of qualifications. For example, considering an undergraduate degree holder with a grade point average of between 3.00 and 3.49, an average aptitude test score, and less than 3 years of related work experience, three of the five experts recommended acceptance (60%), one recommended hold (20%), and one recommended rejection (20%).

It is obvious that if we had a larger number of experts providing these judgments, our calculated probabilities would have been far more reliable. With only five experts involved, if one expert changes his or her mind, our probabilities can change by 0.20. Unfortunately, we just did not have a larger number of recognized experts, because we wanted the remaining experts to provide their decisions on the validating test cases. Table 4 presents the complete expert–opinion-based decision table for a candidate holding an undergraduate degree.

The five experts were also asked to respond to a questionnaire on a candidate with a graduate degree. Table 5 shows the expert–opinion-based decision table for the 64 possible combinations of qualifications of a candidate with a graduate degree.

The final form of knowledge acquisition entailed a detailed examination of 3,864 historical campus interview records of the firm over the last 8 years. Although we were told that our expert system must not simply mimic the historical decisions (which were deemed to be arbitrary, inconsistent, and outdated), because we had the access to this rich data base, wherever possible, we did calculate the historical probabilities of acceptance, hold, and rejection for each one of the 128 possible combinations of qualifications (64 for the undergraduate degree holders, and 64 for the graduate degree holders). Tables 6 and 7 present these historical probabilities for a candidate with an undergraduate and a graduate degree, respectively. As can be noted, for certain combinations (e.g., #G7 in Table 7), we did not have an adequate sample size (a minimum of 5 case histories) to report the probabilities, and they are simply listed as "NA."

A comparison of Tables 4 and 6, as well as Tables 5 and 7, shows that although overall the results were somewhat similar, there were certain discrepancies and, at times, significant departures between the two sets of probabilities. For example, the probabilities of accep-

TABLE 4
Expert–Opinion-Based Decision Table (Applicants Holding an Undergraduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
U01	A	E	M	100	0	0	U33	C	E	M	80	20	0
U02	A	E	R	100	0	0	U34	C	E	R	80	20	0
U03	A	E	U	100	0	0	U35	C	E	U	60	40	0
U04	A	E	N	100	0	0	U36	C	E	N	60	40	0
U05	A	G	M	80	20	0	U37	C	G	M	60	40	0
U06	A	G	R	80	20	0	U38	C	G	R	60	40	0
U07	A	G	U	60	20	20	U39	C	G	U	60	20	20
U08	A	G	N	60	20	20	U40	C	G	N	60	20	20
U09	A	A	M	80	20	0	U41	C	A	M	60	20	20
U10	A	A	R	80	20	0	U42	C	A	R	60	20	20
U11	A	A	U	40	20	40	U43	C	A	U	40	0	60
U12	A	A	N	20	20	60	U44	C	A	N	20	0	80
U13	A	F	M	20	20	60	U45	C	F	M	20	0	80
U14	A	F	R	20	0	80	U46	C	F	R	0	0	100
U15	A	F	U	0	20	80	U47	C	F	U	0	0	100
U16	A	F	N	0	0	100	U48	C	F	N	0	0	100
U17	B	E	M	100	0	0	U49	D	E	M	80	20	0
U18	B	E	R	80	20	0	U50	D	E	R	60	40	0
U19	B	E	U	80	20	0	U51	D	E	U	60	40	0
U20	B	E	N	60	20	20	U52	D	E	N	60	40	0
U21	B	G	M	80	20	0	U53	D	G	M	60	40	0
U22	B	G	R	80	20	0	U54	D	G	R	60	40	0
U23	B	G	U	60	20	20	U55	D	G	U	40	20	40
U24	B	G	N	60	20	20	U56	D	G	N	40	20	40
U25	B	A	M	80	20	0	U57	D	A	M	60	0	40
U26	B	A	R	60	20	20	U58	D	A	R	40	20	40
U27	B	A	U	40	20	40	U59	D	A	U	20	0	80
U28	B	A	N	20	0	80	U60	D	A	N	20	0	80
U29	B	F	M	20	0	80	U61	D	F	M	0	0	100
U30	B	F	R	20	0	80	U62	D	F	R	0	0	100
U31	B	F	U	0	0	100	U63	D	F	U	0	0	100
U32	B	F	N	0	0	100	U64	D	F	N	0	0	100

Note. Grade point average: A = 3.50–4.00; B = 3.00–3.49; C = 2.50–2.99; D = 2.00–2.49. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

TABLE 5
Expert-Opinion-Based Decision Table (Applicants Holding a Graduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
G01	A	E	M	100	0	0	G33	C	E	M	80	20	0
G02	A	E	R	100	0	0	G34	C	E	R	80	20	0
G03	A	E	U	100	0	0	G35	C	E	U	60	40	0
G04	A	E	N	100	0	0	G36	C	E	N	60	40	0
G05	A	G	M	100	0	0	G37	C	G	M	80	20	0
G06	A	G	R	80	20	0	G38	C	G	R	60	40	0
G07	A	G	U	60	20	20	G39	C	G	U	60	20	20
G08	A	G	N	60	20	20	G40	C	G	N	60	20	20
G09	A	A	M	80	20	0	G41	C	A	M	60	20	20
G10	A	A	R	80	20	0	G42	C	A	R	60	20	20
G11	A	A	U	80	0	40	G43	C	A	U	40	20	40
G12	A	A	N	20	20	60	G44	C	A	N	20	0	80
G13	A	F	M	20	20	60	G45	C	F	M	20	20	60
G14	A	F	R	20	20	60	G46	C	F	R	20	0	80
G15	A	F	U	0	0	100	G47	C	F	U	0	0	100
G16	A	F	N	0	0	100	G48	C	F	N	0	0	100
G17	B	E	M	100	0	0	G49	D	E	M	80	20	0
G18	B	E	R	80	20	0	G50	D	E	R	60	40	0
G19	B	E	U	80	20	0	G51	D	E	U	60	40	0
G20	B	E	N	80	20	0	G52	D	E	N	60	40	0
G21	B	G	M	80	20	0	G53	D	G	M	60	40	0
G22	B	G	R	80	20	0	G54	D	G	R	60	40	0
G23	B	G	U	60	20	20	G55	D	G	U	40	20	40
G24	B	G	N	60	20	20	G56	D	G	N	40	20	40
G25	B	A	M	80	20	0	G57	D	A	M	60	0	40
G26	B	A	R	60	20	20	G58	D	A	R	60	0	40
G27	B	A	U	40	20	40	G59	D	A	U	40	0	60
G28	B	A	N	20	20	60	G60	D	A	N	20	0	80
G29	B	F	M	20	20	60	G61	D	F	M	0	20	80
G30	B	F	R	20	20	60	G62	D	F	R	0	0	100
G31	B	F	U	0	0	100	G63	D	F	U	0	0	100
G32	B	F	N	0	0	100	G64	D	F	N	0	0	100

Note. Grade point average: A = 3.75–4.00; B = 3.50–3.74; C = 3.25–3.49; D = 3.00–3.24. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

tance in Tables 4 and 6 were within 0.10 of each other in 35 of the 64 cases, and within 0.20 for 54 of the 64 cases. However, we did have four cases with a difference greater than 0.20 and six cases where such a comparison was not possible because of a lack of historical data (represented by the "NA" in Table 6). More significantly, as will be discussed in the validation section below, the historical probabilities (Tables 6 and 7) seemed to be closer to the validating decisions than the expert-opinion-based probabilities (Tables 4 and 5). We exploited this situation by integrating the use of these two data bases.

4. VALIDATION

The validation process of an expert system ensures that the system performs with an acceptable level of accuracy (Liebowitz, 1986; O'Keefe, 1987). In our case, the experts had indicated that historical screening decisions were often arbitrary and inconsistent and that the company's criteria and policies for screening had changed in recent years. As such, we could not use historical data to validate our system. Therefore, as suggested by O'Leary (1987) we had divided our experts into two groups, one to develop the knowledge base and the other to develop case-by-case validating decisions. The Philadelphia-based experts were used to develop the knowledge base. Then we asked the four experts in New York, who were not aware of the knowledge base developed, to carefully examine a total of 462 cases of new applicants over an 8-month period. Ideally, we would have liked each of the four experts to evaluate each one of our 462 test cases to obtain the probabilities that could then be compared to the probabilities of the rule-based system. However, this would have been a monumental task, and it would have seriously jeopardized the day-to-day operations of the company. Hence, the experts were given complete files (including all the pertinent information collected on the candidate, and not just the information on the six

TABLE 6
Historical Frequency-Based Decision Table (Applicants Holding an Undergraduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
U01	A	E	M	88	02	10	U33	C	E	M	74	22	04
U02	A	E	R	80	07	13	U34	C	E	R	67	27	06
U03	A	E	U	67	14	19	U35	C	E	U	58	29	13
U04	A	E	N	63	14	23	U36	C	E	N	49	38	13
U05	A	G	M	83	02	15	U37	C	G	M	NA	NA	NA
U06	A	G	R	73	07	20	U38	C	G	R	59	28	13
U07	A	G	U	54	14	32	U39	C	G	U	48	26	26
U08	A	G	N	50	14	36	U40	C	G	N	41	30	29
U09	A	A	M	74	16	10	U41	C	A	M	62	24	14
U10	A	A	R	66	21	13	U42	C	A	R	46	27	27
U11	A	A	U	51	07	42	U43	C	A	U	43	08	49
U12	A	A	N	26	14	60	U44	C	A	N	19	11	70
U13	A	F	M	NA	NA	NA	U45	C	F	M	17	20	63
U14	A	F	R	19	23	58	U46	C	F	R	23	20	57
U15	A	F	U	11	03	86	U47	C	F	U	15	02	83
U16	A	F	N	10	09	81	U48	C	F	N	07	04	89
U17	B	E	M	82	11	07	U49	D	E	M	NA	NA	NA
U18	B	E	R	84	07	09	U50	D	E	R	63	24	13
U19	B	E	U	70	14	16	U51	D	E	U	54	39	07
U20	B	E	N	67	22	11	U52	D	E	N	46	38	16
U21	B	G	M	77	10	13	U53	D	G	M	60	23	17
U22	B	G	R	69	14	17	U54	D	G	R	54	30	16
U23	B	G	U	51	26	23	U55	D	G	U	31	26	43
U24	B	G	N	44	26	30	U56	D	G	N	29	34	37
U25	B	A	M	NA	NA	NA	U57	D	A	M	NA	NA	NA
U26	B	A	R	60	30	10	U58	D	A	R	50	07	43
U27	B	A	U	47	04	49	U59	D	A	U	22	04	74
U28	B	A	N	36	10	54	U60	D	A	N	17	12	71
U29	B	F	M	20	21	59	U61	D	F	M	24	19	57
U30	B	F	R	17	17	66	U62	D	F	R	16	23	61
U31	B	F	U	16	04	80	U63	D	F	U	NA	NA	NA
U32	B	F	N	09	04	87	U64	D	F	N	04	03	93

Note. Grade point average: A = 3.50–4.00; B = 3.00–3.49; C = 2.50–2.99; D = 2.00–2.49. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

criteria used by the expert system) on approximately 115 applicants each, and asked to make a decision as to whether they would have accepted, held, or rejected the application. Thus, each of our 462 cases was examined by one of the four recruiting experts. The experts reported back the results of their screening via the questionnaire in Figure 4.

As we have indicated before, for all practical purposes (98% of the cases), a "hold" decision really amounts to a "reject" decision. Hence, we decided that when for a test case, the expert system reports that an "acceptance" would be the correct decision with a probability of greater than or equal to 0.50, and our New York-based expert reports an acceptance, then there is a "match" between the expert system and the validating data. Similarly, if for a test case, the system reports that an acceptance would be the correct choice with less than 0.50 probability, and the New York-based expert either rejects or holds the case, again there is a "match" between the system and the validating data. In all other cases, there is a mismatch between the two. The greater the percentage of matches between the expert system decisions and the case-by-case expert decisions, the more valid the expert system is. To our great surprise, only 283 (61.3%) of the 462 cases showed a "match."

At this stage, we shared these results with all nine of our experts, and asked them to see if the expert-opinion-based rules had to be revised, or if the case-by-case decisions should be done again. Once again, we were surprised to find that the New York-based experts found practically nothing wrong with the rules developed by the Philadelphia experts, nor did the Philadelphia experts consider the case-by-case decisions of the New York experts to be unreasonable, when all pertinent information on a candidate is considered. One conclusion of this entire review was that clearly, the six criteria we had isolated for the expert system screening decision were insufficient. Of course, one would never be able to include all possible criteria in

TABLE 7
Historical Frequency-Based Decision Table (Applicants Holding a Graduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
G01	A	E	M	94	02	04	G33	C	E	M	NA	NA	NA
G02	A	E	R	84	12	04	G34	C	E	R	70	26	04
G03	A	E	U	73	21	06	G35	C	E	U	66	27	07
G04	A	E	N	66	28	06	G36	C	E	N	57	33	10
G05	A	G	M	82	07	11	G37	C	G	M	66	20	14
G06	A	G	R	80	11	09	G38	C	G	R	69	17	14
G07	A	G	U	NA	NA	NA	G39	C	G	U	50	26	24
G08	A	G	N	57	23	20	G40	C	G	N	49	24	27
G09	A	A	M	79	17	04	G41	C	A	M	61	20	19
G10	A	A	R	69	21	10	G42	C	A	R	52	21	27
G11	A	A	U	49	14	37	G43	C	A	U	47	03	50
G12	A	A	N	29	14	57	G44	C	A	N	22	11	67
G13	A	F	M	27	13	60	G45	C	F	M	20	17	63
G14	A	F	R	20	17	63	G46	C	F	R	16	19	65
G15	A	F	U	14	17	69	G47	C	F	U	11	10	79
G16	A	F	N	NA	NA	NA	G48	C	F	N	NA	NA	NA
G17	B	E	M	83	13	04	G49	D	E	M	67	31	02
G18	B	E	R	82	14	04	G50	D	E	R	NA	NA	NA
G19	B	E	U	73	20	07	G51	D	E	U	56	41	03
G20	B	E	N	63	31	06	G52	D	E	N	52	41	07
G21	B	G	M	81	10	09	G53	D	G	M	66	21	13
G22	B	G	R	73	17	10	G54	D	G	R	57	29	14
G23	B	G	U	57	21	22	G55	D	G	U	46	30	24
G24	B	G	N	50	27	23	G56	D	G	N	NA	NA	NA
G25	B	A	M	67	19	14	G57	D	A	M	53	09	38
G26	B	A	R	66	21	13	G58	D	A	R	47	13	40
G27	B	A	U	52	04	44	G59	D	A	U	31	16	53
G28	B	A	N	26	14	60	G60	D	A	N	19	11	70
G29	B	F	M	NA	NA	NA	G61	D	F	M	NA	NA	NA
G30	B	F	R	17	17	66	G62	D	F	R	17	23	60
G31	B	F	U	13	13	74	G63	D	F	U	NA	NA	NA
G32	B	F	N	NA	NA	NA	G64	D	F	N	NA	NA	NA

Note. Grade point average: A = 3.75–4.00; B = 3.50–3.74; C = 3.25–3.49; D = 3.00–3.24. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

an expert system. However, a related conclusion was also that experts' judgments about a "hypothetical" candidate (as required in developing the rules for our system) are often at variance with their case-by-case decisions based on more complete information. We believe these are two important lessons of our study.

Although the experts had indicated that perhaps historical frequencies were not relevant, because we had the data, we decided to see how many "matches" there were between the case-by-case expert decisions and the historical frequencies. Once again, we defined a "match" in a manner similar to our earlier definition. Thus, if historical frequency of "acceptance" is greater than or equal to 0.50 for a test case, and our New York-based expert reports an acceptance, then there is a "match." Similarly, if historical frequency of acceptance is less than 0.50 for a test case, and the New York expert reports a reject or hold decision, then there is a match. In all other cases, there is a mismatch. In this comparison, there were 406 (87.9%) matches. This result surprised our experts, and they conceded that perhaps their company's historical decisions were not as arbitrary and inconsistent as they had thought. Given this concession on their part, we found a way of integrating the expert opinions with the historical data for developing a knowledge base that would yield the greatest number of matches with the case-by-case validating decisions.

Our method is to calculate the integrated probability of acceptance of each case as a weighted average of the historical frequency of acceptance in that case, and the probability of acceptance in that case based on Philadelphia experts' opinions. Similarly, the integrated probabilities of hold and rejection are the weighted averages (using the same weights as those used to calculate the integrated probability of acceptance) of the respective historical frequencies and the probabilities based on Philadelphia experts' opinions. The method of choosing the most suitable weights for this calculation is as follows.

File No.: ________ Date: ________

Candidate's Name: ________________________

1. What is the Major Field of Study?

_____Accounting _____ Others

2. How are the References?

	Excellent (4)	Good (3)	Average (2)	Poor (1)
Reference No. 1:	____	____	____	____
Reference No. 2:	____	____	____	____
Reference No. 3:	____	____	____	____

_____Favorable (More than 7 Points)
_____Unfavorable (7 Points or Less)

3. What is the Highest Degree Completed?

_____Graduate _____ Undergraduate _____Others

4. What is the Grade Point Average?

_____Between 3.75 & 4.00 _____ Between 3.50 & 4.00
_____Between 3.50 & 3.74 _____ Between 3.00 & 3.49
_____Between 3.25 & 3.49 _____ Between 2.50 & 2.99
_____Between 3.00 & 3.24 _____ Between 2.00 & 2.49

5. What is the Result of the Aptitude Test?

_____Excellent _____ Good _____ Average _____ Fail

6. What is the Level of Prior Work Experience?

_____Related/3 years and more _____ Unrelated
_____Related/less than 3 years _____ None

Additional Comments: ________________________

RECOMMENDATION

_____ACCEPT _____ HOLD _____REJECT

Recruiter's Name Signature

FIGURE 4. Screening questionnaire

Consider that a weight of $\alpha (0 \le \alpha \le 1)$ is used on the historical frequencies whereas a weight of $1\text{-}\alpha$ is used on probabilities based on Philadelphia experts' opinions. Using a particular value of α, we calculated the integrated probabilities of acceptance, hold, and rejection for each of the 128 cases and checked the number of matches the integrated decision table yielded with the 462 test cases. Table 8 presents various values of α and corresponding numbers of matches.[4] For example, when $\alpha = .40$, there were 356 (77.1%) matches. As can be seen, the number of matches (427, or 92.4%) were maximized at $\alpha = .70$. Hence, using $\alpha = .70$, we constructed our final integrated decision tables, which are now actually used by our expert system. Tables 9 and 10 present these integrated decision tables for candidates with undergraduate and graduate degrees, respectively. Our experts from both the Philadelphia and the New York offices now believe that these tables represent the best use of the historical data and the expert opinions.

Observe that, in a sense, we have combined our validation with the process of development of the knowledge base. Although Politakis (1983) also recommends such an integration of the two processes, he really recommends integrating the test cases themselves with the domain knowledge. We believe that Politakis' (1983)

[4] We tried all possible two-decimal values of α, and found $\alpha = 0.70$ yielded the maximum number of matches. For simplicity, Table 8 presents only selected values of α.

TABLE 8
The Matching Rate of the Knowledge Base Decisions With the Case-by-Case Expert Decisions for Various Values of α

Weight for the Historical Frequency	Weight for Expert Opinion $(1 - \alpha)$	Number of Matching Responses out of 462	Effective Rate (%)
.00	1.00	283	61.3
.10	.90	302	65.4
.20	.80	315	68.2
.30	.70	332	71.9
.40	.60	356	77.1
.50	.50	398	86.2
.60	.40	412	89.2
.70	.30	427	92.4*
.80	.20	418	90.5
.90	.10	407	88.1
1.00	.00	406	87.9

* Maximum matching.

approach may be dangerous insofar as it contaminates the validating data. We do not integrate the validating decisions with the domain knowledge. Instead, we integrate two sources of domain knowledge to obtain the best possible fit with the validating decisions.

5. CONCLUSION

Our expert system is implemented using an expert system shell. The dialogue interface interacts with the recruiter through a series of questions concerning the candidate's qualifications. The system then computes the probabilities of acceptance, hold, and rejection being the experts' choice. These probabilities are intended to be an aid to the recruiters, and not a substitute for their own judgment. The recruiters are unequivocally advised that management recognizes that the expert system does not capture all pertinent information, and that to gather the information on the six criteria used by the expert system, a campus interview is unnecessary. The recruiters are encouraged to use other pertinent information from these interviews, along with the system's probabilistic recommendations, and make as many definitive decisions as possible without consulting a recruiting expert. When in doubt, the recrui-

TABLE 9
Integrated Decision Table (Applicants Holding an Undergraduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
U01	A	E	M	92	01	07	U33	C	E	M	76	21	03
U02	A	E	R	86	05	09	U34	C	E	R	71	25	04
U03	A	E	U	77	10	13	U35	C	E	U	59	32	09
U04	A	E	N	74	10	16	U36	C	E	N	52	39	09
U05	A	G	M	82	07	11	U37	C	G	M	60	40	0
U06	A	G	R	75	11	14	U38	C	G	R	59	32	09
U07	A	G	U	56	16	28	U39	C	G	U	52	24	24
U08	A	G	N	53	16	31	U40	C	G	N	47	27	26
U09	A	A	M	76	17	07	U41	C	A	M	61	23	16
U10	A	A	R	70	21	09	U42	C	A	R	50	25	25
U11	A	A	U	48	11	41	U43	C	A	U	42	06	52
U12	A	A	N	24	16	60	U44	C	A	N	19	06	73
U13	A	F	M	20	20	60	U45	C	F	M	18	14	68
U14	A	F	R	19	16	65	U46	C	F	R	16	14	70
U15	A	F	U	08	08	84	U47	C	F	U	11	01	88
U16	A	F	N	07	06	87	U48	C	F	N	05	03	92
U17	B	E	M	87	08	05	U49	D	E	M	80	20	0
U18	B	E	R	83	11	06	U50	D	E	R	62	29	09
U19	B	E	U	73	16	11	U51	D	E	U	56	39	05
U20	B	E	N	65	21	14	U52	D	E	N	50	39	11
U21	B	G	M	78	13	09	U53	D	G	M	60	28	12
U22	B	G	R	72	16	12	U54	D	G	R	56	33	11
U23	B	G	U	54	24	22	U55	D	G	U	34	24	42
U24	B	G	N	49	24	27	U56	D	G	N	32	30	38
U25	B	A	M	80	20	0	U57	D	A	M	60	0	40
U26	B	A	R	60	27	13	U58	D	A	R	47	11	42
U27	B	A	U	45	09	46	U59	D	A	U	21	03	76
U28	B	A	N	31	07	62	U60	D	A	N	18	08	74
U29	B	F	M	20	15	65	U61	D	F	M	17	13	70
U30	B	F	R	18	12	70	U62	D	F	R	11	16	73
U31	B	F	U	11	03	86	U63	D	F	U	0	0	100
U32	B	F	N	06	03	91	U64	D	F	N	03	02	95

Note. Grade point average: A = 3.50–4.00; B = 3.00–3.49; C = 2.50–2.99; D = 2.00–2.49. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

TABLE 10
Integrated Decision Table (Applicants Holding a Graduate Degree)

	Condition			Action (%)				Condition			Action (%)		
Rule	GPA	Test	Exp	Accept	Hold	Reject	Rule	GPA	Test	Exp	Accept	Hold	Reject
G01	A	E	M	96	01	03	G33	C	E	M	80	20	0
G02	A	E	R	89	08	03	G34	C	E	R	73	24	03
G03	A	E	U	81	15	04	G35	C	E	U	64	31	05
G04	A	E	N	76	20	04	G36	C	E	N	58	35	07
G05	A	G	M	87	05	08	G37	C	G	M	70	20	10
G06	A	G	R	80	14	06	G38	C	G	R	66	24	10
G07	A	G	U	60	20	20	G39	C	G	U	53	24	23
G08	A	G	N	58	22	20	G40	C	G	N	52	23	25
G09	A	A	M	79	18	03	G41	C	A	M	61	20	19
G10	A	A	R	72	21	07	G42	C	A	R	54	21	25
G11	A	A	U	52	10	38	G43	C	A	U	45	08	47
G12	A	A	N	26	16	58	G44	C	A	N	21	08	71
G13	A	F	M	25	15	60	G45	C	F	M	20	18	62
G14	A	F	R	20	18	62	G46	C	F	R	17	13	70
G15	A	F	U	10	12	78	G47	C	F	U	08	07	85
G16	A	F	N	0	0	100	G48	C	F	N	0	0	100
G17	B	E	M	88	09	03	G49	D	E	M	71	28	01
G18	B	E	R	81	16	03	G50	D	E	R	60	40	0
G19	B	E	U	75	20	05	G51	D	E	U	57	41	02
G20	B	E	N	68	28	04	G52	D	E	N	54	41	05
G21	B	G	M	81	13	06	G53	D	G	M	64	27	09
G22	B	G	R	75	18	07	G54	D	G	R	58	32	10
G23	B	G	U	58	21	21	G55	D	G	U	44	27	29
G24	B	G	N	53	25	22	G56	D	G	N	40	20	40
G25	B	A	M	71	19	10	G57	D	A	M	55	06	39
G26	B	A	R	64	21	15	G58	D	A	R	51	09	40
G27	B	A	U	48	09	43	G59	D	A	U	34	11	55
G28	B	A	N	24	16	60	G60	D	A	N	19	08	73
G29	B	F	M	20	20	60	G61	D	F	M	0	20	80
G30	B	F	R	18	18	64	G62	D	F	R	12	16	72
G31	B	F	U	09	09	82	G63	D	F	U	0	0	100
G32	B	F	N	0	0	100	G64	D	F	N	0	0	100

Note. Grade point average: A = 3.75–4.00; B = 3.50–3.74; C = 3.25–3.49; D = 3.00–3.24. Aptitude test: E = Excellent; G = Good; A = Average; F = Fail. Work experience: M = Related/3 & More; R = Related/Less Than 3; U = Unrelated; N = None.

ters are to put a candidate on "hold," and only these hold decisions are to be reviewed by the recruiting experts.

Preliminary results indicate that both the recruiters and the recruiting experts are satisfied with the system. However, it is too early to provide a full evaluation of the system's effectiveness and success.

As has been discussed, our work built on Extejt and Lynn's (1988) hypothetical proposal for such an expert system. Extejt and Lynn's (1988) philosophy of an expert system being only a decision aid, and their concept of providing only the probabilistic recommendations were found to be very useful. However, our work improved their concepts of probability assessment on both the theoretical and practical level. We found that rational experts require that the probabilities of mutually exclusive and collectively exhaustive choices must add up to 1.00. We also found that individually, experts are not comfortable in estimating such probabilities.

For the development of their knowledge base, Extejt and Lynn (1988) proposed the use of expert opinions only. We found that historical data can be also very valuable. In a sense, our work shows the deficiencies associated with using a single knowledge acquisition approach. While Extejt and Lynn (1988) did not concern themselves with the issues and the processes of validation, we learned important lessons from our validation efforts. For example, we found that experts' judgments about "hypothetical" candidates are often at variance with their case-by-case decisions. We also found that our experts had overestimated the degree of arbitrariness and inconsistency in the historical data.

Finally, by integrating two sources of domain knowledge, namely expert opinions and historical data, we were able to obtain "the most valid" knowledge base.

Needless to say, there are a number of limitations associated with our study. In retrospect, several of our academic colleagues have suggested a number of methodological improvements including:

1. We should have incorporated the top 8 or 10 screening criteria in our expert system.
2. We should have used a larger number of experts in obtaining our probabilities in Tables 4 and 5.

3. Instead of asking the experts for their most likely decisions in the 128 cases, we should have constructed a model, perhaps using the Analytic Hierarchy Process (Saaty, 1990), which scores the various categories of each criterion and weighs the various criteria in relation to one another.
4. Using the historical data, we should have done a discriminant analysis to understand the factor weights historically given to the three variables, namely, GPA, aptitude test, and experience, and the scores implied by a candidate's membership in one or another category of these variables. These historical weights could then serve as a discussion point with the experts for a normative revision of the weights to be given in the future.

Nevertheless, we believe that we did the best we could to build an operational system within the constraints of available data, personnel, time, and other resources. We also believe that ours is one of the few studies in which the development and validation of an expert system for a human resources decision is adequately documented. Our study clearly establishes the feasibility and usefulness of such an expert system. We believe that our work contributes important practical lessons for those who may be involved in the development of expert systems in other companies or in related fields.

REFERENCES

Briggs, S., & Doney, L. (1989). Eight HR expert systems now. *Computers in Personnel,* **4**(1), 10–14.

Dalkey, N.C. (1969, June). *The delphi method: An experimental study of group opinion.* Santa Monica, CA: The RAND Corporation. (Research Paper, PM-5888-PR)

DeSanctis, G. (1986). Human resource information systems. *MIS Quarterly,* **10,** 8–18.

Extejt, M.M., & Lynn, M.P. (1988). Expert systems as human resource management tools. *Journal of Systems Management,* **3**(11), 10–15.

Krebs, V. E. (1988). Can expert systems make HR decisions? *Computers in Personnel,* **3**(2), 4–8.

Liebowitz, J. (1986). Useful approaches for evaluating expert systems. *Expert Systems,* **3**(2), 86–96.

Liebowitz, J. (1988). *Introduction to expert systems.* Santa Cruz, CA: Mitchell Publishing.

Liebowitz, J. (1990). *The dynamics of decision support and expert systems.* Orlando, FL: The Dryden Press.

Lin, E. (1986). Expert systems for business applications: Potentials and limitations. *Journal of Systems Management,* **1,** 18–21.

Linder, J. C. (1985). Computers, corporate culture and change. *Personnel Journal,* **64**(9), 49–55.

O'Keefe, D. E. (1987). Validating expert system performance. *IEEE Expert,* **2,** 81–90.

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

Olson, J.R., & Rueter, H.H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems,* **4**(3), 152–168.

Perry, M., More, J., & Parkison, N. (1987). Does your appraisal system stack up? *Personnel Journal,* **66**(5), 82–87.

Politakis, P.G. (1983). *Using empirical analysis to redefine expert system knowledge bases.* Doctoral dissertation. Newark, NJ: Rutgers University.

Saaty, T.L. (1990). Multicriteria decision making: The analytic hierarchy process. Pittsburgh, PA: RWS Publications.

Simon, S.H. (1983). The HRIS: What capabilities must it have? *Personnel,* **60**(5), 36–49.

Staff. (1989, January). How the government created HR systems. *Software Magazine.*

Walker, A.J. (1986). New technologies in human resource planning. *Human Resource Planning,* **9**(4), 149–159.

Whaley, G.L. (1989). What's delaying HR expert systems? *Computers in Personnel,* **4**(2), 16–20.

Wright, G., & Ayton, P. (1987). Eliciting and modeling expert knowledge. *Decision Support Systems,* **3,** 13–26.