Early Expert Systems: Where Are They Now?

By: T. Grandon Gill
    College of Business
    Florida Atlantic University
    Boca Raton, FL 33431
    U.S.A.
    t_gill@acc.fau.edu

Abstract

Expert systems (ES) were among the earliest branches of artificial intelligence (AI) to be commercialized. But how successful have they actually been? Many well-publicized applications have proven to be pure hype, numerous AI vendors have failed or been completely reorganized, major companies have reduced or eliminated their commitment to expert systems, and even Wall Street has become disillusioned—a predicted $4 billion market proving to be smaller by an order of magnitude. Yet, in spite of these setbacks, there are many companies who remain enthusiastic proponents of the technology and continue to develop important ES applications.

This paper explores how the first wave of commercial expert systems, built during the early and mid-1980s, fared over time. An important subset of these systems, identified in a catalog of commercial applications compiled in 1987, was located through a telephone survey, and detailed information on each system was gathered. The data collected show that most of these systems fell into disuse or were abandoned during a five-year period from 1987 to 1992, while about a third continued to thrive. Quantitative and qualitative analysis of the data further suggests that the short-lived nature of many systems was not attributable to failure to meet technical performance or economic objectives. Instead, managerial issues such as lack of system acceptance by users, inability to retain developers, problems in transitioning from development to maintenance, and shifts in organizational priorities appeared to be the most significant factors resulting in long-term expert system disuse.

Introduction

Expert systems, one of the earliest branches of artificial intelligence (AI) to achieve widespread commercial viability, present managers with a paradox. The technology, which applies AI-derived specialized symbolic reasoning techniques to solve difficult problems (Luconi, et al., 1986), produced a series of resounding successes in the early and mid-1980s. Well known systems, such as Digital's *XCON*, Coopers and Lybrand's *ExpertTax*, and American Express' *Authorizer's Assistant*, have amply demonstrated the technology's capability both to generate huge financial returns and to contribute to the strategic goals of the firm (Sviokla, 1986). By the late 1980s, however, another attitude toward expert systems began to surface in the AI and management communities. Critics argued that expert systems, as a class, rarely succeed or, perhaps, cannot succeed in delivering expert performance.[1] Wall Street, once enthusiastic about the prospects of expert systems, became suspicious of a technology that repeatedly failed to deliver on its promises. As the *Wall Street Journal* reported:

> The AI industry, which many market researchers had predicted would reach $4 billion annual sales by now, remains nascent. Generous estimates of the market today are closer to $600 million. After swallowing up hundreds of millions of dollars in venture capital and exciting some of the brightest professors at top technical schools with visions of riches, hundreds of AI start-ups have yielded only a few profitable public companies (Bulkeley, 1990, Section B, p. 1).

Even industry participants have voiced serious concerns. In a recent survey, 60 percent predicted that the AI industry would either remain flat or decline between 1993 and 1999 (Coleman, 1993).

At the present time, considerable divergence of opinion exists regarding how well expert systems have fared. The travails of AI vendors, who are important suppliers of expert system tools, suggest that demand for ES technology is not exact-

ly thriving. Some of the most influential hardware and tool companies (e.g., Gold Hill, Intellicorp, Inference, Teknowledge, and Symbolics) have been forced to reorganize, enduring major cutbacks in the process. Other participants have simply gone out of business (e.g., Palladian and Lisp Machines, Inc.). Broadly based companies that formerly maintained AI divisions or products (e.g., Texas Instruments, Xerox, Borland, Microsoft, and Radian) have refocused their efforts elsewhere.

Not all indicators of ES technology are bleak, however. Conversations with senior managers indicate that a number of major companies, such as Digital Equipment Corp., Coopers & Lybrand, and American Express, persist in actively developing and maintaining expert systems. Some, in fact, assert that their key businesses are strategically dependent on these systems and are likely to remain so in the future. Furthermore, a variety of new products, ranging from tax preparation software to music and language instruction systems to college search software, now incorporate ES technology.

While numerous opinions exist regarding how successful commercial expert systems have been, the basis for these opinions is largely anecdotal. What is almost completely lacking is quantitative data addressing the question of how well commercial expert systems, as a group, have fared. Specifically, we know little about how the first wave of commercial ES applications performed, or how many are still in use today. Further, almost no information has been systematically gathered to identify those factors that influenced usage.

This paper describes a field study that examined how these early systems fared. In particular, the study acquired and analyzed data relating to two important measures of ES use:

- **User penetration**—the degree to which potential users became actual users, and

- **Longevity**—the period of time over which the system was used.

Its specific goals were to better understand how system performance, system economics, and organizational factors contributed to early ES use. The approach entailed locating approximately 80 expert systems that were developed in the early and mid-1980s and then gathering data relating to the following questions:

1. To what extent did issues of system performance appear to affect the levels of user penetration and longevity of the systems examined?

2. To what extent did the economics of development and maintenance appear to affect the levels of user penetration and longevity of the systems examined?

3. To what extent did issues of individual and organizational adoption, such as user acceptance and fit with organizational priorities, appear to affect the levels of user penetration and longevity of the systems examined?

These questions were deemed particularly relevant to managers who oversee the development of expert systems and must therefore decide how to prioritize their company's efforts and resources.

Expert Systems Development: Alternative Perspectives

The question of how to build and implement expert systems has been studied extensively. Two distinct perspectives have emerged from the literature:

1. A technical perspective, which emphasizes the technological, managerial, and economic issues associated with constructing ES applications that deliver appropriate performance in a timely and cost-effective manner.

2. An organizational perspective, which particularly concerns itself with the challenges of managing the process of deploying and using systems within an organizational setting.

The two perspectives are reviewed in the next two sections.

ES development: technical perspective

Unlike conventional systems, which have existed since the mid-1950s, commercial expert systems have been around for a very short time—few were constructed more than a decade ago. As a consequence, much of the research on building these systems has focused on technical and software development issues. This emphasis is reflected in the types of research that have re-

ceived the greatest attention in the ES literature, including:

- Identifying task and domain areas that are suitable for ES (e.g., Buchanan, et al., 1983; Harmon, et al., 1988; Harmon and King, 1985; Prereau, 1985; Silverman, 1987; Stefik, et al., 1983; Walker and Miller, 1990; Waterman, 1986).

- Deciding whether or not to build an ES (e.g., Harmon and Sawyer, 1990; Silverman, 1987; Turban, 1992; Waterman, 1986).

- Selecting appropriate ES tools (e.g., Gevarter, 1987; Gill, 1991; Harmon, et al., 1988; Harmon and Sawyer, 1990; Stefik, et al., 1983; Stylianou, et al., 1992; Waterman, 1986; Waterman and Hayes-Roth, 1983).

- Roles and stages in ES development (e.g., Buchanan, et al., 1983; 1986; Harmon and Sawyer, 1990; Turban, 1992; Waterman, 1986).

- Working with domain experts and knowledge acquisition (e.g., Bobrow, et al., 1986; Harmon and King, 1985; Harmon and Sawyer, 1990; Prereau, 1985; Slatter, 1987; Waterman, 1986).

- Integrating AI/ES and conventional technologies (e.g., Freedman, 1987; Freundlich, 1990; Harmon and Sawyer, 1990; Stapleton, 1988; Turban and Watkins, 1986).

Beyond the ES development literature, there is also an immense computer science-grounded ES literature, addressing topics such as knowledge and uncertainty representation, algorithms and inference engine design, and automated knowledge acquisition.

Faithful to its focus on design and development, the technical perspective emphasizes, as its number one objective, building systems that exhibit high performance. This is loosely defined to mean systems that "successfully solve the problems to which they are applied" (Brachman, et al., 1983, p. 44). Many different criteria reflecting overall system performance exist, however. Among these are:

1. **Consistency:** The system's ability to produce task solutions with a level of consistency as great or greater than that of the expert. Where expertise exists in multiple individuals, such performance may be reflected in greater con-

sistency than previously existed between experts. For example, one of the performance characteristics cited for Texas Instruments' *Capital Expert* (Gill, 1987) was its ability to enforce consistency in the preparation of capital expenditure proposals across the company.

2. **Quality:** The system's ability to produce task solutions whose quality rivals, or exceeds, that of solutions produced by the expert. For example, one of the performance characteristics of Digital's *XCON* (Sviokla, 1986) was that it produced VAX configuration designs of higher quality than was possible with manual approaches.

3. **Error Rates:** The system's ability to avoid errors in its solutions. Waterman (1986, p. 30) notes that "expert systems, like their human counterparts, will make mistakes." Over time, however, he proposes that the level of errors encountered can be made to decline to below that of human experts, particularly "with the help of skillful users." Such an ability to avoid errors is often an important component of overall system quality, sometimes referred to as a "blunder stopper" role.[2] For example, the *Gatekeeper* system (Gill, 1991), which was used to help controllers assign gates to flights at Houston airport, used color coding to indicate possible gate assignment errors to controllers.

4. **Speed:** The speed at which the system performs the task. Feigenbaum (Harmon and Sawyer, 1990) suggests that one of the primary benefits from expert systems is the increase in problem-solving speed, often by several orders of magnitude. For example, one of the primary benefits cited for American Express' highly successful *Authorizer's Assistant* (Davis, 1987) was enhanced speed in authorizing AMEX card purchases. Waterman (1986) and others warn, however, that declining speed can also become a serious problem as an ES grows. In fact, slow performance was cited as a reason that *MYCIN,* one of the earliest expert systems, was never commercialized (Jackson, 1986).

5. **Learnability:** The system's effect on the rate at which individuals can learn to perform the task. One of the primary benefits of expert systems is proposed to be "leveraging expertise" (Waterman, 1986), meaning users of the

system can learn to perform a task faster than they otherwise could have. One feature that may increase learnability is explanation capability. Another factor is the codification of knowledge that occurs when an expert system is developed, sometimes cited to be one of the primary benefits of a system (Hayes-Roth, et al., 1983, p. 28). For example, the *MACSYMA* system, built to encompass the very large domain of symbolic integration, serves three main purposes: to perform integration, to act as a repository for a broad variety of integration techniques (far beyond what a typical individual could learn), and to aid mathematicians and engineers in learning new symbolic integration techniques in a timely manner.

The relative importance of each of these performance measures varies considerably, depending upon the task situation.

Because the development of expert systems requires significant resources, a second objective of systems development, economic consequences, is sometimes discussed in the technical perspective. Commonly referred to as payback or payoff, a system's ability to generate sufficient economic return is perceived to be a prerequisite for its development. MYCIN developer and MIT professor Randall Davis asserts, for example, that:

> The task domain must have a high payoff because the investment to get useful performance will be great. In the academic environment, we talk about the intellectual payoff—that is, attack problems because we think they will teach us something interesting. In the commercial environment, it is the economic payoff that matters. Make sure the payoff is substantial, because the effort certainly will be (Davis, 1984, p. 38).

A number of techniques have been proposed for estimating the economic return of ES applications. These range from formal techniques for cost-benefit analysis (e.g., Harmon and Sawyer, 1990; Turban, 1992) to informal discussions of how to assess return on investment (e.g., Harmon, et al., 1988). Three important contributors are considered by all techniques:

1. **Start-up costs:** The costs associated with the initial development of the system, which include both direct costs, such as hardware, software, and programmer time, and indirect

costs, such as facilities use and management, user and expert time (Harmon, et al., 1988).

2. **Ongoing costs:** The direct and indirect costs associated with ongoing maintenance and upgrading of an ES. For conventional systems, such costs are estimated to be as high as 70-80 percent of total costs over a system's lifetime (Pressman, 1982, p. 326). For expert systems, such maintenance can be particularly critical, as knowledge is continually added to many domains. For XCON, for example, the costs of ongoing maintenance have proven to be many times that of the initial development, forcing DEC to develop new systems specifically aimed at maintenance (Harmon and Sawyer, 1990).

3. **Ongoing benefits:** The benefits that may take either the form of cost savings (e.g., reduction in personnel, material savings) or revenue enhancements (e.g., improved sales), that are experienced as a consequence of using an ES. Such benefits may also include avoided costs, as was the case for Digital's XCON where quality improvements in VAX configuration eliminated the need to construct multiple Final Assembly and Test plants (Sviokla, 1986).

Thus, even after development costs have been sunk, it is possible that an ES will fail on economic grounds if ongoing costs (2) prove to be greater than ongoing benefits (3).

In summary, the technical perspective that dominates most ES research makes an important assumption: *the objectives of the ES developer can be met by constructing a system that demonstrates both suitable performance and suitable economics.* Such an assumption, however, often fails to take into account many of the subtleties of introducing a new technology or application into an organization. As a result, an organizational perspective of ES development has also emerged.

ES development: organizational perspective

The organizational perspective on ES implementation is deeply rooted in the more general study of information systems (IS) implementation and innovation research (e.g., Keen and Scott Morton, 1978). Such research is premised upon the

belief that both the organizational conditions existing prior to implementation and the approach taken in the implementation process will prove critical determinants of ultimate application success. Indeed, the organizational perspective holds that such organizational factors are often more important than either performance or economic justification in deploying a successful system.

One way to contrast technical and organizational perspectives involves the terms of their success. As DeLone and McLean (1992) observe, defining the term "success" has proven an extraordinarily elusive goal in the MIS field. They propose a taxonomy that consists of six distinct forms of success:

1. **System Quality:** The degree to which the system performs the task it was designed for, including such factors as accuracy, response time, reliability, and completeness.

2. **Information Quality:** The quality of the output produced by the system, such as report format, appearance, currency, relevance, and information value to decision makers.

3. **Use:** Recipient consumption of the output of a system. Within the category of use, there are two important subcategories: *user penetration,* the degree to which potential users become actual users, and *longevity,* the length of time over which a system remains in use.

4. **User Satisfaction:** How positively users react to the operation of and output from a system.

5. **Individual Impact:** The degree to which a system affects the behavior of the individuals using that system and its output.

6. **Organizational Impact:** The effect of the system on organizational performance.

The DeLone and McLean (1992) taxonomy suggests a key way in which the technical perspective differs from the organizational perspective: *while the technical perspective focuses almost entirely on system and information quality in its view of success, the organizational perspective emphasizes the remaining measures—especially satisfaction, user impact, and organizational impact.*

Several recent investigations suggest the importance of considering organizational and im-

plementation issues in ES development. In a study of 45 ES applications (Tyran and George, 1993), systems managers reported the five most important factors for ES success to be:

1. Assessment of user needs,
2. Commitment of human expert to the project,
3. Ease of ES use,
4. Commitment of the user to the project, and
5. Top management support.

Interestingly, the two technical factors included in their survey—technical expertise with ES and ES software/development tools—were perceived to be significantly less important to respondents. Similarly, in a survey of AI participants (Coleman, 1993) non-technical issues (e.g., oversell, lack of direction, business issues) were perceived to be far more serious roadblocks to AI success than technical problems (e.g., failure to deliver).

Implications for managers

The technical and organizational perspectives of ES implementation offer two very different prescriptions for developing ES. The technical perspective—emphasizing problem selection, tool selection, and knowledge acquisition, and managing the economics of development— advocates devoting resources to improving the organization's knowledge of the task to be performed and is very sensitive to the availability, capabilities, and economics of ES developers and tools. The organizational perspective, on the other hand, emphasizes the need for greater attention to managing the individuals associated with development, the organizational environment, and the overall process of implementation. Obviously, neither the technical nor the organizational aspects of development can be ignored. The question a manager must therefore answer is: How do I prioritize them?

Research Method

The empirical research reported in this paper represents one component of a broader study examining a series of issues relating to expert systems. The study employs a quasi-experimental design that proceeds in the following stages:

• **Sample Selection:** A catalog of 111 commercial expert systems, all built prior to 1988, was

chosen as the sample for detailed study. Thirteen systems that were built and used abroad and one classified system were eliminated, bringing the total sample down to 97 systems.

- **Data Gathering:** For each system, an individual knowledgeable in the use and development of the system was identified and a phone survey—to determine the status and the task characteristics of the system—was conducted.

- **Data Analysis:** Upon completion of data gathering, both statistical and qualitative analysis of the data was performed.

The experimental design is described in the next section.

Sample selection

The sample of expert systems investigated was taken from a catalog (called the HMM catalog in later references) of 111 commercial expert systems (Harmon, Maus, and Morrissey, 1988). There were a number of theoretical and practical reasons for the particular sample:

- **Representativeness:** The applications in the HMM catalog were originally selected because they represented the universe of commercial expert system applications, circa 1987, containing systems identified from a broad range of sources, including conference proceedings, vendor marketing materials, and extensive personal contacts within the AI community (Harmon, et al., 1988).

- **Success Criteria:** One of the major frustrations in researching systems can be the unwillingness of participants to discuss their less successful efforts. Only systems over five years old were included in the survey, serving two objectives: (1) participants were willing to talk about systems that didn't work out in the "distant past," and (2) both user penetration (e.g., extent of maximum usage) and longevity (e.g., degree of current use and degree of current development/maintenance) measures could be collected.

- **Practical:** Phone numbers and addresses were included for nearly all the systems in the HMM catalog. For some percentage of the systems, it was expected that these would facilitate determining system status. Further-

more, many of the systems in the database were quite well known, meaning that secondary published sources were often available to facilitate the locating of developers, managers, and users of the systems.

The greatest concern regarding sample selection was that the catalog supposedly contained only successful systems. It would therefore seem reasonable to suppose that such a sample would exhibit only limited variation in usage measures. There were, however, two reasons for anticipating that reasonable variations in usage would be observed:

- Concerns about the marketing "hype" that has permeated AI have been widely expressed (e.g., Buchanan, 1986; Winston, 1984). Since marketing materials were an important source of information for the HMM catalog, it was therefore plausible to expect that many—even most—of the announced systems might never have materialized. Even if the systems were completed, many were new enough at the time the catalog was published that the maximum levels of usage (i.e., user penetration) had not been determined.

- Between the cataloging of the systems and the performance of the survey, considerable time had passed. Thus, even if user penetration measures did not exhibit great variability, it was reasonable to expect that longevity variations across the sample could be substantial.

Perhaps the most compelling reasons for using the catalog were that: (1) it was intrinsically interesting, containing many of the best known systems of the 1980s, and (2) of the several other potential sources of ES applications available, there did not appear to be any with a comparable focus on commercial applications that also contained such a broad range of applications.

Data gathering

The data gathering process, described in Appendix A, was designed both to ensure that the maximum number of systems were located from the HMM catalog and to enhance the accuracy of the data gathered. It included gathering background data on each system in the catalog, making extensive efforts to locate suitable respondents and sending out information that had been gathered for further verification. Figure 1 contains a sum-

Percent of Systems
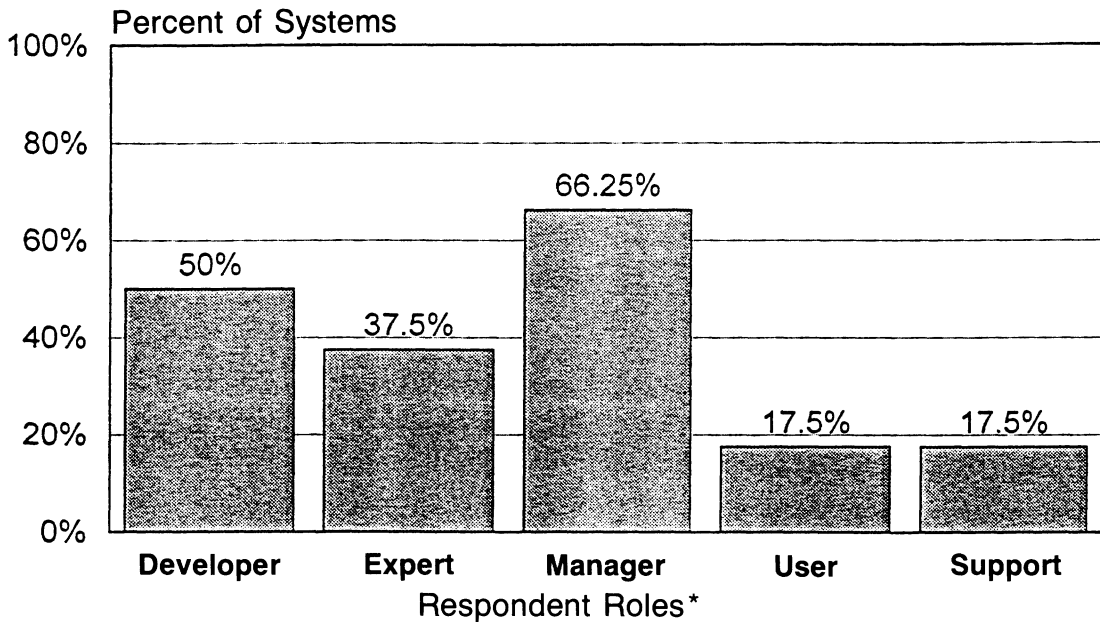
* Respondents can have more than one role.

Figure 1. Summary of Respondent Roles in Survey

mary of respondent roles within their respective firms.[3] Table 1 contains the final list of systems for which maximum level of usage data could be determined.

A particularly critical aspect of the data-gathering process was determining the degree of use achieved by each system. Three different survey questions were employed, aimed at measuring both user penetration and longevity of use:

1. **Level of User Penetration:** Respondents reported maximum degree of usage (interpreted as sales for systems sold as software packages), indicating how actual system usage compared with original expectations.

2. **Longevity (Current):** Respondents reported current level of system usage, indicating whether the system was still actively being used, and the level of use.

3. **Longevity (Prospective):** Respondents reported status of current development and maintenance activities. Because maintenance represents an important part of the ES life cycle, level of maintenance is indicative of the long-term prospects of the system.

These measures were chosen in place of DeLone and McLean's (1992) other measures of success

because: (a) the two IS quality measures of success were, for the most part, already being captured using performance-related questions posed in the survey, and (b) the remaining measures of success—user satisfaction, individual impact, and organizational impact—were, all at once, ambiguously defined, difficult to isolate from the implementation context and difficult to measure (DeLone and McLean, 1992). As a consequence, there appeared to be no way to obtain consistent estimates for the latter measures.

Results and Discussion

For the 97 systems that met survey criteria, the data gathering process yielded the following results:

- Sixty-five systems for which all parts (1 through 6) of the survey were collected.

- Ten systems for which only Parts 1, 2, and selected sections of Part 3 (status, performance, and descriptions) were collected. These represented incomplete and unused systems where the gathering of usage data (Parts 4-6) was deemed inappropriate by the investigator.

Table 1. List of Systems for Which Status Was Determined
[Survey Question 2.1]

Unfinished Systems	Completed Prototypes	Limited-Use Systems	Moderate-Use Systems	Widespread-Use Systems
Cementing Expert System	Ash Mixer	Class	AquaRef	AALPS
Corporate Financial Advisor[x]	Corrosion Expert	DELTA	Brush Designer	ACE
	ExMarine	DIAG8100	COMPASS	Authorizer's Assistant
Foreign Exchange Advisor[x]	Expert Probe	Diagnostics[x]	Cocomo 1	CBT Analyst
	Hotline Helper	Dipmeter Advisor	ESPm	Capital Expert
Portfolio Management Advisor[x]	IPECAC	FAIS	Exnut (formerly Peanut/PEST)	DASD Advisor
	NAVEX	Genesis		DustPro
	PTE Analyst	Grain Marketing Advisor	Hoist Diagnoser	ExperTax
SEATS	Planting[x]	IPT	MASK	GEMS TTA
Unit Commitment Advisor	Pump Pro	LISP-ITS	Mentor	(Renamed TERESA)
	TAX	Management Advisor	MudMan	HP4760AI
	TQMSTUNE		PRESS	Electrocardiograph
	Titan	Metals Analyst	PlanPower	Hazardous Chemical Advisor
	Waves	ONCOCIN	Rotating Eq. Vibration Anal.	Help
	Weld Scheduler	Ocean	SYSCON	IMP
		PowerChart	TIMM/TUNER	MACSYMA
		SNAP	TurboMac	Micro Genie
		SpinPro	Weld Selector	Microprocessor Electrophoresis
		Underwriting Advisor		PERMAID
		Welder Qual. Test Selector		Page-1
				Pulmonary Consult
				Requirements Analyst
				Source Rock Advisor
				TOGA
				XCON

Note: X—Refused to participate, status unverified.

- Five systems for which only partial status data (Part 2) and descriptions could be collected. These represented the refusals in the sample.

- One system for which performance and usage data were collected, but for which status data (Part 2) could not be determined.

- Sixteen systems that could not be located and whose status was therefore undetermined.

The next section profiles the findings and considers the research questions presented in the introduction.

Profile of systems in the sample

How did the early expert systems fare? The results of the survey of use-measures are summarized in Figure 2, Figure 3, and Figure 4 and are given both in terms of system counts and as percentages of systems for which usage data could be determined.

The survey results show considerable variation in both levels of user penetration and longevity of the applications surveyed. On the negative side, under a third of the systems ever achieved widespread or universal levels of usage. Furthermore, only about one-third of the systems managed to avoid declines in usage or total abandonment, a figure that is certainly optimistic given that all or nearly all of the 17 systems that could not be located have presumably been abandoned. On a positive note, however, almost three-fourths achieved some usage. And, for over a third of the applications, development dollars continue to flow for maintenance and/or enhancement.
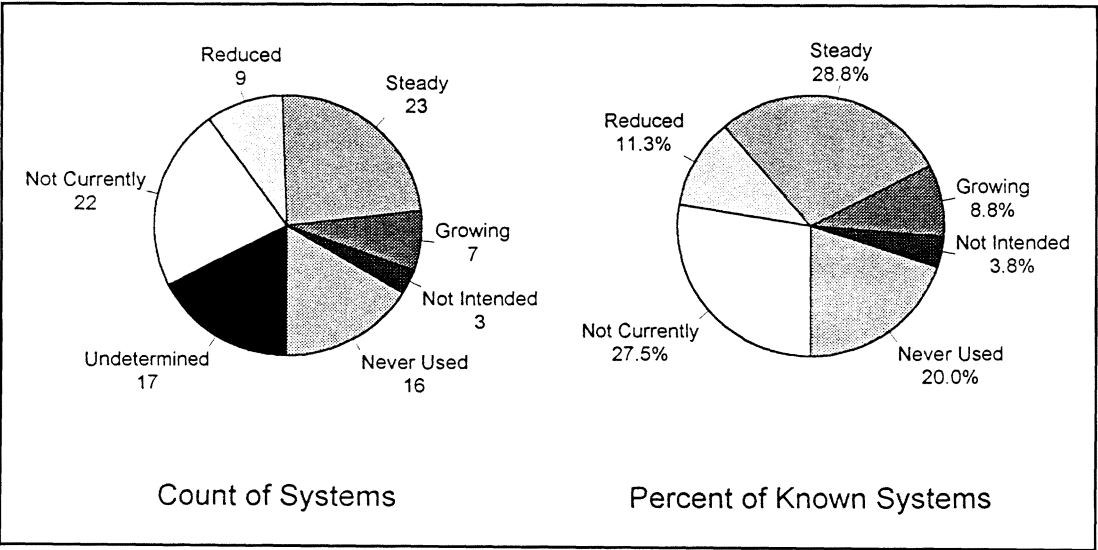
Research questions

The results of the survey also provide interesting insights into the three research questions presented in the introduction: how technical per-
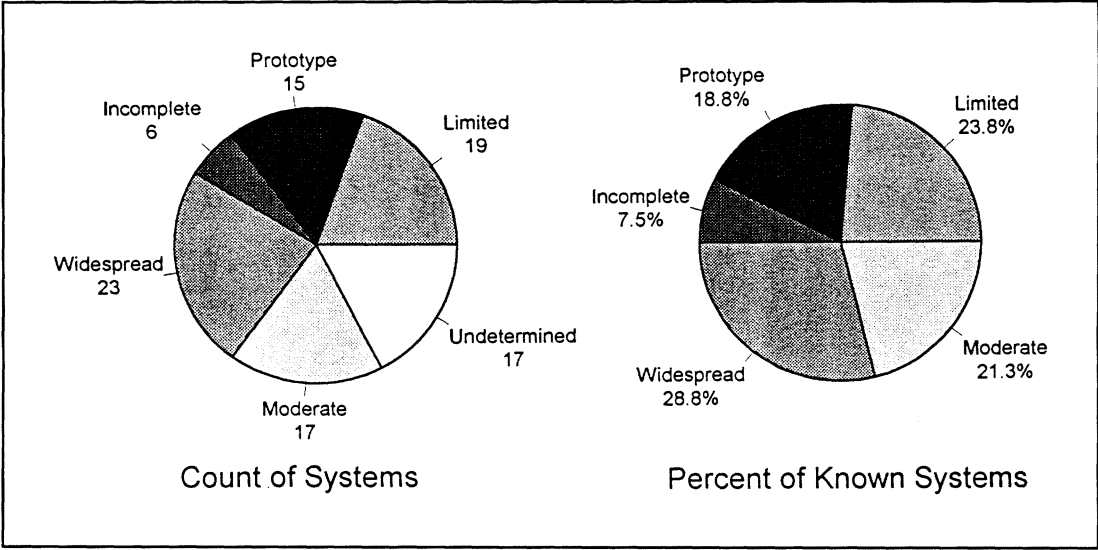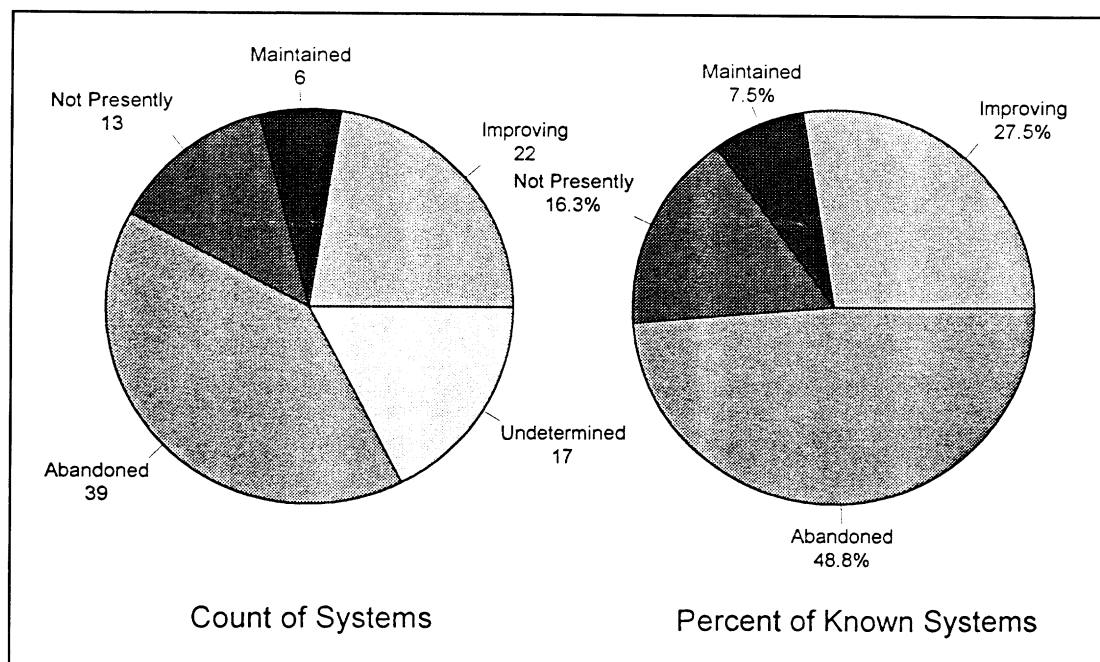
Note: Percentages may not add to 100 percent due to rounding errors.

Figure 2. Maximum Usage Achieved by Systems in Sample
[Survey Question 2.1]

Notes: Percentages may not add up to 100 percent due to rounding errors; "Not Intended" systems represent applications not intended for actual use, such as demonstration systems.

Figure 3. Current Usage Status of Systems in the Sample
[Survey Question 2.3]

Note: Percentages may not add to 100 percent due to rounding errors.

Figure 4. Status of Current Development for Systems in the Sample
[Survey Question 2.2]

formance, system economics, and organizational factors influenced both levels of user penetration and system longevity.

Performance

Did the systems in the sample routinely fail to achieve acceptable task performance? Direct evidence of such routine failure of expert systems to deliver on their performance promises would have come from two sources in the survey:

• Respondents were asked to comment on how using their systems changed task performance on a series of dimensions: consistency, quality, frequency of errors, performance time, and learning time.

• Respondents were asked to indicate which of a series of factors proved most limiting to the potential of their systems.

The respondents' replies to questions about system performance are presented in Table 2. The results suggest that virtually none of the systems surveyed exhibited serious inability to perform their assigned task (i.e., response of

Much Worse). In fact, of the 73 systems for which the data were gathered (omitting five refusals systems and two incomplete systems), only 17 (23 percent) of the systems had one or more responses indicating any performance problems, and there was only one system where either quality, consistency, or frequency of errors appeared to decline from the presystem performance of the task.

Because of the nature of the sample, some problems in achieving acceptable performance for commercial systems could have been masked by the manner in which systems were selected (i.e., HMM's stated criteria for inclusion in the catalog was that systems be in use or ready to use). It is unlikely, however, that *all* consistency problems would be hidden in this way. Regardless of the intent of the HMM sample, the fact remains that over one-fourth of the systems included never evolved beyond the prototype stage, either remaining as prototypes (18.8 percent) or not even reaching the prototype level (7.5 percent). If inability to perform was the primary cause of expert system non-use, some indication of this

Table 2. Perceived Performance Characteristics of Completed Systems (Counts of Systems)
[Survey Question 3.1, a-f]

Performance Criteria	Much Worse	Worse	No Change	Better	Much Better
Consistency	0	0	3	26	44
Quality	0	0	10	41	22
Frequency of errors	0	1	8	36	28
Performance time	0	10	14	19	30
Learning time	1	5	8	29	30
Costs of performance*	1	7	24	23	18

* Considered an economic, not a performance, factor.

fact should have emerged from these incomplete systems in the sample.

Another attempt to measure the degree to which systems routinely experienced performance failures was made by asking respondents to rank those factors that were "most limiting to the system's potential." Of the nine possible factors to choose from, four were clearly indicative of performance problems (i.e., not competent, too slow, too cumbersome, didn't enhance performance). If these four factors were commonly cited, a routine inability of the technology to meet task demands would have been indicated. As illustrated in Figure 5, however, only 10.8 percent of the respondents felt that performance-related limitations were "most limiting."

A final piece of evidence that performance-induced failures were not *de rigueur* for early expert systems comes from the qualitative side of the study, particularly from the free-form discussions with respondents that occurred after the formal protocol was completed. While there was a clear consensus among participants—particularly those still working in AI—that expert systems had fallen upon hard times, not a single respondent suggested that a common contributor was inability to build systems exhibiting acceptable performance. In fact, a number of respondents reported that it was their interest in this very enigma—that so many technically successful systems did not succeed in achieving long-term user-penetration—that induced them to participate in the survey.

Thus, the formal and informal findings compiled in this study suggest the following conclusion:

that *inability to perform tasks competently had only a minor impact on the user penetration and longevity of the applications in the survey.*

Economics

Another reason expert systems in the sample might have failed to achieve high levels of user penetration and longevity could have been inability to meet economic objectives. In the survey, such an effect would most likely manifest itself as a reluctance in managers to continue long-term maintenance, as all initial development costs would have already been invested for those systems in the sample that had evolved beyond the prototype stage (roughly three-fourths of the known systems).

To assess the economics of the systems in the catalog, respondents were specifically asked to agree or disagree with the statement that the expenses of development and maintenance significantly limited the potential of their system. They were also asked how use of the system affected the overall costs of task performance, which was intended to capture the new costs or savings associated with performing the task using the system, exclusive of development expenses.

There is some evidence suggesting that systems development and maintenance expense may have impacted user penetration and longevity for a number of the systems in the survey. Although, as noted earlier in Figure 5, only 13.5 percent of systems surveys listed such expense as the *primary* limitation to system potential, a substantial fraction of respondents (about one-third) agreed that expense was *a* factor that limited the
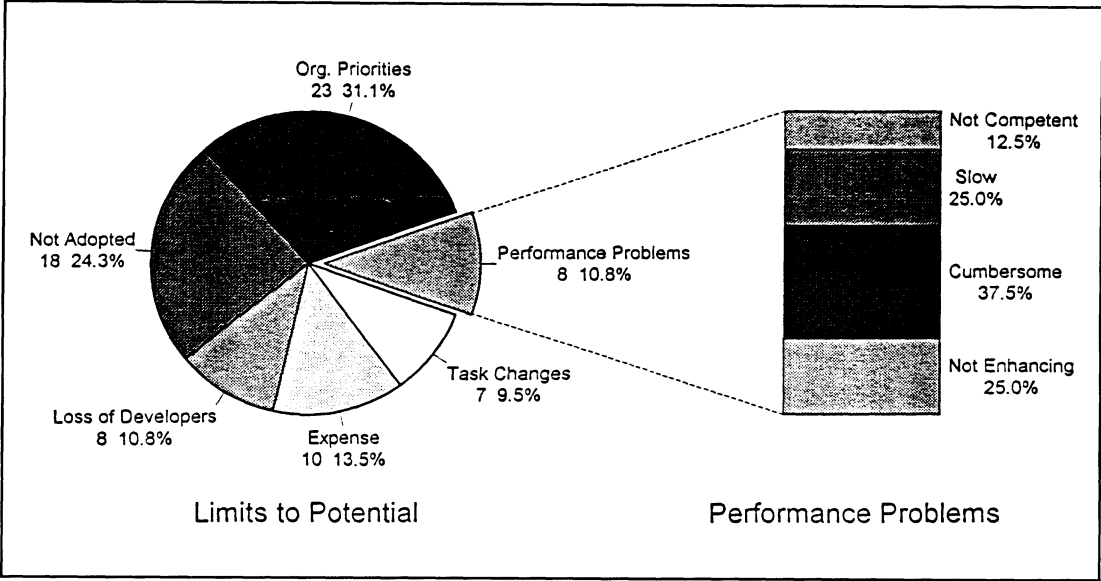
Org. Priorities
23 31.1%

Not Adopted
18 24.3%

Performance Problems
8 10.8%

Loss of Developers
8 10.8%

Task Changes
7 9.5%

Expense
10 13.5%

Not Competent
12.5%

Slow
25.0%

Cumbersome
37.5%

Not Enhancing
25.0%

Limits to Potential   Performance Problems

Figure 5. Maximum Limitations on Potential of Systems
[Survey Question 2.4, a-i]

system's potential. The breakdown of expense responses is presented in Figure 6. On the other hand, Table 2 shows that for only eight of 73 systems was post-system cost performance perceived to be worse than presystem performance.

The relationship between expense of maintenance and level of maintenance for the systems in the sample (the 72 for which both status and limitations to potential data could be gathered) is illustrated in Table 3. The results suggest that high expense is a good predictor that a system will *not* be maintained: of 29 systems for which maintenance expenses were thought to be a limitation, only four systems continued to be actively maintained or improved. In contrast, the fact that expenses did not appear to limit system potential was not a good predictor that a system *would* be maintained. Specifically, of the 43 systems where expenses did not appear to limit potential, 20 were not being maintained, and 23 were being maintained. Such a pattern suggests that achieving economic justification is a necessary, but not sufficient, condition for system success, consistent with the view presented in the literature (e.g., Davis, 1984).

The practical difficulties in acquiring accurate financial data on five-10 year-old systems precluded explicitly attempting to gather and quan-

tify system cost and benefit data. Using the available data, however, only about one-third of the systems appear to have been significantly affected by economic factors. Thus, there is evidence for concluding that *inadequate cost-benefit justification had a measurable impact on the rates of user penetration and longevity for some systems in the sample, but did not appear to be a significant factor for the majority of systems.*

Table 3. System Use vs. Expense
of Maintenance

	Not Being Maintained*	Being Maintained
Expense limited potential**	25	4
Expense didn't limit potential	20	23

* System has either been abandoned or is not being maintained at present.
** Either agreed or strongly agreed system was too expensive to maintain.
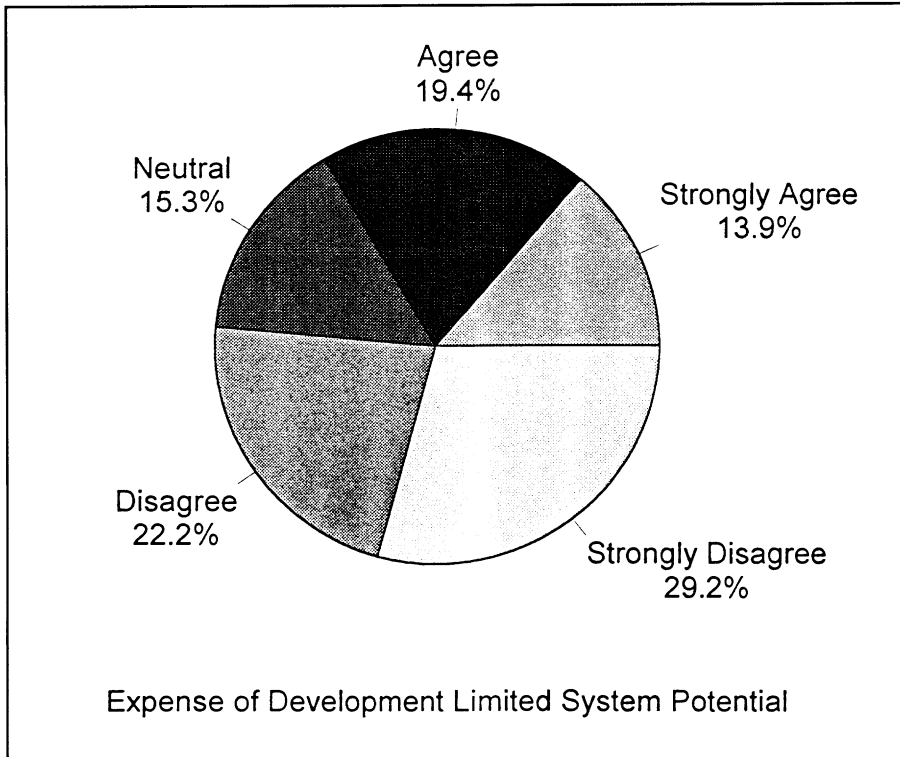Significance: $p < 0.01$ using chi-squared test.

**Figure 6. Limitations to System Potential Based
on Cost of Development and Maintenance**

Organizational and Individual Acceptance

With systems performance and economic factors not appearing to be critical determinants of usage for the systems surveyed, what of organizational issues? The most direct evidence that these issues were critical determinants of penetration and longevity can be gleaned from the question relating to the factor most limiting to system potential. As previously illustrated in Figure 5, failure of users to adopt the system as expected and changes in organizational priorities were cited as most critical by over 55 percent of the respondents. In addition, loss of developers— indicative of an organizational failure to supply a needed resource—added another 11 percent. Thus, two-thirds of all respondents felt they were most limited by factors under the control of the organization and users, rather than by the technology itself or cost/benefit issues.

The importance of the two factors in determining system usage is visibly evident when systems that achieved moderate or universal use are contrasted with those that achieved no use or limited use. As illustrated in Table 4, where respondents agreed that "failure of users to adopt" or "changes in organizational priorities" limited system potential, maximum levels of usage (i.e., user penetration) were almost universally weak. Where they disagreed, on the other hand, maximum levels of usage were much higher. (A very similar pattern was present for longevity measures, which—not surprisingly—proved to be highly correlated with user penetration measures.) Consequently, the evidence in the survey leads to the conclusion that *the most common barrier to achieving high levels of user penetration and longevity was the inability to achieve acceptance by the organization and users of the ES.*

A qualitative analysis of the survey responses further underscores the importance of organizational factors in long-term ES usage, as well as

Table 4. Failure of User Adoption and Change in Organizational Priorities Responses for Systems

Level of Agreement	System was limited by failure of users to adopt*		System was limited by change in organizational priorities?	
	# Weak Penetration**	# Strong Penetration	# Weak Penetration**	# Strong Penetration
Strongly Disagree	1	13	2	14
Disagree	2	15	2	13
Neutral	8	4	6	1
Agree	17	5	7	6
Strongly Agree	10	1	18	4

* Three systems exist for which "failure to adopt" responses were gathered, but not "changes to organizational priorities." These represent three of the systems refusing to participate, which were abandoned for lack of demand in the marketplace.
** Based on limited or no use being achieved.
Significance: $p < 0.001$ for both tables, using chi-squared test.

provides insights into their specific nature. Within the sample, 38 systems were fully constructed (i.e., were not classified as prototypes), but had either (a) never been used, (b) ceased being used, or (c) experienced significant declines in usage. The survey responses, system descriptions, and investigator notes were examined for each of these systems in order to determine the contributors to usage failures/declines. For each system, a primary contributor was then subjectively assessed. The examination led to the identification of 10 distinct types of primary contributors:

1. **Change in task:** For three systems, a change in the nature of the task eliminated or reduced need for the system. GTE's *COMPASS* system was designed to diagnose problems in an electronic switch that was being discontinued. General Research Corporation's *TIMM/TUNER* was designed to aid in adjusting sysgen parameters for a VAX/VMS system, a task rendered increasingly unnecessary as a result of hardware improvements in VAX computers. NORCOMM's *MASK* system was designed to help users with problems in using the company's screen I/O package, a task that was made unnecessary by improvements made in the subsequent release of the package.

2. **Costs of ongoing maintenance too expensive:** There were four systems for which the cost of ongoing maintenance was a major factor in the decline of use. Two of these systems, *Aquaref* and *ONCOCIN*, were public-sector systems, dependent upon grants and budgetary funding that proved insufficient (although, in the case of ONCOCIN—which is used to aid doctors in designing cancer treatments—the system has been resurrected several times, as funding sources and willing developers became available). For two other systems, Hewlett Packard's *IPT*, which diagnosed hard disk problems, and Honeywell's *PRESS*, which attempted to identify duplicate problems in a database of troubleshooting requests, the growth and dynamics of the task domain made ongoing maintenance too costly.

3. **System became misaligned with the company computing environment:** There were three systems that conflicted with the company's MIS environment. General Dynamic's *FAIS* system, developed to aid in scheduling numerically controlled machine equipment, was inconsistent with the MRP II system the company eventually implemented. Travellers Insurance's *DIAG8100*, intended to aid helpline operators in diagnosing IBM 8100 series computer faults, was rendered unusable when helpline terminal hardware was changed immediately before its intended deployment date. The *CLASS* system was

built to help users in determining the appropriate classification levels for Department of Energy (DOE) documents. Although the PC-based system was used briefly in the Albuquerque DOE office, that usage ceased when the agency headquarters in Washington, D.C. developed and mandated the use of a somewhat less sophisticated mainframe-based system.

4. **Change in company focus or industry outlook:** Three systems were impacted by changes in the company or industry-wide business situation. Radian's *TITAN*, designed to help technicians diagnose faults in TI-990 minicomputers, was never used because the company decided to discontinue servicing that type of computer just before the ES was completed. Beckman Instrument's *SPINPRO* system, designed to help scientists configure an ultra-centrifuge, was dropped from the company's product line because, as an inexpensive software product, it was inconsistent with the company's other products—mainly high-priced test equipment. KSI's *Mudman,* a system used to analyze oil drilling mud, was adversely affected by a dramatic decline in the oil industry in the late 1980s (a decline that also contributed to the discontinuation of development for several other systems in the sample).

5. **Failure to recognize size of task domain:** For two systems, the size of the task domain did not become apparent until initial development of the system was completed. The developers of TI's *Hotline Helper,* which diagnosed PC printer problems, were stymied by the sudden increase in the number and variety of different printers that were introduced in the mid-1980s. Palladian's *Management Advisor,* a tool designed to help CFOs in managing the corporate capital budgeting and treasury functions, found that the particular needs of each new client entailed major reworking of the system, making it nearly impossible (and prohibitively expensive) to sell and keep current.

6. **Solved a problem that wasn't perceived as critical by users:** Three systems appeared to solve problems for which there was little user demand. Two of these systems, Honeywell's

SYSCON and NCR's *Ocean,* were modelled after Digital's *XCON* and performed computer system configuration. Unlike Digital's VAX computers—which were extremely difficult to configure—the Honeywell and NCR systems were quite straightforward to configure, meaning that human configurors were already doing a nearly perfect job. As a result, the systems did not yield the magnitude of quality improvement benefits experienced when XCON was introduced, and both systems were eventually discontinued. Another system, Hartford Steam Boiler's (HSB) *TurboMac,* which diagnosed problems with turbo machinery of various types, was given away free to HSB customers. Demand was sufficiently low, however, that the company ceased to update the product.

7. **Subjected developer to potential liability:** One system was not used, in large part, because of potential liability concerns. The *PTE Analyst,* which was intended to aid attorneys in identifying prohibitions and possible exemptions for pension transactions covered under the ERISA act of 1974, was perceived to subject its developer to potential legal liabilities, leading to a decision not to market it.

8. **User resistance to externally developed systems (e.g., Not-Invented-Here syndrome):** For six systems, unwillingness of users to depend on systems developed elsewhere appeared to be the primary contributor to non-usage. APEX's *PlanPower* system was intended to help certified financial planners create detailed plans for high net-worth individuals. Because financial planning is highly subjective, and no clear "right" way to build a plan exists, potential customers resisted the expensive system, use of which would have forced them to change the way they did business. For three systems—the *Grain Marketing Advisor,* the *NAVEX* system, and the *Weld Scheduler*—respondents asserted that user resistance was largely a consequence of the systems having been developed outside of the intended user organization. All three were eventually abandoned. There is also evidence that such user resistance was particularly sensitive to the amount of expertise incorporated into a given system. Two systems in particular—the

GENESIS system and Schlumberger's *Dipmeter Advisor*—originally contained both high levels of embedded expertise *and* procedural routines to aid the users with the mechanics of data entry and processing. In both cases, the expertise components of the systems were ultimately removed as a means of gaining user acceptance.

9. **Unwillingness to take on development responsibilities:** For five systems, no group within the organization could be found to take on development responsibilities. In the case of three of these systems—Infomart's *SNAP,* GE's *Metals Analyst,* and GE's *DELTA*—users were intended to take over maintenance after initial development was completed. In each case, however, no user groups could be found willing to take on the responsibility, so the system was never used. In two other cases, Livermore's *TQMS/TUNE* and the Carnegie Group's *LISP-ITS,* developers were unwilling to take on, or became weary of, continued maintenance responsibilities, leading to a moratorium on maintenance and enhancements that ultimately translated into disuse.

10. **Loss of key development personnel:** For eight of the systems, turnover among development personnel was the primary reason that system use declined or ceased. In some cases, e.g., Dupont's *Ash Mixer,* Teknowledge's *WAVES,* Sperry's *Expert Probe,* and *Diagnostics,* the loss of personnel prevented the system from ever achieving significant usage. In other cases, e.g., Park Row's *CBT Analyst,* Honeywell's *PERMAID,* AIG's *Underwriter's Advisor,* and TI's *Capital Expert,* the loss of personnel precipitated the abandonment of a system that had been in active use. The figure of eight likely understates the impact of personnel loss on system success, however. Although not necessarily cited as the primary contributor, developer turnover was explicitly mentioned by respondents as a source of problems in 19 of the 38 cases.

Of these 10 explanations, the first two can largely be characterized in terms of performance and economics. The remaining eight, however, represent problems with a strong organizational com-

ponent. Specifically, they demonstrate the importance of coordinating ES development with the IT and business strategies of the firm (3 and 4), understanding the full scope and importance of the task to be performed by the system (5 and 6), recognizing the legal implications of a system (7), identifying user concerns and expectations (8), and managing developers and development responsibilities (9 and 10).

Analysis: Generalizability

The finding that managerial and organizational concerns outweighed technical and economic concerns for the systems in the sample could have significant implications for managers trying to prioritize their development efforts—*provided these results apply to today's systems.* There is, therefore, a strong need to consider the potential generalizability of the findings. Of particular concern are the answers to the following questions:

- How representative of today's commercial expert systems are the systems contained in the catalog?

- What was the sensitivity of use measures to specific technologies employed within the sample itself?

These concerns are addressed in the next sections.

Changes in commercial expert systems: mid-1980s to the present

How similar are the technologies present in the sample to today's technologies? If the question is posed purely in terms of "the art of the possible," then the answer would appear to be "not very." Great progress has been made in developing tools and applications for a number of new areas, such as case-based reasoning (e.g., Helton, 1991), model-based reasoning (e.g., Biswas, et al., 1993), and machine learning (e.g., Irani, et al., 1993). In addition, technologies that were in their infancy in the mid-1980s, such as neural networks (e.g., Keyes, 1991; O'Brien, 1993), genetic algorithms, and fuzzy logic (e.g., Karr, 1991), have now begun to produce commercial successes. Thus, it is reasonable to be concerned that today's systems may bear little resemblance to the systems in the period studied.

A recent field survey of 271 AI professionals (Stylianou, et al., 1992) directly addresses these concerns. The AI professionals responded to a questionnaire that contained a listing of features and other characteristics of ES tools, with respondents rating the relative importance of each. The findings of the survey suggest that today's *commercial* applications may not be so different in character from the applications contained in the HMM catalog. For example, the seven tool characteristics deemed most critical by the respondents were: (1) embeddability, (2) rapid prototyping, (3) backward chaining, (4) explanation facility, (5) ability to customize explanations, (6) linkage to databases, and (7) documentation comprehensiveness and readability. All of these represent capabilities that were not only available in the mid-1980s but were also represented in the catalog—which was specifically compiled with breadth of applications/technologies in mind (Harmon, et al., 1988).

The perceived importance of "traditional" features in no way suggests that emerging ES capabilities are not potentially important, or are not presently being used in *any* commercial systems. Rather, they suggest that *the majority* of commercial systems are still being developed using technologies comparable to those included in the catalog. Thus, the technologies in the catalog appear to remain directly relevant for many mainstream ES developers and managers.

Sensitivity of sample to technology issues

Despite the evidence that mid-1980s expert systems technologies continue to be utilized, a broader question can also be posed: How sensitive has the success of ES applications been to the specific technologies employed? If the answer is very sensitive, then the results of the first wave of commercial expert systems reported in this paper would not be particularly useful in understanding the systems of the future. If the results were not particularly technology-sensitive, on the other hand, then the findings for the first wave should not be invalidated solely by the introduction of new ES technologies.

As a test of sensitivity to technology, the correspondence between various technology-related variables provided in the HMM catalog and the three use measures gathered in the survey were

measured. The result of the analysis, detailed in Appendix C, showed only two significant types of relationship:

- **Task Type:** Several significant relationships were found between the type of task performed by the system and usage. The implications are that the type of task performed by a system (e.g., diagnosis, configuration, scheduling, monitoring) may indeed influence usage of the system.

- **Embedded Systems:** Expert systems embedded within conventional systems (i.e., enhanced conventional systems) showed significantly higher rates of maximum usage and current development than those of stand-alone systems.

Beyond these relationships, however, there were no significant relationships between the hardware, development tools, or knowledge representation method employed and any of the three usage measures.

Based on these findings, it is reasonable to conclude that sensitivity to specific technologies is not sufficiently great to render the survey results irrelevant to today's systems. Looking toward the future, the major caveat would be that *where a new technology facilitates performance of task-types that are very different from those permitted by earlier technologies, then the potential relevance of the current results is likely to be lower.* In other words, considerable caution would be required in extending the results of this paper to programming paradigms capable of performing tasks very different from those performed by traditional ES approaches. For example, neural nets and genetic algorithms—both of which employ computational techniques very different from symbolic reasoning—would not appear to be good candidates for generalization. On the other hand, many emerging technologies—such as CBR, model-based reasoning, and various uncertainty representation techniques—would seem better candidates for generalization, representing extensions to, rather than radical departures from, existing approaches.

Conclusions

Where are the early expert systems now? Of the systems for which information was available at the time of the survey:

- about one-third were being actively used and maintained,

- about one-sixth were still available to users but were not being maintained, and

- about one-half had been abandoned.

Regardless of the ultimate outcome experienced by each system, respondents perceived their respective systems had performed admirably. Ninety-six percent of respondents felt their systems improved the consistency of task performance. Eighty-six percent felt overall task quality had been improved by use of the system. Sixty-seven percent felt using the system improved task performance speed. On the economic side, perceptions were also positive, on the whole. Fifty-six percent of the respondents felt using the system reduced the costs of performing the task, and only 33 percent felt that the expenses associated with developing and maintaining their system significantly limited its potential usage.

Perhaps the key conclusion is that success of an ES in the technical or economic sense does not guarantee high levels of adoption or long-term use. Of the systems that did fall into disuse, many experienced problems of a non-technical, non-economic nature. Indeed, the systems in the sample appeared to be vulnerable to a host of managerial and organizational issues, such as:

- **Coordinating ES development with the IT and business strategies of the firm.** Developing expert systems typically entails a substantial investment in the performance of a specific task and often involves the use of technologies that are novel to the organization. It is therefore incumbent upon the manager to ensure that both task and technology are consistent with the organization's strategy and IS environment. Six systems in the survey were ultimately abandoned because the degree of fit was not adequate.

- **Understanding the task to be performed by the system.** Expert systems are frequently employed to perform tasks that are too complex for conventional technologies. As a consequence, particular care must be taken to ensure that the task is doable and that the benefits of using an ES justify its long-term maintenance. Five systems in the survey were ultimately abandoned because either the size

of the task domain was not recognized or because the system did not appear sufficiently useful to justify its continued maintenance.

- **Recognizing the legal implications of systems.** Another aspect of the complexity of expert systems is that "expert systems make mistakes" (Waterman, 1986, p. 29). Managers must be cognizant of the legal implications of developing and using systems that can never be fully tested. At least one system in the survey was abandoned because its developers feared potential legal liability if the system's advice was wrong.

- **Identifying user concerns and expectations.** Conventional systems, such as accounting applications, often automate tasks that are already routine and are performed at a fairly low level in the organization. Expert systems, in contrast, tend to be applied to the types of tasks performed by individuals with greater skills and higher positions in the organization. As a consequence, potential users may be particularly sensitive to a technology they perceive is intruding on their task domain, and, from a purely pragmatic standpoint, may be in a strong position to resist such technology. Six systems in the survey were ultimately abandoned because potential users were concerned about using systems they had not helped to develop.

- **Managing developers and development responsibilities.** Construction of expert systems typically requires substantial knowledge of both task domain and specialized development tools. As a consequence, developing and maintaining such systems entails acquiring and retaining individuals with very specific skills. The practical implications of being dependent upon such skills is that loss of even a single developer can mean that entire portions of a project have to be reworked. If managers do not take their vulnerability to developer turnover into account, serious delays or abandonment of a project can result. Of the systems in the survey, 13 were abandoned either because user organizations were unwilling to take over maintenance responsibilities or because key developers left the project.

While these issues clearly apply to the development of conventional systems as well, they appear to be particularly critical for expert systems.

Failure to address them, as the survey results clearly indicate, can lead to systems that fall into rapid disuse, if they are used at all.

Acknowledgements

Endnotes

[1] E.g., Buchanan (1986) describes some of the major obstacles to systems ever becoming truly expert. Prietula and Simon (1989) argue that true expertise may prove impossible to capture through the straightforward application of technology.

[2] Source: Author conversation with Patrick H. Winston.

[3] The fact that there were so many different respondent roles was viewed to be a potential source of survey error. An analysis of potential respondent bias was conducted and is described in Appendix C. The results of the analysis suggest that respondent role did not significantly impact survey responses.

References

Biswas, G., Manganaris, S., Yu, X. "Extending Component Connection Modeling for Analyzing Complex Physical Systems," *IEEE Expert* (8:1), February 1993, pp. 48-57.

Bobrow, D.G., Mittal S., and Stefik, M.J. "Expert Systems: Perils and Promise," *Communications of the ACM* (29:9), September 1986, pp. 880-894.

Brachman, R.J., Ameral, S., Engleman, C., Engelmore, R.S., Feigenbaum, E.A., and Wilkins, D.E. "What are Expert Systems?" in *Building Expert Systems,* F. Hayes-Roth, D.A. Waterman, and D.B. Lenat (eds.), Addision-Wesley, Reading, MA, 1983, pp. 31-57.

Buchanan, B.G. "Expert Systems: Working Systems and the Research Literature," *Expert Systems* (3:1), January 1986, pp. 32-51.

Buchanan, B.G., Barstow, D., Bechtal, R., Bennett, J., Clancey, W., Kulikovski, C., and Mitchell, T. "Constructing an Expert System," in *Building Expert Systems,* F. Hayes-Roth, D.A. Waterman, and D.B. Lenat (eds.), Addison-Wesley, Reading, MA, 1983, pp. 127-167.

Bulkeley, W. M. "Technology: Bright Outlook for Artificial Intelligence Yields to Slow Growth and Big Cutbacks," *Wall Street Journal,* July 5, 1990.

Coleman, K., "The AI Marketplace in the Year 2000," *AI Expert* (8:1), January 1993, pp. 34-38.

Davis, D. "Artificial Intelligence Goes to Work," *High Technology* (7:4), April 1987.

Davis, R. "Amplifying Expertise with Expert Systems," in *The AI Business,* P.H. Winston and K.A. Prendergast (eds.), MIT Press, Cambridge, MA, 1984, pp. 17-40.

DeLone, W.H. and McLean, E.R. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), 1992, pp. 60-95.

Freedman, D.H. "AI Meets the Corporate Mainframe," *Infosystems* (34:2), February 1987, pp. 32-37.

Freundlich, Y. "Transfer Pricing: Integrating Expert Systems in MIS Environments," *IEEE Expert* (5:1), February 1990, pp. 54-62.

Gevarter, W.B. "The Nature and Evaluation of Commercial Expert System Building Tools," *Computer* (20:5), May 1987, pp. 24-41.

Gill, T.G. "Texas Instruments: Capital Expert System," Harvard Business School Case Study, Case No. N-9-188-050, Cambridge, MA, 1988.

Gill, T.G. *Expert Systems: A Mapping Between Symbol Susceptible Tasks and Software Tools,* unpublished Ph.D. dissertation, Harvard Business School, Boston, MA, 1991.

Harmon, P. and King, D. *Expert Systems: Artificial Intelligence in Business,* Wiley, New York, NY, 1985.

Harmon, P. and Sawyer, B. *Creating Expert Systems for Business and Industry,* Wiley, New York, NY, 1990.

Harmon, P., Maus, R., and Morrissey, W. *Expert Systems: Tools and Applications,* Wiley, New York, NY, 1988.

Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. *Building Expert Systems,* Addison-Wesley, Reading, MA, 1983a.

Hayes-Roth, F., Waterman, D.A., Lenat, D.B. "An Overview of Expert Systems," in *Building Ex-*

pert Systems, F. Hayes-Roth, D.A. Waterman, and D.B. Lenat (eds.), 1983b, pp. 3-29.

Helton, T. "The Need for Case Based Reasoning: CBR Express," *AI Expert* (6:10), October 1991, pp. 55-59.

Irani, K.B., Cheng, J., Fayyad, U.M., and Qian, Z. "Applying Machine Learning to Semi-Conductor Manufacturing," *IEEE Expert* (8:1), February 1993, pp. 41-47.

Jackson, P. *Introduction to Expert Systems,* Addison-Wesley, Reading, MA, 1986.

Johnson, T., Hewett, J., Guilfoyle, C., Jeffcoate, J., and Goodwin, C. *Knowledge-Based Systems: Markets, Suppliers and Products,* Ovum Ltd., London, U.K., 1989.

Karr, C. "Applying Genetics to Fuzzy Logic," *AI Expert* (6:3), March 1991, pp. 55-59.

Keen, P.G.W. and Scott Morton, M.S. *Decision Support Systems: An Organizational Perspective,* Addison-Wesley, Reading, MA, 1978.

Keyes, J. "Getting Caught in a Neural Network," *AI Expert* (6:7), July 1991, pp. 44-49.

Luconi, F.L., Malone, T.W., and Scott Morton, M.S. "Expert Systems: The Next Challenge for Managers," *Sloan Management Review* (27:4), Summer, 1986, pp. 3-14.

O'Brien, L. "Throwing a Net over Wall Street," *AI Expert* (8:4), April 1993, pp. 19-23.

Prereau, D.S. "Selection of an Appriopriate Domain for an Expert System," *The AI Magazine* (5:3), Summer 1985, pp. 26-30.

Pressman, R.S. *Software Engineering: A Practitioner's Approach,* McGraw-Hill, New York, NY, 1982.

Prietula, M.J. and Simon, H.S. "The Experts in Your Midst," *Harvard Business Review* (67:1), January-February 1989, pp. 120-125.

Silverman, B.G. (ed.). *Expert Systems for Business,* Addison-Wesley, Reading, MA, 1987.

Silverman, B.G. "Should a Manager Hire an Expert System," in *Expert Systems for Business,* B.G. Silverman (ed.), Addison-Wesley, Reading, MA, 1987, pp. 5-23.

Slatter, P.E. *Building Expert Systems: Cognitive Emulation,* Ellis Horwood, LTD, Chichester, U.K., 1987.

Stapleton, L. "Embedding Intelligence: The New AI Paradigm," *Computer Language* (5:8), August 1988, pp. 97-98.

Stefik, M., Aikins, J., Balzer, R., Benoit, J., Birnbaum, L., Hayes-Roth, F., and Sacerdoti, E. "The Architecture of Expert Systems," in

Building Expert Systems, F. Hayes-Roth, D.A. Waterman, and D.B. Lenat (eds.), 1983, pp. 89-126.

Stylianou, A.C., Madey, G.R., and Smith, R.D. "Selection Criteria for Expert Systems Shells: A Socio-Technical Framework," *Communications of the ACM* (35:10), October 1992, pp. 30-48.

Sviokla, J.J. *XCON, Planpower and Mudman: An In-Depth Analysis,* unpublished Ph.D. dissertation, Harvard Business School, Cambridge, MA, 1986.

Turban, E. *Expert Systems and Applied Artificial Intelligence,* Macmillian Publishing, New York, NY, 1992.

Turban, E. and Watkins, P.R. "Integrating Expert Systems and Decision Support Systems," *MIS Quarterly* (10:2), June 1986, pp. 121-136.

Tyran, C.K. and George, J.F. "The Implementation of Expert Systems: A Survey of Successful Implementations," *Data Base* (24:4), Winter 1993, pp. 5-15.

Walker, T.C. and Miller, R.K. *Expert Systems 1990,* SEAI Technical Publications, Madison, GA, 1990.

Waterman, D.A. *A Guide to Expert Systems,* Addison-Wesley, Reading, MA, 1986.

Waterman, D.A. and Hayes-Roth, F. "An Investigation of Tools for Building Expert Systems," in *Building Expert Systems,* F. Hayes-Roth, D.A. Waterman, and D.B. Lenat (eds.), 1983, pp. 169-215.

Winston, P.H. "Perspective," in *The AI Business,* P.H. Winston and K.A. Prendergast (eds.), 1984, pp. 1-13.

Winston, P.H. and Prendergast, K.A. *The AI Business,* The MIT Press, Cambridge, MA, 1984.

About the Author

T. Grandon Gill is an assistant professor in the Decision and Information Systems Department of the College of Business at Florida Atlantic University, Boca Raton, FL. He received his D.B.A. from Harvard Business School in 1991. His areas of research interest are expert systems, task activities, and complexity. He has also served as a nuclear submarine officer, managed a small agricultural chemicals firm, consulted extensively in the agribusiness sector, and developed a number of commercial software products, including expert systems.

Appendix A
Data Gathering Protocol

Overview

Identifying and gathering data on the systems surveyed occurred in a series of stages consisting of:

- Gathering background data
- Locating a knowledgeable respondent
- Administering the questionnaire
- Respondent verification

Each of these stages is described below.

Gathering Background Data

The first stage of data gathering involved locating sources of background information relating to each system. The purpose of acquiring such data was to ensure that the researcher was as knowledgeable as possible when it came time to administer the questionnaire and also to provide information useful in locating each system. The specific techniques employed were the following:

- **Database search:** A search was conducted using the ABI/Inform database for each of the 111 systems in the catalog. All relevant articles were acquired.

- **Books on expert systems:** Approximately 25 books on expert systems and five general MIS text-books were examined, and references to the systems in the sample were copied.

- **Application surveys and catalogs:** References to each system were identified in a series of commercial application surveys and catalogs, including Walker and Miller's *Expert Systems '90,* Ovum's *Knowledge-based Systems: Markets, Suppliers and Products* (1989), Waterman's *Guide to Expert Systems* (1986), and Buchanan's "Expert Systems: Working Systems and the Research Literature" (1986).

- **Vendor marketing materials:** Application summaries, written during the mid-1980s, were gathered from all relevant AI/ES vendors who were still in business.

Copies of all references were filed in separate folders prepared for each system. Where a system was found to be classified or developed abroad, no further action was taken.

Identifying Respondents

Once references had been identified for as many systems as possible, attempts were made to identify individuals with sufficient knowledge of task and system to respond to the survey. The techniques used to identify respondents consisted of:

- Contacting the company where the system was developed or used
- Contacting the expert system tool vendor, where a tool was specified
- Tracking down developers or contact people in the AAAI directory, where a name was given in HMM or in the other references
- Contacting authors of journal articles where references to missing systems were found

- Enlisting help of other survey participants using the phone and Internet and through a mailing sent to all participants near the end of the survey.

Using these techniques, appropriate respondents for 81 of the 97 non-classified U.S. systems in the survey were identified.

Administering the Questionnaire

Once a respondent agreed to participate, a time was scheduled for administering the phone survey questionnaire (see Appendix B), a process that usually took from 30 minutes to an hour. Where possible, a written description of the system was prepared (from published sources) prior to the interview. During the course of the interview, after a series of preliminary multiple-choice questions relating to the system's current status and the respondent's relationship to the system, the nature of the system was discussed, and a written description of the task and system was prepared (typically 250-500 words). Obtaining the system description from the respondent served two purposes: to ensure an accurate write-up of the system was collected and to better acquaint the investigator with the nature of the system. Upon completion of the system description, the balance of the formal interview involved answering a series of multiple-choice questions relating to the nature of the system and the task it performed. Variations from the protocol occurred for some systems that had never been completed or never been used (i.e., prototypes). In such cases, usage questions were omitted where the investigator felt there was insufficient experience to reasonably answer the questions.

Throughout the interview, respondents were encouraged to discuss their answers to the questions, and the investigator prompted the respondents to explain any answers that seemed inconsistent with the task and system description. In some cases, respondents would change their answers as a result of these discussions. Upon completion of the formal survey questions, respondents were encouraged to further discuss the state of expert systems in their company, both at the time the specific system was developed and at the present time.

Verification

After each questionnaire was completed, the investigator entered the system information into a database, which automatically prepared a copy of the system descriptions and, later, a copy of investigator-prepared responses to additional survey questions. These were then mailed to respondents for verification, and corrections and additions were made as required. Upon completion of the survey, the investigator prepared two additional items for verification:

- A 40-page written report summarizing the findings;

- A stand-alone PC-based database program, written in Clarion, which allowed users to access system descriptions and access summary statistics for the systems surveyed.

Respondents were also provided with a form that allowed them to offer further feedback. Where appropriate, such feedback was incorporated into the survey database.

Appendix B
Expert Systems Status Questionnaire

Part One: Respondent Questions

System:

Code:

Name of Respondent:

Phone Number:

Can it be used in data base?   Yes No
Callback OK?   Yes No

Company Address:

(At time of acquaintance with system)

Which of the following describes the role in which you became familiar with the system *(all that apply)*:
- D.   Software developer during construction   E.   Task expert
- M.   Manager (development or current)   S.   Software support role
- U.   User of finished system
- R.   Observed system as a researcher
- L.   Researched system in the literature
- O.   Other (or comments): _____

How many years were you involved with the system?: _____
How many years has it been since your last experience with the system: _____
Copy of database (dBase and ASCII formats)?   3.5"   5.25"   None

Part Two: System Status Questions

1. Which of the following best describes the maximum degree of usage which was achieved by the system:
   a. Unfinished prototype
   b. Completed prototype of application, never adopted within the organization
   c. Application completed and achieved only limited adoption by the organization or targeted customers
   d. Application completed and achieved moderate levels of adoption
   e. Application completed and achieved widespread or universal adoption among intended users
   f. Other (or comments): _____

1a. *[If completed and used (c., d. or e.)]* What was the first year the system was used? _____

2. Which of the following best describes the current development and maintenance status of the system:
   a. Abandoned
   b. Active development not being pursued at the present time, although further development is possible
   c. Application maintenance being performed but no efforts to upgrade system are in progress
   d. Application continues to be maintained and improved
   e. Other (or comments): _____

2a. *[If abandoned or not being developed (a. or b.)]* In what year did development stop? _____

3. Which of the following best describes the current usage status of the system:
   a. Never intended to be used
   b. Never been used
   c. Not presently in use
   d. In use by number of users which has declined significantly from an earlier maximum level
   e. In use by a stable number of users at or near the maximum level
   f. In use by a number of users which continues to grow
   g. Other (or comments): _____

3a. *[If not presently in use (c)]* In what year did usage stop? _____

4. You will now be given a series of possible explanations for why the system may not have reached its maximum potential use. For each statement, indicate whether you:
   Strongly Disagree (1), Disagree (2), Are Neutral (3), Agree (4) or Strongly Agree (5)
   that the explanation is valid for your system.
   a. System was unable to perform or support the task competently   ____
   b. System too slow to meet the demands of task   ____
   c. System proved too difficult or cumbersome to operate   ____
   d. Changes in the nature of the task rendered the system obsolete   ____
   e. System proved too difficult or expensive to maintain   ____
   f. Loss of development personnel forced canceling the project   ____
   g. System did not enhance performance of the task to the degree expected ____
   h. Users failed to adopt (or purchase) system as expected   ____
   i. Organizational priorities ceased to support system development and use ____
   j. Other (or comments): _____
   (Circle explanation which is most important)

5. Approximately how much developer time has been expended on the system to date?
   a. < 3 Mo.   b. 3 Mo. to < 1 year   c. 1 year to <5 years   d. 5 years to 20 years   e. >20 years

6. Approximately how long would it take an individual to acquire the expertise embedded in the system during the normal course of training and task performance?
   a. < 3 Mo.   b. 3 Mo. to < 1 year   c. 1 year to <5 years   d. 5 years to 20 years   e. >20 years

Part Three: System Objectives Questions

1.   For the following factors which describe task performance, describe how using the system affected the task on the following scale:

Made performance much worse (1), Worse (2), Unchanged (3), Better (4), or Much better (5)

a.	Level of consistency for task output or solutions	____
b.	Average quality of task output or solutions	____
c.	Number of errors made while performing the task	____
d.	Time or effort required to perform the task	____
e.	Time or effort required to train someone to perform the task	____
f.	Costs associated with performing the task	____
g.	Amount of boring or repetitive work associated with the task	____
h.	Amount of work required of key personnel who are task experts	____

2.   Are there any other major objectives of the system?

3.   Briefly describe the task performed by the system

Part Four: System and User Roles

You will now be given a series of ten statements, which will help us distinguish the activities of the users and the system. For each, indicate whether you strongly disagree, disagree, are neutral, agree or strongly agree that it describes your system and the task performed.

1.   The task performed by the system is similar to the task performed by humans before the system was built

Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

2. In the course of performing the task, the system is in control and the user-- if any-- acts primarily in the subsidiary role of information provider and monitor .

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

3. The individuals who performed the task before the system was developed became the users of the system when it was completed.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

4. The availability of the system significantly increases the number of individuals who can successfully perform the task.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

5. While performing the task, the user must continually direct the system to perform whatever specific activities are desired.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

6. Most of the original users of the system were also involved in its development.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

7. The system applies techniques or algorithms that would not or could not have been used by experienced human task performers

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

8. Use of the system significantly reduces the skills or knowledge required to successfully perform the task.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

9. The system performs the task using the same type of approaches that an experienced individual would use to perform the task

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

10. Experienced task performers were instrumental members of the team which developed the system

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

Part Five: Task Characteristics Questions

The following statements, which will help us understand how the system has changed the task being performed, come in pairs. The first describes the task before the system was built, the second, the changes brought about by using the system.

1. Prior to using the system: a substantial percentage of the task performer's time was spent on routine procedures which did not vary much from task to task.

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

2. After the system was adopted: how was the percentage of time spent on routine procedures affected?

     Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

3. Prior to using the system: time pressures played a significant role in determining the outcome of the task

     Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

4. After the system was adopted: how was the degree to which time pressures impacted the task affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

5. Prior to using the system: completing the task required the cooperative participation of several individuals

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

6. After the system was adopted: how was need for cooperation between different task performers affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

7. Prior to using the system: the task performer often had to make important decisions regarding "what to do next" in the course of performing the task

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

8. After the system was adopted: how was the number of important "what to do next" decisions to be made by the task performer affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

9. Prior to using the system: experienced task performers almost always came up with results very similar or identical to those produced by other experienced performers

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

10. After the system was adopted: how was the variability in the results achieved by different task performers affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

11. Prior to using the system: task performers could readily gauge the quality of their task performance

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

12. After the system was adopted: how was the task performer's ability to gauge the quality of his or her performance affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

13. Prior to using the system: even after an acceptable conclusion to the task had been reached, the task performer could choose to expend additional effort to improve upon the results

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

14. After the system was adopted: how was the task performer's ability to improve results by applying additional effort affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

15. Prior to using the system: the task was arranged such that a single task performer would not be involved in the task from beginning to end

    Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

16. After the system was adopted: how was the ability of an individual to be involved in the task from beginning to end affected?

    Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

17.   Prior to using the system: The key elements of the task solution, such as decisions which were made and conclusions which were reached, were at the discretion of the task performer

         Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

18.   After the system was adopted: how was the degree to which the task performer could control the task solution affected?

         Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

19.   Prior to using the system: how well an individual performed the task had a significant impact upon the work of others or upon the quality of the product or service being delivered.

         Strongly Disagree(1)----Disagree(2)---Neutral(3)---Agree(4)---Strongly Agree(5)

20.   After the system was adopted: how was impact of how well an individual performed affected?

         Significantly Reduced(1)---Reduced(2)---Unchanged(3)---Increased(4)---Significantly Increased(5)

Part Six: Researcher Coded Questions

These questions were coded by the researcher, during the course of entering the data on your system into the database. The answers coded for your system are in **bold** type.

If you could take a moment to ensure that no great inaccuracies have been introduced, this would be greatly appreciated.

1. For which of the following purposes was the system developed?
   For the internal use of the organization which initiated the system
   For sale by the organization which initiated the system
   For the internal use and for sale by the organization which initiated the system
   *** Note: "For Sale," in this context, includes providing the system to customers free of charge.

2. Who was the original developer of the system?
   The organization which initiated the system
   A party contracted by the organization which initiated the system
   Jointly, the organization which initiated the system and a contracted party

3. Who was responsible for maintaining the system?
   The organization primarily responsible for initial system development
   An organization which used the system, but was not the initial primary developer
   A third party which neither initiated nor uses the system
   Jointly, the organization which used or initiated the system and a contracted party
   No significant maintenance was ever performed after development

4. Where did the system get the information it needed for a specific problem while it was running?
   The information was supplied by the user
   The information was supplied by connected systems
   Jointly, by the user and by connected systems

5. What types of databases were accessed by the system?
   No databases, or internal databases requiring minimal updating
   Internal databases requiring routine updates
   Databases kept current by external systems
   Both internally and externally updated databases
*** Note: Databases which are relatively static in nature (e.g., a periodic table of the elements) or databases used as a convenient alternative to writing rules are differentiated from databases containing time sensitive information in this question.

6. What automated data inputs were received by the system (excluding database accesses in-
itiated by the system)?
   None
   Real-time (or data) inputs from other computer systems
   Real-time (or data) inputs directly from other equipment
   Real-time (or data) inputs from both computers and directly from equipment

7. What automated data or signal outputs were provided by the system?
   None, only outputs were to user
   Real-time (or data) outputs to other computer systems
   Real-time (or data) outputs to other equipment
   Real-time (or data) outputs to both computers and equipment

8. About how often would your "typical" intended user have consulted, or received information
from, the system?
   Many times a day
   About once a day
   Once a week or more
   Once a month or more
   Once a quarter or more
   Once a year or more
   Less than once a year

*** Note 1: Where a system is primarily intended for one-time use (e.g., a tutoring system), the frequen-
cy of use should reflect the period over which the system would be used. The same applies for project-
oriented systems. Otherwise, it should reflect an average access frequency over a sustained period of
user-time.

*** Note 2: The values should reflect the usage of an individual user—not the system—if many individuals
use the same copy of the system.

Directions: If no values are in bold, then the researcher could not judge an appropriate response to the
question. In such circumstances, a call or note would be greatly appreciated. Otherwise, for inaccurate
information, please contact:

[Omitted for purposes of publication]

Appendix C
Statistical Tests of Generalizability

Respondent Bias

Respondent bias would be present if the nature of the respondent significantly influenced survey
responses. Of particular concern in this regard was the possibility that different categories of respondents
might perceive performance of their systems differently (e.g., users might systematically regard soft-
ware quality as lower than developers did). As a test for such bias, each of the performance questions
in the survey (Question 3.1, a-h) was regressed against the five respondent-type dummy variables. The
results are presented in Table C1.

The analysis shows that respondent type appears to bear little relationship to perceived performance,
as indicated by survey responses. Of all the regression coefficients, only two exhibit significance at the
five percent level: developers appear to perceive a slightly lower system quality (-0.27), while managers
perceive quality to be slightly higher (+0.28). Thus, there is no evidence that variations in respondent
type exerted a major influence on the perceptions of system performance reported in the survey.

Table C1. Coefficients of Regression of Respondent Type Against Performance Responses (Questions 3.1 a-h)

	Developer	Expert	Manager	User	Support
Consistency	− .08	+ .15	− .07	− .12	+ .19
Quality	− .27*	+ .05	+ .28*	+ .20	+ .01
Error frequency	− .18	− .05	− .20	− .18	+ .04
Performance time	+ .06	− .54	+ .44	+ .39	− .56
Training time	+ .12	− .03	+ .14	− .23	+ .00
Costs	+ .11	+ .36	+ .21	+ .13	+ .34
Boring/repetitive work	− .22	+ .30	+ .15	+ .17	+ .15
Demands on key personnel	+ .12	− .03	+ .14	− .23	+ .00

*Results of 1-tailed t-test significant at .05%.

Note: Positive coefficient indicates respondent perceived performance measure to be better, on average.

Technological Sensitivity

Technological sensitivity would be present if the specific nature of the technology employed in building an application had a significant impact on the use of the application. A test of such sensitivity can be made using the usage data gathered in the survey and the technology classifications provided for all systems in the HMM catalog. The test is conceptually simple: the more sensitive user penetration and longevity are to the specific nature of the development technology, the less generalizable the results are likely to be to future technologies.

Task and technology variables provided in the HMM catalog were coded and OLS regression performed against the three usage measures provided in the survey (i.e., maximum usage, current development status, and current usage status). The specific independent variables used were as follows:

- **Application Size:** Coded as 1 = Small, 2 = Medium, 3 = Large.

- **Task Type:** Four task types—monitoring, configuration, planning, and scheduling—were coded as dummy variables. The fifth, diagnosis, was omitted for reasons of linear independence and therefore constituted the base case.

- **Development Shell:** Three development shells—shell, language, hybrid environment—were coded as dummy variables.

- **Representation:** Three representations— rules, induction and frames—were coded as dummy variables.

- **Function:** Two interfaces—front end to existing application and enhanced conventional system (i.e., embedded)—were coded as dummy variables. The third, standalone, was omitted for reasons of linear independence and therefore constituted the base case.

- **Input:** One input source—signal processing—was coded as a dummy variable. The second, dialog, was omitted for reasons of linear independence and therefore constituted the base case.

- **Delivery Hardware:** Three types of delivery hardware—minicomputer, AI workstation, and mainframe—were coded as dummy variables. The fourth, PC, was omitted for reasons of linear independence and therefore constituted the base case.

A summary of regression coefficients and significance is presented in Table C2.

Table C2. Regression Coefficients for HMM Task and Technology Variables

Independent Variable	Maximum Use	Current Development	Current Use
Size (1 = Small, 2 = Medium, 3 = Large)	− .09	− .18	− .49
Monitoring	− .16	+ .86	+ .55
Configuration	+ .73	+ 1.04*	+ 1.20*
Planning	+ .33	+ .44	+ .22
Scheduling	− 1.35*	− .80	− 1.01
Shell	− .29	− .15	− .98
Language	+ .19	+ .19	− .49
Environment	+ .31	− .48	− .68
Rules	− .79	− .31	− .83
Induction	− .67	− .37	− 1.33
Frames	− 1.11	− .59	− .88
Front-end to existing system	+ .85	+ .94	+ .79
Enhanced conventional system	+ 1.12*	+ 1.25*	+ 1.06
Signal processing	+ .09	+ .11	+ .22
Minicomputer/workstation	+ .27	+ .62	+ .58
AI Workstation	− .34	− .12	− .20
Mainframe	+ .91	+ .93	+ 1.22

* − 5% significance, using 1-tailed t-test.

The analysis of coefficients shows the following significant ($p \leq .05$) relationships exist in the sample:

- **Task Type:** Several significant coefficients were found for task types configuration (positive) and scheduling (negative).

- **Embedded Systems:** Expert systems embedded within conventional systems (i.e., enhanced conventional systems) showed significantly higher rates of maximum usage and current development than those of standalone systems.

Beyond these relationships, however, there were no significant relationships between the hardware, development tools, or knowledge representation method employed and any of the three usage measures.