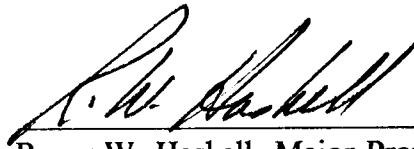


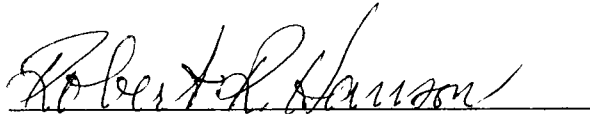
To the Graduate Council:

I am submitting herewith a thesis written by J. David Collier entitled "Kirkpatrick's Level Three Evaluation Applied to a Maintenance Apprenticeship Training Program." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Human Resources Development.



Roger W. Haskell, Major Professor

We have read this thesis
and recommend its acceptance:



Robert R. Hanson



Gerald Cheek

Accepted for the Council:



Associate Vice Chancellor and
Dean of The Graduate School

KIRKPATRICK'S LEVEL THREE EVALUATION APPLIED TO A
MAINTENANCE APPRENTICESHIP TRAINING PROGRAM

A Thesis

Presented for the Master of Science

Degree

The University of Tennessee, Knoxville

J. David Collier

December, 1996

DEDICATION

This thesis is dedicated to my daughter

Jamie L. Collier

in the hope that she will find the pursuit of knowledge
an admirable and worthwhile lifelong goal.

ACKNOWLEDGEMENTS

My thesis committee deserves a note of thanks for their assistance in my efforts to complete this assignment. I am grateful for the guidance provided by Dr. Gerald Cheek and Dr. Robert Hanson whose early contributions to my graduate study were invaluable and also for their assistance during the previous four years of undergraduate and graduate work. Special appreciation is extended to Dr. Roger Haskell, who, as major professor was instrumental in guiding the completion of my thesis and bringing this stage of my academic career to a successful completion.

I thank my supervisor, Jim Freeman, and my department manager, Edd Baldock, for interest in, and support of my graduate work. I was allowed the freedom and flexibility necessary for conducting the research on which my thesis was based. I also thank Joe Keith, Ed.D., for his guidance and assistance in completing this project.

But most of all, I wish to thank my wife, Pamela, and daughter, Jamie, for their patience and support during these past eight years of undergraduate and graduate work. Without them, all of this effort would have been for nothing.

Abstract

As American business and industry prepare to meet the challenges of the twenty-first century, units within organizations will come under ever-increasing scrutiny to demonstrate their value to the parent organization. Training and human resource development departments will be under increased pressure to direct their efforts toward meeting the goals, objectives, or desired outcomes of their parent organizations. Professionals in the human resource development field will be expected to provide managers and other decision makers with data to assist in the decision-making process concerning human resource development programs.

The purposes of this study were to measure the effectiveness of a maintenance apprenticeship training program and provide feedback to training designers for program improvement. Products of the study included a model to evaluate other technical training programs and data for management to use in the decision-making process related to technical training programs.

The objectives of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program and identify specific modifications and/or improvements which could be made to the apprenticeship training program. Data obtained and analyzed by the researcher were shared with management and other decision makers to assist in the decision-making process for the selected apprenticeship training program.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	2
Purpose of the Study	3
Objectives of the Study	3
Hypotheses	4
Significance of the Study	5
Assumptions and Limitations of the Study	5
Glossary	6
Research Design	9
Organizational Plan	11
Organization of the Study	13
II. REVIEW OF RELATED RESEARCH AND LITERATURE	14
Evaluation and Instructional Design Processes	14
Training Evaluation	18
Kirkpatrick's Four Levels of Evaluation	26
Summary	32
III. RESEARCH METHOD	33
Subjects	34
Instrument	35
Collection of Data	36
Data Analysis	38
Secondary Data Sources	39
Summary	39
IV. ANALYSIS OF STATISTICAL FINDINGS	41
Study Sample	43
Summary Statistics	44
Null Hypotheses	46
V. SUMMARY, FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS	57
Summary	58
Major Findings	60
Conclusions	61
Recommendations	65
Future Research	69

REFERENCES	73
APPENDICES	77
Appendix A	78
Appendix B	85
Appendix C	87
Appendix D	91
Appendix E	93
Appendix F	97
Appendix G	100
Appendix H	103
VITA	106

FIGURES

FIGURE	PAGE
1. Evaluation Model	10
2. Modified Evaluation Model	10
3. Summary of Candidates'/Graduates'/Supervisors' Performance Evaluations	45
4. Summary of Graduates'/Supervisors' Performance Evaluations	52
5. Workforce Development Model	68
6. Level III Evaluation Model for Existing Programs	71

CHAPTER I

INTRODUCTION

American business and industry are coming to the realization that the quality, competence, and productivity of the United States work force are impacting their ability to meet the demands of global competition. New technology, participative management, and just-in-time production are a few ways organizations are changing the way they do business. Reports on United States productivity rates indicate that foreign competitors are producing goods faster and cheaper than American companies. American corporate success depends on the skills of all employees for improvements in efficiency, quality, customer service, and the development of new applications for existing products and services.

Successful companies across America consider work force training a good investment. Research shows that training conducted in companies is effective and beneficial to workers and is one of the most important predictors of business success (Study, 1994). Successful companies have learned that by investing in human capital, they are building a work force that makes better use of technology, manages more effectively, solves problems more readily, thinks more creatively, and increases its ability to learn as jobs change (Lee, 1990).

To be successful in today's competitive marketplace, organizations must train their employees to respond quickly and effectively to new responsibilities and demands. A major part of organizational competitiveness is a well-structured and adaptable employee training system. Organizations must realize that although

training programs are often well intended, it is not always clear if they are achieving desired outcomes for the organization.

Statement of the Problem

As organizations downsize, re-engineer, re-structure, or become organizationally flatter due to global competition, training departments are under increased pressure to direct their efforts toward meeting the goals, objectives, or desired outcomes of their parent organizations. This translates into efforts by training departments to evaluate their training programs to determine if they are achieving the desired results.

According to Goldstein (1993), the past history of evaluation of training programs indicated that much more effort is necessary to acquire the information needed for the decision-making process. Goldstein (1993) cited a study conducted in 1953 which found that only one company in forty made any scientific evaluation of supervisory-training programs. There is no question that progress has been made since that time but it is still not possible to conclude that evaluation is considered a norm in the development of training programs. One survey of 611 companies indicated that 42% of companies did not evaluate the effectiveness of MBA programs and 32% did not evaluate short-course programs (Saari, Johnson, McLaughlin, and Zimmerle, 1988). However, the positive side of those figures indicated a substantial number of organizations did evaluate programs. Of the companies surveyed, 92% conducted evaluations of company-specific programs. The less positive side of those

numbers is that most of the evaluations focused on trainee reactions to the program rather than determining whether learning had taken place and job performance had been positively impacted.

Purpose of the Study

According to Lerda and Cross (1962), Mager (1988), and Goldstein (1993), there are two important considerations to keep in mind when evaluating training programs: (a) to get information that will be useful to training organizations in determining how effectively the specific training activity is contributing toward improved performance on the job on a broad basis, and (b) to identify specific modifications and improvements which can be made to the training program.

The purposes of this study were to measure the effectiveness of a maintenance apprenticeship training program and provide feedback to training designers for program improvement. Products of the study included a model to evaluate other technical training programs and data for management to use in the decision-making process related to technical training programs.

Objectives of the Study

The objectives of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program and identify specific modifications and/or improvements which could be made to the apprenticeship training program.

Hypotheses

To meet the objectives of the study, two research hypotheses were postulated.

Specifically, the research hypotheses of this study were:

Research Hypothesis One: There is a significant improvement in the job performance of apprentice graduates when compared to apprentice candidates (who have received no training on tasks related to the specific craft).

The null hypothesis which tested research hypothesis one of this study was:

H₀₁: There is no statistically significant difference between the perceptions of apprentice candidates and apprentice graduates with regard to their abilities to perform 30 selected job tasks to defined standards.

Research Hypothesis Two: Specific improvement and modification opportunities for maintenance apprenticeship training programs can be identified by comparing supervisors' performance evaluations of apprentice graduates and apprentice graduates' performance evaluations of themselves.

The null hypothesis which tested research hypothesis two of this study was:

H₀₂: There is no statistically significant difference between the perceptions of apprentice graduates and their immediate supervisors with regard to the graduates' ability to perform each of 30 selected job tasks to defined standards.

Significance of the Study

The results of the study were used as evidence to support recommendations to the Apprenticeship Committee that improvements and modifications be made in planning, design, and content of the Control System Mechanic Apprenticeship (CSM) training program. The study was also used as a model to evaluate other technical training programs at the manufacturing facility. The study also provided experience for a member of the manufacturing facility's Training Department in conducting Level III (Kirkpatrick, 1959) training evaluation.

Assumptions and Limitations of the Study

Assumptions:

1. The personnel surveyed answered the survey truthfully.
2. If apprentice graduates could meet performance expectations on the surveyed tasks, it was assumed they could meet performance expectations on less difficult tasks included in the 1988 CSM job analysis.
3. It was assumed that the population of apprentice graduates, prior to beginning the training program, was at an equivalent level of performance as the population of apprenticeship candidates surveyed for this study.

Limitations:

1. This study was limited to opinions of apprentice graduates and their supervisors to measure the performance level of apprentice graduates on the related tasks.

2. This study was limited to opinions of apprenticeship candidates to measure their performance level on the related tasks.
3. This study was delimited in scope to the training program associated with the manufacturing facility for which the study was conducted.

Glossary

To assist in the clarification of terminology utilized in this study, the following terms and definitions were provided:

Apprenticeship - A company plan containing all terms and conditions for the qualification, selection, training, and progression of apprentices, including the requirement for a written agreement. Apprenticeships are registered with the Bureau of Apprenticeship Training, U. S. Department of Labor.

Apprenticeship Candidate - An employee who has been selected to participate in a Department of Labor, Bureau of Apprenticeship Training registered apprenticeship training program at the manufacturing facility. Selections are partially based on the results of tests administered in the areas of reading, mathematics, and general mechanical aptitude.

Apprenticeship Committee - The apprenticeship program operates under the general direction of an Apprenticeship Committee whose members are appointed by the Manufacturing Vice-President of the manufacturing facility. The Apprenticeship Committee is the decision making body of the apprenticeship program and exercises overall control of the apprenticeship program.

Control Systems Mechanic (CSM) - A skill-based job which combines the crafts of instrument mechanic and maintenance electrician.

CSM Job Analysis - An analysis which was conducted in 1988 to determine the tasks comprising the job of CSMs at the manufacturing facility.

Evaluation - The collection of information necessary to make training decisions related to various instructional activities.

CSM Task Analysis - An analysis conducted in 1988 of the CSM Job Analysis to determine the skills and knowledge needed by CSMs to perform the tasks which constitute the CSM job.

General Supervisor - Normally an electrical or mechanical engineer who supervises the work of two or more first-line supervisors. Work is conducted at a strategic level rather than a tactical level. Reports to a department head.

Individual Training Plan - An employee development plan in which an employee and his or her supervisor agree on a training plan to improve the employee's job performance.

Job Certification - The process for determining and documenting that an employee's knowledge and skills meet established performance standards.

Job Related Education (JRE) - A planned learning activity to help the learner gain job related knowledge and skills for future job responsibilities. JRE teaches basic/advanced principles and concepts in a self-study, classroom, or laboratory environment.

Job Specific Training (JST) - Training which is focused on the tasks (including the

specific knowledge and skills) which must be learned in order for a person to successfully do his/her job. Typical JST would include documented on-the-job training, adequate opportunity to practice and apply training, and performance tests to assess trainees' abilities to perform the tasks actually associated with the job in the workplace.

Perceived Performance - An employee's opinion of his/her or a subordinate's ability to perform a task or job to established standards.

Performance - The ability of an employee to perform tasks associated with his/her job to established standards as measured in quality of work, timeliness in performing the task, and the safety of personnel, equipment, and the environment. For this study, attitude was not included.

Structured Training Program - A company plan containing all terms and conditions for the qualification, selection, training, and progression of trainees, including the requirement for a written agreement. Structured Training Programs are not registered with the Department of Labor.

Supervisor - Normally a craft-person who has been promoted to a first-line supervisor position. Supervises the work of 10-20 craftsmen and apprentices. Works at a tactical level and reports to a general supervisor.

Training - Activities focused on the acquisition of knowledge and skills designed to improve human performance on the job.

Training Program - Training activities in which courses and/or course work are directed at current job responsibilities and which have a definite beginning and end.

Research Design

The research design for this study was based upon a combination of Kirkpatrick's (1959) four-step hierarchy of training program evaluation and the modification of a training evaluation model suggested by Phillips (1991, p. 122). To conduct an evaluation of the transfer of knowledge and skills to the job, Kirkpatrick (1994, chap. 6) suggested measuring performance both before and after training. Phillips' (1991) pretest and post-test model design (See Figure 1) addresses the issue of measuring performance before and after training.

A modified version of Phillips' (1991) model was used to conduct this study (See Figure 2). A four-phase study was conducted to obtain and analyze the data. The initial phase of the study measured the perceived performance of 43 apprentice candidates on a representative sample of tasks which comprise the job of control system mechanic. The measurement was taken prior to any training and was used as a pretest step to establish baseline data. The researcher conducted the measurement with the use of surveys which were administered during orientation sessions for the apprentice candidates prior to beginning the training program.

Phase two of the study measured the perceived performance of 42 apprentice graduates on the same representative sample of tasks used in phase one of the study. The measurement was taken after completion of the apprenticeship training program and was used as a post test step to establish performance data upon completion of the apprenticeship training program. The researcher conducted the measurement with the use of surveys which were mailed via intra-plant mail to each participant. The

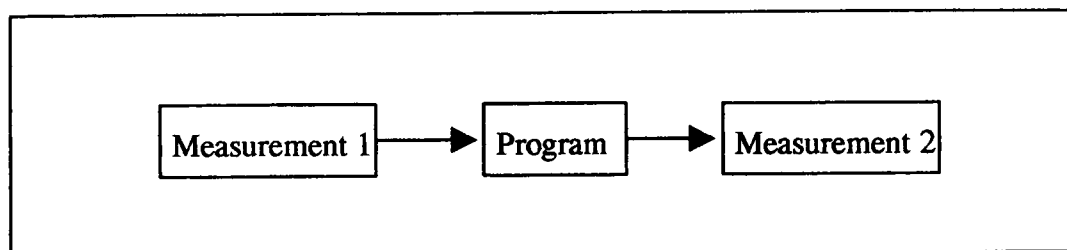


Figure 1. Evaluation Model

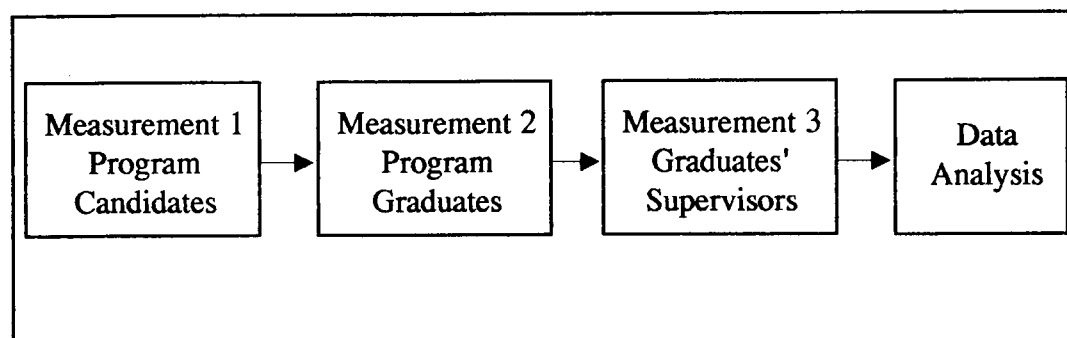


Figure 2. Modified Evaluation Model

selected program graduates were requested to participate in the survey and were provided examples of how the information they contributed would be used in the study.

Phase three of the study measured the perceived performance of the 42 apprentice graduates in phase two of the study by the supervisors of the graduates. The same representative sample of tasks used in phases one and two of the study were used in phase three. The researcher conducted the measurement with the use of surveys which were mailed via intra-plant mail to each supervisor. The

supervisors were requested to participate in the survey and were provided examples of how the information they contributed would be used in the study.

Phase four of the study consisted of data analysis of the measurements taken in phases one, two, and three. The collection of quantitative data were analyzed using descriptive and advanced statistics which are discussed in detail in the data analysis section.

Organizational Plan

In order to conduct this study, the researcher obtained permission from the Manufacturing Training Supervisor, the Training Department Manager, and a representative of the Community and Public Affairs Department of the Human Resources Division of the manufacturing facility. After obtaining permission to conduct the study, the researcher obtained a list of apprentice candidates for 1996 from the Manufacturing Training Supervisor. This list identified the population for phase one of the study. The entire population of apprentice candidates were included in the study ($n=43$).

The researcher also generated a list of graduates of the CSM apprenticeship program between the dates of May 1, 1994 and May 1, 1996 by using the Training Information Management System of the manufacturing facility. This list identified the total population ($n=63$) of most recent graduates of the CSM apprenticeship program. The researcher then generated a list of supervisors of the 63 graduates using the Employee Data Base System of the manufacturing facility. Due to the

distribution of graduates among supervisors of CSMs, 47 graduates were selected from the population to participate in the study. Three factors determined the selection criteria: (a) no more than six apprentice graduates were to be evaluated by any one supervisor, (b) no less than two apprentice graduates were to be evaluated by any one supervisor, and (c) the most recent apprentice graduates were to be selected to participate in the study. This list identified the sample ($n=47$) of CSM apprentice graduates selected to participate in phase two of the study. Information packets were sent to each potential participant which explained the nature of the study, a request from the researcher for the graduate to participate in the study, instructions for completing the survey, the survey, and examples of how the information provided by the participants would be used in the study.

The list of supervisors of CSM apprentice graduates identified above also identified the population of supervisors ($n=12$) selected to participate in the study. The researcher received 42 completed surveys from CSM apprentice graduates. The supervisors of these graduates were selected to participate in phase three of the study. Information packets were sent to each supervisor which explained the nature of the study, a request from the researcher for the supervisor to participate in the study, instructions for completing the survey, the survey, and examples of how the information provided by the participants would be used in the study.

Organization of the Study

Chapter I contains the statement of the problem, purposes of the study, objectives of the study, the hypotheses, the significance of the study, the assumptions and limitations of the study, a glossary, the operational plan, and the organization of the study.

Chapter II contains the review of related literature including research on evaluation and its role in instructional design processes, training evaluation, Kirkpatrick's four levels of evaluation, and the summary.

Chapter III contains information on the subjects, instrument, data collection procedure, method for processing the data, and summary.

Chapter IV contains information on the study sample, descriptive statistics, an analysis of each test, and summary.

Chapter V contains a summary of the procedures, major findings which resulted from the analyzed data, conclusions, and recommendations for the CSM Apprenticeship. The chapter ends with the need for additional research and concluding remarks.

CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

This section includes a review of relevant literature and research pertinent to criteria for evaluating training. To locate literature related to the objective of the study, multi-disciplinary and educational computer data queries served as the primary resource. The following descriptors were cross-referenced: training evaluation, training effectiveness, Kirkpatrick's levels of evaluation, and industrial training. A secondary source of information was a review of recent (1975-1995) professional publications. They included the Journal of European Industrial Training, the Training: The Human Side of Business magazine, the Journal of Applied Psychology, and the Performance Improvement Quarterly Journal.

Evaluation and Instructional Design Processes

To understand the importance of training evaluation, it is necessary to view evaluation within the context of instructional design processes. Instructional design processes are system approaches to the design of instructional material for training and educational programs. Goldstein (1993) referred to the system approach to training and education as instructional technology. Various models of instructional design processes exist to meet the training needs of military, business, and educational systems. Campbell (1984), Goldstein (1993), Mager (1988), Nadler and Nadler (1989), Kirkpatrick (1994), and Sredl and Rothwell (1992) have reviewed various systematic approaches to the development of training and educational programs.

Campbell (1984) and Sredl and Rothwell (1992) reviewed the Instructional Systems Development (ISD) model developed by the U. S. Department of Defense. Campbell (1984, p. 2) described it as a methodical process "for gathering and analyzing job information, writing learning objectives and tests, preparing and validating instructional materials in a variety of media, conducting instruction, and evaluating as well as improving the effectiveness of training." Sredl and Rothwell (1992, p. 350) stated, "ISD lends itself to all jobs, though it is associated with technical training more often than with so-called 'soft skills' instruction (such as communications or most management training)."

The ISD model consists of five phases: (a) analyze, (b) design, (c) develop, (d) implement, and (e) control. Each phase contains a series of steps which, when followed, are intended to lead the training professional through a systematic approach for the development of training and educational programs.

Sredl and Rothwell (1992) reviewed two methods of instructional design using the objectives-centered approach based on the learning and instructional assumptions of the behaviorist theory. According to the behaviorist theory, the purpose of learning is to produce prescribed behaviors—how one should perform (Gilley and Egglund, 1989). The two methods reviewed by Sredl and Rothwell (1992) were the decision-based method and the competency-based method. The decision-based method focused on correcting performance problems by discovering barriers or impediments to performance and designing training or educational plans to alleviate the performance problem. The competency-based method focused on exploiting

opportunities for performance improvement by defining good performance and then facilitating or improving it.

According to Sredl and Rothwell (1992), the decision-based method used careful analysis to discover underlying causes of performance problems so the improvement efforts could focus on the causes of the problems rather than the symptoms. It also required consideration of a range of improvement methods with training being one of the alternatives. Training is an expensive alternative and the decision-based model included cost-benefit analysis.

The competency-based method consisted of the following three models (Sredl and Rothwell, 1992): (a) the instructional systems design (ISD) model, discussed previously, (b) the behavioral skills-outputs (BSO) model, and (c) the developing-a-curriculum (DACUM) model. The BSO model emphasized work outputs which made it "uniquely suited for dealing with professional and managerial jobs" (Sredl and Rothwell, p. 352) because these types of jobs often involve unobservable mental activity. The BSO model also focused on desired performance.

The third competency based model reviewed by Sredl and Rothwell (1992) was the DACUM model. DACUM is based on the assumptions that expert or experienced workers are most competent to describe their own job, that any job can be analyzed according to its tasks, and that each task has direct implications for occupational training. According to Sredl and Rothwell:

DACUM is a cost-effective method to prepare off-the-job or on-the-job training. However, it has the disadvantage of being sponsored by a

governmental body. As a consequence, it is not easy to send HRD staff to DACUM workshops to learn the method, since governmental bodies do not effectively market their workshops. (p. 353)

Goldstein (1993) reviewed a systems approach to instructional technology which emphasized "the specification of instructional objectives, precisely controlled learning experiences to achieve these objectives, criteria for performance, and evaluative information" (p. 19). Goldstein's instructional system model consisted of three phases: (a) assessment, (b) training and development, and (c) evaluation.

Characteristics of Goldstein's model included the use of feedback to continually adapt and modify training programs based on information that indicated whether training programs are meeting their stated objectives. Another characteristic of Goldstein's model was the recognition of the complex interaction of the components of the system. An example of this interaction involves considering the use of different mediums to deliver the training based on the learning variables and characteristics of the individual trainees. Goldstein's model stressed a concern with the total program rather than with the objectives of any single course.

Still another characteristic of Goldstein's model emphasized the use of systematic analysis employing a research approach to provide a frame of reference for planning and evaluating training programs. The model also recognized that training programs "interact and are directly affected by a larger system involving corporate policies (for example, selection and management philosophy)," (Goldstein, p. 20).

Mager (1988) reviewed an instructional system design model which focused on changing human performance capability. The design model consisted of four major phases: (a) analysis, (b) development, (c) implementation, and (d) improvement.

Nadler and Nadler (1989) reviewed the Critical Events Model which focused on identifying and solving organizational problems associated with training needs. The model consisted of the following nine elements: (a) identify organizational needs, (b) specify job performance, (c) identify learner needs, (d) determine objectives, (e) build curriculum, (f) select instructional strategies, (g) obtain instructional resources, (h) conduct training, and (i) evaluate and obtain feedback throughout the entire process.

Kirkpatrick (1994) reviewed an instructional design model which focused on planning and implementing training programs to ensure their effectiveness. The model consisted of the following ten steps: (a) determine training needs, (b) set training objectives, (c) determine subject content, (d) select training participants, (e) determine the best schedule, (f) select appropriate training facilities, (g) select appropriate instructors, (h) select and prepare audiovisual aids, (i) coordinate the program, and (j) evaluate the program.

Training Evaluation

A pattern can be observed in all of the aforementioned system approaches to instructional design processes. The pattern includes the following steps: (a)

analysis, (b) design, (c) development, (e) implementation, and (d) evaluation.

Analysis consists of determining what an employee must know or be able to do to perform a job. This information is acquired through interviews, observations, or questionnaires involving job incumbents, supervisors, or other subject matter experts.

Analysis should also be used to determine if training is the proper treatment for performance deficiencies. If training is identified as the solution to performance deficiencies, information acquired in the analysis is used in training design.

The design step consists of using information acquired from analysis to write behavioral, or terminal objectives. Campbell (1984) suggested that behaviorally stated objectives are statements that specify measurable behavior that a trainee should exhibit upon completion of training, the environment or conditions under which the objective will be carried out, and the minimum standard or criteria to which the objective will be attained. Goldstein (1993) stated that terminal objectives "state the terminal behavior, the conditions under which the terminal behavior is expected, and the standard below which the performance is unacceptable" (p. 150). The design step also involves planning and sequencing the training experience or treatment and determining the media to support the behavioral objectives.

The development step consists of writing learning objectives and selecting or writing instructional materials to support the learning objectives. The learning objectives and the instructional materials should be reviewed at this point by subject matter experts to determine their relevance to the behavioral objectives of the training treatment or experience required of the job. If needed, the instructional

materials should be revised based on the subject matter expert review.

The implementation step consists of delivering the training treatment to the trainees. The instructional materials should be piloted by trainees and revised as needed.

The final step in the pattern of all the instructional design processes, except the DACUM model, is the evaluation step. Campbell (1984) stated that evaluation is "necessary to determine the effectiveness of training operating as a dynamic interrelated entity" (p. 27). According to Campbell, evaluation includes internal evaluation to determine if trainees meet the learning objectives and external evaluation to determine if trainees are applying the training on the job. Depending upon the results of the evaluations, the program and/or instructional materials should be reviewed and revised.

Internal evaluation "seeks to determine if instruction is providing trainees with the knowledge and skills necessary to attain the learning objectives" (Campbell, 1984, p. 28). Measures to acquire this information include trainee test scores, instructor observations, and trainee and instructor opinions concerning instruction and testing.

External evaluation seeks to determine if trainees are employed in the job for which they were trained and whether they can do the job for which they were trained. Measures to acquire this information include surveys, personal and telephone interviews, and questionnaires directed to the program graduates and their supervisors. According to the ISD model, external evaluations should be conducted

at an interval of at least three months after completion of the training program.

Information obtained through internal and external evaluations is used in conjunction with other data obtained during analysis to make decisions concerning the improvement, continuation, or cessation of the particular training or educational program.

Goldstein (1993) defined evaluation as "the systematic collection of descriptive and judgmental information necessary to make effective training decisions related to the selection, adoption, value, and modification of various instructional activities" (p. 147). To perform a complete training evaluation, Goldstein (p. 182) recommended that four questions be addressed:

1. Does an assessment of the various criteria indicate that a change has occurred?
2. If changes occurred, can they be attributed to the instructional program?
3. Is it likely that similar changes will occur for new trainees in the same program?
4. Is it likely that similar changes will occur for new trainees in the same program in a different organization?

Information collected in the evaluation step is used as feedback for the analysis step to use when making decisions concerning the training program.

Within Goldstein's (1993) model, the evaluation process centered around two procedures—establishing measures of success or criteria and using the necessary evaluation design to determine what changes in behavior occurred during the training

process. To establish criteria to use as measures of success, Goldstein recommended using the behavioral or terminal objectives established in the analysis step of the training model. The criteria should describe the behavior required to demonstrate the trainee's skill, the conditions under which the trainee is to perform, and the minimum limit of acceptable performance. Criteria should be established for both the evaluation of trainees at the conclusion of the training program and the evaluation of on-the-job performance.

Goldstein (1993) reviewed three evaluation design models which could be used to obtain information in four areas concerning training programs. The evaluation design models reviewed were: (a) individual-difference, (b) experimental, and (c) content. The four areas of information are: (a) training validity, (b) transfer validity, (c) intraorganizational validity, and (d) interorganizational validity. The individual-difference model relates individual performance in training to individual performance on the job. Goldstein suggested that this model could be used as a measure of training validity. Goldstein defined training validity as a measure of what trainees learned during training. Although this model can be used to select better performers, it does not tell the researcher very much about the training program.

The experimental evaluation model is based on the use of pretests and post tests and uses control groups as a way of accounting for training efforts. Goldstein (1993) suggested that this model could be used as a measure of transfer validity. Goldstein defined transfer validity as a measure to determine if what trainees learned

in training is transferred as improved performance in the work environment. The evaluation of on-the-job performance upon completion of a training program is identified as transfer validity in Goldstein's model.

The content evaluation model reviewed by Goldstein (1993) refers to whether the critical knowledge, skills, and abilities (KSA) identified in the analysis phase are emphasized in the training program. Goldstein (p. 28) stated, "That [content evaluation] does not assure that the material was learned; but then again, if the critical KSAs are not included in the training program, there is not much hope for the [training] effort."

The other two areas of information to be considered when evaluating training programs in Goldstein's (1993) model are intraorganizational validity and interorganizational validity. Goldstein defined intraorganizational validity of a training program as a measure of the degree to which the application of an instructional treatment to new trainees will result in similar on-the-job performance results obtained by the initial group of trainees. Goldstein defined interorganizational validity as a measure of the degree to which a training program used in one organization could be successfully used in another organization.

Goldstein (1993) offered the preceding design models and the information gathered from them as different methods to acquire information about training programs. He acknowledged that researchers, when determining evaluation methods for a specific training program, are limited by organizational constraints but suggested that for training programs to achieve their purpose, obtaining evaluation

information and using the information for program improvement is a must.

Mager (1988) suggested that the following questions be answered to evaluate instruction to determine if improvement was needed:

1. Does the instruction do what it's supposed to do?
2. Is the instruction of value to the trainees and the organization?
3. Is the instruction current (does it use the best available instructional techniques)?

Mager suggested that to be able to improve a program, a measure or measures must be established for some characteristic or characteristics of the program. These established measurements are referred to as standards. Other measurements are then made and compared with the established standard or standards. Judgements can then be made concerning the program. Mager (p. 182) defined this process as evaluation. He stated, "When you make a judgement based on a comparison of a measurement with a standard, you are evaluating. . . . If you have no standard against which to compare a measurement, you cannot tell whether the thing measured is OK or not OK" (p. 182).

Kirkpatrick (1994) stated that the reason for evaluating training is to determine the effectiveness of training programs in four areas. Evaluation should measure changes in behavior that occur as a result of training programs. Evaluation should determine what final results occur because of training programs. Evaluation should measure how participants react to aspects of training programs. Evaluation should measure the learning that takes place during training programs. Measures in

all four areas will provide evidence of the effectiveness of training programs, whether the programs are conducted in education, business, or industry.

Bloom, Madaus, and Hastings (1981) suggested that evaluations of training and education should be conducted. Bloom et al. stated that evaluation is the systematic collection of evidence to determine if certain changes are taking place in learners as well as to determine the amount of change in individual learners. Education is a process to change learners. Evaluation is an attempt to describe, appraise, and influence the changes rather than to analyze all the processes which bring learning about. Evaluation should be used to "relate the actualities of change in students to the formulation of changes sought" (Bloom, et al., p. 5). According to Bloom, et al., evaluation can be summative or formative.

Formative evaluation is the use of systematic evaluation in the development step and is used to determine if the course, curriculum, or program is operating as originally planned or if improvements are necessary. Formative evaluations are conducted during the development step in such a way that revisions of the course, curriculum, or program can be based on evidence gathered from systematic evaluations.

Summative evaluation is the use of systematic evaluation to determine if a curriculum or program is meeting its primary objectives. Summative evaluation is used to judge the overall effectiveness of a program and is based upon outcome criteria established during the analysis step.

Kirkpatrick's Four Levels of Evaluation

In the preceding paragraphs, the views of numerous experts have been offered concerning value, importance, and methods of conducting training evaluations.

Kirkpatrick constructed a four-step hierarchy of program evaluation in 1959 that remains the best known way to address the issue of training evaluation (Alliger and Janak, 1989; Carnevale and Schulz, 1990; Goldstein, 1993; Long, 1990; Plant and Ryan, 1994; Rothwell and Sredl, 1992).

Kirkpatrick (1959) suggested that training organizations use the data from these evaluations to make decisions concerning training and training programs. According to Kirkpatrick, the four levels are arranged in a hierarchy such that results level data are the most useful to the organization followed in order of decreasing usefulness by behavior level, learning level, and reactions level.

The first and lowest level of Kirkpatrick's hierarchy is an evaluation of the participant's reaction to the training. This may be defined as how well the learner liked or disliked the training program. Kirkpatrick (1994) stated that measuring participant reaction to training programs is the same thing as measuring customer satisfaction. According to Kirkpatrick, measuring participant reaction provides the training professional with valuable information concerning improvement opportunities in the areas of facilities, schedules, exercises, audiovisual aids, handouts, and the value participants place on the individual aspects of the program. Questions or surveys related to these areas are usually answered by participants and the answers and evaluations are strictly subjective and usually conducted at the end of the training

program.

The second level of Kirkpatrick's hierarchy is an evaluation of participant's learning. This type of evaluation measures knowledge, skills, or attitudes gained from the training. Kirkpatrick (1994) stated, "It is important to measure learning because no change in behavior can be expected unless one or more of these learning objectives have been accomplished" (p. 42). To conduct an evaluation of participant's learning, Kirkpatrick (1994) suggested the use of the following guidelines: (a) use a control group if practical, (b) evaluate knowledge, skills, and/or attitudes both before and after training, (c) get a 100% response from training participants, and (d) use the results of the evaluation to take appropriate action. The format of the evaluation can be a paper-pencil test or it can be a demonstration of skill.

The third level of Kirkpatrick's hierarchy is an evaluation of the participant's behavior following the training. This type of evaluation measures the participant's application of skills and/or knowledge gained during training on the job. Goldstein stated (1993, p. 165), "Just as favorable reaction does not necessarily mean that learning will occur in the training program, superior training performance does not always result in similar behavior in the transfer setting." To conduct an evaluation of participant's change in job behavior, Kirkpatrick (1994) suggested the use of the following guidelines: (a) use a control group if practical, (b) allow time for behavior change to take place, (c) evaluate both before and after the training program if practical, (d) survey or interview trainees, their immediate supervisor, their

subordinates, or others who observe their behavior, (e) get 100% response or a sampling of the training participants, (f) repeat the evaluation at appropriate times, and (g) consider cost of the evaluation versus benefits. This type of evaluation can be done in an objective or subjective manner. If self-evaluated the results are usually very subjective. If the evaluation is conducted by an independent assessor or by the participant's supervision, the evaluation becomes more objective. This type of evaluation is usually conducted at the work site at a minimum of three months following the training.

The fourth and highest level of Kirkpatrick's hierarchy is an evaluation of the results of training. This evaluation measures the impact training has on the business results of the organization. In most cases it is the toughest to conduct (Gordon, 1987; Mitchell, 1994; Pine & Tingley, 1993; Tannenbaum and Woods, 1992). To conduct an evaluation of the results of a training program, Kirkpatrick (1994) suggested the use of the following guidelines: (a) use a control group if practical, (b) allow time for results to be achieved, (c) measure both before and after the program if practical, (d) repeat the measurement at appropriate times, (f) consider cost versus benefits, and (g) be satisfied with evidence if proof is not possible. The results can be subjective, but for it to be a true evaluation of training results, the evaluation must be as objective as possible.

In addition to the aforementioned four levels of training evaluation, Phillips (1991) proposed a fifth level which attempts to measure return on investment of training programs. This type of measurement would be conducted to determine if

the cost of the program is justified. Phillips (1991) stated, "Overall, this type of evaluation provides management with data to eliminate an unproductive program, to increase support for programs which yield a high payoff, or to make adjustments in a program to increase the benefits" (p. 41).

As stated previously, Kirkpatrick's four-step hierarchy model remains the best-known way to evaluate training. In a study conducted by Long (1990) to test the hierarchical arrangement of Kirkpatrick's four levels by using data collected at all four levels during an actual program evaluation, the results of the study did not agree with Kirkpatrick's hierarchy as far as the importance related to decision making is concerned. The results of the study indicated that evaluations of training programs should be designed to ensure that behavior level evaluation data are collected. Long's study revealed that behavior level data were shown to have the greatest decision making utility, and should therefore be included as a minimum in any training program evaluation.

Studies by others (Alliger and Janak, 1989; Mitchell, 1994; Moore, 1984; Tannenbaum and Woods, 1992; and Waste Isolation Division, Westinghouse Electric Corporation, 1995) present evidence to support Long's contention. Alliger and Janak (1989) conducted a study to examine the validity of three assumptions associated with Kirkpatrick's evaluation model. During the study, a literature search was conducted for any article dealing with evaluation of a training program. Journals were reviewed from 1959 to 1988. The number of articles reporting training evaluation results totaled 203. Of this number, 11 articles pertained to level one evaluations, 61

articles pertained to level two evaluations, 66 articles pertained to level three evaluations, and 11 articles pertained to level four evaluations. The remaining 54 articles pertained to various combinations of the four levels of evaluations with level three included in 37 of these combinations. Thus, in this study, reports of level three evaluations or combinations of the four levels of evaluations including level three made up greater than 50% of the total number of articles pertaining to training evaluation.

A study conducted by the Human Resources Development Group at the U. S. Office of Personnel Management (OPM) and reported by Mitchell (1994) indicated that one of the primary reasons for conducting training was to increase employee job performance. To evaluate the results of training, the OPM evaluation model emphasized performance evaluation and measurement of participant use of skills on the job. Mitchell (p. 213) stated, "Decision makers seek analytic evidence to determine whether skills are being developed and how much employees are learning." Kirkpatrick's level three evaluation was deemed most appropriate for providing decision makers with the evidence they sought concerning various training programs being conducted in numerous agencies throughout the OPM.

A research project conducted by the Massachusetts Mutual Life Insurance Company and reported by Moore (1984) measured the changes in job performance resulting from a self-development training program for almost 1000 secretarial and clerical employees. The research design was modeled on Kirkpatrick's guidelines for evaluating training programs in terms of behavioral changes. According to Moore,

the research revealed important benefits including the ability to provide follow-up training for employees who performed poorly while targeting the additional training with great accuracy. Also, Moore reported that revisions in course content could be made to intensify the program's impact and improve the quality of the program, results of the research increased management's commitment to training, and results of the research demonstrated that training dollars were paying off in increased employee competence levels and improved performance.

Tannenbaum and Woods (1992) reported on an evaluation of a training program for supervisors at a large financial services company. The purpose of the evaluation was to determine if the training program was effective and worth the resources dedicated to it. The report revealed that the use of Kirkpatrick's level two and three evaluations provided the management development group with "defensible, quantitative indicators of learning, behavior, and/or results in order to move management training out of the 'nice to have' mindset and into the 'necessary to do' category" (p. 76).

Research conducted by the Organizational Development (OD) section of the Waste Isolation Division of the Westinghouse Electric Corporation for the U. S. Department of Energy (1995) indicated that it needed an improved mechanism for evaluating the effectiveness of its classroom training. Similar to many training organizations facing reduced budgets, increased scrutiny, and the challenge to continually improve its training programs, OD recognized that Kirkpatrick's level one and two evaluations were not providing the necessary information to make

appropriate decisions concerning its many training programs. Results of the research revealed that the most important aspect of a training program for the decision makers within the organization was the transfer of training from the training setting to the job. OD developed the Transfer of Training Evaluation Model (TOTEM) to provide decision makers within the organization evidence for revising course content to improve the transfer of training. TOTEM is based on Kirkpatrick's level three evaluation.

Summary

Evaluation of training programs is an integral part of systematic training processes. From the literature review, it is apparent that evaluations of training programs should be conducted. Kirkpatrick's (1959) evaluation model is perhaps the most recognized model in the training profession. Long's (1990) study indicates that when time and money are limited, the focus of training program evaluation should be on the behavior level of evaluation. This level of evaluation produces information that is the most useful for making decisions concerning training programs. Studies by others (Alliger and Janak, 1989; Mitchell, 1994; Moore, 1984; Tannenbaum and Woods, 1992; and Waste Isolation Division, Westinghouse Electric Corporation, 1995) present evidence to support Long's study. The other evaluation levels of reactions, learning, and results should be added to the program evaluation design as time and money permit.

CHAPTER III

RESEARCH METHOD

This chapter presents the methodology used in the study. The following sections of the chapter describe the subjects, the instrument, the data collection process, the method for processing the data, and the summary.

To test the two research hypotheses of this study and to meet the criteria of both Kirkpatrick's (1959) level three evaluation and Phillips' (1991) pretest and post-test model design, a four-phase study was necessary to obtain and analyze the data. The initial phase of the study measured the perceived performance of 43 apprentice candidates on a representative sample of tasks which comprise the job of control system mechanic.

Phase two of the study measured the perceived performance of 42 apprentice graduates on the same representative sample of tasks used in phase one of the study. Phase three of the study measured the perceived performance of the 42 apprentice graduates in phase two of the study by the supervisors of the graduates. The same representative sample of tasks used in phases one and two of the study were used in phase three.

Phase four of the study consisted of data analysis of the measurements taken in phases one, two, and three. Quantitative data collected during phases one, two, and three were analyzed and information was extracted using the statistical capabilities of Excel version 5.0. The purposes of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance

apprenticeship training program and to identify specific modifications and/or improvements which could be made to the training program. The following sections describe the methods of the study in detail.

Subjects

The population consisted of employees of the manufacturing facility. After obtaining permission to conduct the study, the researcher obtained a list of apprentice candidates for 1996 from the Manufacturing Training Supervisor. This list identified the population for phase one of the study. The entire population of apprentice candidates were included in the study ($n=43$).

The researcher also generated a list of graduates of the CSM apprenticeship program between the dates of May 1, 1994 and May 1, 1996 by using the Training Information Management System (TIMS) of the manufacturing facility. This list identified the total population ($n=63$) of most recent graduates of the CSM apprenticeship program. The researcher then generated a list of supervisors of the 63 graduates using the Employee Data Base System (EDBS) of the manufacturing facility. Due to the distribution of graduates among supervisors of CSMs, 47 graduates were selected from the population to participate in the study. Three factors determined the selection criteria: (a) no more than six apprentice graduates were to be evaluated by any one supervisor, (b) no less than two apprentice graduates were to be evaluated by any one supervisor, and (c) the most recent apprentice graduates were to be selected to participate in the study. This list

identified the sample ($n=47$) of CSM apprentice graduates selected to participate in phase two of the study.

The list of supervisors of CSM apprentice graduates identified above also identified the population of supervisors ($n=12$) selected to participate in the study. The researcher received 42 completed surveys from CSM apprentice graduates. The supervisors of these graduates were selected to participate in phase three of the study.

Phase one of the study was comprised of 43 apprentice candidates. The candidates were 100% of the population of candidates for apprenticeships at the manufacturing facility during 1996. The sample for phase two of the study was comprised of 47 of the most recent graduates of the CSM apprenticeship from a pool of 63 CSM apprentice graduates. The sample for phase three of the study was comprised of 12 supervisors of CSMs from a pool of 15 supervisors of CSMs.

Instrument

The researcher developed the instrument for the study. The instrument was a survey (see Appendix A) which consisted of 30 tasks and performance standards for the tasks established by the 1988 CSM job analysis and a modified Likert scale for rating of performance on each task. The instrument was validated by a panel of experts of the manufacturing facility consisting of three CSM general supervisors, one maintenance department head, and the maintenance division personnel representative. The instrument was also approved by a panel of college professors.

Both the panel of experts and the panel of college professors provided feedback on the survey items, construction, and appearance from different perspectives and were crucial to good instrument development.

The instrument gathered information on perceived performance of apprentice candidates, apprentice graduates (by both the graduates and their supervisors) on 30 significant tasks identified in a 1988 job analysis of the CSM craft. The tasks were a representative sample of work which comprised the job of CSM at the manufacturing facility. The selected tasks were drawn from five areas of work related to the specific job of CSM. These areas were: (a) distributive control systems (DCS) and programmable-logic controllers (PLC's) maintenance, (b) electrical maintenance, (c) electronic maintenance, (d) electronic instrumentation maintenance, and (e) pneumatic instrumentation maintenance.

Collection of Data

During phase one of the study, the researcher collected data by administering surveys during orientation sessions for apprentice candidates prior to the beginning of the training program.

During phase two of the study, the researcher sent an electronic note to all of the potential participants which explained the nature of the study and requested the graduates to participate in the study (See Appendix B). The researcher also mailed via intra-plant mail information packets which included instructions for completing the survey (See Appendix C), the survey (See Appendix A), and an addressed

envelop for return of the completed survey. Each participant in the study was requested to provide his/her name and insurance number as a means of identification. A list of participants was maintained by the researcher to identify non-respondents. The confidentiality of individual responses was stressed.

During phase three of the study, the researcher sent an electronic note to all of the selected supervisors of CSMs which explained the nature of the study and requested the supervisors to participate in the study (See Appendix D). The researcher also mailed via intra-plant mail information packets which included instructions for completing the survey (See Appendix E), the survey (See Appendix A), and an addressed envelop for return of the completed survey. The researcher provided the name of the apprentice graduate to be evaluated and the name of the supervisor evaluating the graduate on each survey mailed to the selected supervisors. A list of participants was maintained by the researcher to identify non-respondents. The confidentiality of individual responses was stressed.

The researcher obtained information on the participants' perceived ability to perform each of 30 significant tasks to the stated standards. The participants' perceived ability to perform each of the tasks was rated on a modified Likert scale with ratings of one through four. A rating of one indicated the participant could not perform the task. A rating of two indicated the participant needed improvement to perform the task and could not perform the task to all stated standards. A rating of three indicated the participant could meet expectations and perform the task to all stated standards. A rating of four indicated the participant exceeded expectations and

could perform the task to all stated standards.

The tasks and standards were identified in a 1988 job analysis of the CSM craft. The tasks were a representative sample of work which comprised the job of CSM at the manufacturing facility. The selected tasks were drawn from five areas of work related to the specific job of CSM. These areas of work were designated as job groups and the five job groups were: (a) Job Group A = distributive control systems (DCS) and programmable-logic controller (PLC) maintenance, (b) Job Group B = electrical maintenance, (c) Job Group C = industrial electronics maintenance, (d) Job Group D = electronic instrumentation maintenance, and (e) Job Group E = pneumatic instrumentation maintenance.

Data Analysis

Information was extracted from the data using statistics capabilities of Excel version 5.0. Summary statistics for the apprentice candidates' and graduates' performance evaluations of the tasks for each job group were computed by calculating frequencies and percentiles. Chi-square tests of homogeneity were used to detect the presence or absence of statistically significant differences between the apprentice candidates' and graduates' performance evaluations for each job group at the .05 level of significance. In the study, one null hypothesis was stated for research hypothesis one and was denoted H_{o1} .

Summary statistics for the apprentice graduates' performance evaluations of the tasks for each job group and their supervisor's performance evaluations of the

apprentice graduates for each job group were computed by calculating frequencies and percentiles. Chi-square tests of homogeneity were used to detect the presence or absence of statistically significant differences between the apprentice graduates' performance evaluations for each task and their supervisor's performance evaluations of the apprentice graduates for each task at the .05 level of significance. In the study, one null hypothesis was stated for research hypothesis two and was denoted H_{o2} .

In cases where null hypotheses failed to be rejected, the calculated values of chi-square were less than the values obtained from a chi-square distribution chart. In cases where null hypotheses were rejected, the calculated values of chi-square were greater than the values obtained from a chi-square distribution chart.

Secondary Data Sources

Library material, interviews, and articles were also researched. Attention was paid to the advice given on statistical testing by Dr. Gerald Cheek, Dr. Robert Hanson, and Dr. Roger Haskell.

Summary

The purposes of this study were to measure the effectiveness of a maintenance apprenticeship training program and provide feedback to training designers for program improvement. Products of the study included a model to evaluate other technical training programs and data for management to use in the decision-making

process related to technical training programs.

The objectives of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program and to identify specific modifications and/or improvements which could be made to the apprenticeship training program.

Summary statistics and chi-square tests of homogeneity were used to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program by comparing performance evaluations of apprentice candidates and apprentice graduates on a representative sample of tasks which comprised the job of CSM at the manufacturing facility.

Summary statistics and chi-square tests of homogeneity were used to identify specific modification and/or improvement opportunities for the training program by comparing performance evaluations of apprentice graduates and supervisors' performance evaluations of the apprentice graduates on a representative sample of tasks which comprised the job of CSM at the manufacturing facility.

This chapter presented and discussed the methodology used in the study. As described, a four phase study was conducted and described in detail to provide the foundation for the results that will be presented in Chapter IV.

CHAPTER IV

ANALYSIS OF STATISTICAL FINDINGS

To test the two research hypotheses of this study and to meet the criteria of both Kirkpatrick's (1959) level three evaluation and Phillips' (1991) pretest and post-test model design, a four-phase study was necessary to obtain and analyze the data. A representative sample of tasks which comprised the job of CSM was identified. Study participants were requested to evaluate their ability to perform each task to defined standards or in the case of supervisors, they were asked to evaluate their subordinates' ability to perform each of the tasks to defined standards (See Appendix C).

The purposes of this study were to measure the effectiveness of a maintenance apprenticeship training program and to provide feedback to training designers for program improvement. Products of the study included a model to evaluate other technical training programs and data for management to use in the decision-making process related to technical training programs.

The objectives of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program and to identify specific modifications and/or improvements which could be made to the training program.

This study evaluated the performance of apprentice candidates and apprentice graduates on a representative sample of tasks which comprised the job of CSM at the manufacturing facility. Apprentice candidates were self-evaluated. Apprentice

graduates were self-evaluated and were also evaluated by their supervisors.

This chapter presents and discusses results of the measurements proposed in the model. This chapter also presents and discusses results of the tests of the hypotheses put forth in Chapter I. The chapter begins by presenting the results of the study, which, as discussed in Chapter III, used surveys to collect quantitative data. An analysis of the sample in the study is discussed. Overall summary statistics for each job group's performance evaluations are reported. These summary statistics include graphical presentations which display percentages. In the last section, results of the testing are discussed.

Research hypothesis one of the study was tested using a chi-square test of homogeneity. This statistical test was used to detect the presence or absence of statistically significant differences between the apprentice candidates' and graduates' performance evaluations of each job group at the .05 level of significance. In the study, one null hypothesis was stated for research hypothesis one.

Research hypothesis two of the study also was tested using a chi-square test of homogeneity to detect the presence or absence of statistically significant differences between the apprentice graduates' performance evaluations for each task and their supervisor's performance evaluations of the apprentice graduates for each task at the .05 level of significance. In the study, one null hypothesis was stated for research hypothesis two.

Study Sample

The subjects were chosen from a population of employees of the manufacturing facility which consisted of maintenance apprenticeship training program candidates, graduates of the CSM maintenance apprenticeship training program, and supervisors of graduates of the CSM maintenance apprenticeship training program. The population contained 43 apprentice candidates, 63 apprentice graduates who had completed the program in the last two years, and 15 supervisors. Forty-three apprentice candidates (100% of the population of apprentice candidates) were selected to participate in the study.

Due to the distribution of apprentice graduates among supervisors of CSMs, 47 graduates (75% of the population of CSM apprentice graduates) were selected to participate in the study. The following three factors determined the selection criteria for apprentice graduates: (a) no more than six apprentice graduates were to be evaluated by any one supervisor, (b) no less than two apprentice graduates were to be evaluated by any one supervisor, and (c) the most recent apprentice graduates were to be selected to participate in the study.

Supervisors of the sample of apprentice graduates were selected to participate in the study. Twelve supervisors (80% of the population of supervisors) were selected to participate in the study.

The study sample consisted of 43 apprentice candidates, 47 apprentice graduates, and 12 supervisors. The researcher received 43 completed surveys from apprentice candidates for a return rate of 100%. The researcher received 42

completed surveys from CSM apprentice graduates for a return rate of 89%. The researcher received 37 completed surveys from 11 supervisors for a return rate of 88% of the surveys from 92% of the supervisors.

Summary Statistics

As discussed in Chapter III, the researcher obtained information on participants' perceived ability to perform each of 30 significant job tasks to stated standards. Participants' perceived ability to perform each of the tasks was rated on a modified Likert scale with ratings of one through four. A rating of one indicated participants could not perform the task. A rating of two indicated participants needed improvement to perform the task and could not perform the task to the stated standards. A rating of three indicated participants could meet expectations and perform the task to all stated standards. A rating of four indicated participants exceeded expectations and could perform the task to all stated standards. See Appendix F for a summary of apprentice candidates' performance evaluations, Appendix G for a summary of apprentice graduates' performance evaluations, and Appendix H for a summary of supervisors' performance evaluations of graduates.

Due to the absence of counts in several of the rating cells, the researcher collapsed the four rating cells into two rating cells to conduct tests of the research hypotheses. Ratings one and two were collapsed into one rating designated as "No, could not meet the stated standards for task performance". Ratings three and four were collapsed into one rating designated as "Yes, could meet or exceed the stated

standards for task performance". For brevity, these ratings are represented as "No" and "Yes" throughout the remainder of this study.

A graphical representation of the ratings is presented in Figure 3. The ratings are the percentage of "Yes" ratings for all Job Groups by candidates, graduates, and supervisors. The ratings are displayed by candidates', graduates', and supervisors' replies for each Job Group.

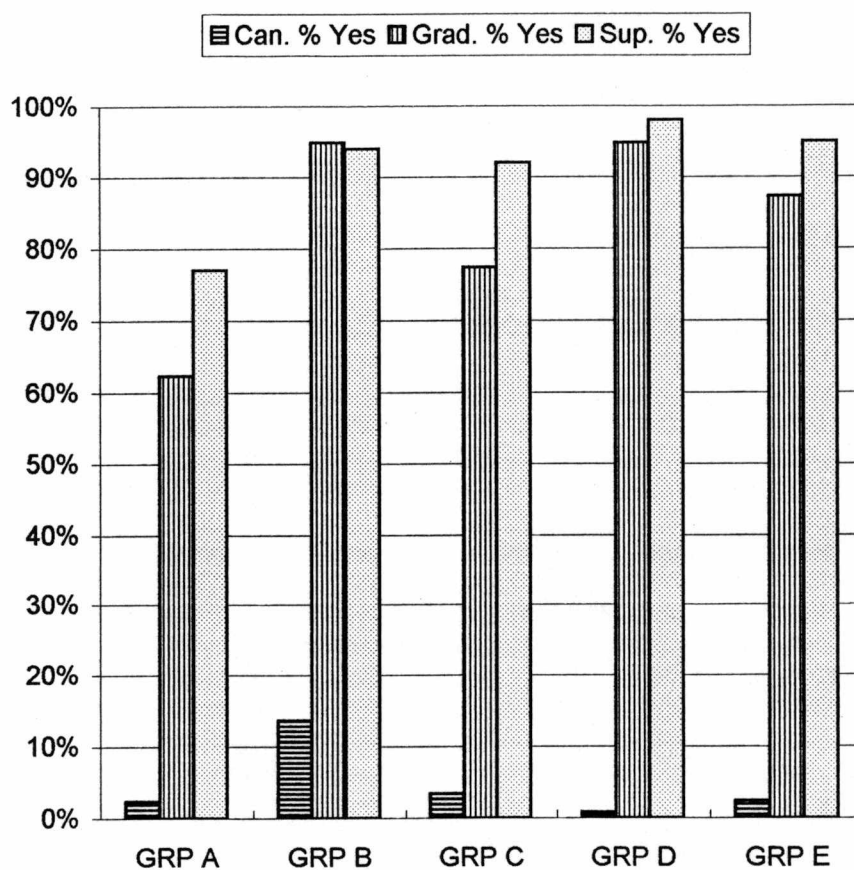


Figure 3. Summary of Candidates'/Graduates'/Supervisors' Performance Evaluations

Null Hypotheses

H_{o1} : There is no statistically significant difference between the perceptions of apprentice candidates and apprentice graduates with regard to their abilities to perform the 30 selected job tasks to defined standards.

As presented in Table 1, 85 participants rated their perceived performance for each of the five tasks of Job Group A. Among those, 43 replies were received from 43 apprentice candidates and 42 replies were received from 42 apprentice graduates. Approximately 2% of the candidates indicated the tasks of Job Group A could be performed to standards. Approximately 62% of the graduates indicated the tasks of Job Group A could be performed to standards (See Figure 3).

As a result of the analysis, a critical value of chi-square with 9 degrees of freedom and $\alpha = .05$ was obtained from the chi-square distribution table (See Appendix G). The critical value of chi-square was 16.9. The calculated value of chi-square for Job Group A was 176.4. Consequently, the null hypothesis was rejected. As one would expect, apprentice graduates' perceptions regarding their ability to perform each of the tasks in Job Group A to standards were significantly higher than the perceptions of apprentice candidates. Moreover, this finding held true for each of the Job Groups A, B, C, D, and E as well as for the specific tasks that made up each of these Job Groups.

Table 1. Comparison of Candidates'/Graduates' Task Evaluations

Candidates'/Graduates' Task Evaluations					
Task/ Job Group	Candidates' Yes Replies (N = 43)	Graduates' Yes Replies (N = 42)	Difference	chi- square	p
A1	2%	64%	62%	37	p < .05
2	2%	48%	46%	23.6	p < .05
3	5%	74%	69%	42.8	p < .05
4	2%	71%	69%	43.9	p < .05
5	0%	55%	55%	32.1	p < .05
Job Group A	2%	62%	60%	176.4	p < .05
B1	16%	90%	74%	47.2	p < .05
2	14%	100%	86%	64.1	p < .05
3	12%	88%	76%	49.9	p < .05
4	26%	98%	72%	46.4	p < .05
5	5%	93%	88%	67.1	p < .05
6	12%	100%	88%	67.7	p < .05
7	12%	95%	83%	59.9	p < .05
Job Group B	14%	95%	81%	402.2	p < .05
C1	0%	93%	93%	73.6	p < .05
2	2%	93%	91%	69.6	p < .05
3	7%	83%	76%	49.9	p < .05
4	5%	81%	77%	50.6	p < .05
5	8%	45%	37%	16.1	p < .05
6	0%	69%	69%	45.2	p < .05
Job Group C	3%	77%	74%	305.1	p < .05
D1	0%	98%	98%	81.7	p < .05
2	0%	95%	95%	77.1	p < .05
3	0%	93%	93%	73.8	p < .05
4	0%	83%	83%	60.9	p < .05
5	0%	100%	100%	84.6	p < .05
6	5%	100%	95%	77.4	p < .05
Job Group D	1%	95%	94%	455.5	p < .05
E1	0%	90%	90%	70.2	p < .05

Table 1. Comparison of Candidates'/Graduates' Task Evaluations, continued

Candidates'/Graduates' Task Evaluations					
Task/ Job Group	Candidates' Yes Replies (N = 43)	Graduates' Yes Replies (N = 42)	Difference	chi- square	p
E2	2%	83%	81%	57	p < .05
3	0%	83%	83%	60.9	p < .05
4	2%	93%	91%	69.6	p < .05
5	5%	86%	81%	56.3	p < .05
6	5%	88%	83%	59.4	p < .05
Job Group E	2%	87%	85%	373.4	p < .05

H₀₂: There is no statistically significant difference between the perceptions of apprentice graduates and their immediate supervisors with regard to the graduates' ability to perform each of the 30 selected job tasks to defined standards.

Data related to null hypothesis two are presented in Table 2. Supervisors' perceptions of skill levels of apprentice graduates and apprentice graduates' perceptions of their own skill levels were significantly different for 20% of the tasks. Chi-square tests of the null hypothesis for each of these tasks identified statistically significant differences between performance evaluations of CSM apprentice graduates by supervisors and by graduates on the following six job tasks: A2, A3, C5, C6, E2, and E3. Therefore, the researcher rejected the null hypothesis associated with each of these tasks. In each case, supervisors of apprentice graduates were

Table 2. Comparison of Supervisors'/Graduates' Task Evaluations

Supervisors'/Graduates' Task Evaluations					
Task	Supervisors' Yes Replies (N = 11)	Graduates' Yes Replies (N = 42)	Difference	chi-square	p
A1	79%	64%	15%	2.090	ns
2	71%	48%	23%	4.067	p < .05
3	92%	74%	18%	4.411	p < .05
4	81%	71%	10%	1.003	ns
5	53%	55%	2%	0.020	ns
B1	86%	90%	2%	0.310	ns
2	95%	100%	5%	2.537	ns
3	92%	88%	4%	0.312	ns
4	100%	98%	2%	0.892	ns
5	92%	93%	1%	0.026	ns
6	97%	100%	3%	1.150	ns
7	95%	95%	0%	0.017	ns
C1	100%	93%	7%	2.747	ns
2	100%	93%	7%	2.747	ns
3	95%	83%	12%	2.471	ns
4	86%	81%	5%	0.438	ns
5	78%	45%	33%	9.062	p < .05
6	92%	69%	23%	6.356	p < .05
D1	100%	98%	2%	0.892	ns
2	100%	95%	5%	1.808	ns
3	100%	93%	7%	2.674	ns
4	92%	83%	9%	1.303	ns
5	100%	100%	0%	-	-
6	97%	100%	3%	1.150	ns
E1	97%	90%	10%	1.470	ns
2	97%	83%	14%	3.910	p < .05
3	100%	83%	17%	6.416	p < .05
4	89%	93%	4%	0.328	ns
5	92%	86%	6%	0.744	ns
6	97%	88%	9%	2.373	ns

significantly more confident that graduates could perform the tasks to the identified standards than were the graduates themselves.

Supervisors' perceptions of skill levels of apprentice graduates and apprentice graduates' perceptions of their own skill levels were congruent for the remaining 80% of the tasks. Chi-square tests of the null hypothesis for each of these tasks did not identify significant differences. Therefore, the researcher failed to reject the null hypothesis associated with each of these tasks.

No other significant disagreements between supervisors of apprentice graduates and apprentice graduates were found in the data analysis. However, data presented in Table 2 provided the researcher with necessary information to establish criteria for identifying modification and improvement opportunities for the CSM apprenticeship program. As postulated in research hypothesis two, specific improvement and modification opportunities were identified by comparing supervisors' performance evaluations of apprentice graduates and apprentice graduates' performance evaluations of themselves.

Based on a minimum performance standard of 80% skill levels, the following specific job tasks were found to be deficient:

Job tasks A1, A2, A5, and C5 were each identified by supervisors as problem areas. Additionally, graduates concurred that they lacked adequate skill levels with each of these job tasks. Graduates also identified three other specific job tasks (A3, A4, and C6) where they felt a deficiency existed.

Specific improvement and modification opportunities are discussed in Chapter V.

Graphical representations of the ratings for the tasks of all Job Groups are displayed in Figure 4a, 4b, 4c, 4d, and 4e. Figure 4a displays the ratings for Job Group A. Figure 4b displays the ratings for Job Group B. Figure 4c displays the ratings for Job Group C. Figure 4d displays the ratings for Job Group D. Figure 4e displays the ratings for Job Group E. The ratings are the percentage of "Yes" replies by graduates and supervisors for the tasks of all Job Groups.

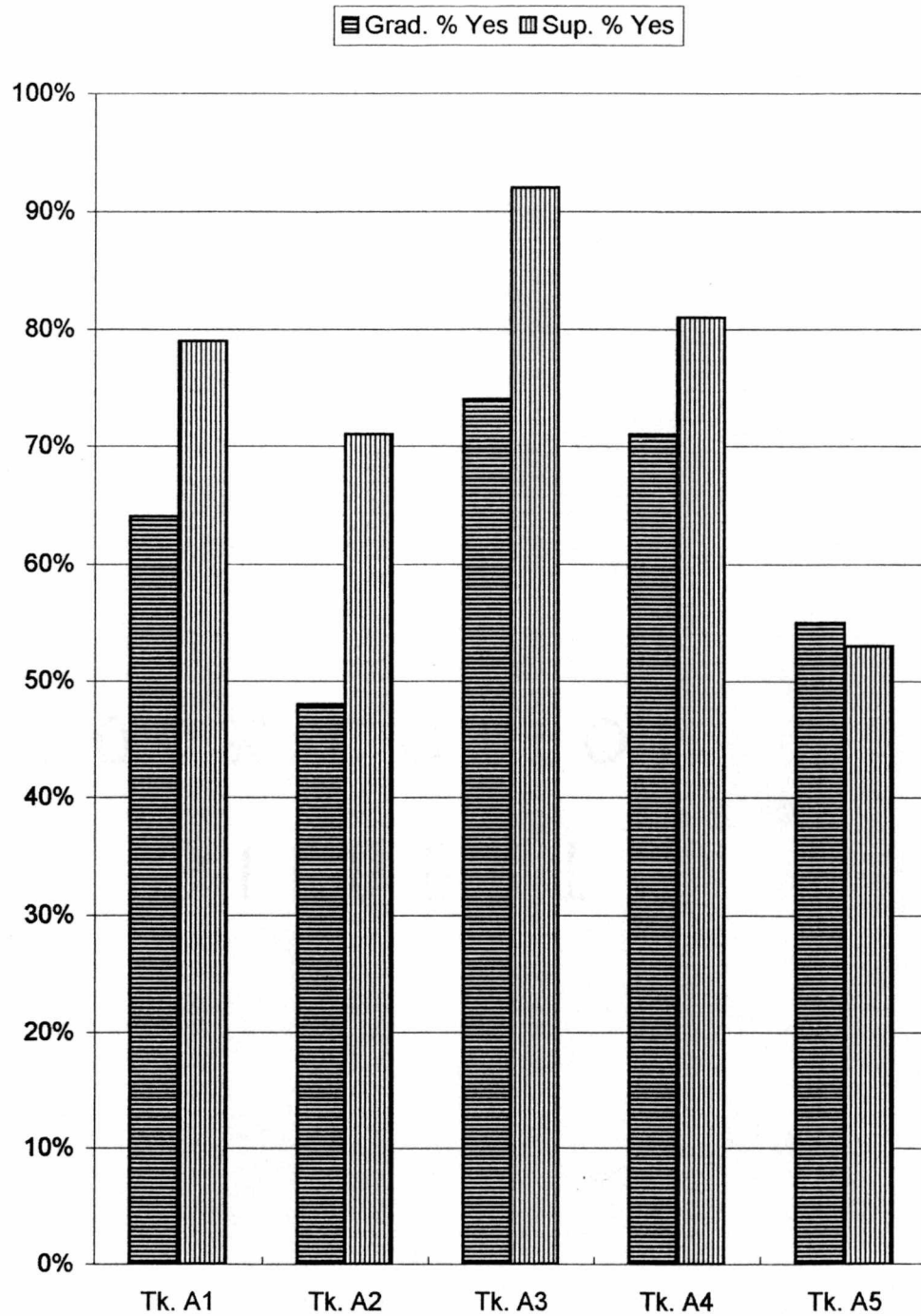


Figure 4. Summary of Graduates'/Supervisors' Performance Evaluations: (a) Job Group A (DCS and PLC Maint.)

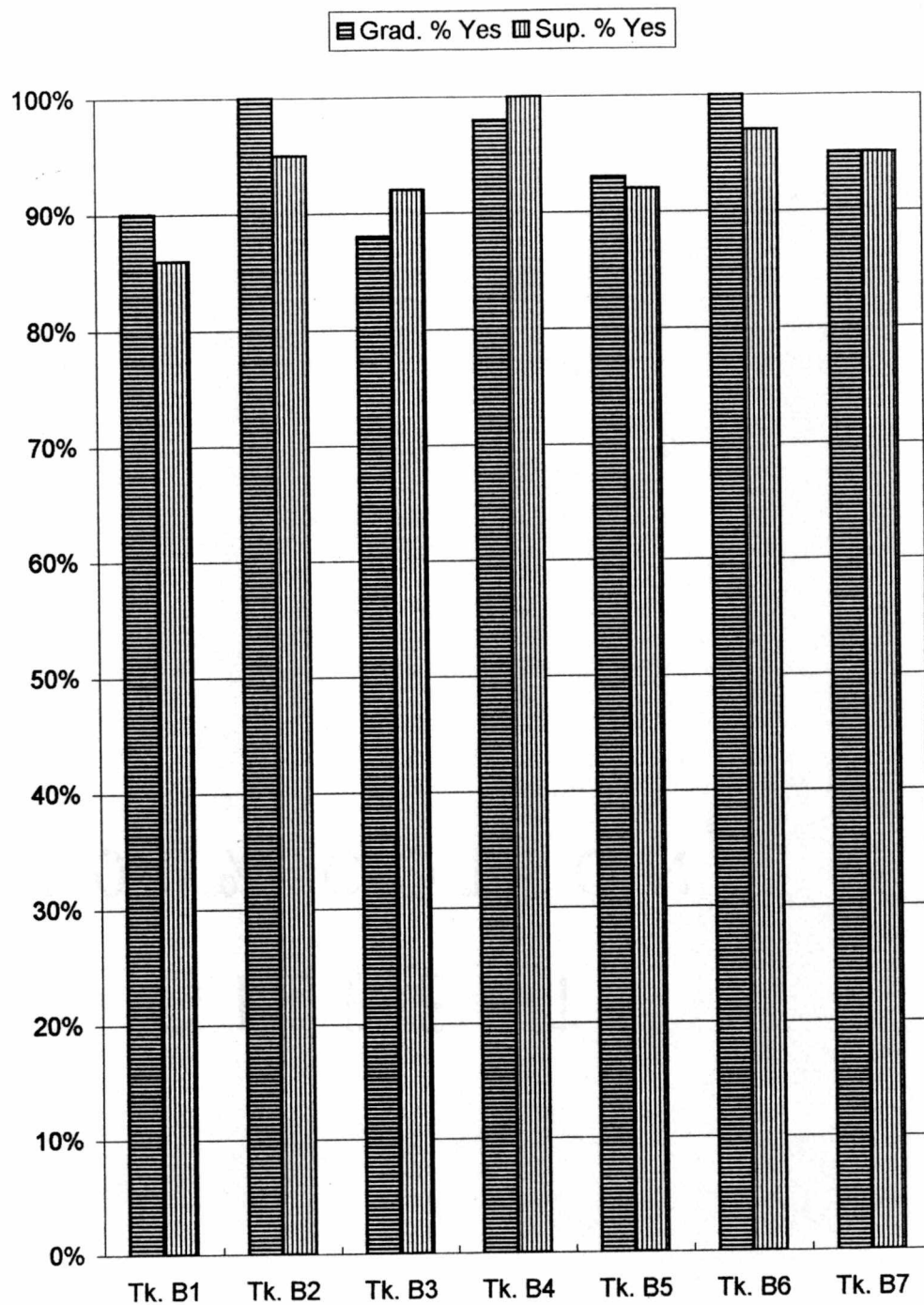


Figure 4. (continued)

Summary of Graduates'/Supervisors' Performance Evaluations: (b)
Job Group B (Electrical Maint.)

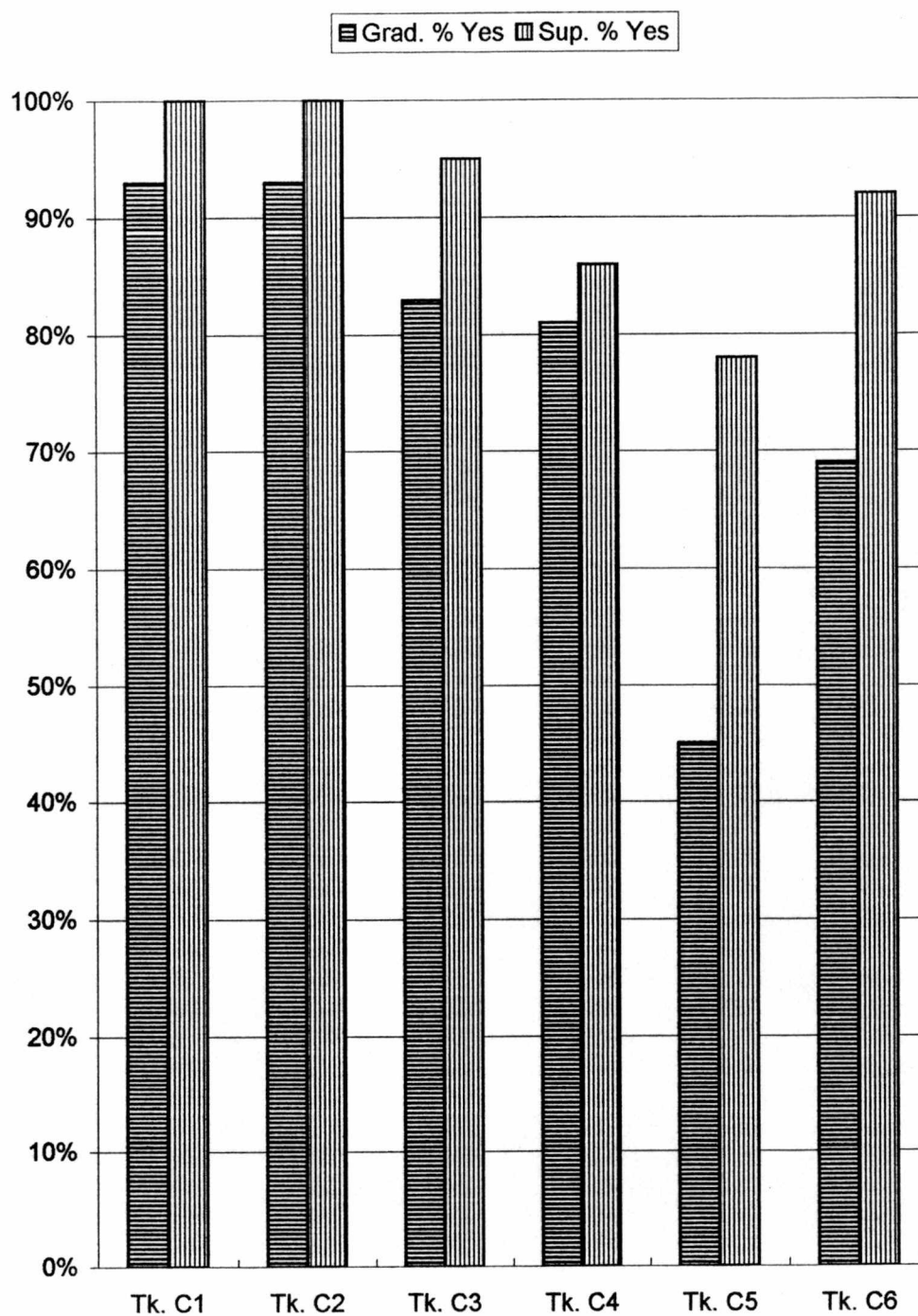


Figure 4. (continued)
Summary of Graduates'/Supervisors' Performance Evaluations: (c)
Job Group C (Electronic Maint.)

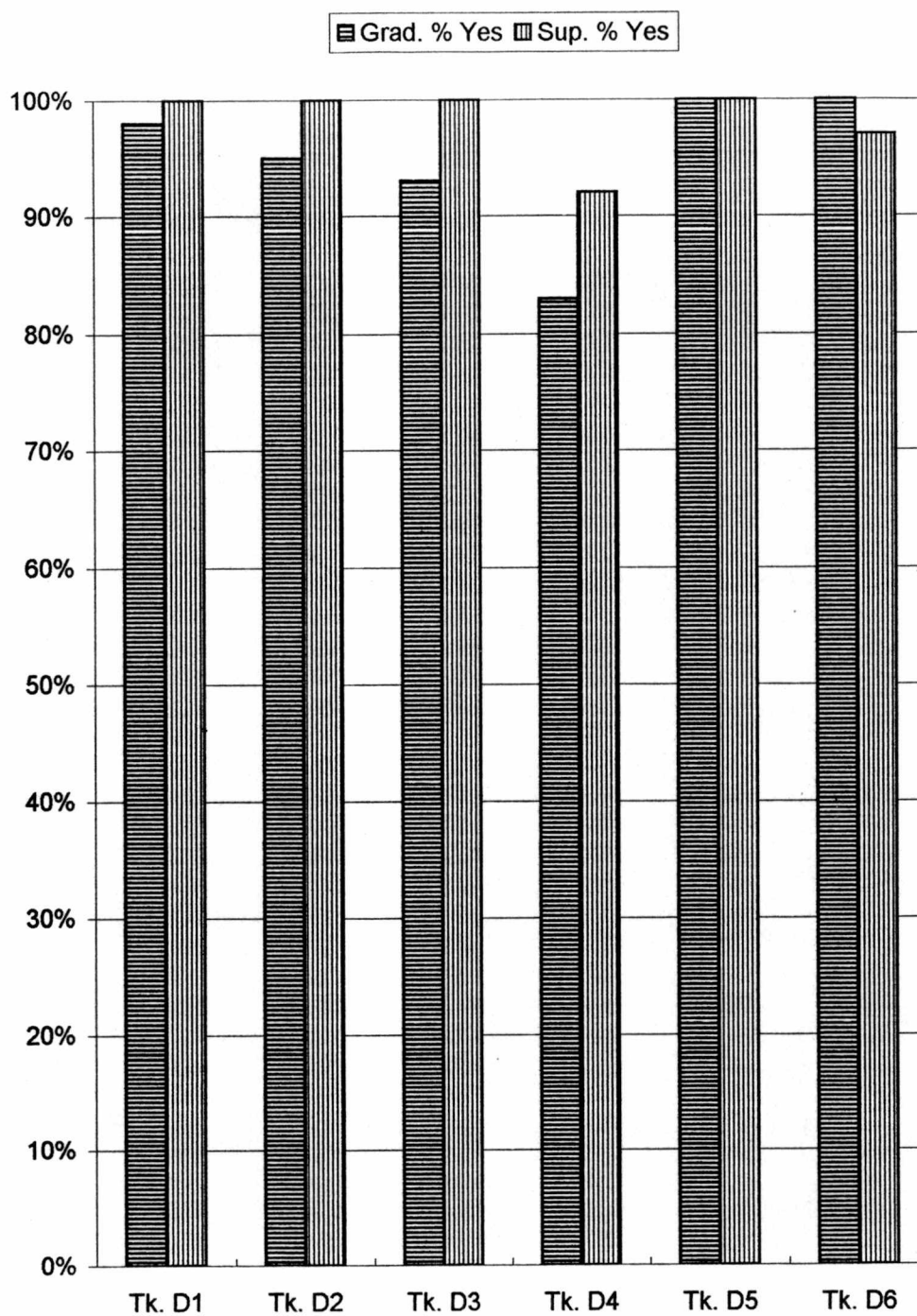


Figure 4. (continued)

Summary of Graduates'/Supervisors' Performance Evaluations: (d)
Job Group D (Elect. Instr. Maint.)

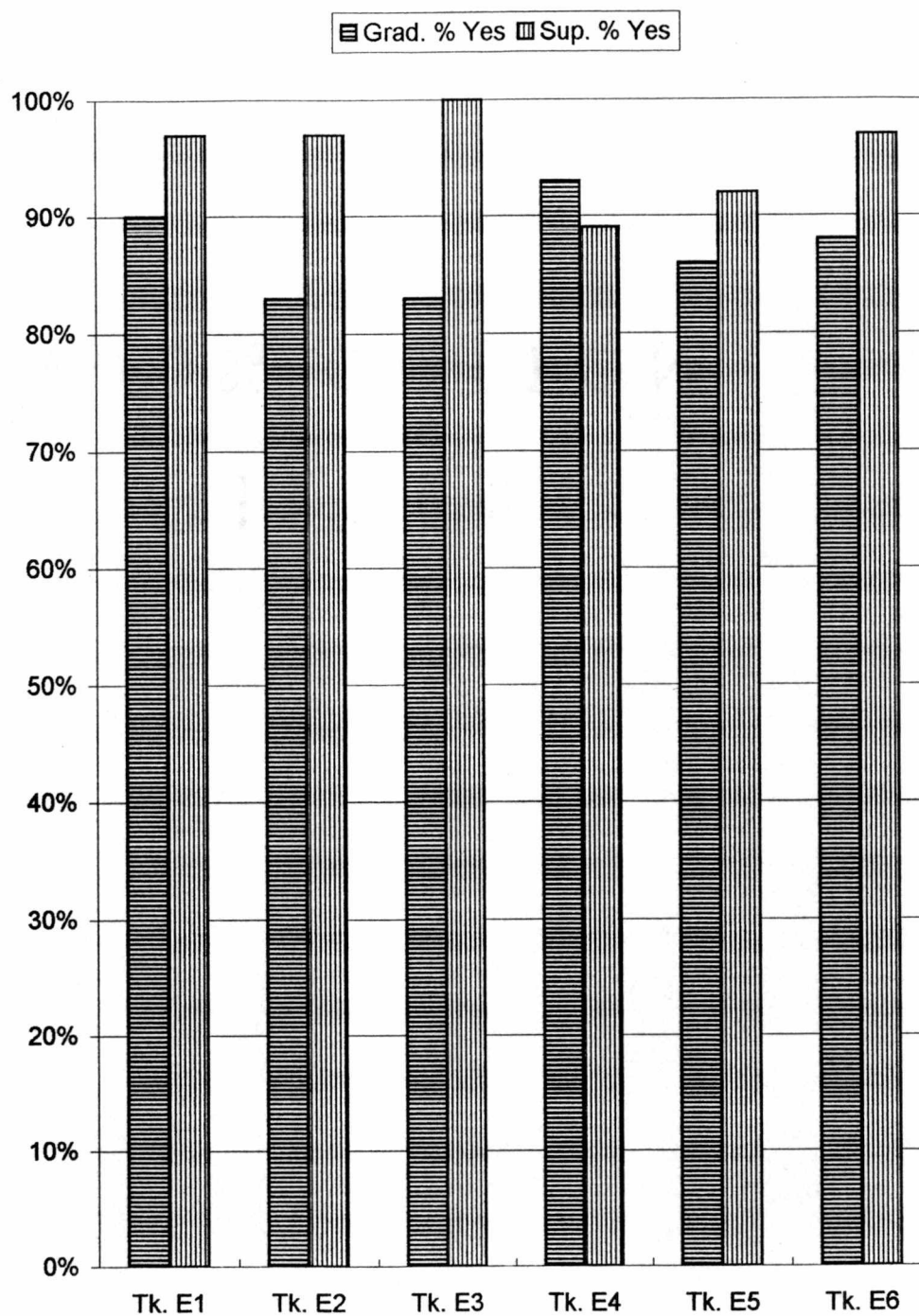


Figure 4. (continued)
Summary of Graduates'/Supervisors' Performance Evaluations: (e)
Job Group E (Pneu. Instr. Maint.)

CHAPTER V

SUMMARY, FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

As American business and industry prepare to meet challenges of the twenty-first century, units within organizations will come under ever-increasing scrutiny to demonstrate their value to the parent organization. Training and human resource development departments will be under increased pressure to direct their efforts toward meeting the goals, objectives, or desired outcomes of their parent organization. Professionals in the human resource development field will be expected to provide managers and other decision makers with data to assist in the decision-making process concerning human resource development programs.

This study was an attempt to establish a precedence in training program evaluation for the manufacturing facility. Data obtained and analyzed by the researcher have been shared with management, corporate trainers, and other decision makers to assist them in improving the CSM apprenticeship training program.

The results of the study are discussed in the following pages. This chapter provides a summary of the research study, the major findings and recommendations related to the research hypotheses, and a model to evaluate other technical training programs at the manufacturing facility. The final section focuses on the conclusions and recommendations for future studies.

Summary

The purposes of this study were to measure the effectiveness of a maintenance apprenticeship training program and provide feedback to training designers for program improvement. Products of the study included a model to evaluate other technical training programs and data for management to use in the decision-making process related to technical training programs.

The objectives of this study were to determine if job performance, on a broad basis, was positively impacted by a maintenance apprenticeship training program and to identify specific modifications and/or improvements which could be made to the apprenticeship training program.

The subjects were chosen from a population which consisted of employees of the manufacturing facility. The sample was drawn from a population which consisted of apprentice candidates, CSM apprentice graduates, and supervisors of CSM apprentice graduates. The entire population of apprentice candidates were included in the study. Due to the distribution of CSM graduates among supervisors of CSMs, 47 graduates were selected from the population to participate in the study. Three factors determined the selection criteria: (a) no more than six apprentice graduates were to be evaluated by one supervisor, (b) no less than two apprentice graduates were to be evaluated by one supervisor, and (c) the most recent apprentice graduates were to be selected to participate in the study. Supervisors of the 47 CSM apprentice graduates were also selected to participate in the study.

To test the two research hypotheses of this study and to meet the criteria of

both Kirkpatrick's (1959) level three evaluation and Phillips' (1991) pretest and post-test model design, a four-phase study was necessary to obtain and analyze the data. A representative sample of tasks which comprised the job of CSM was identified. Study participants were requested to evaluate their performance or their subordinates' performance on each of the tasks to defined criteria.

The initial phase of the study measured the perceived performance of 43 apprentice candidates. The second phase of the study measured the perceived performance of 42 CSM apprentice graduates. The third phase of the study measured supervisors' performance ratings of 42 CSM apprentice graduates who worked directly under them. During the final phase of the study, the data were analyzed and information was extracted through the use of statistical capabilities of Excel version 5.0.

The study tested the following research hypotheses:

Research Hypothesis One: There is a significant improvement in the job performance of apprentice graduates when compared to apprenticeship candidates (who have received no training on tasks related to the specific craft).

Research Hypothesis Two: Specific improvement and modification opportunities for maintenance apprenticeship training programs can be identified by comparing supervisors' performance evaluations of apprentice graduates and apprentice graduates' performance evaluations of themselves.

Major Findings

The major findings of this study were based on the null hypotheses and were reported using the data analysis described in Chapter IV.

1. Performance evaluations of CSM apprentice graduates for the tasks of Job Groups A, B, C, D, and E were significantly different when compared to performance evaluations of apprentice candidates. As one would expect, apprentice graduates' perceptions regarding their skill levels associated with the tasks of Job Groups A, B, C, D, and E were significantly higher than the perceptions of apprentice candidates. This was the case for each of the five job groups as well as for each of the 30 specific job tasks that comprised these groups.

2. Performance evaluations of CSM apprentice graduates by supervisors and by graduates were corresponding. Supervisors' perceptions of skill levels of apprentice graduates and apprentice graduates' perceptions of their own skill levels were congruent for 80% of the tasks.

3. Supervisors' perceptions of skill levels of apprentice graduates and apprentice graduates' perceptions of their own skill levels were significantly different for 20% of the tasks.

4. Supervisors' and apprentice graduates' perceptions of skill levels indicated that approximately 77% of the job tasks included in the survey could be performed to defined standards by apprentice graduates.

5. Major improvement and modification opportunities for the CSM apprenticeship program were identified by the study.

Conclusions

The researcher's conclusions were based on: (a) the data analysis described in Chapter IV, (b) follow-up conversations with general supervisors and supervisors of CSMs and CSM apprentice graduates, (c) the researcher's knowledge of the manufacturing facility's maintenance apprenticeship training programs, and (d) discussions with the supervisor of the training section responsible for the job-related education aspect of maintenance apprenticeships. According to general supervisors of CSMs, supervisors of the manufacturing facility are assuming additional management roles and responsibilities while delegating many of the technical job responsibilities to their crews due to changing organizational culture and management philosophy fostered by total quality management. Also, performance evaluations of employees, including maintenance apprentices, at the manufacturing facility have been eliminated due to total quality management efforts.

Many supervisors of CSMs indicated they assumed apprentices were provided opportunities to practice or apply training on the job through job assignments because their crew's work assignments were being completed. The supervisor who chose not to participate in the study indicated that he seldom had an opportunity to observe performance of individuals within his crew due to increasing management roles and responsibilities and relied on feedback from customers, crew meeting leaders, and data from measures established by the crew to determine the crew's performance.

Several CSM apprentice graduates stated the opportunity to apply training on

the job in many instances was not advantageously timed to the training they were receiving in the apprenticeship program. Job assignments were determined by the workload of the crew and often were not directly linked to the job related training they were receiving in the apprenticeship training program.

Maintenance apprenticeship programs have been utilized as a method of training employees at the manufacturing facility since 1936. The maintenance apprenticeship program has been registered with the federal government since 1946, first with the Federal Apprenticeship Committee and later with the Bureau of Apprenticeship Training. Approximately 1600 employees have completed maintenance apprenticeship training programs at the manufacturing facility since 1936.

Current maintenance apprenticeships at the manufacturing facility vary with content due to the subject matter being taught but the format among all maintenance apprenticeships is similar. The programs are divided into manageable progression parts (Part I, II, and III) and each part is divided into individual courses. The apprenticeship training programs are self-paced with limits placed on minimum and maximum time allowed per part. Apprenticeship programs consist of job related education. Job related education (JRE) consists of related studies with written tests and hands-on labs with performance tests. A team consisting of representatives from the manufacturing facility's Human Resources Department, Training Department, and the President's Manufacturing Team has recommended that pass/fail scores for all testing be set at 80%.

The philosophy of the Apprenticeship Committee concerning on-the-job training for maintenance apprentices is that apprentices will be provided opportunities to apply skills and knowledge obtained during JRE through job assignments monitored by supervisors. Currently, a review process exists which monitors job performance and training progress but does not monitor task performance. While the current system is satisfactory for monitoring job performance and training progress, there is room for improvement in monitoring task performance and on-the-job training.

Discussions with the Manufacturing Training Supervisor, who is responsible for the JRE training aspect of maintenance apprenticeships at the manufacturing facility, concerning the data presented in Table 2 and Figure 4 assisted the researcher in establishing criteria for identifying modification and improvement opportunities for the CSM apprenticeship program. Criteria for identifying successful program graduates are based on pass/fail scores of 80% for all testing in maintenance apprenticeship programs. It was the opinion of the Manufacturing Training Supervisor that maintenance apprenticeship programs at the manufacturing facility should be held to the same criteria for evaluation as apprentices.

An indication of a successful program would be that 100% of a program's graduates could perform 80% of the significant tasks to defined standards. Another indication of a successful program would be agreement between supervisors and graduates on perceived performance levels of program graduates. Criteria for modification and/or improvement opportunities for the CSM apprenticeship program

from the data presented in Table 2 and Figure 4 consisted of identifying tasks with a "Yes" rating of less than 80% by either supervisors or program graduates and identifying tasks with significant differences in perceived level of performance between supervisors and apprentice graduates. It should also be noted that the task rating method utilized for this study was a combination of quality, safety, and timeliness (See Appendix C) and no one single factor was differentiated in the analysis.

The study provided evidence which suggested the training program contributed to major improvement in overall job performance of apprenticeship program graduates when compared to apprenticeship program candidates.

The study also provided evidence which suggested the following modification and improvement opportunities for the CSM apprenticeship training program:

1. As supported by data analysis for task A1, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to install, remove, or replace distributive control system hardware components.
2. As supported by data analysis for task A2, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to troubleshoot and/or repair distributive control system hardware components.
3. As supported by data analysis for task A3, the training program did not adequately address the skills, knowledge, and application of training needed by

program graduates to install, remove, or replace programmable logic controllers.

4. As supported by data analysis for task A4, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to troubleshoot and/or repair programmable logic controllers at a module level.

5. As supported by data analysis for task A5, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to modify existing ladder logic programs in programmable logic controllers.

6. As supported by data analysis for task C5, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to maintain intercom systems.

7. As supported by data analysis for task C6, the training program did not adequately address the skills, knowledge, and application of training needed by program graduates to maintain variable speed drives.

8. As supported by data analysis for tasks A2, A3, C5, C6, E2, and E3, there were significant differences in the opinions of supervisors and apprentice graduates as to the ability of graduates to perform the job tasks.

Recommendations

The researcher recommends that research be conducted to determine if differences in perceptions of supervisors and CSM apprentice graduates are, in fact,

a gap in skill levels or if the gap is truly one of perception. If research determines that a true gap in skill levels exist, the researcher recommends the Apprenticeship Committee of the manufacturing facility consider changes and enhancements to the CSM apprenticeship program to improve job performance of CSM apprentice graduates. These recommendations are as follows:

1. Revision of job related education currently being delivered by the Training Department to adequately address the skills and knowledge needed by program graduates for each of the specific job tasks where skill levels have been identified as being deficient.
2. Implementation of job specific training (JST) which is focused on the tasks that must be learned in order for an apprentice to be successful on the job. Typical JST would include documented on-the-job training, opportunity for practice, and performance tests to assess trainees' ability to perform the critical and essential tasks actually associated with the job in the work place. To reduce significant differences in perceptions of performance between supervisors and apprentice graduates, implementation of JST should directly involve apprentice supervision and management with training, opportunities for practice, and performance testing of apprentices.
3. Implementation of JST trainers for CSM apprentices (by crew, maintenance department, or manufacturing department) as deemed appropriate by the Apprenticeship Committee. Job duties of JST trainers would include analysis, design, development, implementation, and evaluation of crew specific training with

input and coaching from the manufacturing facility's Training Department and the appropriate level of apprentice supervision.

4. Linkage of JRE and JST throughout the CSM apprenticeship program.

Providing trainees opportunities to apply newly acquired skills and knowledge is a basic concept of adult learning theory. Current training practices in the JRE portion of the CSM apprenticeship advance from training objectives based upon relatively simple tasks to training objectives based upon more complex tasks. Job specific training should be delivered in conjunction with the objectives of JRE as apprentices progress through the program (See Figure 5).

5. Kirkpatrick's (1959) Level I and Level II evaluations are conducted on apprenticeship programs at the manufacturing facility by the Training Department. The Training Department of the manufacturing facility should evaluate its process for design and development of training programs to also incorporate Level III evaluations. As proposed by Nadler and Nadler (1989) and Phillips (1991), training evaluation of any type or level should be considered during the design and development phases of training courses and programs.

The final two recommendations were directed to management of the Training Department of the manufacturing facility. The Training Department should implement Level III evaluation for all apprenticeship training programs at the manufacturing facility. This study has identified major improvement opportunities in the CSM apprenticeship training program. The manufacturing facility has five other apprenticeship programs where possible improvement opportunities may exist.

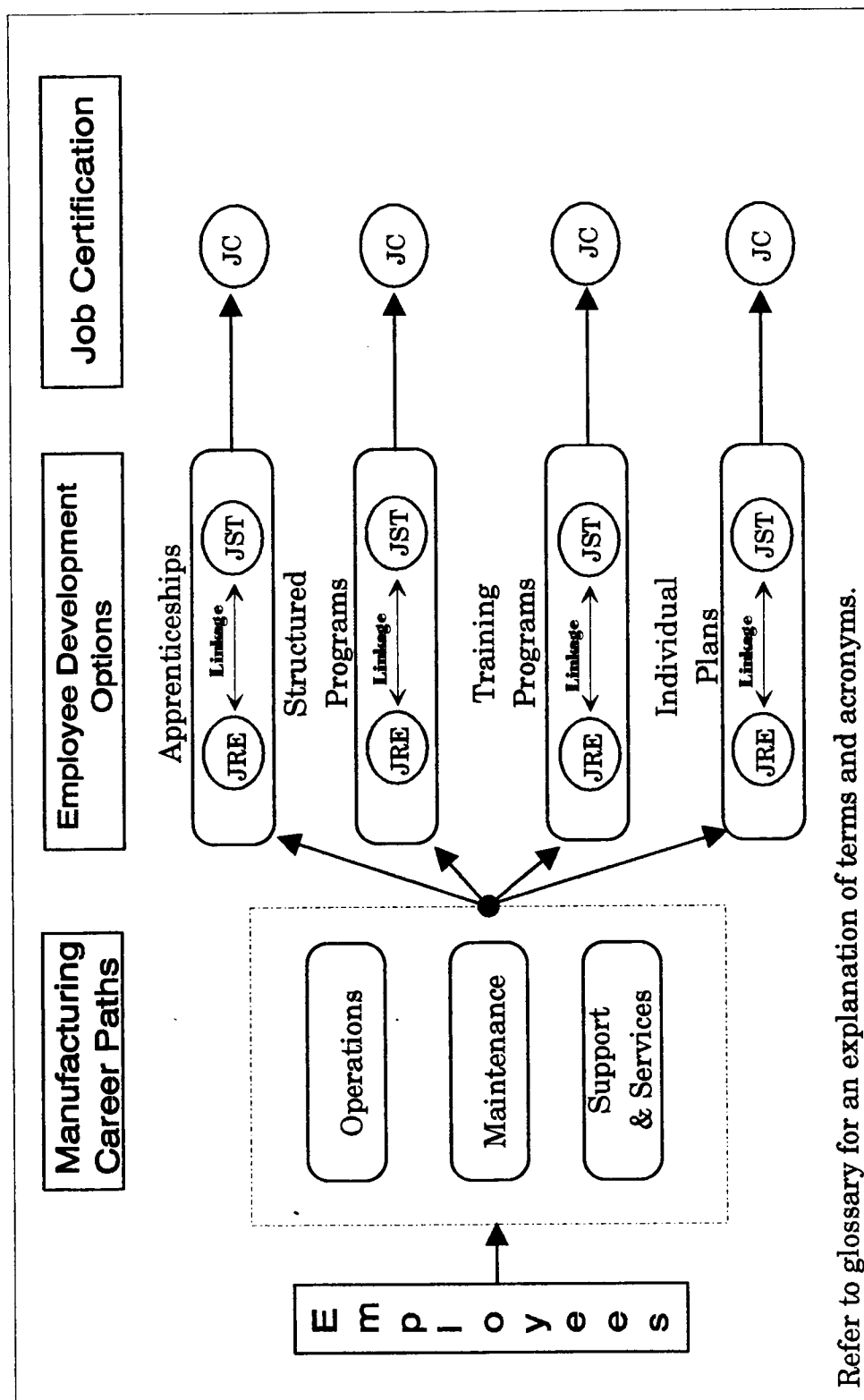


Figure 5. Workforce Development Model

Future Research

Justification of resources committed to any project and the total quality management concept of continuous improvement are major issues for American business and industry. Organizational management wants to know if the resources committed to training and human resource development are contributing to organizational goals. Training organizations must provide factual evidence to support resource allocations and they must also conduct efforts to improve their products and services.

In order to provide organizational decision makers with information which is useful and accurate concerning training, several suggestions for future research follow from the results of this study.

1. Conduct research to solicit information comparing current job analysis data to results of the CSM apprenticeship evaluation to determine why significant differences in perceived performance appear between supervisors of CSM apprentice graduates and apprentice graduates.
2. Upon implementation of the above CSM program recommendations, conduct a follow-up evaluation to determine the effectiveness of program changes.
3. Implement and conduct Level I (Kirkpatrick, 1959) evaluations for every course delivered by the Training Department. The evaluation results should be provided to Training Department personnel to use as a formative evaluation tool to improve the course.
4. Implement and conduct Level II (Kirkpatrick, 1959) evaluations for every

course delivered by the Training Department. The evaluation results should be provided to Training Department personnel and organizational decision makers (when appropriate) to use as formative and summative evaluation tools to improve the course and also as factual evidence for decision makers to use for resource allocation. Maintain databases to identify trends over long periods of time or if large numbers of employees receive the training. Such data may also be used in cases of litigation.

5. Implement and conduct Level III (Kirkpatrick, 1959) evaluations on courses or programs in which considerable organizational resources are devoted. The evaluation results should be provided to Training Department personnel and organizational decision makers to use as summative evidence that the course or program is or is not meeting program objectives. Also use the data to identify major improvement or modification opportunities within the course or program. See Figure 6 for the researcher's recommended model for Level III evaluation of apprenticeship training programs at the manufacturing facility.

6. Level IV (Kirkpatrick, 1959) evaluations should be conducted on selected courses or programs in which considerable organizational resources are devoted and as training personnel become adept in conducting training evaluations. The evaluation results should be made available to organizational decision makers and should be used to provide summative evidence to determine if the course or program is meeting organizational goals and providing benefits to the organization.

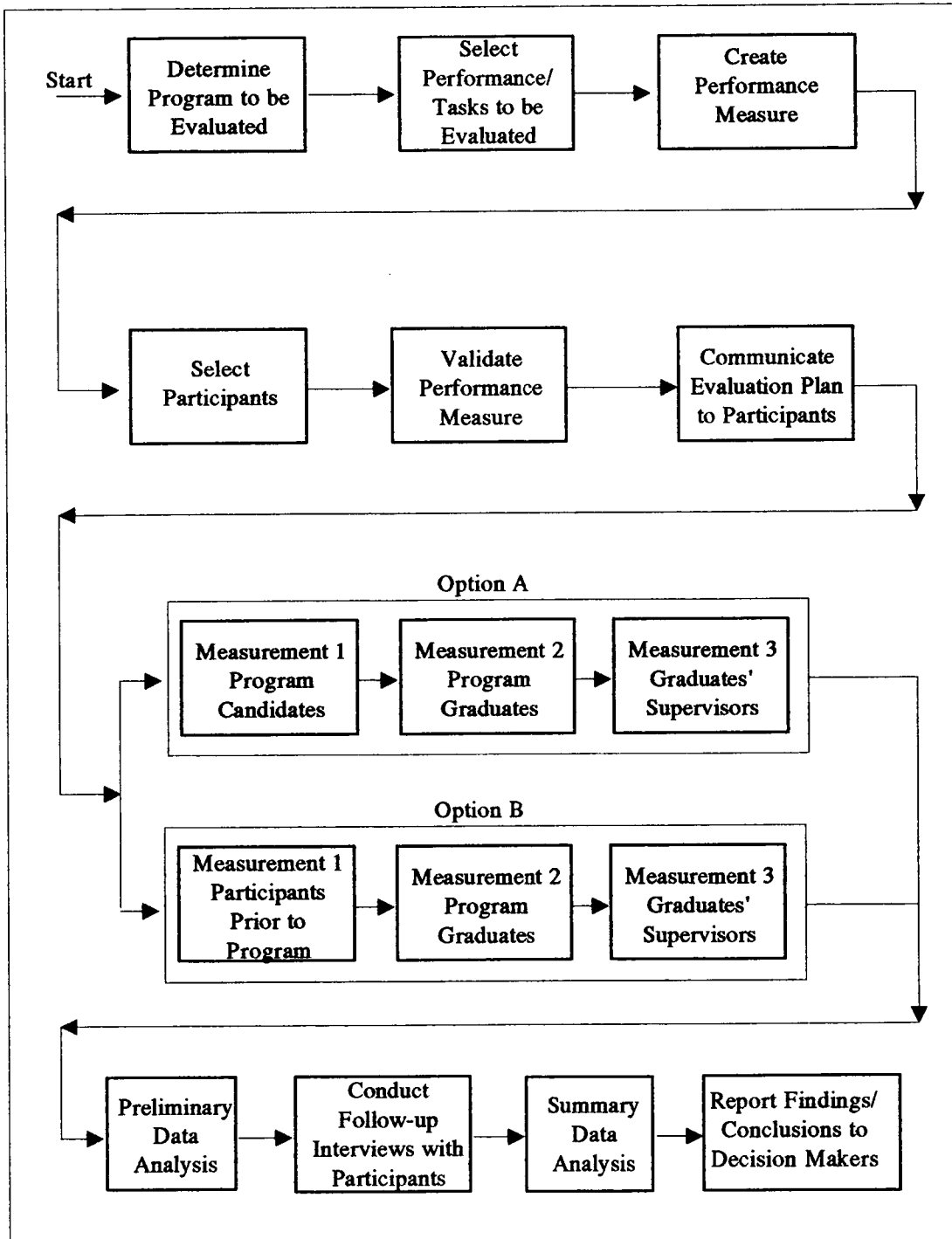


Figure 6. Level III Evaluation Model for Existing Programs

7. Level V (Phillips, 1991) evaluations should be conducted on selected courses or programs in which considerable organizational resources are devoted and as training personnel become adept in Level IV evaluations. The evaluation results should be made available to organizational decision makers and should be used as summative evidence to determine if the course or program is providing return on investment.

REFERENCES

References

- Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. Personnel Psychology, 42, 331-342.
- Bloom, B. S., Madus, G. F., & Hastings, J. T. (1981). Evaluation to improve learning. New York: McGraw-Hill.
- Campbell, C. P. (1984). A digest on instructional systems development. (Available from The University of Tennessee [Knoxville] Library, Knoxville, TN)
- Carnevale, A. P., & Schulz, E. R. (1990, July). Evaluation practices. Training and Development Journal, 44 (Suppl. 1), 23-29.
- Gilley, J. W., & Egglund, S. A. (1989). Principles of human resource development. Reading, MA: Addison-Wesley.
- Goldstein, I. L. (1993). Training in organizations. Pacific Grove, CA: Brooks/Cole.
- Gordon, J. (1987, June). Romancing the bottom line. Training: The Human Side of Business, 25 (6), 31-42.
- Hinkle, D. E., Wiersma W., & Jurs, S. G. (1988). Applied statistics for the behavioral sciences. Boston, MA: Houghton Mifflin.
- Kirkpatrick, D. L. (1959). Techniques for evaluating training programs [Monograph]. Journal of the American Society for Training and Development, 1975, 1-17.
- Kirkpatrick, D. L. (1994). Evaluating training programs. San Francisco, CA: Berkett/Koehler.

Lee, C. (1990, October). Industry report 1990. Training: The Human Side of Business, 27 (10), 29-32.

Lerda, L. W., & Cross, L. W. (1962, August). Performance oriented training- results measurement and follow-up [Monograph]. Journal of the American Society for Training and Development, 1975, 210-215.

Long, L. N. (1990). A test of Kirkpatrick's training evaluation levels in terms of utility for decisionmaking. Unpublished doctoral dissertation, The University of Tennessee, Knoxville.

Mager, R. F. (1988). Making instruction work. Belmont, CA: Lake.

Mitchell, K. D. (1994, Winter). Putting evaluation to work for human resources development. Public Productivity & Management Review, 18 (2), 199-215.

Moore, E. R. (1984, November). Competency-based training evaluation. Training and Development Journal, 38 (11), 92-94.

Nadler, L. & Nadler, Z. (1989). Developing human resources. San Francisco, CA: Jossey-Bass.

Phillips, J. J. (1991). Handbook of training and evaluation and measurement methods. London, England: Kogan Page.

Pine, J., & Tingley, J. C. (1993, February). ROI of soft-skills training. Training: The Human Side of Business, 30 (8), 55-60.

Plant, R. A., & Ryan, R. J. (1994). Who is evaluating training? and Kirkpatrick's mirror. Journal of European Industrial Training, 18 (5), 27-32.

Abstract from: DIALOG File: ABI/INFORM(R) Item: 15813.00

Rothwell, W. J., & Sredl, H. J. (1992). The ASTD reference guide to professional human resource development (2nd ed., Vol. 1). Amherst, MA: HRD Press.

Saari, L. M., Johnson, T. R., McLaughlin, S. D., & Zimmerle, D.M. (1988). A survey of management training and education practices in U.S. companies. Personnel Psychology, 41, 731-743.

Study compares training at small and large firms (1994, January/February). National report on human resources. American Society for Training and Development. Alexandria, VA.

Tannenbaum, S. I., & Woods, S. B. (1992). Determining a strategy for evaluating training: Operating within organizational constraints. Human Resource Planning, 15 (2), 63-81.

Waste Isolation Division, Westinghouse Electric Corporation. (1995). Transfer of training evaluation model (Document No. WZ: 95: 02321). Carlsbad, NM: Author.

APPENDICES

Appendix A
CSM Apprenticeship Survey

CSM Apprenticeship Survey

Task #	Task Description	Performance			
		1	2	3	4
A1	Using proper test equipment, tools, and prints or drawings, install, remove, or replace DCS hardware components to manufacturer's specifications and TED Division 16 Specifications/Details.	1	2	3	4
A2	Using proper test equipment, tools, and prints or drawings, troubleshoot and/or repair DCS hardware components at a board level to manufacturer's specifications and TED Division 16 Specifications/Details.	1	2	3	4
A3	Using proper test equipment, tools, and prints or drawings, install, remove, or replace PLCs to manufacturer's specifications.	1	2	3	4
A4	Using proper test equipment, tools, and prints or drawings, troubleshoot and/or repair PLCs at a module level to manufacturer's specifications.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

A5	Using proper test equipment, tools, and prints or drawings, modify existing ladder logic programs in PLCs to manufacturer's specifications.	1	2	3	4
B1	De-Energize/Re-Energize circuit breakers in Power Panels and MCCs and air circuit breakers in substations applying TED Lock-Out/Tag-Out procedure.	1	2	3	4
B2	Using proper test equipment, tools, and prints or drawings, install, remove, and/or replace electrical cable and/or wiring rated at or below 600 volts to TED Division 16 Specifications/Details.	1	2	3	4
B3	Using proper tools and prints or drawings, install, remove, replace, and/or repair conduit to TED Division 16 Specifications/Details and/or NEC specifications.	1	2	3	4
B4	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, install, remove, or replace incandescent and fluorescent lighting to TED Division 16 Specifications/Details.	1	2	3	4
B5	Using proper test equipment, tools, and prints or drawings, troubleshoot electrical motors (AC and DC) according to TED Division 16 Specifications/Details.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

B6	Using proper test equipment, tools, and prints or drawings, install, remove, or replace process switches (pressure, limit, temperature, etc.) to TED Division 16 Specifications/Details and manufacturer's specifications.	1	2	3	4
B7	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate process switches (pressure, limit, temperature, etc.) to TED Division 16 Specifications/Details and manufacturer's specifications.	1	2	3	4
C1	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate analog electronic instruments to manufacturer's specifications, TED Division 16 Specifications/Details, and ISA standards.	1	2	3	4
C2	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, install, remove, or replace photoelectric devices to TED Division 16 Specifications/Details and manufacturer's specifications.	1	2	3	4
C3	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, install, remove, or replace electric/electronic relays to TED Division 16 Specifications/Details and manufacturer's specifications.	1	2	3	4
C4	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, install, remove, or replace electric/electronic counters and/or timers to manufacturer's specifications.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

C5	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, install, remove, or replace intercom systems to manufacturer's specifications and TED Division 16 Specifications/Details.	1	2	3	4
C6	Using proper test equipment, tools, and prints or drawings, troubleshoot and/or repair variable speed drives to TED Division 16 Specifications/Details and manufacturer's specifications.	1	2	3	4
D1	Using proper test equipment, tools, and prints or drawings, install, remove, replace electronic transmitters (analog & digital) to manufacturer's specifications, TED Division 16 Specifications/Details, and ISA standards.	1	2	3	4
D2	Using proper test equipment, tools, and prints or drawings, install, remove, replace electronic controllers (analog & digital) to manufacturer's specifications, TED Division 16 Specifications/Details, and ISA standards.	1	2	3	4
D3	Using proper test equipment, tools, and prints or drawings, install, remove, replace electronic recorders (analog & digital) to manufacturer's specifications, TED Division 16 Specifications/Details, and ISA standards.	1	2	3	4
D4	Using proper test equipment, tools, and prints or drawings, troubleshoot, install, remove, or replace signal converters to TED PID Design Handbook guidelines, TED Division 16 Specifications/Details, and ISA Standards.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

D5	Using proper test equipment, tools, and prints or drawings, troubleshoot, install, remove, or replace transducers to TED PID Design Handbook guidelines, TED Division 16 Specifications/Details, and ISA Standards.	1	2	3	4
D6	Using proper test equipment, tools, and prints or drawings, troubleshoot, install, remove, or replace temperature measuring devices (thermocouples and RTDs) to TED PID Design Handbook guidelines and ISA Standards.	1	2	3	4
E1	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate pneumatic transmitters to manufacturer's specifications and ISA standards.	1	2	3	4
E2	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate pneumatic controllers to manufacturer's specifications and ISA standards.	1	2	3	4
E3	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate pneumatic recorders to manufacturer's specifications and ISA standards.	1	2	3	4
E4	Using proper test equipment, tools, and prints or drawings, troubleshoot, and/or repair control valves to manufacturer's specifications and ISA Standards.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

E5	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate control valve actuators to manufacturer's specifications and ISA standards.	1	2	3	4
E6	Using proper test equipment, tools, and prints or drawings, troubleshoot, repair, and/or calibrate pressure gauges and/or indicators to manufacturer's specifications and ISA standards.	1	2	3	4

RATINGS: Please mark or circle the response which best describes your ability to complete the task. Please indicate only one response per task.

- 1 - Cannot perform the task.
- 2 - Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
- 3 - Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
- 4 - Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).

Appendix B

Electronic Note to CSM Apprenticeship Graduates

To:

From: J. David Collier, Bldg. 226, Room 226
Manufacturing Training Products and Services Team

Subject: CSM APPRENTICESHIP SURVEY

The Training Department is preparing to conduct an evaluation of the effectiveness of maintenance apprenticeships on the job performance of apprentice graduates on tasks related to their specific craft. The survey information will enable us to identify opportunities for improving the program's effectiveness and value. A survey packet has been mailed to you separately. Your input is critical to the success of this study.

This is often referred to as Kirkpatrick's level three evaluation of a training program. I have been assigned the task of preparing and conducting the pilot study. The pilot study will be conducted for the CSM Apprenticeship and will use surveys to gather that data.

The survey consists of a performance evaluation of a representative sample of tasks (30) identified by the job analysis for the CSM craft. The tasks are broken down into five job areas. The job areas are:

A = DCS & PLC maintenance	B = Electrical maintenance
C = Electronic maintenance	D = Electronic Instrumentation maint.
E = Pneumatic Instrumentation maintenance	

You are being requested to rate your perceived performance on each of the listed tasks according to the provided standards. Your individual responses will remain strictly confidential. Once the results of the surveys are tallied, all individual surveys will be destroyed.

An example of how the data will be used is included in the hardcopy packet which will be mailed to you. This survey is to be used ONLY to evaluate the training program and there is NO INTENTION of sharing individual results with anyone. The results of the study will be shared with anyone interested in them.

I am requesting that you spend about 15-20 minutes of your time to carefully read each task and rate your ability to perform each task to the identified standards. Please return the completed survey to me in the enclosed envelope.

Thanking you in advance for your assistance in this study.
Dave Collier

Appendix C

Information Packet to CSM Apprenticeship Graduates

CSM Apprenticeship Graduate Survey

**Survey
Information**

The purpose of this survey is to evaluate the CSM Apprenticeship Program by determining the performance of recent CSM Apprenticeship graduates on tasks related to the CSM craft compared to apprenticeship candidates not trained on tasks related to the CSM craft.

This evaluation will have no bearing on your job. All individual responses will remain strictly confidential. After the responses are tabulated, the individual surveys will be destroyed. The purpose of this study is to evaluate the effectiveness of the CSM Apprenticeship. The study is not being conducted to evaluate individuals.

Recent CSM Apprenticeship graduates, CSM supervisors, and apprenticeship candidates are being requested to complete the following survey in an effort to evaluate the CSM Apprenticeship.

**Survey
Instructions**

Instructions on how to complete the survey are on the following pages.

**Survey
Completion**

Upon completing the survey, please return it to Dave Collier, Bld. 226, Training. An envelope with my name and address which has been stamped "Open by Addressee Only" has been enclosed .

Your timely response to this survey will be greatly appreciated!

I thank you in advance,
Dave Collier

CSM Survey

Survey Instructions

Research Information:

On the first page of the survey, please list the following information:

- Name
- Insurance number

Answering the Survey:

On the survey, circle the appropriate response which best describes your ability to complete each indicated task. Please rate each task in terms of your ability to complete the task in the areas of quality, safety, and timeliness.

Quality, Safety, and Timeliness are defined as follows:

- Quality:* Quality is defined as the standard to which the task must be completed. It may be to TED Standards, others (such as ISA, NEC, manufacturer's specifications, etc.) or to combinations of the above.
- Safety:* Safety standards refer to the ability to perform a task without endangering yourself, co-workers, equipment, facilities, and the environment.
- Timeliness:* Timeliness is based upon the ability to complete a task in an amount of time which an identified subject matter expert in your work area would take to complete the task.

Continued on next page

CSM Survey, Continued

**Performance
Rating**

The ratings are as follows (please mark each task with only one rating):

- 1 Cannot perform the task.
 - 2 Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
 - 3 Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
 - 4 Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).
-

Appendix D

Electronic Note to Supervisors of CSM Apprenticeship Graduates

To: Supervisors of CSMs

From: J. David Collier, Bldg. 226, Room 226
Manufacturing Training Products and Services Team

Subject: CSM APPRENTICESHIP SURVEY

I need your help!

The Training Department is preparing to conduct an evaluation of the effectiveness of maintenance apprenticeships on the job performance of apprentice graduates on tasks related to their specific craft. The evaluation information will enable us to identify opportunities for improving the program's effectiveness and value.

This is often referred to as Kirkpatrick's level three evaluation of a training program. I have been assigned the task of preparing and conducting the pilot study. The pilot study will be conducted for the CSM Apprenticeship and will use surveys to gather the data.

A hardcopy survey packet has been mailed to you. The packet consists of surveys, instructions for filling out the surveys, examples of how the data will be used, and an envelope to mail the completed surveys back to me. Your input is critical to the success of the study!

The survey has been validated by a panel of one maintenance department head, three E&I general supervisors, and the APR for CM&S Division.

The survey consists of a performance evaluation of a representative sample of tasks (30) identified by the job analysis for the CSM craft. The tasks are broken down into five job areas. The job areas are:

A = DCS & PLC maintenance	B = Electrical maintenance
C = Electronic maintenance	D = Electronic Instrumentation maint.
E = Pneumatic Instrumentation maintenance	

You are being requested to rate selected apprentice graduates in your area on each of the tasks. Using the survey provided, please rate each apprenticeship graduate on his/her performance on each of the listed tasks according to the provided standards.

Your individual responses will remain strictly confidential. Once the results of the surveys are tallied, all individual surveys will be destroyed.

An example of how the data will be used is included in the hardcopy packet which has been mailed to you. This survey is to be used ONLY to evaluate the training program and there is NO INTENTION of sharing individual results with anyone. The results of the study will be shared with anyone interested in them.

I am requesting that you spend about 15-20 minutes of your time to carefully read the survey and rate each apprenticeship graduate's ability to perform the tasks to the identified standards. If at all possible, please try to have the completed surveys back to me by July 15, 1996.

Thanking you in advance for your assistance in this study,
Dave Collier

Appendix E

Information Packet to Supervisors of CSM Apprenticeship Graduates

CSM Supervisor Survey

**Survey
Information**

The purpose of this survey is to evaluate the CSM Apprenticeship Program by determining the performance of recent CSM Apprenticeship graduates on tasks related to the CSM craft compared to apprenticeship candidates not trained on tasks related to the CSM craft.

This evaluation will have no bearing on anyone's job. All individual responses will remain strictly confidential. After the responses are tabulated, the individual surveys will be destroyed. The purpose of this study is to evaluate the effectiveness of the CSM Apprenticeship. The study is not being conducted to evaluate individuals.

Please rate _____ performance on the following tasks.

Recent CSM Apprenticeship graduates, CSM supervisors, and apprenticeship candidates are being requested to complete the following survey in an effort to evaluate the CSM Apprenticeship.

**Survey
Instructions**

Instructions on how to complete the survey are on the following pages.

**Survey
Completion**

Upon completing the survey, please return the survey to Dave Collier, Bld. 226, Training. An envelope with my name and address which has been stamped "Open by Addressee Only" has been enclosed .

Your timely response to this survey will be greatly appreciated!

I thank you in advance,
Dave Collier

CSM Survey

Survey Instructions

Research Information:

On the first page of the survey, please list the following information:

- Name
- Insurance number

Answering the Survey:

On the survey, circle the appropriate response which best describes your ability to complete each indicated task. Please rate each task in terms of your ability to complete the task in the areas of quality, safety, and timeliness.

Quality, Safety, and Timeliness are defined as follows:

- Quality:* Quality is defined as the standard to which the task must be completed. It may be to TED Standards, others (such as ISA, NEC, manufacturer's specifications, etc.) or to combinations of the above.
- Safety:* Safety standards refer to the ability to perform a task without endangering yourself, co-workers, equipment, facilities, and the environment.
- Timeliness:* Timeliness is based upon the ability to complete a task in an amount of time which an identified subject matter expert in your work area would take to complete the task.

Continued on next page

CSM Survey, Continued

**Performance
Rating**

The ratings are as follows (please mark each task with only one rating):

- 1 Cannot perform the task.
 - 2 Can perform the task but need improvements in one or more areas (quality, safety, and timeliness).
 - 3 Can perform the task and meet expectations in all areas (quality, safety, and timeliness).
 - 4 Can perform the task and exceed expectations in all areas (quality, safety, and timeliness).
-

Appendix F

Apprenticeship Candidates' Performance Evaluations

Apprenticeship Candidates' Performance Evaluations

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
	A1	37	5	1	0	43
	A2	37	5	1	0	43
	A3	29	12	1	1	43
	A4	32	10	1	0	43
	A5	29	14	0	0	43
Frequency		164	26	4	1	215
Percent		76	21	2	1	100*
	B1	27	9	7	0	43
	B2	24	13	6	0	43
	B3	24	14	4	1	43
	B4	12	20	9	2	43
	B5	27	14	2	0	43
	B6	24	14	4	1	43
	B7	29	9	4	1	43
Frequency		167	93	36	5	301
Percent		55	31	12	2	100*
	C1	38	5	0	0	43
	C2	39	3	1	0	43
	C3	31	9	3	0	43
	C4	33	8	2	0	43
	C5	28	12	2	1	43
	C6	36	7	0	0	43
Frequency		205	44	8	1	258
Percent		79	17	3	1	100*

Apprenticeship Candidates' Performance Evaluations (continued)

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
	D1	35	8	0	0	43
	D2	32	11	0	0	43
	D3	32	11	0	0	43
	D4	34	9	0	0	43
	D5	33	10	0	0	43
	D6	32	9	2	0	43
Frequency		198	58	2	0	258
Percent		77	22	1	0	100*
	E1	37	6	0	0	43
	E2	35	7	1	0	43
	E3	39	4	0	0	43
	E4	31	11	1	0	43
	E5	35	6	2	0	43
	E6	28	13	1	1	43
Frequency		205	47	5	1	258
Percent		80	18	1	1	100*

NOTES:

* Actual value may be slightly more or less due to rounding.

Appendix G

Apprenticeship Graduates' Performance Evaluations

Apprenticeship Graduates' Performance Evaluations

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
A1		0	15	26	1	42
A2		2	20	19	1	42
A3		1	10	28	3	42
A4		0	12	29	1	42
A5		1	18	20	3	42
Frequency		4	75	122	9	210
Percent		2	36	58	4	100*
B1		0	4	25	13	42
B2		0	0	33	9	42
B3		0	5	29	8	42
B4		0	1	27	14	42
B5		0	3	28	11	42
B6		0	0	29	13	42
B7		0	2	30	10	42
Frequency		0	15	201	78	294
Percent		0	5	68	27	100*
C1		0	3	29	10	42
C2		0	3	29	10	42
C3		1	6	30	5	42
C4		0	8	33	1	42
C5		1	22	19	0	42
C6		0	13	24	5	42
Frequency		2	55	164	31	252
Percent		1	22	65	12	100*

Apprenticeship Graduates' Performance Evaluations (continued)

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
	D1	0	1	34	7	42
	D2	0	2	36	4	42
	D3	0	3	35	4	42
	D4	0	7	31	4	42
	D5	0	0	33	9	42
	D6	0	0	33	9	42
Frequency		0	13	202	37	252
Percent		0	5	80	15	100*
	E1	0	4	34	4	42
	E2	0	7	32	3	42
	E3	0	7	32	3	42
	E4	0	3	36	3	42
	E5	0	6	32	4	42
	E6	0	5	33	4	42
Frequency		0	32	199	21	252
Percent		0	13	79	8	100*

NOTES:

* Actual value may be slightly more or less due to rounding.

Appendix H

Supervisors' Performance Evaluations of Apprenticeship Graduates

Supervisors' Performance Evaluations of Apprenticeship Graduates

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
A1		0	7	24	6	37
A2		0	10	22	5	37
A3		0	3	28	6	37
A4		0	7	25	5	37
A5**		3	12	14	3	32
Frequency		3	39	113	25	180
Percent		1	22	63	14	100*
B1		0	5	15	17	37
B2		0	2	22	13	37
B3		0	3	21	13	37
B4		0	0	20	17	37
B5		0	3	21	13	37
B6		0	1	23	13	37
B7		0	2	24	11	37
Frequency		0	16	146	97	259
Percent		0	6	56	38	100*
C1		0	0	26	11	37
C2		0	0	26	11	37
C3		0	2	24	11	37
C4		0	5	22	10	37
C5		1	7	20	9	37
C6		0	3	27	7	37
Frequency		1	17	145	59	222
Percent		< 1	8	65	27	100*

Supervisors' Performance Evaluations of Apprenticeship Graduates (continued)

Frequency/ Percent	Tasks	Ratings				Totals
		#1	#2	#3	#4	Totals
D1		0	0	22	15	37
D2		0	0	23	14	37
D3**		0	0	23	13	36
D4		0	3	22	12	37
D5		0	0	26	11	37
D6		0	1	21	15	37
Frequency		0	4	137	80	221
Percent		0	2	62	36	100*
E1**		0	1	28	7	36
E2**		0	1	27	7	35
E3**		0	0	29	6	35
E4		0	4	27	6	37
E5		0	3	27	7	37
E6		0	1	31	5	37
Frequency		0	10	169	38	217
Percent		0	4	78	18	100*

NOTES:

* Actual value may be slightly more or less due to rounding.

** Denotes tasks that one or more supervisors indicated were not performed by employees in their work areas.

VITA

J. David Collier was born in Kingsport, Tennessee on August 2, 1949. He attended public schools in Sullivan and Hawkins counties and was graduated from Surgoinsville High School in 1967. After graduation from high school, David attended the University of Tennessee at Knoxville for one quarter in 1967. He then joined the United States Marine Corps, serving in southeast Asia from June, 1969 until April, 1970.

Upon successful completion of his military duties, David accepted an entry level position with Tennessee Eastman Company. He completed an apprenticeship as an electrician in 1977. In 1988, David began attending night classes at East Tennessee State University and Northeast State Technical and Community College. In 1989, David was assigned to the Training Department of Tennessee Eastman on a one-year temporary assignment. In 1990, he was assigned as a permanent member of the training staff. In 1992, David transferred to the University of Tennessee and graduated in June, 1994 with a Bachelor of Science degree in Education and entered Graduate school at the University of Tennessee in August, 1994. David received the Master of Science degree in Human Resource Development in December, 1996.

He is presently employed at Tennessee Eastman Division of Eastman Chemical Company as a training representative in the Training Department of Human Resources Division. David is married and very dedicated to his wife Pam and their daughter, Jamie.