Quantitative Measurement of Successful Performance from the Project Manager's Perspective

Cindy L. Menches, P.E., M.ASCE[1]; and Awad S. Hanna, P.E., M.ASCE[2]

Abstract: This paper describes a process for converting a project manager's *qualitative* evaluation of "successful performance" to a *quantitative* measurement. In a recent study, randomly selected electrical companies throughout the United States submitted two projects for analysis: (1) one project that was well planned and performed *successfully*; and (2) one project that was poorly planned and performed *less-than-successfully*. Overall, 55 projects were evaluated. Success, however, is a subjective concept that often depends on the team member providing the definition. Consequently, during the interview process, project managers were asked to provide their *own* definition of successful performance, and this definition was matched to the collected data to identify appropriate variables for inclusion in a performance measurement index. A reliability analysis was performed, and six variables were ultimately used to construct the index, including actual percent profit, percent schedule overrun, amount of time given, communication between team members, budget achievement, and change in work hours. A cross-validation technique was used to evaluate the performance measurement index, and a misclassification rate of 2% was observed, which indicated that the model was useful for quantitatively measuring successful performance based on the project managers' definitions of success.

DOI: 10.1061/(ASCE)0733-9364(2006)132:12(1284)

CE Database subject headings: Planning; Project management; Performance characteristics; Measurement; Construction industry.

Introduction

A recent article published by Civil Engineering magazine (Bernstein 2003) recognized that there are few consistent measures of project success. Further, although numerous studies have focused on the *identification* of critical success factors and key performance factors, few have actually developed techniques for *measuring* successful performance.

A study, sponsored by the Foundation for Electrical Construction, aimed to develop a model preconstruction planning process that can be used by electrical contractors to improve their planning practices and achieve more successful outcomes. The study consisted of three distinct phases that critically evaluated the relationship between preconstruction planning and successful project performance. As the fundamental goal of the research was to evaluate and quantify the relationship between planning and performance, the analysis began with the development of a quantitative measure of *successful performance*.

[1]Assistant Professor, Dept. of Civil, Architectural and Environmental Engineering, The Univ. of Texas at Austin, 1 University Station C1752, Austin TX 78712-0273 (corresponding author). E-mail: menches@mail.utexas.edu

[2]Professor and Chair of Construction Engineering and Management Program, Dept. of Civil and Environmental Engineering, Univ. of Wisconsin-Madison, 2314 Engineering Hall, 1415 Engineering Dr., Madison, WI 53706. E-mail: hanna@engr.wisc.edu

Note. Discussion open until May 1, 2007. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on September 29, 2005; approved on June 7, 2006. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 132, No. 12, December 1, 2006. ©ASCE, ISSN 0733-9364/2006/12-1284–1293/$25.00.

As part of the data collection process, participating contractors were asked to submit two projects for analysis: (1) one project that was well planned and performed *successfully*; and (2) one project that was poorly planned and performed *less-than-successfully*. In general, owners, designers, general contractors, and subcontractors *each* have their own definition of a successful project. However, the concept of success is highly subjective, constantly evolving, and requires continuous reassessment to ensure the proper factors are being used to measure success. Therefore, although contractors were asked to subjectively select projects for inclusion in the study, it was acknowledged that the most appropriate way to ensure that the projects submitted for analysis were properly classified was to ask the project managers to provide their own definition of successful performance and to match their definition with quantitative data that could be used to retrospectively evaluate performance. This paper describes the process that was used to define *successful performance*, match the definition with the data collected, and develop a model that measured each project's *actual* performance against the project managers' definitions of success. The model essentially converted the project managers' qualitative opinions about their project's success to a quantitative measurement.

Problem Statement

Intuitively, most contractors *believe* better planning can lead to more successful project performance but the evidence has been mostly anecdotal. One professional association has specifically acknowledged the connection between planning and performance by identifying the greatest benefits of preconstruction planning as "the project control and organization that led to increased productivity, fewer accidents, and increased profitability" [Plumbing-Heating-Cooling Contractors (PHCC)

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

National Association 2002]. However, little *quantitative* evidence has been provided to support the claim that better planning contributes to successful performance. As a result, one goal of the research project was to quantify the relationship between planning and performance. This entailed measuring both *planning effort* as well as *project performance*. This paper focuses on the development of a tool to measure project performance.

Methodology

Planning and performance data were collected from 27 randomly selected electrical contractors throughout the United States that responded to an initial survey. All contractors were members of the National Electrical Contractors Association (NECA). Participating companies were asked to submit information about two recently completed projects: (1) a project that the company considered *successful* primarily because of its outstanding planning effort; and (2) a project that the company considered *less-than-successful* primarily because of its poor planning effort. Characteristics of *successful* and *less-than-successful* projects were provided to assist participants in their selection of appropriate projects. Each participant completed a questionnaire and participated in a four-hour interview about the planning and performance of their submitted project. The interviews were conducted with the project manager who had responsibility for the planning and management of the submitted project.

During each interview, the participant was asked to provide his or her own definition of successful project performance. The question was open-ended, and the participant could provide any characteristic they felt was necessary for success. Likewise, during the interview, data was collected on the actual performance of their submitted project.

Data analysis involved first assembling a prioritized list of the project managers' definitions of success, where the priority was determined by the frequency a characteristic was named as a success factor. This list was compared to the data collected on actual performance, and those performance variables that matched the contractor's definition of success were selected for further evaluation. The intention was to use quantitative measures of actual performance, such as actual profit, to determine how well each project *actually* performed relative to the project managers' definitions of success.

After conducting exploratory data analyses, an index was ultimately selected as the most appropriate model to measure performance. The decision to use an index was grounded in theory, and, when compared to other techniques, it produced the most accurate and consistent results. Other techniques, such as logistic regression and classification tree development, were evaluated but rejected because the variables were highly correlated and produced biased estimated parameters. Further, principal components analysis was also rejected because it weighted several variables equally rather than assign weights based on the frequency each was named as a success factor. Hence, an index was constructed that used quantitative performance variables that matched the project managers' definitions of success, and these variables were weighted according to their importance to project managers (as evidenced by the frequency they were identified as a success characteristic). The index produced a *measure of successful performance*.

The index was subsequently converted to a *probability of successful performance* by inserting the index equation into a logistic regression function. Two benefits of converting to a probability were identified, including (1) the performance was placed on an intuitive scale; and (2) the probability could be used as the response variable in a regression equation at a later stage of the research.

The final stage of the performance analysis involved validating the results. Leave-one-out cross validation was used to evaluate the performance measurement index and its subsequent conversion to a probability. Overall, a misclassification rate of 2% was observed, which indicated that the model was useful for predicting performance.

Review of Previous Research

Numerous studies have been conducted to identify factors that influence project success. These studies have found that success can be determined by budget and schedule performance, safety record, customer satisfaction, previous work experience, or numerous other subjective and objective performance factors. However, no other study specifically measured *actual* performance with respect to the project manager's own definition of successful performance. What follows is a small sample of the large number of previous projects that were reviewed. The field of *success* research is quite extensive; consequently, a limited number of projects are reported here to provide insight into the various types of projects that have been undertaken.

A two-phase research project was conducted by Ashley and Jaselskis on the determinants of construction project success. The results from the first phase of the study identified six factors that were used most frequently to measure success, including budget performance, schedule performance, client satisfaction, functionality, contractor satisfaction, and project manager/team satisfaction (Ashley et al. 1987). During phase two, Jaselskis (1988) developed a series of predictive models for achieving construction project success. Three models were developed to identify how to achieve (1) outstanding project performance; (2) better than expected schedule performance; and (3) better than expected budget performance.

Sanvido et al. (1992) developed critical project success factors (CPSF) and developed a tool to predict the probability of having a successful project. A list of typical success factors was assembled, including: *meet schedule, under budget, minimal construction problems*, and *profit*. Using a model called the integrated building process model as the framework for identifying the CPSFs, a questionnaire was developed to capture information about the success of a project within the context of the model.

Chua conducted a study to identify the critical success factors for construction projects based on expert opinion (Chua et al. 1999). An analytic hierarchy process was used to break the problem down into increasingly detailed elements. At the top level of the hierarchy was the goal: Success. Success was further broken down into budget success, schedule success, and quality success. Ultimately, 67 factors were evaluated, and the top 10 critical success factors for each project objective (budget, schedule, and quality success) were identified.

Finally, Cox et al. (2003) completed a recent study on key performance indicators used by construction executives and project managers to measure construction performance at the project level. Their study identified a substantial difference between construction executives' perception of successful performance and the perception of project managers. However, they found six common measures used by both parties: quality control,

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

Table 1. Percentage of Companies among Categories of Average Annual Sales

Revenue categories	Revenue range	University of Wisconsin survey (%)	NECA financial report (%)
1	Less than $2,000,000	18	14
2	$2,000,000–$4,999,999	18	25
3	$5,000,000–$9,999,999	21	19
4	$10,000,000–$19,999,999	14	17
5	Over $20,000,000	23	24
	Missing data	7	0
	Total	100	100

Note: NECA=National Electrical Contractors Association.

on-time completion, cost, safety, dollars per unit, and units per work hour.

Eight common success factors that were identified through the review of previous research included: (1) budget performance; (2) accurate cost estimate; (3) schedule performance; (4) profit achievement; (5) planning effort; (6) management of labor and work hours; (7) customer satisfaction; and (8) total team performance and communication. As a result, the questionnaire and interview guide were developed to capture data on these eight factors, with the additional intention of matching these factors to the project managers' own definition of successful performance.

Results of the Data Collection Process

For practical and economic reasons, the sampling plan employed a two-stage random sampling technique in which electrical companies were grouped geographically, and then geographic regions were randomly selected for inclusion during the first stage of sample selection and companies from within those regions were randomly selected in the second stage of sample selection. Overall, 27 companies from 11 states covering seven geographic regions agreed to participate in the questionnaire and interview process, and data was collected on 55 projects.

Each company that participated in the research reported their average annual sales. The sales categories included: (1) Less than $2,000,000; (2) $2,000,000–$4,999,999; (3) $5,000,000–$9,999,999; (4) $10,000,000–$19,999,999; (5) $20,000,000–$49,999,999; (6) $50,000,000; and greater. Table 1 identifies the percentage of companies among the categories of average annual sales. Table 1 also identifies how this distribution compares to the distribution of companies identified from the NECA Annual Financial Report. The two distributions are sufficiently consistent, with less than 10 percentage point's difference between the two, confirming that the sample of contractors from the current research is approximately representative of the larger NECA population.

Seven companies (26%) reported that they had a formal planning process, 15 companies (56%) reported they had a semiformal process, four (15%) reported having an informal process, and one company (4%) indicated they used some other process. Overall, companies that had a formal or semiformal planning process were much more willing to participate in the research project than those companies that had an informal process. Nearly all of the companies that declined to participate indicated they did not feel they could significantly contribute to the research because their planning process was very informal.

Each company that participated in the research was mailed a questionnaire and participated in an interview. Overall, 55 usable interviews were conducted on 28 *successful* projects and 27 *less-than-successful* projects, and 47 project managers participated in the interviews.

Limitations: Projects that were investigated in this study: (1) Ranged in work hours from 600 hours to 265,000 h, (2) ranged in contract value from $60,000 to $23 million, and (3) ranged in duration from 12 to 250 weeks. However, extra caution should be exercised when evaluating projects with (1) work hours less than 1,000 or greater than 75,000, (2) contract values less than $140,000 or greater than $10 million, and (3) a duration longer than 110 weeks. Very little data were available in these marginal ranges. Additional limitations that might have influenced the outcome of the research include: (1) The size of the sample, where additional projects might have permitted a larger degree of confidence, (2) missing data that might have modestly impacted the results of the analysis, and (3) data that was not requested but was later found to be useful to the data analysis and possible results.

How Project Managers Defined Successful Performance

In an earlier phase of the research, contractors were asked to rank order a series of success factors. Topping the list were (1) satisfied customer; (2) within-budget completion; (3) accident-free job; (4) on-time completion; and (5) achieving the estimated productivity. From these factors, criteria were developed to assist contractors with the selection of a *successful* and *less-than-successful* project for submission during the final phase of the research. The following is an excerpt from the questionnaire, which describes the criteria for selecting a *successful* project.

"To be considered *successful*, the project should meet the following criteria:

1. The project made money (was profitable),
2. The project met its target work-hours or was completed under the estimated work-hours, *and/or*
3. The project was completed on, or before, the contract completion date (the contract completion date is the original completion date plus all time extensions).

"The *successful* project that is selected for this survey should also meet the following additional criteria, when possible: (1) The project procurement method was design/bid/build (traditional hard bid or negotiated lump sum); (2) the project size was greater than 2,000 work hours; (3) the project did not have *unusually* high levels of complexity or specialized work (for example, clean rooms or security vaults would *not* be appropriate for this survey); and (4) the project did not experience any *major* accidents or disruptive events. *Further, the project you select for this survey should be considered successful primarily because of its outstanding planning effort*."

The criteria for the *less-than-successful* project were similar, with the exception that the project should have *failed to meet* the first three criteria. These criteria were developed to help contractors select projects that were appropriate for the study, to minimize the variability of the projects, and to reduce the amount of subjectivity involved in the selection process.

Success tends to be a subjective concept that varies widely throughout the construction industry and depends heavily on the team member defining *success*. Therefore, it was necessary to ask project managers to define *successful* performance during the interview. Further, it was determined that it would be necessary to

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

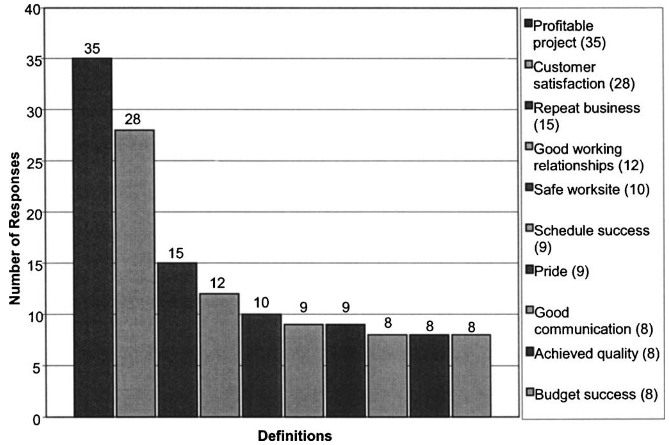

Fig. 1. Electrical contractors' top ten definitions of *successful* performance

Table 2. Comparison of Contractor's Definition of *Successful* Performance and *Success* Factors from Previous Research

Contractors' top ten definitions of success	Success factors from previous research
Profitable project	Profit achievement
Customer satisfaction	Customer satisfaction
Repeat business	—
Good work relationships	—
Safe worksite	—
Schedule success	Schedule performance
Pride	—
Good communication	Total team performance and communication
Achieved quality	—
Budget success	Budget performance
—	Accuracy of the cost estimate
—	Management of labor and work hours
—	Planning effort

re-evaluate each project to rate the *actual* performance of the project in quantitative terms. Logically, since the selection criteria was *successful* performance, it was necessary to base the measurement of *actual* performance on the contractors' *own* definition of *success*.

Hence, two primary objectives of the data collection process were to have contractors define successful performance and, at the same time, collect data on performance that could be used to develop a model to measure *actual* performance. One limitation to the research was that the electrical project managers'definition of successful performance was not available prior to conducting the interviews, and, as a result, the performance data that was collected was based on factors identified through a review of previous research. However, the match between the contractors' definition of success and the collected data, while not perfect, *did* permit a meaningful model of performance to be developed.

Data Analysis and Results: Participants were asked to provide their definition of successful performance, and 47 interviewees provided a total of 190 responses, which were analyzed using qualitative data analysis techniques. This process involved entering all definitions into a database and initially searching the database for common themes and trends in the responses. A full coding framework was developed and the responses were "coded," which involved assigning at least one code to each response. A hierarchical coding structure was developed, and frequencies of responses were tallied and summarized in a table (Appendix I) and chart (Fig. 1).

Project managers cited (1) profitable project; (2) satisfied customer; (3) repeat business; (4) good working relationships; and (5) a safe job as their top definitions of successful performance. Furthermore, good communication, schedule success, pride in performance, and budget success were also identified as important to a successful outcome.

Measuring Actual Performance Using the Project Managers' Definitions

Recall that eight success factors were identified through the review of previous research, and that data on these eight factors were collected during the interviews with the additional intention of matching these factors to the project managers' own definitions of successful performance. Table 2 identifies the factors from the collected data that could be matched to the definitions. Ultimately, 45 variables were carefully evaluated for potential inclusion in a model to measure *actual* project performance.

Development of a Theoretical Model for Measuring Successful Performance

Projects were submitted as either *successful* or *less-than-successful*, and each project was assigned a binary number—either 1 for *successful* or 0 for less-than-successful. The performance modeling process had two goals: (1) to measure *actual* performance based on the contractors' definitions of success; and (2) to assign each project a continuous number, rather than a binary number, that indicated its overall performance relative to the other projects in the data set. This performance measurement would be used in a later stage of the research to quantify the relationship between planning and performance. Since the concept of *successful performance* cannot be measured directly, an index was selected as the best method to capture the concept.

The theoretical concept of *successful performance* is a latent variable that can be caused by numerous factors and, likewise, have numerous consequences (Coleman 1964). One of the functions of a latent variable is that it serves to link many causes with many consequences. For example, "spent less money than expected" might be a cause of *successful performance*, which might likewise result in the consequence of having a profitable project. Comparatively, the theoretical concept of *success* might have exactly the same importance to a construction researcher as temperature does to a chemical researcher (Coleman 1964). However, although temperature can be measured by an associated theoretical construct, such as the mercury thermometer, the concept of *successful performance* has no such other theoretical construct by which it can be "measured." As a result, *success* is often measured by those particular variables that it ultimately gives meaning to. For example, the variable "percent profit" might have the numeric value of 15%, which alone has no specific meaning. However, if the theoretical concept of *successful performance* can be defined, in part, as "a project whose percent profit is greater than 0%," then the value of 15% now leads one to conclude that a project whose percent profit was 15% was also *successful*. Thus, quantitative variables can be combined to form an index that represents the latent variable *successful performance*, and the index can then be used to quantitatively measure perfor-

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

Table 3. Interitem Correlation Matrix for Final Six Variables

Variable	Actual percent profit	Percent schedule overrun	Amount of time given to complete your work	How would you rate the communication between team members	Did you exceed your budgeted cost	Percent change in work hours
Actual percent profit	1.000	0.281	0.335	0.558	0.728	0.518
Percent schedule overrun	0.281	1.000	0.185	0.269	0.347	0.215
Amount of time given to complete your work	0.335	0.185	1.000	0.459	0.380	0.218
How would you rate the communication between team members	0.558	0.269	0.459	1.000	0.678	0.376
Did you exceed your budgeted cost	0.728	0.347	0.380	0.678	1.000	0.438
Percent change in work hours	0.518	0.215	0.218	0.376	0.438	1.000

mance. According to Coleman (1964), by matching up the meaning of *successful performance* as a theoretical concept with the dimensions of the index used to measure it, the likelihood of obtaining theoretical relevance will be increased.

Index Construction

Index construction generally consists of two steps: (1) deciding *which* variables will be used to measure success; and (2) deciding *how* to combine these variables. Further, defining an explicit procedure for selecting the variables in the index, so that its "meaning" can be analyzed, will help ensure that the index will closely represent the true meaning it was intended to have (Coleman 1964).

Each variable was carefully examined for potential inclusion in a performance measurement index. The variables were evaluated using two primary criteria: (1) how closely the variable matched the project managers'definitions of success; and (2) whether the variable was a quantitative or qualitative measurement of performance. It was preferable to use quantitative measurements whenever possible. It was also desirable to seek out those variables that matched the most frequently named success factors. Hence, *profitable project* was the top success factor identified by project managers, and, as a result, it was desirable to include a variable that could measure the profitability of each project.

One of the limitations to the data analysis was lack of a complete match between the project managers' definitions of *successful* performance and the actual performance data that were collected. Although customer satisfaction rated high on the list of project managers' definitions of success, there was insufficient data to reliably include any customer satisfaction variable in the index construction. Further, the most reliable *satisfaction* data would necessarily need to come from the customers themselves, who were unavailable for this research project. As a result, the index was constructed from the *best available data* on actual performance.

By comparing the project managers' definitions of successful performance with the variables in the data set, ten variables were initially selected for potential inclusion in the performance measurement index. These ten variables included: (1) Actual percent profit; (2) How would you rate the owner's overall satisfaction with your work? (3) How would you rate the general contractor/ construction managers (GC/CM's) overall satisfaction with your work? (4) How would you rate your relationship with the owner on this project? (5) How would you rate your relationship with the GC/CM on this project? (6) Percent schedule overrun; (7) Describe the amount of time you were given to complete your work; (8) How would you rate the communication between team members on this project? (9) Did you exceed your budgeted cost? (10) Percent change in work hours.

Before proceeding with the index construction, all ten variables were first converted to z scores in order to transform variables that have different scales (dollars versus days versus percents) to the same scale and place them on a "level playing field." A z score is essentially a rescaling of a variable so that the mean is 0 and the standard deviation is 1.

To construct the theoretical *performance measurement* index, the reliability and validity of the index was carefully evaluated in order to assess the usefulness and accuracy of the index. *Reliability* is the ability of the measurement tool to accurately measure the true value (Ward 2005). Reliability computes the proportion of variability in a measured value that is due to the *true* value. For example, a reliability of 95% would indicate that 95% of the variability of the observed or measured value is "true" whereas only 5% of the variability is due to random error. A reliability of 0.80 (80%) or greater is desirable and can be evaluated using Chronbach's alpha, which measures the internal consistency of the factors in the index.

Chronbach's alpha evaluates the extent to which the items in the index measure the same concept, such as project performance. It is used to evaluate whether a variable should be included or excluded from the index. Chronbach's alpha should be used in concert with the bivariate correlations between the variables in the index, where high correlation among the variables is desirable. High correlations indicate that the items are closely related to one another and to the concept upon which the index is constructed, such as project performance. Hence, the variables in the index should be closely related to *performance*.

Validity is the ability of the tool to measure what it was intended to measure—specifically, it must measure *performance*. Three types of validity should be evaluated: (1) face validity; (2) content validity; and (3) criterion-related validity. Face validity is the evaluation of the assumption that there is a logical connection between the variables in the index and the purpose of the index. Content validity is the evaluation of the variables in the index to ascertain whether the variables adequately represent the construct proposed. Criterion-related validity evaluates the relationship between how well the subject (project) is rated by the index versus how well the subject (project) actually performed.

After the variables were transformed to z scores, the ten transformed variables were entered into the reliability analysis function of the statistical program Statistical Package for the Social Sciences (SPSS). The output identified the inter-item correlation coefficients, the measure of internal consistency (Chronbach's

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

alpha), and also identified the change to Chronbach's alpha if each variable were to be the next one eliminated from the index. The Chronbach's alpha was 0.757, which is adequate but not great. A review of the correlation coefficients revealed that the two *satisfaction* and two *relationship* variables correlated poorly with most of the other six success variables. These variables were eliminated due to their very low correlations to other variables and their qualitative nature. The remaining six variables were used in the second iteration of the reliability analysis. Chronbach's alpha increased to 0.80, which indicated very good reliability. The correlation coefficients were evaluated, and all were found to be very good as well (Table 3). As a result, six factors were selected for the performance measurement index, including:

1. Actual percent profit;
2. Percent schedule overrun;
3. Describe the amount of time you were given to complete your work;
4. How would you rate the communication between team members on this project?
5. Did you exceed your budgeted cost?
6. Percent change in work hours.

Clearly, all variables (success factors) were not rated by contractors as being equally important characteristics of successful performance—profitability was identified more frequently than schedule success and, likewise, schedule success was identified more frequently than budget success. As a result, the accuracy of the index depended on properly weighting each of the variables included in the index to properly reflect their importance to project managers. The "weights" were calculated by summing up all responses for the six variables, which included 60 total responses that included one or more of the six performance variables (success definitions), and dividing the individual variables frequencies (f_i) by the total number of responses (n). Table 4 identifies how this calculation was performed.

Table 4. Assignment of Weights for Index Construction

Definition	Responses	Weight
Profitable project	35	35/60=0.583
Schedule success:		
On-time completion	7	7/60=0.117
Realistic schedule	2	2/60=0.033
Good communication	8	8/60=0.133
Budget success:		
Achieved budget	5	5/60=0.083
Achieved work hours	3	3/60=0.05
Total	60	60/60=1.00

Therefore, the initial performance measurement index was calculated for each project by multiplying the z scores for the six variables by their appropriate weights. The following model identifies the performance measurement index:

$$\begin{aligned}\text{performance measurement } (y) = {} & (0.583 \times \text{actual percent profit } z \text{ score}) + (0.117 \times \text{percent schedule overrun } z \text{ score}) \\ & + (0.033 \times \text{amount of time given to complete your work } z \text{ score}) \\ & + (0.133 \times \text{how would you rate communication between team members } z \text{ score}) \\ & + (0.083 \times \text{did you exceed your budgeted cost } z \text{ score}) + (0.05 \times \text{percent change in work hours } z \text{ score})\end{aligned} \tag{1}$$

Transformation to a Probability of Successful Performance

The performance measurement was calculated from the z scores of the raw variables, and, as a result, the values, which ranged from −2 to +2, were not intuitive. Further, the long-term goal of the research was to quantify the relationship between planning and performance, and the theoretical argument was put forth that it was more appropriate to predict the *probability* of *successful performance* than to definitively predict performance outright. Consequently, the performance measurement (y) was converted to a probability by incorporating it into a logistic response function. The logistic response function is represented by the following equation:

$$\begin{aligned}\pi &= \Pr(Y = 1 | X_1 = x_1, \ldots, X_p = x_p) \\ &= (e^{b_0+b_1x_1+\cdots+b_px_p})/(1 + e^{b_0+b_1x_1+\cdots+b_px_p})\end{aligned} \tag{2}$$

where π=probability of $Y=1$ (successful performance) for a given set of x_i variables. Typically, the variables of interest (x_i) are used directly in the logistic regression model, a logit transformation is performed to linearly transform the parameters, and then the parameters (b_i) are estimated. However, the variables that were used to define *successful performance* were highly correlated, which resulted in biased estimates of the parameters, and, consequently, the decision was made to construct an index, instead, to measure performance. By using the index equation [Eq. (1)] to replace $b_0+b_1x_1+\cdots+b_px_p$ in Eq. (2), the probabilities could be modeled in a manner similar to an ordinary logistic regression.

Although parameter estimation was not necessary as index weights were used as coefficients, the logistic response function provided a useful tool for modeling the response probabilities (i.e., the probability of achieving a successful outcome). These response probabilities were essentially a transformation of the Performance Measurement to a corresponding probability of successful performance. Probabilities greater than 0.50 corresponded to a successful outcome ($Y=1$) and to a higher performance measurement. The probability of achieving a successful outcome was calculated for every project in the data set and is shown in Appendix II, which also identifies whether the project was originally submitted as successful ($Y=1$) or less-than-successful ($Y=0$) and the corresponding misclassification designation. Overall, four projects were misclassified, for a misclassification rate of 7%.

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

Brief Example of How to Use the Performance Measurement Index

Project 2369-S was a complete upgrade to a fire alarm system in a postal processing center. The project was submitted as a "successful" project ($Y=1$). The first step to constructing the performance measurement (Y) was to identify the raw variables used in the index:

1. Actual percent profit=0.344 (34.4%).
2. Percent schedule overrun=0 (0%).
3. Describe the amount of time you were given to complete your work=3
   - Significantly more than actually needed=1;
   - Moderately more than actually needed=2;
   - About the same as actually needed=3;
   - Compressed, requiring extra levels of labor and materials=4; and
   - Unrealistic, requiring extraordinary efforts to achieve the duration=5.
4. How would you rate the communication between team members on this project?=4
   - Below average=1;
   - Slightly below average=2;
   - Average=3;
   - Slightly above average=4; and
   - Above average=5.
5. Did you exceed your budgeted cost? No=0.
6. Percent change in work hours=−0.331 (−33.1%).

To ensure consistency in meaning, several transformations were needed; specifically, so that all larger values conveyed "positive" meaning (i.e., where 1 is better than 0, and 2 is better than 1, and so on), the following reversals were made: (1) Percent schedule overrun reversal: $1-0=1$; (2) amount of time given reversal: $6-3=3$; (3) exceeded budget reversal: $1-0=1$; (4) percent change in work hours reversal: $1-(-0.331)=1.331$. Essentially, all raw variables were rescaled to a *larger-is-better* scale. Following the rescaling, all variables were transformed to their z scores. The equation for a z-score transformation is

$$z = (y - \mu)/\sigma \tag{3}$$

where z=z score, y=value of the raw variable, μ=theoretical mean of the variable in the data set, and σ=theoretical standard deviation. The z-score transformations resulted in the following values:

1. Actual percent profit z score=1.28;
2. Percent schedule overrun reversal z score=0.35;
3. Amount of time given reversal z score=0.056;
4. Communication between team members z score=0.61;
5. Did you exceed your budgeted cost reversal z score=0.75; and
6. Percent change in work hours z score=1.25.

The z scores were entered into the performance measurement index to calculate the following:

$$\begin{aligned}&\text{performance measurement } (y)\\&\quad= (0.583 \times 1.28) + (0.117 \times 0.35) + (0.033 \times 0.056)\\&\quad+ (0.133 \times 0.61) + (0.083 \times 0.75) + (0.05 \times 1.25) = 0.99\end{aligned}$$

The performance measurement, as previously mentioned, is not very intuitive. Therefore, to convert the performance measurement to a *probability of successful performance*, the measurement (y) was plugged into the logistic regression function

$$\text{probability of successful performance} = e^{0.99}/(1 + e^{0.99}) = 0.73$$

It is clear, now, that the probability of successful performance is 0.73 or 73%. As projects that have a *probability of successful performance* greater than or equal to 50% would be classified as "successful" ($Y=1$), Project 2369-S would be classified as a successful project. This project was submitted as a successful project by the project manager, and this qualitative designation was converted to a quantitative value, which demonstrated that, according to the project managers' definitions of success, that the project, indeed, could be classified as a success.

Cross Validation of the Measurement Tool

Cross validation is a model evaluation technique that uses a portion of the existing data to build a model while leaving out one or more cases or data points for future prediction. The purpose of cross-validation is to validate a constructed model. In leave-one-out cross validation, one case is left out of the data set, and the remaining $n-1$ cases are used to construct the model. The model is then used to predict the outcome for the one case that was left out, and the predicted outcome is compared to the actual outcome to determine whether the case was misclassified by the model. Upon completion of the cross-validation procedure, a misclassification rate can be calculated. Generally, a misclassification rate of 10% or less is desirable.

Cross validation was used to validate the performance measurement index and subsequent transformation to a probability of success. To perform the cross validation, the following steps were followed:

1. A spreadsheet was constructed that identified the project number, interviewee identification, whether each of the six variables in the index was stated as a definition of success, and the variable z score.
2. If two projects had the same interviewee (whose definitions of success were included *only once* in the data set, not twice), a coin was flipped to assign the definitions to one of the two projects, and the other project was eliminated.
3. The spreadsheet was coded with the interviewee's definitions of successful performance. Essentially, if the interviewee, for example, stated "made money" as a characteristic of a successful project, then the factor "Profit" was coded with a 1. If "achieved the budget" was *not* mentioned by the interviewee as a characteristic of a successful project, then the factor "achieved budget" was coded with a 0. In this manner, definitions of success were recorded for every project and its associated project manager.
4. The first project was removed from the data set, and set aside.
5. The spreadsheet was programmed to calculate the new frequency (f_i) that each success factor (definition) was selected by those projects in the remaining data set, as well as a new summed total number of responses (n).
6. The factor frequencies (f_i) and the total number of responses (n) were used to calculate new weights for each factor as f_i/n.
7. A new performance measurement index was developed using

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

the factor weights as coefficients and z scores as the variables, similar to Eq. (1).

8. The new performance measurement index was used to calculate a performance measurement score for the project that was left out in Step 4.
9. The performance measurements score for the left out project was used in the logistic regression function [Eq. (2)] to calculate a probability of a successful outcome.
10. The predicted outcome was compared to the actual outcome to determine whether the project was properly classified by the model or misclassified.
11. Steps 4 through 11 were repeated for every project in the data set.
12. Upon completion of the cross-validation procedure, the number of misclassified cases was tallied and a misclassification rate was determined.

Ultimately, 47 projects and their associated project managers were used in the cross-validation procedure, including 25 successful projects and 22 less-than-successful projects. Four of those projects had varying degrees of missing data that limited their utility in the cross-validation process. Overall, 43 projects were classified, and of those, only one was misclassified, for a misclassification rate of 2% (Appendix III). It should be acknowledged that, of the four previously misclassified projects, only one was included in the analysis based on random chance. However, if the four had again been misclassified, the misclassification rate would be 9%. This low misclassification rate (less than 10%) demonstrated that the model was a valid and useful tool for measuring performance and ultimately predicting the probability of a successful outcome.

Conclusion and Practical Applications

This paper described the process for defining *successful project performance*, matching the definition to collected data, and developing a model that measured a project's *actual* performance against the project manager's definition of success. The utility of the model is that it essentially converted the project managers' *qualitative* concept of "successful performance" to a *quantitative* measurement.

The performance measurement index, and its subsequent conversion to a probability, was constructed primarily to advance the current research project; however, the index has several other useful academic and practical applications.

Specifically, this *process* for developing a performance measurement index can be used by academic professionals to convert qualitative concepts into quantitative values and to convert binary measurements into continuous numeric measurements in the specific context of the subject being measured. The index can also be used to measure project performance on other research projects, where the quantitative index value could be used as a response variable in a modeling process. For example, the index could be used as the response variable on a research project that uses a *design of experiments* approach to investigate the impact of numerous variables on the performance of projects.

Likewise, electrical construction professionals might find this tool useful for retrospectively assessing project performance, which will allow the contractor to compare the performance of numerous completed projects across the same factors. Further, contractors can use the index to compare *expected* versus *actual* performance by initially plugging in estimated values, and, after completion, plugging in actual values to compare estimated to actual probability of successful performance. Contractors can also use the index to improve project planning and management by evaluating the impact of each index variable on the probability of success. For example, a contractor might expend greater effort on managing the schedule or improving team communication in order to increase the probability of achieving a successful outcome.

Finally, both academic and construction professionals might find the index useful for tracking performance trends among projects with various characteristics. For example, a contractor could track the probability of successful performance on elementary schools and compare it to the probability of success on office buildings.

Appendix I

See Table 5.

Table 5. Contractors' Definition of *Successful* Performance

Definition	Subtotals	Totals
Profitable project		35
Customer satisfaction		28
Satisfied customer	17	
Satisfied owner	8	
Satisfied general contractor	3	
Repeat business		15
Good work relationships		12
Good relationship with owner	3	
Good relationship with general contractor	4	
Good relationship with team members	2	
Good relationship with workforce	3	
Safe worksite		10
Schedule success		9
On-time completion	7	
Realistic schedule	2	
Pride		9
Good communication and cooperation		8
Good communication	3	
Good cooperation/coordination	5	
Achieved quality		8
Quality product	3	
Functional product	5	
Budget success		8
Achieved budget	5	
Achieved work hours	3	
Good work environment		7
Minimum stress or conflicts	7	
Good project management		6
Good management/change management	3	
No delays or stoppages or difficulties	3	
Workforce efficiency		5
Team satisfaction		5
Satisfied team members	2	
Satisfied PM/field supervisor/crew	3	
Good material/equipment management		4
Project went well		4
Increased marketability		4
Good pay process		3
Successful recovery		3
Design success		3
Good planning		2
Favorable project characteristics		2
Total		190

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

Appendix II

See Table 6.

Table 6. Probability of Achieving a Successful Outcome

Project ID	Outcome (success=1) (less than success=0)	Performance measurement (*Y*)	Probability of successful performance	Misclass ID
182-L	0	−0.07	0.48	0
182-S	1	0.33	0.58	0
194-L	0	−1.23	0.23	0
194-S	1	0.66	0.66	0
199-L	0	−0.05	0.49	0
199-S	1	0.11	0.53	0
2109-L	0	−1.42	0.19	0
2109-S	1	0.85	0.70	0
2146-L	0	−0.69	0.33	0
2146-S	1	1.11	0.75	0
2207-L	0	−1.16	0.24	0
2207-S	1	0.21	0.55	0
2239-L	0	−1.46	0.19	0
2239-S	1	0.12	0.53	0
2369-L	0	−0.65	0.34	0
2369-S	1	0.99	0.73	0
243-S	1	0.02	0.51	0
2501-L	0	−1.42	0.20	0
2501-S	1	0.71	0.67	0
2710-L	0	0.60	0.65	Misclassified
2710-S	1	1.77	0.85	0
276-L1	0	−0.57	0.36	0
276-S1	1	0.89	0.71	0
283-L	0	−0.87	0.30	0
283-S	1	0.26	0.56	0
309-L	0	−0.70	0.33	0
309-S	1	0.01	0.50	0
436-L	0	−0.24	0.44	0
436-S	1	0.45	0.61	0
47-L	0	0.35	0.59	Misclassified
47-S	1	0.96	0.72	0
568-L	0	−1.22	0.23	0
568-S	1	0.10	0.53	0
571-L1	0	−0.75	0.32	0
571-L2	0	−0.50	0.38	0
571-S1	1	0.21	0.55	0
571-S2	1	0.23	0.56	0
616-L	0	−0.35	0.41	0
616-S	1	1.02	0.74	0
666-L	0	*	*	*
666-S	1	0.64	0.66	0
669-L	0	*	*	*
669-S	1	0.33	0.58	0
741-L	0	−1.13	0.24	0
741-S	1	0.74	0.68	0
831-L	0	0.08	0.52	Misclassified
831-S	1	1.00	0.73	0
833-L	0	−1.23	0.23	0
833-S	1	0.39	0.60	0
853-L	0	−1.60	0.17	0
853-S	1	0.79	0.69	0
913-L	0	0.10	0.53	Misclassified
913-S	1	0.95	0.72	0
96-L	0	*	*	*
96-S	1	*	*	*

Note: An asterisk denotes a missing value.

Appendix III

See Table 7.

Table 7. Results of the Cross Validation

Project ID	Outcome (success=1) (less than success=0)	Performance measurement (*Y*)	Probability of successful performance	Misclass ID
182-L	0	−0.07	0.48	0
182-S	1	0.33	0.58	0
194-L	0	−1.22	0.23	0
194-S	1	0.67	0.66	0
199-L	0	−0.05	0.49	0
199-S	1	0.10	0.52	0
2109-L	0	−1.42	0.19	0
2109-S	1	0.84	0.7	0
2146-L	0	−0.70	0.33	0
2146-S	1	1.12	0.75	0
2207-L	0	−1.16	0.24	0
2207-S	1	0.22	0.55	0
2239-L	0	−1.46	0.19	0
2239-S	1	0.12	0.53	0
2369-L	0	−0.61	0.35	0
2369-S	1	1.00	0.73	0
243-S	1	0.03	0.51	0
2501-L	0	−1.41	0.2	0
2501-S	1	0.71	0.67	0
2710-S	1	1.77	0.85	0
246-L1	0	−0.58	0.36	0
276-S1	1	0.89	0.71	0
283-L	0	−0.83	0.3	0
283-S	1	0.26	0.57	0
309-L	0	−0.70	0.33	0
436-L	0	−0.24	0.44	0
436-S	1	0.45	0.61	0
47-S	1	0.95	0.72	0
568-L	0	−1.21	0.23	0
568-S	1	0.11	0.53	0
571-L1	0	−0.74	0.32	0
571-L2	0	−0.50	0.38	0
616-S	1	1.03	0.74	0
666-L	0	*	*	*
666-S	1	0.65	0.66	0
669-L	0	*	*	*
669-S	1	0.34	0.58	0
741-L	0	−1.11	0.25	0
741-S	1	0.75	0.68	0
831-L	0	0.08	0.52	Misclassified
831-S	1	1.00	0.73	0
833-S	1	0.39	0.6	0
853-L	0	−1.58	0.17	0
853-S	1	0.79	0.69	0
913-S	1	0.95	0.72	0
96-L	0	*	*	*
96-S	1	*	*	*
			Number analyzed	43
			% Misclassified	0.02

Note: An asterisk denotes a missing value.

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293

References

Ashley, D. B., Lurie, C. S., and Jaselskis, E. J. (1987). "Determinants of construction project success." *Proj. Manage. J.*, 18(2), 69–79.

Bernstein, H. M. (2003). "Measuring productivity: An industry challenge." *Civ. Eng. (N.Y.)*, 73(12), 46–53.

Chua, D. K. H., Kog, Y. C., and Loh, P. K. (1999). "Critical success factors for different project objectives." *J. Constr. Eng. Manage.*, 125(3), 142–150.

Coleman, J. S. (1964). *Introduction to mathematical sociology*, The Free Press of Glencoe, London.

Cox, R. F., Issa, R. R. A., and Ahrens, D. (2003). "Management's perception of key performance indicators for construction." *J. Constr. Eng. Manage.*, 129(2), 142–151.

Jaselskis, E. J. (1988). "Achieving construction project success through predictive discrete choice models." Doctoral dissertation, The Univ. of Texas at Austin, Austin, Tex.

Plumbing-Heating-Cooling Contractors (PHCC) National Association. (2002). ⟨http://www.phccweb.org/ContentItem.cfm?ContentItemID=59⟩ (April 23, 2003).

Sanvido, V., Grobler, F., Parfitt, K., and Coyle, M. (1992). "Critical success factors for construction projects." *J. Constr. Eng. Manage.*, 118(1), 94–111.

Ward, J. (2005). "Measurement theory, reliability estimation model." Clinical investigator's corner, ⟨http://www.bamc.amedd.army.mil/DCI/articles/dci0400d.htm⟩ (June 5, 2005).

J. Constr. Eng. Manage., 2006, 132(12): 1284-1293