A MULTI-DIMENSIONAL APPROACH TO PERFORMANCE EVALUATION FOR I/S DEVELOPMENT

Jay G. Cooprider
John C. Henderson

December 1989

CISR WP No. 197
Sloan WP No. 3103-89

Center for Information Systems Research
Sloan School of Management
Massachusetts Institute of Technology

A Multi-Dimensional Approach to Performance Evaluation for I/S Development

Jay G. Cooprider
John C. Henderson

ABSTRACT

Information systems are a major force in today's organization. Unfortunately, the development of information systems remains very labor-intensive and quite difficult to manage. It is not surprising that I/S management is searching for improved methods for evaluating the performance of I/S developers. This paper makes an initial step in the evaluation of the I/S function by presenting an approach for evaluating the performance of I/S development units and describing its use in a large international technology manufacturing firm.

We first present a diagnostic, behavior-based model for measuring software development from a management perspective. This model proposes measures of development processes and products from the task, social, and organizational levels of analysis. We then describe the application of this model in a large international technology manufacturing firm. Data Envelopment Analysis (DEA) is used as a technique for applying the model to the firm's performance data. The results of the DEA analysis is then used to investigate the performance impacts of various management policies. This evaluation approach provides a systematic method for evaluating development performance. It highlights the importance of using a range of behavior-based measures for evaluating performance, and it illustrates a methodology for examining performance based on such measures.

A Multi-Dimensional Approach to Performance Evaluation for I/S Development

1.0 Introduction

Information systems are a major force in today's organization. They impact how business is conducted and directly affect organizational success. They enable problem-solving capabilities that could not have been imagined even a few years ago. It is not surprising that management is searching for improved methods for evaluating the processes used in information system (I/S) development and the resulting products. Enhanced evaluation capabilities are required both to identify opportunities for improved development practices and to establish the feedback necessary to manage the development process.

The development of information systems remains very labor-intensive and, unfortunately, quite difficult to manage. Despite over thirty years of experience, the management and measurement of software development remains problematic (Boehm 1987, Brooks 1987). Many have argued that a first step in mastering the management of I/S development is to identify and develop methodologies to allow its description, measurement, and understanding (Conte et al. 1986, Boehm 1987). There are a number of reasons why such methodologies have been slow to win acceptance, but perhaps the most significant is that a definitive causal model of I/S development does not exist (Brooks 1987, Curtis et al. 1988, Elam and Thomas 1989). The development environment is rich in alternative processes for converting I/S resources into usable applications. However, these various processes are not well understood. There is no systematic method for examining the relationship between the resources used and the products delivered (Boehm 1987). The very concept of I/S "productivity" is not well understood (Brancheau and Wetherbe 1987).

One factor that has limited the ability of organizations to effectively evaluate the I/S function is the lack of a systematic approach to measurement using multiple inputs and outputs (Boehm 1987). Many organizations have gathered numerous productivity measures but are finding it difficult to analyze the collected data in a meaningful way. This research makes an initial step in the evaluation of the I/S function by presenting an approach for evaluating the performance of I/S development units and relating its use in a large international technology manufacturing firm. We first present a model for measuring software development from a management perspective. We then demonstrate the use of Data Envelopment Analysis (DEA)[1] to operationalize the model, using data collected from a group of software development sites in a large technology firm. The potential of this measurement approach for evaluating I/S management practices is then discussed. Finally, we discuss the contributions of this research and recommend directions for further study.

2.0 Background

> "You cannot measure what is not defined. You also cannot tell whether you have improved something if you have not measured its performance." (Strassman 1985, page 100)

Previous research attempting to evaluate I/S development performance has suggested a wide range of possible measures (Conte et al. 1986, Boehm 1987, Banker et al. 1987). However, this research has not generally provided the management insight necessary to improve the development process. Hirschheim and Smithson (1987) state that I/S evaluation has been misdirected toward tools and techniques for measurement and away from understanding. For example, many researchers have based their productivity measurements on the number of lines of code produced for a given amount of effort. This approach has long been criticized because of a

[1] DEA is a linear programming-based technique that converts multiple input and output measures into a single comprehensive measure of production efficiency. See Section 5.0 for a description of DEA and its use.

number of obvious problems. For example, each line of source code does not have the same value and does not require the same amount of time or effort to produce. However, most software productivity researchers still use lines of code as a primary focus (Boehm 1987). While performance measures such as "source lines of code per month" are relatively easy to gather and may provide some insight into the basic production processes of software development, their narrow focus may distort the overall productivity picture and cause managers to overlook promising avenues for performance improvement. The ultimate goal of I/S development is to provide applications that benefit the organization, not to create lines of code.

3.0 A Performance Evaluation Model

One potentially significant measurement approach is to provide a mechanism for performance analysis that is *diagnostic* and *behavior-based.* Since no definitive causal model of the software development process exists, a model of software development performance should not be expected to provide management with unambiguous signals. Rather, a development performance model should be used as a means for management to "diagnose" the development process: to gain awareness of trends, relationships, and useful classification schemes. A diagnostic measurement model for management serves to meet these needs by tending to provoke interest, raise issues, reveal preferences, and help surface assumptions (Einhorn and Hogarth 1982, Epstein and Henderson 1989).

A management measurement model should also provide insights into the behaviors underlying the processes being measured. Mason and Swanson (1979) recommend an approach toward measurement for management that extends traditional "scientific" measurement to include behavioral dimensions. Boehm (1981) demonstrates that the perceptions of upper management, middle management, and programmers are very different concerning the factors that have the most leverage for affecting development productivity. These perceptions can be important motivational factors for affecting I/S development productivity and they should be reflected in any performance measurement system. Buss (1987) asserts that top executives, I/S managers, and users often have conflicting views of which project proposals should get approval. These conflicting views can be mitigated to some degree if the language (or measures, in this context) used by each of these perspectives is the same, or is at least translatable. In order to improve the management of the I/S development process, a measurement system tied to management of the actual behaviors of the process is required.

3.1 Process and Product

Figure 1 presents a general measurement model for the performance of I/S development units that reflects both process and product measures. Many researchers have suggested the usefulness of separate measures for the products and processes of I/S development. For example, Agresti (1981) uses an industrial engineering perspective to propose process and product measures of software development. Since a primary goal of performance measurement is to improve the production processes involved, not only must the final product of the development effort be evaluated, but the processes used to obtain that product must also be considered. Case (1985) argues that it is important to use *both* types of measures because there is a potential conflict between the efficiency of the process and the quality of the product.

This process-product dimension reflects the more general traditions of measurement in organizational control theory. Ouchi (1979) and Eisenhardt (1985), for example, have categorized control measures as either behavior (process) based or outcome (product) based. The measurement model presented in Figure 1 can be viewed as an application of these ideas to the I/S development arena.

	Process	Product
Task Measures	Efficiency	Quality
Social Measures	Dependability	Validity
Organizational Measures	Responsiveness	Value

Figure 1: A Model for Measuring I/S Development Performance

3.2 Level of Analysis

A second major issue in performance evaluation is the level or unit of analysis represented in the model. I/S research in general can be characterized in terms of its unit of analysis (Curley and Henderson, 1989). In essence, this characterization of the development process leads to different behavioral perspectives that must be reflected by different measures. Curtis et al. (1988) conducted a field study of large software development projects as part of an effort to understand critical performance factors. They found that behavioral processes in such projects can be analyzed from three perspectives: cognitive, social, and organizational.

Curtis et al.'s cognitive perspective reflects a production view of I/S development. We model this perspective as a *task* dimension. Task measures (Figure 1) are indicators of the accomplishment of the developer's defined tasks. They are generally individual oriented and are concerned, from a management perspective, with "doing things right." *Task process* measures are primarily used to evaluate the efficiency of production. As such, they represent measures typical of the economic perspective of efficiency (Packer 1983). Source lines of code or function points per work-month are typical I/S measures of this perspective. *Task product* measures evaluate the output of this production process. These measures often reflect the quality of the software product. Quality may be measured in terms of defects per unit output (e.g. bugs per thousand lines of code). Alternatively, quality measures may reflect the product's overall technical performance rather than its defect count. Examples of this form of task product measure include run-time speed and object code size.

These task-oriented measures dominate the I/S performance literature (Hirschheim and Smithson 1987). However, there are a number of problems with focusing exclusively on these types of measures. In general, they are measures of output (in an economic sense) rather than outcome (the total impact of the process). They are much more closely aligned with efficiency (how well input is converted to output) than with effectiveness (how the input is used to accomplish the goals of the organization). The addition of social and organizational perspectives (see Figure 1) provides a more complete performance evaluation framework.

Hirschheim and Smithson (1987) argue that both technical (task-oriented) and non-technical (social) criteria must be included for I/S evaluation to be meaningful. Social measures generally evaluate how well the development team performs as a whole. They are significant because members of development teams typically work as groups rather than as separate individuals. These measures are group oriented and, from a management perspective, are concerned with "doing the right things." *Social process* measures indicate how well members of the development team function together. From the standpoint of I/S development, the primary issue is process dependability. Internal to the team, these measures involve team maintenance (Gladstein 1984) and include such things as commitment, quality of work life, and satisfaction with the team. From an external perspective, the issue is the perceived dependability of the team. This is largely a function of the group's track record -- its record of success or failure on past projects. Frequently used measures include the ability to meet delivery schedules and financial commitments.

Social product measures evaluate design validity. That is, does the development team meet project requirements that actually satisfy a real business need? These measures can be critical in providing what Bjorn-Andersen (1984) refers to as a "challenge to certainty." He states that I/S developers tend to spend more and more time and use ever more refined technological tools for solving the wrong problem more precisely. Social product measures must provide the ability to test design validity, and the best source of these tests comes from communication with non-team stakeholders (Churchman 1971, Henderson 1987). Measures for this perspective include the users' evaluation of I/S importance to their jobs and actual system usage. System usage has been operationalized in several ways and has long been advocated as an indicator of the extent to which I/S is meeting business requirements.

Performance measurement occurs within an organizational context, and, hence, should include explicit evaluation from an organizational perspective. Churchman (1968) refers to this as the "systemic" level of measurement: placing a manager's problem situation in a larger context. Very little progress has been made in discovering definitive organizational measures of performance (Crowston and Treacy 1986, Packer 1983). In this paper, the measures of I/S development performance from an organizational perspective involve the contribution of I/S to business success.

Organizational process measures are primarily concerned with the ability of I/S to quickly respond to changes in business need. The motivation for this view is the need to manage organizational resources in the face of significant environmental changes. There are two common approaches for these measures. The.first approach evaluates user satisfaction. Cyert and March (1963) argue that satisfaction is a reasonable surrogate measure for how individuals perceive that their needs are being met. Nolan and Seward (1974), Elam and Thomas (1989), and others have argued that user satisfaction is one way to measure the responsiveness dimension of I/S performance.

The second perspective focuses on the organizational resource of *time*. The main concern in this perspective is the speed with which the organization reacts to changes in the environment. A fundamental measure in this perspective is time-to-break-even: the length of time it takes an organization to convert an I/S concept into a viable product that has earned back its development costs for the organization. This measure highlights the criticality of rapid response to environmental changes while directly incorporating aspects of quality and maintainability. It is rapidly becoming a key measure of I/S performance (Boddie 1987).

The *organizational product* issue is fundamentally one of business value. It is concerned with whether I/S products contribute to the success of the organization and/or improve its competitive position. Two common measures of this dimension are market share (did the system result in a position of competitive advantage for the organization?) and benefit/cost ratios (what was the payoff of the application?). These measures generally result from business and competitive analysis.

4.0 Operationalization of the Model

The performance evaluation model illustrated in Figure 1 was used as a guide for analyzing data collected during a major measurement initiative of a large international technology manufacturing and marketing organization. The firm had collected data on over sixty measures of I/S development productivity at each of its I/S development sites for a number of years. However, I/S management had no systematic way to use this data to evaluate development performance. The measurement model described above provided management with a framework for analyzing the collected data.

The data used in this research had been collected by the firm from 67 I/S development sites in 1986 and 1987. The sites for which data was collected represent I/S functional groups operating at a business unit/staff level. The average size of each development site is 114 employees, ranging from 9 to 513. The developers at each site report to a distinct management team and have a clear user clientele. These sites vary in terms of user type (i.e., marketing, R&D, corporate staff, etc.) and geographic region (a significant number of non-US sites are included). The data was collected as part of a corporate measurement initiative and involved a management review, clarification, and sign-off procedure.

As might be expected, all sites did not report every data element. Further, due to organizational changes, some sites were not represented for both 1986 and 1987. The data reported here includes only those sites that reported data for both years. An average of each site's data for these two years is used to reduce the effects of random variation across years (analysis of data for units reporting data in at least one of these two years results in a substantial increase in sample size, and the results are consistent with this more conservative approach, thereby indicating the analysis is fairly robust).

The inputs and outputs used for the performance analysis of each site are described below and are illustrated in Figure 2. While one may well imagine indicators that would enhance the operationalization of this model, these measures reflect the data that could be obtained for a reasonable cost. The potential gains from enhancing these measures is discussed in Section 7.0 (Conclusions).

INPUTS: 1. Direct Development Costs
2. Operations Support Costs

OUTPUTS:

	Process	Product
Task Measures	Efficiency Function Points per Work-Month	Quality 1/(Defect Months per Installed Function Point)
Social Measures	Dependability Percentage of Commitments Met	Validity Perception of I/S Importance
Organizational Measures	Responsiveness User Satisfaction	Value Benefits/ Costs

Figure 2: Operationalization of a Model for Analyzing I/S Development Performance

4.1 Inputs

The first input is the site's direct development costs. Each site was asked to report its annual I/S expense for development activities. Development was defined as activities that introduce business change, including studies and experiments, and enhancements that deliver new function or capabilities to improve business processes. This amount included all consulting and application delivery expense, including vendor expenses but excluding operations support for I/S development.

The second input is the site's operations support costs. Each site was asked to report its annual I/S operations expense for supporting I/S development activities. This amount included all of the production support and computer operations expense (people, facilities, and machines) required to support I/S development activities.

Both inputs were normalized for site size by dividing the reported costs by the number of employees at the site working on I/S development activities. The major portion of the direct development expense is labor costs, while the major share of the operations support expense is facility and machine costs. These two inputs represent surrogate measures for management's choice between investments in labor or capital. In this context, each site's manager must decide to invest in development people (labor: direct development costs) or in additional machine power or software tools (capital: operations support costs).

4.2 Outputs

Efficiency. The first output is function points per work-month. Each site was asked to report the total function points delivered during the year and the work-months expended to develop/deliver those function points. Projects to be included were new development, enhancements, and the selection, modification, and installation of shared or purchased applications delivered during the year. The output was calculated by dividing the total function points delivered by the work-months expended for them.

Quality. The second output is 1/(Defect Work-Months per Installed Function Point), and is used as a measure of the quality of the products that have been produced by the site. Each site was asked to report the work-months spent during the year to correct application defects. Application defects were defined as deviations from user approved specifications that require unique corrections to the applications in the installed application base (including purchased and shared applications). This excluded work to implement suggested changes and user errors. The reported work-months were divided by the total function points installed at the site at year-end to normalize for site size. The inverse of this result was then calculated so that this output would positively correlate with the models' inputs (an assumption of the Data Envelopment Analysis technique described below).

Dependability. Direct measures of quality of work life were not available. As an indicator of this social process variable, the percentage of the development commitments met is used. The rationale for this is that positive social process characteristics within a development group should result in the group's increased dependability: a better track record. Each site was asked to report the total work-months spent on I/S development work (whole projects, releases, or phases) that met its schedule, cost, and function commitments. This reported value was divided by the site's reported total work-months for all I/S development work with commitments scheduled for completion in the current year.

Validity. The fourth output measures the importance of the I/S applications to the user community. If developers are solving the right business problems -- if their solutions have validity -- they will be providing applications that are important to the user. Each user was asked the question "How important is information processing to you in doing your job?" The value reported for this output is the percentage of respondents for the site that answered "of utmost importance" or "of considerable importance" to this question.

Responsiveness. The fifth output measures the level of satisfaction of the users with their I/S applications. The motivation for this output is that user satisfaction should be higher for more responsive I/S units. Each user was asked the question "Overall, how satisfied are you with the Information Processing services you use?" The value reported for this output is the percentage of respondents for the site that answered "very satisfied" and "satisfied" to this question.

Value. The final output measures the contribution of I/S to the organization by measuring the ratio of expected gross benefits to related I/S development expenses for projects completed during the year. Each site was asked to report a summary of the benefits and expenses of its business cases for projects completed during the current year. The benefits and expenses included applicable business case adjustments (e.g., the time value of money), and they were reviewed and approved by the controller of the reporting site. The benefits and expenses reported are the planned total aggregate gross benefits and the planned total aggregate related I/S expenses for the full period of the business cases for the projects completed in the current year. The benefits are divided by the expenses to calculate the ratio used for this output.

5.0 Analysis of Performance Data: DEA

The data collected by the firm for 1986 and 1987 was analyzed using Data Envelopment Analysis (DEA). DEA is a linear programming-based technique that converts multiple input and output measures into a single comprehensive measure of productive efficiency (Charnes, Cooper, and Rhodes 1978). This methodology uses observed outputs and inputs for each unit in an analysis to derive an empirically-based efficiency frontier rather than using a production function to generate a theoretical one. The choice of DEA for analyzing the performance of I/S development units is motivated by the need to simultaneously consider multiple inputs and outputs and to not impose an arbitrary parametric form on the underlying software production processes. In addition, a distinguishing property of DEA is that it optimizes on each unit in the analysis. This is in contrast with ratio analyses which use no optimizing principle and regression-based techniques where optimization is across all observations and results in an average (rather than extremal) value. This attribute makes DEA a meaningful tool for management use and for efficiency evaluation purposes. DEA has been applied to a number of I/S-related settings (Elam and Thomas (1989), Banker, Datar, and Kemerer (1987), Chismar and Kriebel (1986), Kauffman and Kriebel (1988)).

DEA accomplishes its analysis by the construction of an empirically based production frontier and by the identification of peer groups. Each unit is evaluated by comparisons against a composite unit that is constructed as a convex combination of other decision-making units (DMUs) in its peer group (Banker and Morey 1986), and is assigned a DEA efficiency rating between zero and one. The efficiency rating is a ratio of a weighted sum of outputs to a weighted sum of inputs, with the weights associated with the inputs and outputs determined uniquely by solving a linear programming model that assigns the most favorable weights possible. If the computed weights result in an efficiency ratio of one, the unit is on the frontier. The reader should review Charnes, Cooper and Rhodes (1978), Bessent, Bessent, Elam, and Clark (1988), and Epstein and Henderson (1989) for a description of the details of DEA analysis.

5.1 Applying DEA to the Performance Models

The evaluation model operationalization illustrated in Figure 2 above was utilized to analyze the performance of the reporting sites. In order to accomplish this, three separate DEA models were used. Each model reflects one of the levels of analysis of the overall performance model. By separating the models in this way, it is possible to analyze the performance and management practices of each site for each dimension independently. Figure 3 illustrates the inputs and outputs of each of the three models.

Model Name:	Outputs:
Task Model	1. Function Points per Work-Month (Efficiency) 2. 1/(Defect-Months per Installed FP) (Quality)
Social Model	1. Percentage of Commitments Met(Dependability) 2. User Opinion of I/S Importance (Validity)
Organizational Model	1. User Satisfaction With I/S (Responsiveness) 2. Benefits/Costs Ratio (Value)
All models have 2 inputs: 1. Direct Development Costs 2. Operations Support Costs	

Figure 3: Inputs and Outputs for Three DEA Performance Models

Initial analysis of the collected data revealed significant variations for a given site between years. This is due in part to the nature of the data collected from the sites. The function point data and the benefit/cost data, for example, were reported for a project only when the entire project was completed. Therefore, a two-year project would have no data reported during its first year and would have all of its data reported in year two. Therefore, an average of the data for the two years was used to "smooth" the values for the DEA analysis.

If a site had data for a variable for both 1986 and 1987, then the data for both years were averaged. If the site were missing data for either 1986 or 1987 (or both), it was considered to have missing data and was excluded from the analysis. A total of thirty sites had enough variables with both 1986 and 1987 data to be used in at least one of the three models described in Figure 3. Effective response rates for each model ranged from .28 to .42. While this is reasonable for survey research, it clearly suggests the generalization of this analysis, even to this single firm, must be limited.[2] Follow-up interviews with the firm's senior management indicated that the response rate was not the result of a systematic performance bias. Rather, the consistent feedback was that the organization's extensive data collection initiative had not previously been combined with an analysis process usable by site managers to improve performance. As a result, site management viewed the collection effort strictly as a *cost* rather than a benefit, and, therefore, many sites did not participate.

A specialized software package was used to perform the DEA analyses for this study. Using the measures described above, each of the models was run for the I/S development sites for which data was collected. Table 1 shows the results of the analyses for sites with both 1986 and 1987 data. Two points should be made about the results illustrated in this table. First, it should be noted that extremely few sites in the models have values that are inefficient and enveloped. Though there are a number of approaches for analyzing unenveloped sites with DEA (Bessent et al. 1988), there is some debate concerning the ability of DEA to reliably assess the performance of such unenveloped sites (Epstein and Henderson 1989). Second, in each of the models there are generally a relatively few sites which are efficient or nearly efficient, and a large gap exists between the efficiency ratings of those sites and the ratings of the other sites. Banker, Gadh, and Gorr (1989) suggest that DEA analysis can suffer greatly when large measurement errors are present in observed data, especially for data sets with a small sample size. There was concern that such errors might be artificially raising the efficiency ratings of the few firms on the DEA frontier.

---

[2] As indicated earlier, additional runs were made using sites that had data from a single year. This greatly increased the sample size but did not lead to significant changes in the analysis results. Therefore, this conservative approach is taken.

Table 1: DEA Efficiency Ratings
All Sites With Data Are Included

Site No.	Task Model	Social Model	Organization Model
1	0.51*	0.40*	0.53*
2	**	0.46*	0.39*
3	0.10*	0.69*	0.55*
4	0.08*	0.43*	**
5	0.03*	0.25*	0.23*
6	0.07*	0.39*	**
7	0.20*	0.39*	0.42*
8	**	0.81*	1.00
9	**	0.43*	**
10	0.37*	**	**
11	0.27*	1.00	1.00
12	0.23*	0.27*	0.22*
13	0.16*	0.55*	0.48*
14	1.00	0.55*	**
15	0.12*	0.57*	**
16	0.13*	0.39*	**
17	0.48*	0.48*	**
18	0.10*	**	1.00
19	0.12*	0.25*	0.30*
20	0.34*	1.00	0.94*
21	0.33*	0.46*	0.69*
22	0.84*	1.00	1.00
23	**	0.32*	**
24	**	0.52*	0.44*
25	**	0.36*	0.29
26	**	0.50*	0.47
27	**	0.23*	**
28	**	0.14*	**
29	1.00	1.00	1.00
30	**	0.29*	0.70*

* denotes a site which is inefficient and not enveloped
** denotes a site which lacked complete data for the model

Because of these factors, there was concern that the derived DEA frontier was not truly representative of the larger set of sites. Data outliers can have a sizable adverse impact on a DEA analysis (Epstein and Henderson 1989). After review with the organization's senior I/S management, it was determined that the efficient sites had unique characteristics (generally, they were exceptionally small sites and/or had exceptionally low wage rates, for example). To address this situation, these unique sites were removed for further analyses. For the purposes of this study, sites with efficiency ratings greater than .8 were removed from the data set. Twenty-seven sites were used in at least one of the DEA models, and the results of this second DEA analysis are presented in Table 2. It should be noted that a larger percentage of sites are enveloped and that there is a more even distribution of efficiency ratings. These results are used for all further analyses. It should be noted that some recent developments in DEA (stochastic DEA (Banker 1989), for example) have addressed this issue of outliers and measurement error, but they were not used here due to the study's exploratory nature and small sample size.

The significance of the different dimensions of performance are highlighted by the results shown in Tables 1 and 2. In Table 2, for example, Site 1 has high efficiency ratings for each of the models. This can be interpreted as indicating that Site 1 has high performance on each of the dimensions of performance. Site 13, on the other hand, has high efficiency for every model but Task. Management attention at that site might be directed toward improving basic production capabilities -- through the use of CASE tools or reusable code modules, for example. Site 5 has relatively low performance on each of the dimensions, and requires more general management attention.

Table 2: DEA Efficiency Ratings
"Outlier" Sites Removed From Each Model

Site No.	Task Model	Social Model	Organization Model
1	1.00	0.81*	0.96
2	**	0.76	0.79*
3	0.41	1.00	1.00
4	0.25	0.81*	**
5	0.09*	0.42*	0.44*
6	0.20	0.71*	**
7	0.56*	0.70	0.85
9	**	0.81*	**
10	1.00	**	**
11	1.00	**	**
12	0.80*	0.41*	0.42*
13	0.34*	1.00	1.00
14	**	1.00	**
15	0.40*	1.00	**
16	0.41*	0.74*	**
17	0.95*	1.00	**
18	0.31	**	**
19	0.30	0.47*	0.52
20	1.00	**	**
21	0.65*	0.96*	1.00
23	**	0.60*	**
24	**	1.00	0.91*
25	**	0.56*	0.62
26	**	0.82*	1.00
27	**	0.43*	**
28	**	0.29*	**
30	**	0.47*	1.00

* denotes a site which is inefficient and not enveloped
** denotes a site which lacked complete data for the model

6.0 Analysis of Management Practices

Further analyses were carried out to provide management with additional information for judging management practices. DEA efficiency models directly relate inputs to outputs. What is missing from such analyses is any evidence of direct links between performance ratings and specific management practices. To establish this link, data was gathered on the investment in education and strategic planning processes found at each site. In essence, these two variables are treated as controlled policy variables by the local I/S manager. The analysis is intended to assess how these management practices relate to the overall performance of the I/S function. Specifically, the two policy variables are:

Education. The annual number of student days invested in formal I/S technical and career development education was collected for each site. Formal education included classroom, guided self-study, and remote satellite education. This measure was collected in two parts: technical education and career education. Technical education included all formal education on I/S subjects. Career education included fundamental business process education not related to I/S (e.g., advanced management education, manufacturing, sales training, quality, etc.). These values were divided by the total number of I/S employees at the site to obtain a value for average technical student days per employee, career student days per employee, and total student days per employee.

Strategic Linkage. The general or line manager(s) at each site reported an evaluation of the involvement of the I/S organization in strategic planning activities. This ranking was on a scale from 0 to 4, with '0' meaning that no I/S strategy existed, and '4' indicating that an I/S strategy existed that was developed by joint "pro-active" participation of I/S and business function participants.

In order to analyze the impact of these policy variables on the different perspectives of performance analysis, a simple correlation was calculated between the values of each policy variable and of the DEA efficiency ratings for each of the three models. From this analysis, the relationships between the policy variables and the development performance dimensions can be discussed.

Figure 4 shows the correlations between the DEA efficiency ratings of the models and the education policy variables. Several tentative conclusions can be drawn from these correlations. First, education significantly correlates with performance. This significance holds for the task and social models, and the organizational model has a large positive correlation but with a sample size too small to establish statistical significance. This is clearly indicative of the important role that education can play in improving the performance of I/S development units. This result is consistent with a range of research reporting on the effect of education on I/S performance (Chrysler 1978, Lutz, 1984, Jones 1986, Banker et al. 1987). It also bears mentioning that *career education*, in particular, strongly correlates with each dimension of performance. This result highlights the value of educating I/S developers in business functional areas as well as technical topics, corroborating similar findings by Banker et al. (1987).

Figure 5 shows the correlations between the DEA efficiency of the sites and their reported values for strategic linkage. While all correlations are positive, it is somewhat surprising that the only relationship that is statistically significant is between the line managers' assessment of strategic linkage and the task model efficiency ratings. It might have been expected that the social and, especially, organizational models would show significant correlations instead. Given the limited ability to obtain multiple measures, one must be careful not to eliminate potential problems in the measures as an explanation for these low correlations. Regardless, the positive correlations at all levels are indicative of the significance of strategic linkage to overall I/S development performance, even if the statistical significance of the relationship is not established.

	Correlations Between the Models' DEA Efficiency and the Average Number of Days per Employee of:		
Model Name	**Technical Education**	**Career Education**	**Total Education**
Task Model	0.53 * (n=17)**	0.52 * (n=13)	0.53 * (n=17)
Social Model	0.48 * (n=23)	0.53 * (n=21)	0.54 * (n=23)
Organizational Model	0.41 (n=13)	0.38 (n=11)	0.41 (n=13)

* $p<.05$

** Sample eliminates missing data for every variable included in this model.

Figure 4: Correlations between DEA Efficiency & Education

Correlations between the Models' DEA Efficiency Rating and the Site's Line Managers' rating of:

Model Name	**Strategic Linkage**
Task Model	0.69 * (n=16)
Social Model	0.11 (n=20)
Organizational Model	0.25 (n=13)

* $p<.05$

Figure 5: Correlations between DEA Efficiency & Strategic Linkage

7.0 Conclusions

This paper has presented an approach for evaluating the performance of I/S development units. The approach begins by establishing an overall model for analyzing developing performance. This model uses two dimensions: process-product and level of analysis (task, social, and organizational). DEA is used as a technique for applying the model to an actual business situation. In order to apply DEA, the overall model is divided into separate models by level of analysis, and each model is analyzed. The results of these analyses were then correlated with various management policy variables to investigate the impact of specific policies on the various dimensions of performance.

This approach provides a systematic method for evaluating development unit performance and establishes a way to assess the performance of individual units as well as the impact of particular management practices. It highlights the importance of using a range of behavior-based measures for evaluating performance, and it illustrates a technique for examining performance based on such measures.

Since this approach is comparative, the issue of appropriate input and output measures is critical. While the measures selected for this study provide useful examples, additional effort is clearly required to refine these measures and to develop additional measures. In particular, measures from an organizational perspective need further development, particularly in the area if I/S responsiveness. Similarly, more attention could be given to the effective measurement of the policy variables. Still these results suggest that efforts to obtain better measures could be further justified by the ability to incorporate them in a meaningful assessment approach.

Generalizations about the specific results of this study or about the effectiveness of the approach have to be limited given the small sample size and the fact that only a single organization was involved. However, the I/S senior management team did evaluate and use these findings. As a result, management has committed to incorporate this approach in future performance analysis efforts. In addition, the results have served as a major motivation for future data collection efforts in the organization. This response is evidence that the approach can be a major element of larger programs to measure and improve I/S processes.

Bibliography

Agresti, W. "Applying Industrial Engineering to the Software Development Process," in *Proceedings: IEEE Fall COMPCON*, IEEE Computer Society Press, Washington D.C., 1981, pp. 264-270.

Banker, R. "Stochastic Data Envelopment Analysis," *Management Science*, forthcoming.

Banker, R., Datar, S., and Kemerer, C. "Factors Affecting Software Maintenance Productivity: An Exploratory Study," *Proceedings of the Eighth International Conference on Information Systems*, Pittsburgh, PA, December 1987, pp. 160-175.

Banker, R., Gadh, V., and Gorr, W. "A Monte Carlo Comparison of Efficiency Estimation Methods," Working Paper, Carnegie Mellon University, 1989.

Banker, R. and Morey, R. "The Use of Categorical Variables in Data Envelopment Analysis," *Management Science* (32:12), December 1986, pp. 1613-1627.

Bessent, A., W. Bessent, J. Elam, and T. Clark. "Efficiency Frontier Determination by Constrained Facet Analysis," *Operations Research* (36:5), May 1988, pp. 785-796.

Bjorn-Andersen, N. "Challenge to Certainty," in *Beyond Productivity: Information Systems Development for Organizational Effectiveness*, T. Bemelmans (ed.), North-Holland, Amsterdam, 1984.

Boddie, J. *Crunch Mode: Building Effective Systems on a Tight Schedule*, Prentice Hall, Englewood Cliffs, NJ, 1987.

Boehm, B. "Improving Software Productivity," in *Proceedings: IEEE Fall COMPCON*, IEEE Computer Society Press, Washington D.C., 1981, pp. 264-270.

Boehm, B. "Improving Software Productivity," *IEEE Computer* (20:9), September 1987, pp. 43-57.

Brancheau, J. and Wetherbe, J. "Key Issues in Information Systems Management," *MIS Quarterly* (11:1), March 1987, pp. 23-45.

Brooks, F. "No Silver Bullet: Essence and Accidents of Software Engineering," *IEEE Computer* (20:4), April 1987, pp. 10-19.

Buss, M. "How to Rank Computer Projects," in *Information Analysis: Selected Readings*, R. Galliers (ed.). Addison-Wesley, Sydney, 1987.

Case, A. "Computer-Aided Software Engineering." *Database* (17:1), Fall 1986, pp. 35-43.

Charnes, A., Cooper, W., and Rhodes, E. "Measuring the Efficiency of Decision Making Units," *The European Journal of Operational Research* (2:6), 1978, pp. 429-444.

Chismar, W. and Kriebel, C. "A Method for Assessing the Economic Impact of Information Systems Technology on Organizations," *Proceedings of the Sixth*

International Conference on Information Systems, Indianapolis, December 1986, pp. 45-56.

Chrysler, E. "Some Basic Determinants of Computer Programming Productivity," *Communications of the ACM* (21:6), June 1978, pp. 472-483.

Churchman, C. "Suggestive, Predictive, Decisive, and Systemic Measurement." Paper presented at the Second Symposium on Industrial Safety Performance Measurement, National Safety Council, Chicago, Dec. 1968.

Churchman, C. *The Design of Inquiring Systems*. Basic Books, 1971.

Conte, S., Dunsmore, H., and Shen, V.. *Software Engineering Metrics and Models*. The Benjamin/Cummings Publishing Co., Menlo Park, CA, 1986.

Crowston, K. and Treacy, M. "Assessing the Impact of Information Technology on Enterprise Level Performance," in *Proceedings of the Seventh International Conference on Information Systems*, San Diego, CA, December 1986, pp. 299-310.

Curley, K. and Henderson, J. "Evaluating Investments in Information Technology: A Review of Key Models," *Proceedings of the 1989 ACM SIGOIS Workshop on the Impact and Value of Information Systems*, Minneapolis, MN, June 1989.

Curtis, B., Krasner, H., and Iscoe, N. "A Field Study of the Software Design Process for Large Systems," *Communications of the ACM* (31:11), November 1988, pp. 1268-1286.

Cyert, R. and March, J. *A Behavioral Theory of the Firm*. Prentice-Hall, Englewood Cliffs, NJ, 1963.

Einhorn, H. and Hogarth, R. "Prediction, Diagnosis, and Causal Thinking in Forecasting," *Journal of Forecasting* (1:1), 1982, pp. 23-26.

Eisenhardt, K. "Control: Organizational and Economic Approaches," *Management Science* (31:2), February 1985, pp. 134-149.

Elam, J. and Thomas, J. "Evaluating the Productivity of Information Systems Organizations in State Government Environments," *Public Productivity Review* (12:3), Spring 1989, pp. 263-277.

Epstein, M. and Henderson, J. "Data Envelopment Analysis for Managerial Control and Diagnosis," *Decision Sciences* (20:1), Winter 1989, pp. 90-119.

Gladstein, D. "Groups in Context: A Model of Task Group Effectiveness," *Administrative Science Quarterly*, Vol. 29, 1984, pp. 499-517.

Henderson, J. "Managing the IS Design Environment," Center for Information Systems Research Working Paper No. 158, Sloan School of Management, M.I.T., Cambridge, MA, May 1987.

Hirschheim, R. and Smithson, S. "Information Systems Evaluation: Myth and Reality," in *Information Analysis: Selected Readings*, R. Galliers (ed.), Addison-Wesley, Sydney, 1987.

Jones, T.C. *Programming Productivity*. McGraw-Hill, New York, 1986.

Kauffman, R. and Kriebel, C. "Measuring and Modeling the Business Value of IT," in *Measuring Business Value of Information Technologies*, ICIT Research Study Team #2, ICIT Press, Washington, D.C., 1988.

Lutz, T. *Foundation for Growth -- Productivity and Quality in Application Development*. Nolan, Norton and Company, Lexington, MA, 1984.

Mason, R. and E. Swanson. "Measurement for Management Decision: A Perspective," *California Management Review* (21:3), Spring 1979, pp. 70-81.

Nolan, R. and Seward, H. "Measuring User Satisfaction to Evaluate Information Systems," in *Managing the Data Resource Function*, R. Nolan (ed.) , West Publishing Co., Los Angeles, 1974.

Ouchi, W. "A Conceptual Framework for the Design of Organizational Control Mechanisms," *Management Science* (25:9), September 1979, pp. 833-848.

Packer, M. "Measuring the Intangible in Productivity," *Technology Review*, February/March 1983, pp. 48-57.

Strassman, P. *Information Payoff. the Transformation of Work in the Electronic Age*. The Free Press, New York, 1985.

Date Due