INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

VALIDATION OF THE 360-DEGREE APPRAISAL PROGRAM
AS AN EFFECTIVE AND FAIR MEASUREMENT OF
INDIVIDUAL PERFORMANCE

by

William Bryan Morgan

Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy
Applied Management and Decision
Sciences

Walden University
May 1998

UMI Number: 9840085

Copyright 1998 by
Morgan, William Bryan

UMI Microform 9840085
Copyright 1998, by UMI Company. All rights reserved.

UMI

300 North Zeeb Road
Ann Arbor, MI 48103

DOCTOR OF PHILOSOPHY DISSERTATION

OF

WILLIAM BRYAN MORGAN

APPROVED:

J. KENT MORRISON
VICE PRESIDENT FOR ACADEMIC AFFAIRS

WALDEN UNIVERSITY
1998

Walden University

APPLIED MANAGEMENT AND DECISION SCIENCES

This is to certify that I have examined the doctoral dissertation by

William Bryan Morgan

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Dr. Joseph Barbeau, Committee Chair
Applied Management and Decision Sciences Faculty
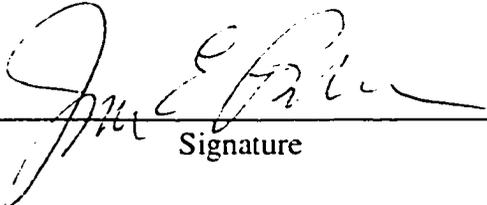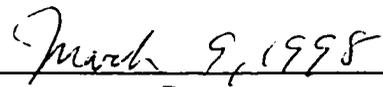
_____
Signature

_____
Date

Walden University

APPLIED MANAGEMENT AND DECISION SCIENCES

This is to certify that I have examined the doctoral dissertation by

William Bryan Morgan

and have found that it is complete and satisfactory in all respects.

Dr. Robert O'Reilly, Committee Member
Applied Management and Decision Sciences Faculty
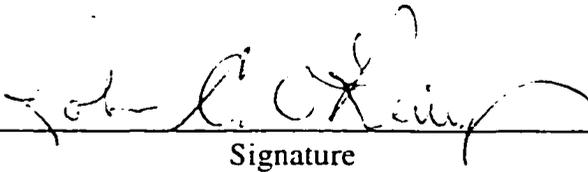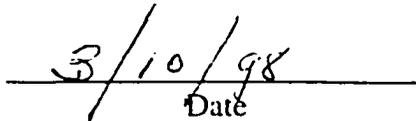
_____
Signature

$3/10/98$
Date

Walden University

APPLIED MANAGEMENT AND DECISION SCIENCES

This is to certify that I have examined the doctoral dissertation by

William Bryan Morgan

and have found that it is complete and satisfactory in all respects.

Dr. Martin Gerstein, Committee Member
Professional Psychology Faculty

_____
Signature
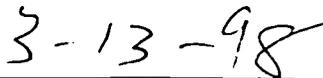
_____
Date

Walden University

APPLIED MANAGEMENT AND DECISION SCIENCES

This is to certify that I have examined the doctoral dissertation by

William Bryan Morgan

and have found that it is complete and satisfactory in all respects.

Dr. Marilyn Simon, Faculty Representative
Education Faculty
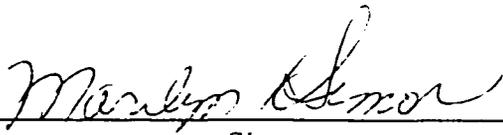
_____
Signature
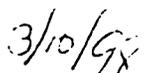
_____
Date

Abstract

VALIDATION OF THE 360-DEGREE APPRAISAL PROGRAM AS AN EFFECTIVE AND FAIR MEASUREMENT OF INDIVIDUAL PERFORMANCE

by

William Bryan Morgan

M. S., Strayer College, 1995
B. S., University of Maryland, 1978

Dissertation Submitted in Partial Fulfillment of
the Requirement for the Degree of
Doctor of Philosophy
Applied Management and Decision
Sciences

Walden University
May 1998

ABSTRACT

The purpose of the study was to develop a solution to the management problem of selecting a performance appraisal that will effectively and fairly measure individual contributions to organizational effectiveness and to meet the demands of social change. The researcher conducted a survey of employees from both government and private organizations who have or are currently using the 360-degree appraisal program. The survey included sampling of management and employee personnel within each of the organizations. The survey was designed to solicit responses related to the perception of effectiveness and fairness of the 360-degree appraisal program to measure individual performance. For the purpose of analysis, the responses to the survey were grouped by organization. Within each organization, the responses were quantitatively analyzed to identify any significant variances based on supplied demographic information such as gender, age, job position, and educational level. The results of the survey sample of all responding organizations were quantitatively analyzed and compared to produce a measure of the perceptions of the effectiveness of the 360-degree appraisal program.

The overall results of the survey were that the respondents' perceptions of the 360-degree appraisal as being an effective, and fair measure of individual performance are slightly negative. However, the study found a number significant demographic relationships useful for profiling differences in these perceptions. A quantitative analysis of these relationships supports the finding that females, who responded to the survey,

were more likely to be younger, less educated, in lower job positions, and to perceive the

360-degree appraisal less positive than men who responded.

ACKNOWLEDGMENTS

To complete a Ph.D. program requires the help and support of a family of friends and supporters. I was blessed with the help and inspiration from a great number of people who fall into both categories. While I can not acknowledge everyone who contributed to my success I would like to single out a few.

I would like to thank Dr. Robert C. O'Reilly and Dr. Martin Gerstein, who both served on my dissertation committee and who willingly served as assessors for my required program writings assignments. Both of these fine professors started with me on my first day at Walden University and were always just a phone call away.

I would especially like to thank Dr. Joseph Barbeau, my faculty advisor and mentor, for his guidance and encouragement from the start of this journey. He provided me with needed truthful and helpful comments at times when I felt like quitting. Every student needs a professor like Dr. Barbeau as their advocate.

I would like to thank a number of key people who helped me conduct the dissertation survey. First, Mr. Paul Wagner, president of U.S. Army Management Engineering College (AMEC). Mr. Wagner provided the critical coordination and introductions to organizations willing to participate in my study. Second, I would like to acknowledge the help I received from Ms. Sharon Gongwer, U.S. Department of Commerce; Mr. Thomas Kelly, U.S. Department of Energy - Golden Field Office; Mr. Ed Whitney, Rock Island Arsenal; Mr. Larry E. Jones, U.S. Army Corps of Engineers - Rock

Island District; and, Ms. Mary Hovden, Norwest Mortgage, Inc. These individuals are each creative leaders who are genuinely interested in improving their organizations. Their interest was evidenced by the efforts they demonstrated in obtaining approvals, making survey distributions, and conducting the follow-ups to insure a high survey return rates.

I would like to single out and thank Dr. David Walonick, who made significant contributions to my dissertation by providing statistical advice for the analysis of the survey. His special area of expertise helped me to select appropriate techniques for the analysis of the data.

Finally, I would like to thank my wife for being able to live with someone who became obsessed with the accomplishment of this specific task. I hope this makes all those who contributed proud. Thank you!

iii

TABLE OF CONTENTS

vi

LIST OF TABLES

vii

CHAPTER 1

Introduction to the Study

Introduction

This chapter provides background and the implications for social change within the workforce of a validated, effective, and fair measurement of individual performance. The statement of the problem investigated by this study and the purpose of the study are presented. This chapter also outlines the general research strategy intended for achieving the study's purpose and states the specific research questions, and identifies the general area of management inquiry related to the problem. The significance of the study and its implication for social change are detailed. Lastly, the assumptions and limitations of the study are presented.

Background

While working on a training program for a government organization which was reorganizing from a matrix (see appendix A - Glossary of Terms) support organizational structure to a project--or systems--team structure, the question "How can management effectively and fairly measure individual job performance under the new management structure of teaming?" was identified. The question of how to measure job performance has long been a matter of interest, debate, and research at most business schools and within many organizations. The answer is now more critical than ever as organizations downsize, reorganize, and become team or systems orientated to meet a global

competitive environment. This paradigm shift creates the need for organizations to be more customer focused and adaptable.

Employees rated by means of the traditional job performance systems were usually rated by only their supervisor. These traditional performance systems were useful and appropriate for organizations having single function departments and a limited span of supervisory control. However, current trends in business appear to opt for less middle-management personnel but toward multifunctional departments. This situation makes it more difficult for supervisors to fully understand and evaluate individual performance.

The paradigm shift of identifying and meeting customer demands is an essential element for competing in today's faster paced global market. The ability to evaluate how well an organization or individual is meeting customer demands and expectations must be measured effectively. Current, traditional, supervisor-only rating systems do not directly measure this critical element.

Much of the related literature suggests the answer to be the use of a multiple-source rating system, often referred to as the 360-degree feedback appraisal system. The 360-degree feedback appraisal system is also known as the full-circle appraisal, multirater assessment or group performance appraisal. It is an approach that gathers behavioral and performance observations from an employee self-assessment and different external and internal sources to evaluate an individual's performance (Bournellis, 1995; Hoffman 1995; Milliman, Zawacki, Norman, Powell, & Kirksey 1994).

The traditional evaluation method of relying on feedback from an employee's supervisor is still used. However, the supervisor's evaluation is only part of an overall

assessment of the employee. By incorporating feedback from the self-assessment, peers, customers, and other sources, it is expected that a more effective and fair evaluation will result. The summary data gathered from the individual 360-degree system survey is compared against organizational strategies, values, and business objectives. Feedback is then given to the rated employee. The objective is to identify areas for both organizational and individual improvement (Hoffman, 1995).

When compared to the supervisor-only form of appraisal system, and based on its multiple input evaluation technique (total view of the individual), the 360-degree system appears to provide the most effective and fair measurement of individual performance. However, some form of proof is needed to substantiate this assumption. For a researcher, this provides an opportunity for such study and analysis--the intent of this dissertation.

The selection of the 360-degree system as the focus of study, and the need for measuring perceptions of effectiveness and fairness, was based on a review of the literature and a comparative analysis of the single versus multirater evaluation systems. The methodology used for the selection was the four-step process presented below. This process revealed a number of interesting strengths and weaknesses for both systems.

The first step was to identify an appropriate evaluation system as a basis for the study and to identify if there was a need for an evaluation system in the first place. The purpose of this process was to identify why it is necessary to evaluate employee performance. Schneier and Beatty (1979) presented a possible answer to this question by defining performance appraisals and outlining the objectives of evaluations. They defined performance appraisals as, "the process of identifying, measuring, and developing human

performance in organizations" (p. 66). They presented the objectives of appraisals as being tied to the following:

1. Giving feedback to employees to improve future performance.

2. Identifying future training needs.

3. Documenting criteria used to allocate rewards.

4. Forming a basis for personnel decisions such as promotions and salary.

5. Providing organizational diagnosis and development information.

6. Facilitating communications between employees and management.

7. Validating selection techniques and human resource policies to meet federal and state Equal Employment Opportunity requirements.

Included in these objectives are two schools of thought pertaining to the appropriate use of evaluations (Brotherton, 1996). First, the evaluations may be used for developmental purposes. This gives the employees candid feedback about their work-related strengths and weaknesses. Second, the evaluation may be used as a tool to help supervisors more objectively assess and reward employees. Regardless of the how evaluation results will be used, Brotherton stresses that trust and confidentiality are vital to insure candid feedback.

Hoffman (1995) also identifies performance appraisals as being a diagnostic and assessment tool for companies to use for aligning employee performance with organizational needs and strategies. Hoffman agrees with Brotherton that there is an important distinction between using the evaluation for developmental purposes as opposed to a performance measure. A review of the comments provided by respondents

to this study's survey (see appendix C) indicates that the distinction between the two purposes, as well as knowing which purpose management was going to use, was a major factor for how they evaluated others.

The use of performance evaluations to validate employer compliance with federal equal employment opportunity laws and policy requirements is identified as another important use of a fair and effective evaluation process. Through the passage of antidiscrimination legislation, such as the Civil Rights Act of 1964, the government attempts to protect categories of employees from employer discrimination in hiring, promotions, pay raises, and other rewards. Management's selection and reward decisions often become legal issues. In fact, according to Patterson (1987), the courts have found organizations in violation of civil rights laws for failing to validate performance appraisal criteria and methods. The risk of litigation makes it essential for organizations to use accurate, valid, and fair appraisals as a defensive action. The same risk of litigation for personnel actions taken by management is also prompting more extensive documentation of poor performance and the need to use scientifically tested employee job performance measurement instruments (Hoffman, 1995; Mayhew 1997; Patterson 1987).

The second step of the process for developing this study was to identify why the current or traditional single-rater evaluation system was no longer an effective evaluation measurement tool. A review of the literature revealed that a number of human resource managers agree that the old single-rater evaluation system does not provide an effective or fair measurement of the employee performance.

Weaver (1996) challenges the effectiveness of the traditional system by stating that it does not lead to productivity and quality improvements. Weaver believes the traditional model places too much emphasis on how employees can improve their individual jobs and not enough emphasis on the customers, the team, and the employee's contribution to improvements. Weaver calls for a performance review process that does the following:

1. Identifies and focuses on the right issues. In the traditional single-rater system the boss is the customer. This situation tends to focus the employee's attention on what the boss wants rather than the customer or broader organizational needs.

2. Uses the right measure. With the traditional rating system, the emphasis is on past performance, past mistakes or successes, and past goals, rather than on present-day problem solving. Performance reviews need to be based on measurable and continuous process improvements that are directly tied to the organization's core business.

Mariotti (1997) challenges the fairness of the traditional single-rater system. He identifies the first problem as being inconsistency between raters, whereby some are easy and others hard graders. This situation makes it possible for one supervisor to rate a mediocre performer as excellent, while another supervisor may reserve high ratings for only exceptional performers. This type of disparity is particularly harmful when there is no process in place for leveling the inequity between raters.

In addition to being plagued by inconsistency, Mariotti identifies lowest-common-denominator and the lack of deviation from the norm scoring as significant problems with the traditional system. Regardless of all the literature and training, many managers

apparently want to avoid the painful, and often confrontational, experience of evaluating subordinate performance. These managers tend to use scoring adjustments to avoid creating barriers between themselves and subordinates. This unfair scoring weakness of the single-rater system can have a very negative effect on employee morale, motivation, and performance.

The third step used to identify an appropriate evaluation system for study was to identify an evaluation system that has the potential for overcoming the problems identified with the traditional single-rater system. The 360-degree performance appraisal system was identified as meeting this requirement. This selection was based on the 360-degree system's ability to provide the following:

1. A shared role of evaluation. The 360-degree system uses input from multiple sources, reducing the severity of any one person's shortcoming as an evaluator (Hoffman, 1995). This is particularly important for reducing the higher potential for leniency, bias, and subjectivity found with the traditional single-rater system. Additionally, using a multirater system reduces the problem of confidentiality by combining the various scores into a composite evaluation.

2. Feedback received from the 360-degree evaluation system is a more current assessment of effectiveness. Unlike the traditional system, the 360-degree system focuses on current skills as perceived across organizational boundaries (Hoffman, 1995).

3. The peer review element of the 360-degree system provides motivational and organizational direction by using peer pressure as an effective force (Mariotti, 1997). As organizations increasingly shift to using work teams, it becomes logical that the team

(peers) is in the best position to fairly and effectively become the evaluator of member performance.

The 360-degree system is not without fault. Logistically, the more employees involved in the evaluation process the higher administrative requirements in terms of cost and time (Brotherton, 1996). Patterson (1997) finds that feedback, obtained by those who have not been trained to give or receive feedback, likely consists of general comments such as "I like . . ." or "I enjoy working with . . ." Patterson also cautions that unsolicited customer feedback tends to come from very pleased or very unhappy customers. If not factored correctly this type of input may impact the overall evaluation process.

The fourth and final step used for developing this study was the selection of a method for measuring the effectiveness and fairness of the 360-degree system. It was decided that a survey instrument would be used as a measure of perceptions of the effectiveness and fairness of the 360-degree system. This decision was based on the results of earlier research conducted by Roberts (1992). Roberts's survey of government officials sought to answer the key question: Is there a significant association between perceived user acceptance and the perceived effectiveness of the performance appraisal system? Based on his analysis of the data, Roberts found a positive association between perceived user attitudes and acceptance of performance appraisals and perceived measures of performance appraisal systems effectiveness. Given this identified association, an inference can be made that a similar association might exist between user perceptions of effectiveness, fairness, and contribution of the 360-degree evaluation

system. The result of this four-step process was the development of a statement of the problem, purpose of the study, research questions, and other elements of this study.

Statement of the Problem

Management needs a validated instrument that can be used to effectively and fairly measure an individual's job performance. Brotherton (1996) identifies the use of team-based management and flatter organizational structures as contributing to this problem. When this type of management philosophy and organizational structure is used, employees spend less time being directly supervised and more time with peers and customers. Additionally, the use of electronic mail and other technology further reduces the amount of face-to-face interaction between employees and supervisors. This situation increases the difficulty for supervisors to fairly and accurately assess an employee's performance or to identify methods for improvement. This study identifies the 360-degree feedback appraisal system as a possible solution to this management need.

Purpose of the Study

The purpose of the study was to conduct a subjective survey study to validate the use of the 360-degree feedback appraisal program as an effective and fair measure of individual job performance as perceived by employees and supervisors.

This purpose clearly states the limited scope of the scientific inquiry for academic purposes. However, it continues to be broad enough to be applied in practice to the

research statement of the problem: A validated instrument is needed by management that can be used to effectively and fairly measure an individual's job performance.

The above purpose represents a topic manageable in size and within the range of the researcher's means. Data are available and the study demonstrates a mastery of the subject. Depending on the method used, the objectives offer an opportunity to make an original and useful contribution to the organization, specifically, and to the literature and business field, in general.

The purpose also contributes to solving the original management question, "How can management effectively and fairly measure individual performance under the new management structure of teaming?" As discussed in detail in chapter 3, evaluation of individual members of functioning teams presents a number of unique management challenges such as (a) competing team and individual goals, (b) decreased hierarchies resulting in less management oversight, (c) interpersonal and cultural aspects, and (d) the ability to operate and interact as a team player.

Research Questions

A survey conducted of employees and supervisors whose job performance had been rated/evaluated (or who have rated/evaluated the job performance of others) using the 360-degree feedback appraisal system sought answers to the following research questions:

1. What is the perceived effectiveness of the 360-degree system to measure individual job performance?

2. What is the perceived <u>fairness</u> of the 360-degree system to measure individual job performance?

3. How do employees and supervisors perceive the 360-degree system as a <u>contributor to improving</u> organizational effectiveness?

The survey, developed to answer the research questions, contained 20 Likert-scale formatted questions and six demographic questions. The 20 questions solicited responses related to the effectiveness, fairness, and contribution to organizational effectiveness of the 360-degree evaluation system. All responses to the 20 questions were considered and scored. The six demographic questions were used to determine what significance age, gender, education, job category, and the amount of experience using the 360-degree system had on individual responses.

A positive response relationship was expected between the three research questions. This was evidenced by those who perceived the 360-degree system as fair also perceiving it as effective and contributory. It was also expected that the majority of responses would be positive. Some correlation between responses to the 20 questions and the demographic elements was predicted.

<u>Significance of the Study</u>

This study provides evidence to validate the 360-degree system as a fair and effective measure of job performance, and its results represent an authoritative source for creating a significant organizational and social change. It is hoped that, by changing to a

new and more effective job evaluation system, a cultural shift within the workforce will take place--a shift directed toward satisfying customer demands.

The significance of this study is also closely related to Madsen's (1992) criteria of originality. Madsen (1992) defines originality as a study or topic that has the "potential to uncover new facts or principles, suggest relationships that were previously unrecognized, challenge existing truths or phenomena, or suggest new interpretations of known facts that can alter people's perceptions of the world around them" (p. 38).

Importance and significance also imply some contribution to knowledge. This contribution might be through the presentation of theoretical (i.e., establishing or verifying theories or models) or practical (the impact of results on the profession or solving a practical problem) evidence considered to be significant (Long, Convey, & Chwalek, 1985).

This study makes an original and useful contribution to the organization being studied and also to the literature and business field in general. Such a contribution is, through the presentation of theoretical evidence, considered to be significant and practical.

The theoretical significance of the study is evidenced by the range of the subjective survey data collected and the resulting quantitative analysis. The practical importance of this study is demonstrated through the presentation of a quantitative, validated, personal evaluation system that is collectively perceived as being an effective and fair measurement of individual contribution. The importance of the results of this

study is of specific importance to those organized under a teaming structure in which individual performance is often influenced and directed by team goals and objectives.

Assumptions

Organizations who permitted their employees and supervisors to participate in this study were the Department of Commerce, U. S. Army Management Engineering College (AMEC), U. S. Department of Energy, and Norwest Mortgage, Inc. The validity of the responses received from those participating was based on the assumption that the perceptions of employees and supervisors would be valid and truthful. The following was also assumed:

1. The participating organizations placed no restrictions on how or when participants answered the survey.

2. The participating organizations would permit the survey to be conducted without making it possible to identify subjects.

3. The organizational unit(s) identified to participate in the study would result in a sufficient sample size.

4. Management would not interfere with the study groups in any way that could influence the survey results.

Limitations

Identified limitations of this study follow:

1. The study only considered employee and supervisor perceptions for evaluation and analysis. No attempt was made to verify these perceptions through the use of other data (i.e., financial statements, production data, or other organizational metrics).

2. Only organizations who used the 360-degree system were surveyed. No attempt was made to survey organizations who were using other appraisal systems to compare the effectiveness of those systems with the 360-degree system.

3. Mostly governmental organizations were surveyed. A total of 180 survey responses were received; of those, only six were from nongovernment workers or supervisors.

4. No attempt was made to identify the administrative use of the data gathered from the 360-degree evaluation of employees.

CHAPTER 2

Literature Review

Introduction

There are many possible ways to solve any research problem, and this review of

the literature provided insight to choose an appropriate method. By reviewing the work

others have done to solve the same or similar problems, the researcher was able to

accomplish the following:

1. place the topic in a historical context,

2. justify selection of the topic,

3. provide an assessment of previous studies, and

4. make a better judgment in selection of the research design (Long et al., 1985,

pp. 71-72).

Historical Context

Performance appraisal research has a relatively long history. The formal study has

been traced to work performed by industrial psychologists during the period just prior to

the Second World War (Scott, Clothier, & Spriegel, as cited in Murphy & Cleveland,

1991, p. 3). Since that time considerable research has taken place to study the

measurement instruments, cognitive and behavioral processes, and the broader contextual

issues of performance appraisal. Wood and Marshall (1993) suggest the most convenient

way to present an overview of this history is to separate studies into four broad schools or

perspectives. These perspectives are labeled the measurement, appraisal interview, social

cognitive processes, and performance management school.

The measurement school or perspective of the 1940s can be traced to the seminal

work of Thorndike on criterion model of measurement, and to other earlier studies about

rating forms and trait psychology (Wood & Marshall, 1993). During this period

performance appraisal methods were studied from a measurement perspective. The

majority of the literature of the time focused on the development of a proper designed

instrument or rating scale to record performance problems.

The appraisal interview school or perspective of the 1950s and 1960s changed the

focus of the research to the communication processes in appraisal interviews and other

managerial dynamics of the process. Among the important work of this time was that of

Maier who identified three communication styles for appraisal interviews: tell and listen,

tell and sell, and problem-solving (Maier, as cited in Wood & Marshall, 1993, p. 12).

With the exception of McGregor and a few of his followers, who identified managers as

judges whose primary role was to evaluate employees, most of the research identified

managers as directly involved with the structure of the interview process, feedback, and

goal-setting (McGregor, as cited in Wood & Marshall, 1993, p. 12).

The social cognitive process school of the 1970s and early 1980s saw the focus of

appraisal research as trying to understand the rater as a motivated decision maker who

processes a range of social cues for making judgements about people (Feldman, 1981;

Landy & Farr, 1980; Wood & Marshall, 1993). Much of the work attempted to identify

the rater's ability and motivation to evaluate accurately. Significant research during this

time centered on understanding rating bias, variance in raters, and the relationship between raters' job knowledge and their ability to evaluate employees (Ilgen, Barnes-Farrell, & McKellin, 1993).

The performance management school of the 1980s identified the appraisal as an integral part of a manager's day-to-day activities and a means of shaping expectations, encouraging, and motivating employees. Performance management was an attempt to transform appraisals from a painful annual event to a continual process (Levinson, 1976; Landy & Farr, 1980; Murphy & Cleveland 1991; Schneirer, Beatty, & Baird, 1986). Within this period the integration of the performance process with other planning and accountability management systems such as the Management by Objectives (MBO) system occurred as well.

In the late 1980s the performance management and collaborative MBO techniques appear to have evolved into a number of comprehensive and customer-focused performance evaluation systems. There is a comprehensive title for these performance evaluation instruments, the 360-degree feedback appraisal system. This appraisal system is also known as the full-circle appraisal, multirater assessment, or group performance appraisal. It is an approach that gathers behavioral and performance observations from an employee self-assessment, and different external and internal sources to evaluate an individual's performance (Bournellis, 1995; Hoffman 1995; Milliman Zawacki, Norman, Powell, & Kirksey 1994).

The traditional evaluation method of relying on feedback from an employee's supervisor is still used. However, the supervisor's evaluation becomes only part of an

overall assessment of the employee. It is expected that by incorporating feedback from the self-assessment, peers, customers, and other sources, a more effective and fair evaluation will result. The summary data gathered from the individual 360-degree system are compared against organizational strategies, values, and business objectives, and feedback is given to rated employee. The objective is to identify areas for both organizational and individual improvement (Hoffman, 1995).

The exception to the majority of the findings in the literature, as well as the first area to resolve before selecting a performance program to measure, was the considerable controversy about the usefulness and fairness of any performance appraisal system. Much of the debate remains centered on the writings of W. Edwards Deming. In his book, Out of the Crisis (1986), Deming discusses some of the negative effects performance appraisals have on performance and why the performance review should be completely eliminated. He writes that the performance appraisal "nourishes short-term performance, annihilates long-term planning, builds fear, demolishes teamwork, and nourishes rivalry and politics" (p. 102). Deming is very clear about his dislike for performance reviews, and identifies them as one of the seven deadly diseases of the corporation that should be eliminated immediately and completely.

According to Weaver (1996) at least one study appears to support Deming. Weaver reports on a study that demonstrated that in a traditional review system, the most reliable predictor of this year's performance rating is last year's rating. This study conducted by Steven Armstrong, manager of manufacturing at Varian Ion Implant Systems in Gloucester, Massachusetts, examined the ratings of 30 employees who had

been with the company 10 years or longer. Armstrong found that their ratings showed virtually no variation over that period although indicators of the companies productivity varied greatly. The inference made from this study, according to Weaver, was that individual performance reviews do not improve productivity or quality.

Many human resource professionals take an opposite point of view from Deming by arguing that modern appraisal systems promote better communication between managers and employees (Gerst, 1995). Even with his negative position on appraisals, Deming appears to support some form of performance feedback. Promoting the Japanese way as better, Deming points out that, in contrast to Americans who depend on one or two evaluation events per year, feedback to the average Japanese employee comes daily (Peters, 1987). Regardless of who is correct in this debate, it can be generally agreed that some form of job performance measure will be found in organizations for the foreseeable future.

Justification of the Topic

The critical question for this inquiry was, "Which evaluation system will provide the most effective and fair assessment of individual job performance?" For the purpose of this study, the 360-degree feedback appraisal system was selected as being the most suited for selection and validation. This selection was strongly supported in the research and writing of accepted leaders in the fields of human resources, management, and organizational effectiveness.

What is measured and how it is measured is considered an important and powerful behavioral tool because it communicates to the workforce, and to customers, what is important to the organization. Kenneth Rose (1995) defines performance measurement as "the process of evaluating performance relative to a defined goal" (p. 63). Some form of metric is needed that creates a process for measuring data, displaying results, and determining subsequent actions. Such a metric should be customer centered and focused on indicators that provide value to customers; measure performance across time; direct information at the level at which they are applied; link the organization's mission, strategies, and actions; and are collaboratively developed by teams of people (Dixon, 1991; Rose, 1995). Although not directly stated by Dixon and Rose in the works referenced, the 360-degree system appears to fit the above metric perfectly.

Rose's definition and call for a customer focused metric was supported in the literature by Peter Drucker (1974). Drucker asks the questions: "What does the worker--unskilled or skilled, manual, clerical, or knowledge worker--need to be able to take the burden of responsibility? What tools does he require? What incentives? What security? and, What do manager and enterprise have to do to be able to ask the worker to take responsibility and to expect him to respond to this demand?" (p. 266). His answers were

1. productive work,

2. feedback information, and

3. continuous learning (p. 267).

Drucker (1974) explained his answers by stating that for workers to become responsible, they must have feedback information on their own performance as it

compares against a standard. This feedback must provide effective, timely, relevant, and operational information. Such information must focus on important job issues. Most important, the information must be provided for the workers' use for self-control and improvement. When these elements exist, the feedback becomes a major reinforcer and is effective in increasing individual job performance. It can be easily deduced that this early explanation by Drucker (1974) of worker needs and feedback may have been the early call for and a significant catalyst in the creation of the 360-degree feedback appraisal system.

A review of the literature related to this research subject found considerable support for organizations to implement the 360-Degree Feedback Appraisal Program. The majority of the contributors to the literature based this support on having identified the 360 as being a valuable instrument for increasing organizational effectiveness. This effectiveness was derived from the 360-degree system's ability to match and then measure individual and organizational goals with customer needs.

It is predicted that the days of traditional supervisor--subordinate performance evaluations are numbered. Companies are turning to 360-degree appraisals. Milliman et al. (1994) found organizations and employees need more accurate and fair employee performance measurements because of the increasingly flatter organizational structures, greater internal changes, and more external competitive pressures. This is supported by G. William Dauphinais (1996), partner-in-charge, U. S. Change Integration Practice, Price Waterhouse LLP, based in New York. Dauphinais (1996) found that feedback from all levels improves performance in a way traditional appraisals never could. They both

suggest that the next generation's appraisal system solution to be the 360-degree feedback appraisal system.

Milliman et al. (1994) also believe the 360-degree system to be extremely effective, fair, and useful. While their beliefs do not appear to be based on any scientific data or study, Milliman et al. support their findings on information from, and statements made by, managers and human resource representatives of Hewlett Packard, Digital Equipment, Hamilton Standard Commercial Aircraft, and Johnson & Johnson Advanced Behavioral Technology.

In another article, Milliman, Zawacki, Schulz, Wiggins, and Norman (1995) present the advantages of 360-degree goal setting as

1. providing greater and more explicit understanding of customer needs and expectations,

2. creating mutual expectations between the supplier of the services (employee) and the customer,

3. providing better goals and feedback for employee development, and

4. measuring goals precisely in 360-degree performance appraisals and making employees more accountable.

These benefits clearly indicate the overarching advantage as providing multiple communication links between employees, supervisors (the organization), and customers. Using these multiple inputs of communication to develop agreeable and measurable goals provides a more precise and fair assessment for employee evaluation (Milliman et al., 1995).

Another advantage and demonstration of the effectiveness of the 360-degree system is a more rapid development of employees, both professionally and personally. This is accomplished because insightful feedback is being provided by the supervisor, a group of peers, and a number of customers in lieu of a single supervisor (Milliman et al., 1995). The depth and breadth of this insight is limited only by the range and type of questions presented in the evaluation--usually those areas considered directly related to the individual's and organization's growth and success.

Hoffman (1995) supports the finding of Milliman et al. (1995) and continues by identifying another advantage of the 360-degree system in that it enables companies to align employee performance with the organization's needs and strategy. Hoffman's research found that those organizations that incorporated a 360-degree performance assessment process gained valuable feedback for decision making. Hoffman, citing Mobil Oil, Digital Equipment of Canada, and Colgate-Palmolive as examples, suggests that when implemented correctly the 360-degree feedback appraisal system is useful for the following reasons:

1. Defines corporate competencies and mission.

2. Increases the focus on customer service.

3. Supports team initiatives.

4. Creates a high-involvement workforce.

5. Decreases hierarchies, promotes streamlining.

6. Detects barriers to success.

7. Assesses developmental needs.

8. Avoids discrimination and bias.

9. Identifies performance thresholds.

10. Is easy to implement.

There are a number of important performance elements that are difficult to quantify. Several of the more obvious examples are (a) interpersonal aspects of work, (b) leadership, (c) communication, (d) service orientation, and (e) the ability to operate as a team player. This is particularly true in the case in which the employee is a member of a team or group and individual objectives may be foregone for a higher team objective or goal. The 360-degree system provides visibility of individual commitment to the organization, recognizes individual contributions to team objectives, and permits peers to critically review and comment on the impact.

Stephen Covey, author of such noted bestsellers as <u>Principle-Centered Leadership</u> (1996) and <u>The 7 Habits of Highly Effective People</u> (1989), found that one of the critical conditions of empowerment necessary to release the enormous capacity within people to meet and exceed the needs of customers and other important stakeholders is accountability through 360-degree feedback. Covey (1996) advocates the 360-degree appraisal system because of the feedback it provides. He states that, without such feedback, the information comes too late, too general, and from the wrong source to be truly empowering (Covey, 1996).

Assessment of Previous Studies

One example of a quantifiable measure (subjective) of the effectiveness of the 360-degree feedback appraisal system was provided by the study performed by Hewitt Associates LLC, a management/consulting company in Lincolnshire, Illinois (Bournellis, 1995). According to their study of publicly traded companies, the use of performance-management programs improved employee performance. Of the 437 companies surveyed, 205 said their performance-management programs resulted in higher profits, better cash flow, stronger stock-market performance, significant gains in financial performance and productivity, higher sales growth for each employee, and lower staff growth rates (Bournellis, 1995).

Evidence of the perception of fairness and effectiveness of the 360-degree system was more recently announced by Texaco Inc. On 18 December 1996, Texaco announced a comprehensive plan to ensure fairness and economic opportunity for its employees and business partners. This plan followed a rigorous review by Texaco of human resource programs accomplished in response to negative events over the past 6 months related to the conduct of a number of Texaco's executives. Among the significant actions to be taken, Texaco reported it would expand the company's 360-degree feedback process to include all managers and supervisors to increase accountability. This plan requires employees, peers, and supervisor of each manager to complete a confidential questionnaire annually to help evaluate how well managers demonstrate expected leadership behavior. Additionally, Texaco plans to align their entire company performance evaluation system to ensure individual and organizational goals are in line

and to create more openness and inclusion. The intended purpose of these efforts will be to foster a fairer and more effective system to improve individual and team performance (Texaco, 1996).

A survey study similar to the one used for this research was conducted by Pierre DuBois and Associates Industrial Psychologists Incorporated (1996). This firm introduced the 360-degree feedback in several Quebec companies involving over 350 managers. Their follow-up evaluation survey found

- 89.8% believed the 360-degree feedback will help them become be better

managers, and

- 87.7% felt the program will definitely help their companies become better.

Furthermore, 63 of the managers, after receiving a second feedback appraisal, were able to measure how their behavior and effectiveness had changed. An evaluation of those managers showed that 86% of the participants clearly improved their behavior at work. These results support the theory that the 360-degree feedback provides an effective tool for measuring and improving performance.

Two other studies have been done that are directly related and similar to this study. The Nevada Operations Office of the United States Department of Energy (1996), and United States Army Engineer District at Rock Island, Illinois (1996), both conducted survey studies to gather feedback regarding the use of the 360-degree evaluation system. While these studies are not replicated here, there is an association that exists. The results of these two surveys were used for verification and validating purposes. This was

possible because both of these organizations also participated in this study by completing and returning 120 surveys.

The Nevada Operations Office of the U. S. Department of Energy (1996) conducted two surveys. The first survey was done in 1995 to determine employees' satisfaction with the supervisor-only rating system in use at that time. After this survey was completed, the Nevada office began using the 360-degree feedback appraisal system. In 1996, another survey was conducted to determine whether the degree of satisfaction had improved with the new appraisal system. The response rates were 141 in the 1995 survey, and 158 in the 1996 survey.

The United States Army Engineer District at Rock Island, Illinois, conducted its survey in 1996. The survey was intended to assess the 360-degree evaluation system they had begun using that year. The survey was sent to 175 ratees and raters, of which 129 responded.

While a more detailed discussion of results is presented in chapter 5, the general conclusion may be inferred from answers given to question 5 of the Nevada survey. Question 5 asked if the 360-degree system represents a fair and accurate reading of performance. During the 1995 survey (supervisor rating), 44% responded positively, while only 33% did so in 1996 (360-degree rating).

Summary

When the original statement of the problem, "A validated instrument is needed by management that can be used to effectively and fairly measure an individual's job

performance," was compared with the relevant work done by others, it became evident that the 360-degree feedback appraisal system was the logical instrument for the researcher to validate. This was also based on the 360-degree system being the instrument most recommended as effective and fair. However, there appears to be a lack of needed scientific and quantitative validation for the 360-degree system in the literature. This need for validation makes this study significant and useful.

CHAPTER 3

Methodology

Introduction

In this chapter the methodology of the study will be explained. As stated in chapter 1, the purpose of the study was to conduct a subjective survey study to validate the use of the 360-degree feedback appraisal program as an effective and fair measure of individual job performance as perceived by employees and supervisors.

The most appropriate method of conducting the research was established by examining the type of research questions posed, the extent of control the researcher had over the events, the degree of focus on contemporary as opposed to historical events the study required, time required, and budget (Yin, 1994). Another major consideration was to accomplish the multiple purpose for conducting the research--academic (dissertation) and business (practical).

The above were considered in a holistic manner, and after careful review and comparison of the different research techniques, it was concluded that a subjective survey would be the best approach for obtaining a measure of the effectiveness and fairness of the 360-degree feedback program. This approach to answering a research question was defined by Balian (1994) as quantitative-descriptive because it uses numbers (statistics) to describe characteristics of a respondent group. The descriptive statistics this study used are presented later in the chapter.

Procedure

A survey study was conducted of employees and supervisors whose job performance was rated/evaluated (or who have rated/evaluated the job performance of others) using the 360-degree system. The study sought to answer three research questions:

1. What is the perceived effectiveness of the 360-degree system to measure individual job performance?

2. What is the perceived fairness of the 360-degree system to measure individual job performance?

3. How do employees and supervisors perceive the 360-degree system as a contributor to improving organizational effectiveness.

Sixteen items on the survey were designed to measure the three constructs (effectively, fairly, and contribution). A factor analysis on these items was expected to support the existence of the constructs. Four items (3, 14, 19, and 20) were designed to measure perceptions of the effectiveness of the 360-degree feedback program. Six items (1, 2, 12, 15, 17, and 18) were designed to measure perceptions about the fairness of the 360-degree feedback program, and six items (4, 8, 9, 10, 13, and 16) were designed to determine the degree to which the 360-degree feedback program contributes to improving organizational effectiveness.

The purpose in conducting the factor analysis was to confirm the existence of the three constructs. The objective was to develop a single factor score for each construct. A factor score is a measure of an individual's attitudes toward the construct as a whole.

Rather than analyzing each of the survey items individually, a factor score enables the researcher to assess a respondent's attitudes towards an entire construct with a single value. Factor scores are calculated by multiplying a person's standardized responses for the items making up a factor by the factor loadings for those items. Thus, factor loadings represent the relative weights of the items in the factor. Items with high factor loadings contribute more weight to the overall factor score than items with low factor loadings. That is, items with high factor loadings are more important (relative to the construct) than items with low factor loadings.

It was expected that three factors would emerge from the factor analysis, and this would result in three factor scores for each individual. Multiple regression analysis was used to explore the relationships between the demographic characteristics of the sample and their factor scores. Multiple regression is a multivariate technique that enables a researcher to explain the variability in a dependent variable using a set of independent variables. It was expected that three regression analyses would be performed (one for each factor). The dependent variable would be the factor score and the independent variables would be the demographic information.

The overall F-ratio is used to test the significance of a regression model. If the overall F-ratio is significant then one or more of the independent variables helped to explain the variability in the factor. The r-squared statistic reveals the proportion of variability in the factor score that is explained by the independent variables. The stepwise multiple regression technique simplifies the job of the researcher by including only those independent variables that help to explain the dependent variable. Independent variables

that are not significantly related to the dependent variable are automatically excluded from the final regression model. Thus, the stepwise technique eliminates all unimportant independent variables.

A one-way analysis of variance was used to explore the possibility of differences between the five organizations included in the sample. It was expected that three ANOVA analyses would be performed (one for each construct). Like multiple regression, the goal was to explain the variability in the dependent variable. In ANOVA, however, the independent variable is called a factor, and it is a nominal variable consisting of categories (called levels). This is in contrast to multiple regression, where the independent variables are interval, ratio, or dichotomous data. For the ANOVA performed in this study, the dependent variable was the factor score, and the factor variable was the organization. Since there were five organizations in the sample, there were five levels in the factor. The F-ratio is used to determine if the variability between groups is significantly different than the variability within groups. When the F-ratio is significant, post-hoc testing is used to determine the nature of the difference(s) (i.e., which levels are different from the others). While there are many different types of post-hoc tests, the least significant difference t-test was chosen because it is a very conservative post-hoc test and unlikely to result in a Type I error where the null hypothesis is wrongly rejected.

Research Outline of Events

A detailed discussion of each of the procedures is presented below. However, in order to provide a logical frame of reference for this research, the following sequential outline of events and procedures was developed.

1. Develop a survey questionnaire consisting of a cover page, introduction letter, instructions, demographics page, and not less than 20 survey questions.

2. Conduct a pilot-test using not less than 20 participants.

3. Based on the results of the pilot-test, revise the survey and/or test procedures.

4. Identify government and private organizations (clusters) necessary to obtain the minimum sample size, and conduct the survey. While one hundred percent sample rate was desirable, the minimum survey sample size required for the survey to be considered valid was 95.

5. Analyze the data obtained from the survey questionnaire.

6. Briefing survey results through both oral and written presentations.

The Survey Instrument

A survey instrument (questionnaire) developed and distributed by mail sought to measure a minimum sample size of 95. The purpose of the survey was to measure the perceptions of employees and supervisors of the effectiveness, fairness, and the contribution to organizational effectiveness of a 360-degree feedback appraisal system.

The mailed questionnaire survey method of research technique was selected based on its

lower costs, better samples, standardization, and respondent privacy (Air University,

1996).

The survey consisted of a cover page, an introduction letter, instruction page,

identification (demographics) pages, and survey questions. A copy of the survey

instrument is found in Appendix B. A different color of paper was used for the cover

page of each organization surveyed. This color coding was used as a method to identify

the number of responses received from each organization. The demographic page was

used to identify key variables related to those responding to the survey. While individual

results were kept confidential, it was necessary to collect demographic information in

order to determine the response rates, to conduct follow-up mailings (to solicit

responses), and to calculate other variances dependent upon demographic data.

Each element or question within the survey was scored using a Likert-type

5-point numerical scale. Selection of this scoring scale was based on Balian's (1994)

evaluation of the Likert-type scale and the use of his "Evaluation of Likert-type Response

Widths." This selection was also based on Sproull's (1995, p. 199) "Major Advantages

and Disadvantages of Different Response Formats." The "ordinal" Likert scale format

used for survey questions was

    1 = Strongly Disagree

    2 = Disagree

    3 = Neutral

4 = Agree

5 = Strongly Agree

The <u>Performance TRAK Expert System and Measurement Library</u>[1] software

program was originally used to compile and format the questions, as well as for

calculating and comparing question responses. Using this computer resource permitted

the researcher to compare and plot individual average survey results for each of the

demographic variables. However, this program does not provide the scientific statistical

testing capability necessary to identify and analysis demographic and organizational

variance. To perform this sophisticated statistical analysis, the advice and assistance of

Dr. David Walonick, President of StatPac Inc., became necessary.

As suggested by Sproull (1995), to avoid bias consideration was given to the

following when writing the questionnaire:

1. Use of language.

2. Avoidance of value laden words and phrases.

3. Avoidance of suggestive wording.

4. Asking one question per item.

5. Specifying the framework for the question.

6. Consideration for adding open-ended items for comment.

7. Limiting the number of items.

---

[1] <u>Performance TRAK Expert System & Measurement Library</u> is produced by Performance TRAK International, Wooster, Ohio 44691. This program permits data to be presented graphics and provides averages, data sorting based on parameters, and plot distribution of responses.

<u>Identification of Sample Size</u>

In addition to knowing the purpose of the study and the population size, three criteria should be specified to determine an appropriate sample size: the level of precision, the level of confidence, and the degree of variability (Miaoulis & Michener, 1976). The criteria may be defined as follows:

- <u>Level of Precision</u> - also referred to as sampling error, is the estimate of the range of the population's true value. Probability theory permits us to estimate the degree of error to be expected for a given sample design (Babbie, 1973). For example, if a researcher using a precision rate of 5% found 60% of those responding were likely to vote in favor of an issue the conclusions could be made that between 55% and 65% of those in the population would vote the same way.

- <u>Level of Confidence</u> - also referred to as the risk level, is based on the central limit theorem. This theorem encompasses the idea that when a population is repeatedly sampled or sampled with a reasonably large sample size, the average value obtained is equal to the true population value. In other words, the distribution will tend toward a normal distribution. The confidence level is the mechanism employed in classical statistical theory to make a statement of inference (Rossi, Wright, & Anderson, 1983). For example, in a normal distribution, approximately 95% of the sample values are within 2 standard deviations of the true population value. This means that if a 95% confidence level is selected, 95 out of 100 samples will have the true population value within a specified range of precision.

- <u>Degree if Variability</u> - this refers to the distribution of attributes in the population. Variability provides a quantitative measure of the degree a distribution is spread out or clustered (Gravetter & Wallnau, 1992). The most frequently used measures of variability are the standard deviation and its square (Sproull, 1995). Since a proportion of .5 indicates the maximum variability in a population, it is often used for determining a conservative sample size.

For this study the total population size could only be estimated. While the population can be defined, their actual numbers cannot be exactly determined and were therefore considered infinite--one for which census in a reasonable time period is impossible (Lapin, 1973). To solve this problem, the population for this study was defined as all workers and supervisors who have used, or are using, the 360-degree feedback appraisal program as a method to evaluate their or other's performance. For calculating appropriate sample size for the population, the number 10,000 was used. Such a high number was selected to minimize its effect on the statistical sample formula selected. The use of an estimate for the population size was considered in the statistical formula used to determine sample size.

<u>Sampling Technique</u>

The sample technique selected for this study was cluster sampling. This was based on a comparative analysis of other methods using Balian's matrix guide worksheet (Balian, 1994). The sampling techniques considered by this analysis were simple,

systematic, stratified proportion, multistage stratified proportion, stratified disproportion, multistage stratified disproportion, cluster, quota, and convenience. The criteria for comparing the sample techniques were population size, accuracy and cost, population listing availability, geographic area, population diversity, prior knowledge of population, and simplicity or complexity of the research issues.

The cluster sampling technique is an economical method for collecting data from a large geographic area where geographic boundaries are considered a survey subarea or "cluster." Rossi et al. (1983) found that while the cluster sampling approach may not possess the same degree of reliability as some other sampling techniques, through appropriate choice of a cluster size, the corresponding decrease in cost will more than make up for the loss. Sproull (1995) defines cluster sampling as a sampling method in which the sampling unit is a group of population rather than a single element. This method is cheaper and faster. However, there is the possibility of a greater sampling error. This disadvantage was overcome by increasing the required sample size and maximizing the number of clusters selected (Babbie, 1973; Babbie, 1992; Rossi et al., 1983; Sproull, 1995).

For this study, the clusters were defined as the individual government or private organizations identified to participate in the survey. Within each cluster, the elements (employees and supervisors) were surveyed. Every attempt was made to sample 100% of the elements of each selected cluster or a calculated minimum sample size based on this study's stated precision, the level of confidence, and the degree of variability. To minimize sampling error, a minimum sample number of 95 was required. The formula for

calculating the study minimum total sample and the minimum sample size for any cluster

is presented below (Air University, 1996):

$$n = \frac{NZ^2 * .25}{[d^2 * [N-1]] + [Z^2 * .25]}$$

Where:

n = sample size required

N = total population size [known or estimated]

d = precision level [usually .05 or .10

Z = number of standard deviation units of the sampling distribution corresponding to the desired confidence level

The minimum sample size of 95 was calculated for this study based on the total

population (N) of 10,000, a 95% confidence level, and a plus (+) or minus (-) of 10%

precision level (d = .10, Z = 1.96). The use of this formula to determine appropriate

sample size is specifically recommend for studies where results will be reported in a

variety of ways, or if there is difficulty estimating percentages or standard deviation of

the attribute of interest (Air University, 1996).

<u>Pretest or Pilot Test</u>

An important step in developing a survey instrument is the pretest or pilot test

(Babbie, 1973; Balian, 1994; Rossi et al., 1983; Sproull, 1995). Babbie makes a

distinction between the two tests by identifying a pretest as an initial testing of one or

more aspects of the study design, and the pilot test as a miniaturized walk-through of the

entire study design. For the purpose of this study the pretest and pilot test were combined and are referred to as a pilot test.

For original instruments, as was the case for this survey, Balian (1994) recommends the pilot test be mandatory. This extra testing is justified because it provides the researcher a full review of the instrument, respondents, and actual test administration. Additionally, and of critical importance, a pilot test provides an opportunity to objectively measure validity and reliability of the instrument as discussed in the validity and reliability sections below. Balian recommends evaluating the pilot test against specific indices. The indices used for evaluating this study were reading of instructions, demonstration of form completion, clarity of questions, validity, reliability, survey completion time, and halo, Hawthorne, or self-fulfilling prophecy effects.

Sproull (1995) highly recommends a pilot study if the research procedures are not familiar or are new. He finds the test useful for uncovering unanticipated events and to avoid disasters or annoyances. Some of the 22 advantages of conducting a pilot study follow:

- Provides information on possible ethical problems.

- Helps determine if the research questions or hypotheses are appropriate.

- Helps determine sample size by allowing estimation of variance from the pilot sample.

- Provides a check on all aspects of the data collection method.

- Provides a check on the validity and reliability of the instruments.

- Provides information for modification of all procedures prior to

- Provides a check on the appropriateness of the statistical test.

- Conducting the actual study.

- Enhances the researcher's reputation for thoroughness (Sproull, 1995, pp. 349-350).

Rossi, Wright, and Anderson (1983) specifically recommend using pilot tests when developing a survey instrument. They suggest an obvious reason as being the difficulty of writing questions that are not confusing or ambiguous. Rossi et al. find a pilot test of 20-50 cases is usually sufficient to discover the major flaws in a questionnaire.

Based on the scientific evidence and recommendations presented above, a pilot test for this research was necessary. After the survey instrument was written and all scoring, evaluating, and administrative requirements were defined, the pilot test was administered. The pilot test was conducted using a selected group of 22 individuals who were representative of the populations being studied. Recommendations from the pilot test were evaluated, and those found to be valid were incorporated into the survey or test design prior to the actual study.

Conducting the Survey Sample

After making pilot test recommended changes, the survey was administered to government and private industry employees and supervisors. As stated earlier, these

groups consisted of participants representing various organizations that have or are using a 360-degree feedback evaluation program. Distribution and receipt of the questionnaire were made using the postal mail.

Every attempt was made to distribute a survey questionnaire to every employee and supervisor within the organization. However, this was dependent on the organization providing a complete and accurate list of eligible respondents. The minimum organizational response rate acceptable for that organization to be considered for comparison with other organizations was the minimum sample rate determined as explained above. (See "Identification and Sample Size and Technique" section.)

The survey distribution and response rates were controlled by using an identification color code on the cover page of each survey. A different cover page color was used for each organization. This provided an accurate response rate count for each surveyed organization. An additional use of the survey identification coding system was to validate the demographic data. Such validation was critical to insure the survey results were representative of the population of the organization. No individual identification coding was done.

Five organizations participated in the survey. A total of 180 surveys were returned which represented a return rate of 60%. The responses from the organizations were Department of Energy (57.8%), Department of Commerce (27.2%), AMEC (8.9%), Corp of Engineers (8.3%), NWM (3.3), and other (4.4%). These organizations started using the 360-degree evaluation program in 1996. Prior to implementing the 360-degree system, the employees were rated using a supervisor-only rating system. Actual experience using

the 360-degree system was solicited as one of the demographic questions. This was done to determine if any variance existed between those new to the 360-degree system and those who had used the system a number of times. A significant interaction between experience level with the 360-degree system and attitudes toward it was found and is explained in chapter 4. Table 1 shows how many times the respondents had been evaluated using the 360-degree feedback performance appraisal system.

Table 1

Experience Being Evaluated Using the 360-Degree System

Times Evaluated	Number	Percent
1-2	120	69.0%
3-4	38	21.8%
5-6	9	5.2%
>6	7	4.0%

The survey also asked those surveyed how many times they had rated others using the 360-degree system. A significant interaction between experience rating others using the 360-degree system and attitudes toward the system was also found, and is explained in chapter 4. Table 2 shows how many times the respondents had rated others using the 360-degree feedback performance appraisal system.

Table 2

Experience Evaluating Others Using the 360-Degree System

Times Evaluated	Number	Percent
1-2	87	50.3%
3-4	41	23.7%
5-6	18	10.4%
7-8	3	1.7%
9-10	2	1.2%
>10	22	12.7%

Analysis of Data

The purpose of the research was to determine the following:

1. Does the 360-degree feedback appraisal effectively measure individual job performance?

2. Does the 360-degree feedback appraisal fairly measure individual job performance?

3. Does the 360-degree feedback appraisal program contribute to improving organizational effectiveness?

To accomplish this measurement, the quantitative-descriptive research method as defined by Balian (1994) was used. The quantitative element of the research is related to the use of numbers (statistics) to describe characteristics of a respondent group. The descriptive statistics and analysis used to measure respondent feedback and to determine variances between groups included standard mean, standard deviation, chi-square, Cronbach alpha, ANOVA, and t-test. This report made use of both descriptive and inferential statistics as defined by Babbie (1973).

The quantitative research approach used the "numeric values" obtained from the questionnaire to first determine mean, mode, and standard deviation calculations of each of the organizations surveyed. Second, an evaluation was done to identify any significant variances within the organizations based on the demographic data provided. Third an evaluation was done to identify any significant variances between the standard statistics. Fourth, all responses were analyzed together as one collective group. Finally, based on the collective results, an inference was made to answer the research questions.

Briefing Survey Results

A primary purpose of this study was to present participating organizations with quantitative data about employee's and supervisor's attitudes toward the 360-degree system. It was felt that such information would be instrumental as a basis for implementing change and developing more effective evaluation systems. Therefore, results were presented to each of the organizations participating in the study.

The survey results were presented in an oral form to each of the organizations participating in the study. The formal written report detailing the results of the study was submitted to each participating organization. The survey result's briefings and written reports included

- The analysis of the results from their organization.

- A comparison of their organization with others who participated in the study.

- Identification of variances, positive and negative, and suggestions on which were significant.

- Summary of the conclusions and recommendations resulting from the study.

Validity

Validity and reliability are critical elements of successful research. Along with an acceptable research strategy, validity and reliability measurements provide the methods to control research and thus insure accurate and consistent results.

Validity: The term simply refers to the question "Does the instrument measure what it was intended to measure?" Babbie, (1992), Balian, (1994), Rosnow and Rosenthal (1996), Sproull (1995), and others provide detailed information related to this measurement. Generally, they each agree there are three types of validity: criterion-related (includes predictive and concurrent validity), construct validity, and content validity.

Criterion-related validity (also referred to as empirical validity) measures the degree to which the test or questionnaire correlates with one or more outcome criteria. Basically, this is an attempt to assess criterion validity by selecting the most sensitive and meaningful criterion in the present (concurrent validity) or future (predictive validity) and then correlate (compare) performance of their test or questionnaire with that criterion (Rosnow & Rosenthal, 1996).

Considered the highest form of validity, construct validity uses both the subjective and objective approaches to determine measurement validity. It uses a multivariate factor analysis to develop factors (or constructs) to measure the degree to which a test or questionnaire relates to expectations formed from those theoretical or hypothetical constructs. Basically, this is the researcher's process of measuring the sensitivity to the correlation between their test or questionnaire and some appropriate criterion, and also the correlation between their test and some other and possibly "inappropriate" criterion. This is done by using variables (constructs) which are not directly observable but may be inferred from other behaviors, and then hypothesizing that, if they are correct, certain behaviors should occur. Construct validity attempts to determine if the test or questionnaire is valid.

The third type of validity, "content validity," is a subjective judgment of the contents of the test or questionnaire to the objectives of using it. Basically, the researcher wants to know if the test items are relevant or represent the kinds of material necessary to achieve the purpose of the research.

In summary, validity is "a judgment of the appropriateness of a measure for specific inferences or decisions that result from the scores generated" (McMillan & Schumacher, 1993, p. 223). This judgment is made subjectively and/or objectively depending on a number of factors. A comparison of the types of validity measurements and examples of the process used to determine the different types of validity is provided in Table 3 (Balian, 1994, Sproull, 1995).

Table 3

Validity Measurement Comparison

Type	Purpose	Advantages	Disadvantages
Criterion-Validity Predictive/Concurrent (OBJECTIVE)	Valid prediction of specific criterion; or, according to already validated measurement.	Statistical objectivity; widely used, accepted, and understood; easy to calculate; fast (concurrent).	Power of statistics depends on criteria used, long time period to establish results, and costly (depending on number of subjects.
Content Validity (SUBJECTIVE)	Valid according to representativeness.	Easy to implement; no statistical analysis; fast; economical; easy to understand.	Not quantitative; process depends on who is involved (author or judge); hard to defend.
Construct Validity (Combination of SUBJECTIVE & OBJECTIVE)	Valid according to support of theory through assessment of various relationships to major variable.	Sophisticated statistics; relatively fast results using analysis; represents synergistic combination of subjective & objective methods.	Requires many instrument questions & subjects; difficult to learn and explain; not easily understood; can be expensive.

The Process For Determining Validity

1. Criterion-related Validity

a. Predictive - Gather scores on predictor variables from validity sample; gather scores on the criterion variable from the same sample at a later time; compute a correlation coefficient between the two sets of scores.

b. Concurrent - Gather scores from the nonvalidated instrument given to a validity sample; gather scores from a previously validated instrument which purports to measure the same variable and which was given to the same sample at approximately the same time; compute a correlation coefficient between the two sets of scores.

2. Content Validity

a. Examine variables of interest and list task, skills, characteristics, etc. involved.

b. Add to the list the important (critical) and frequency of occurrence of each of the tasks, skills, etc. This insures all crucial items are included regardless of frequency of occurrence.

c. Compare each task, skill, characteristic, etc. to the items of the measure to ensure each crucial and frequently occurring task or skill is measured by at least one item. Usually more items (questions) are included for those tasks or skills which are more important or occur frequently.

d. Examine each item (question) to ensure the difficulty level is appropriate for the variable being measured.

3. Construct Validity

a. Examine theories associated with the variable of interest.

b. Select several behaviors the theory indicates will differentiate subjects with differing amounts of the variable.

c. Administer the instrument measuring the variable to the interest to the validity sample.

d. Gather scores for the validity sample on each behavior selected in step 3b.

e. Analyze the data, using statistical test, to determine if subjects scoring high and those scoring low are statistically differentiated on each of the selected criterion variable.

f. Accept evidence of construct validity if each of the statistical tests indicate a significant difference, or a significant relationship, between high and low scorers on the major variable and the criterion variables.

g. Examine reasons if construct validity is not supported (e.g., theory incorrect; instrument not a valid measure of the variable; errors in administration of the instrument, scoring, or analysis of the data).

The survey instrument used for this study was determined valid. Content validity was verified during the pilot-test and construct validity was demonstrated through factor analysis. The factor analysis distinguished one construct which explained 50.3% of the total variance.

Reliability

After determining the validity of an instrument, the researcher must assess its reliability (consistency). Defined by McMillan and Schumacher (1993) as referring to the "consistence of measurement, the extent to which the results are similar over different forms of the same instrument or occasions of data collection" (p. 227). Basically, the researcher must be sure the test or questionnaire, when applied repeatedly to the same subjects, yields the same results each time.

There are several types of reliability estimates for use depending on the type of instrument. The major types are test-retest, split-half, equivalent forms, and Cronbach alpha.

Again, using Balian (1994) and Sproull (1995), a comparison of the different types of reliability estimates and an example of the process used to determine each reliability method is provided in Table 4 (Balian, 1994; Sproull, 1995).

Table 4

Reliability Measurement Comparison

Type	Purpose	Advantage	Disadvantage
Test-Retest	Assess stability over time	Simple to use; easy to understand.	Takes time for two testings; practice effects; extra cost.
Split-Half	Equivalence of two halves (of same instrument)	Administration of one test; easy to understand	Difficulty to split test; not used for quickly-speeded test.
Equivalent Forms (Parallel & Alternate)	Equivalence of forms.	Creates two separate tests.	Time for two testings; difficult to write two equivalent test; and expensive.
Cronbach Alpha	Internal consistency when responses are dichotomous. (Use coefficient alpha when NOT dichotomous)	Recommended for many situations; statistically sophisticated; respected & recognized; provides evaluative information.	Harder to understand & explain; requires computer calculations; not used for quickly-speeded test.

The Process For Determining Reliability

1. Test-Retest Reliability

a. Administer the instrument to the reliability sample at Time 1.

b. Wait a period of time and administer the same instrument to the same sample at Time 2.

c. Correlate the scores from Time 1 and 2.

2. Split-Half Reliability

a. Use an instrument in which the two halves were formulated to measure the same variable.

b. Administer the instrument to the reliability sample.

c. Correlate the summed scores from the first half with (often the odd numbered items) with the summed scores of the second half (often the even numbered items).

d. Compute the Spearman-Brown Prophecy formula to correct for splitting one instrument into halves.

Formula:

$$r \text{ corrected} = \frac{n\,(r_{xx})}{1 + (n - 1)\,r_{xx}}$$

Where: r_{xx} = uncorrected reliability
 n = number of splits (for two halves, n= 2)

3. Equivalent Forms Reliability

a. Administer Form A of the instrument to the reliability sample.

b. Break the sample for a short rest period (10-20 minutes).

c. Administer Form B of the instrument to the same reliability sample.

d. Correlate the scores from Form A and B.

4. Cronbach Alpha

a. Administer the instrument to the reliability sample.

b. Compute the variance of the scores.

c. Compute the proportion of correct responses to each item.

d. Compute the proportion of incorrect responses to each item.

The survey instrument used for this study was determined reliable. The survey

was concluded as being reliable based on the extremely high Cronbach alpha coefficient

obtained. Cronbach's alpha, which is used to measure the reliability (internal

consistency), was .939 for this instrument.

Questionnaire and Research Questions

Implicit in the research questions was the idea that the three constructs

(effectively, fairly, and contribution) were unique, and that respondents would be able to

distinguish between them in their evaluations. The survey was designed around these

constructs; it was assumed that a factor analysis would confirm the existence of the three

constructs. Four items (3, 14, 19, and 20) were designed to measure perceptions of the

effectiveness of the 360-degree feedback program. Six items (1, 2, 12, 15, 17, and 18) were designed to measure perceptions about the fairness of the 360-degree feedback program, and six items (4, 8, 9, 10, 13, and 16) were designed to determine the degree to which the 360-degree feedback program contributes to improving organizational effectiveness.

The demographic questions were used for the following:

1. Age - To determine if any variance in response exists based on the age of respondent.

2. Gender - To determine if any variance in response exists based on gender.

3. Educational Level - To determine if any variance in response exists based on respondent educational level.

4. Job Category/Classification - To identify respondent by separate job categories and classifications (employee, supervisor) to determine if any variance in response exists.

5 and 6. Times 360-Degree System Used - To identify if any variance exists between those new to the 360-degree system as compared to those who have used the system more times.

CHAPTER 4

Results

Characteristics of the Sample

One hundred eighty surveys were returned. This represents a response rate of 60%. Over half (57.8%) were from the Department of Energy. The balance were from the Department of Commerce (17.2%), AMEC (8.9%), Corp of Engineers (8.3%), NWM (3.3%), and others (4.4%). Five respondents did not supply any demographic information and were excluded for the purpose of reporting demographic percentages.

Age was an ordinal variable on the survey. About half (48.6%) of the respondents were less than 46 years of age and 51.4% were 46 years or older. Table 5 shows the distribution of ages for each response category.

Table 5

Counts and Percents for the Age Distribution of the Sample

Age Category	Number	Percent	Cumulative
Less than 25	4	2.3%	2.3%
26-30	11	6.3%	8.6%
31-35	18	10.3%	18.9%
36-40	24	13.7%	32.6%
41-45	28	16.0%	48.6%
46-50	36	20.6%	69.1%
51-55	26	14.9%	84.0%
Over 55	28	16.0%	100.0%

Fifty-six and a half percent of the respondents were male, and 43.5% were female ($\underline{N}$=170). A contingency table was prepared to examine the possibility of an interaction

between gender and age. In order to maximize cell counts, age was classified as a dichotomous variable divided near the median (i.e., less than 46 years of age, and 46 years or older). The chi-square statistic was highly significant, $X^2(1,\underline{N}=170)=13.64$, $\underline{p}<.0001$. Thus, there was a strong significant interaction between gender and age in the sample. This had implications for subsequent analyses in that gender and age should not be treated independently from each other. Table 6 is the contingency table of cell counts. Females in the sample were significantly younger than males in the sample.

Table 6

Contingency Table of Cell Counts Showing the Interaction Between Gender and Age

	Female	Male
Less than 46 years	49	35
46 or more years	25	61

As a whole the sample was very educated. Only 5.1% had only a high school diploma or GED, and 40.0% had at least a master's degree. Table 7 shows the education level of the sample.

Table 7

Education Level of the Sample

	Number	Percent	Cumulative
High School/GED	9	5.1%	5.1%
College, no degree	23	13.1%	18.3%
Associate Degree (A.A., A.S.)	11	6.3%	24.6%
Bachelor's Degree (B.A., B.S.)	62	35.4%	60.0%
Advanced Degree (M.A., M.S., Ph.D, M.D.)	70	40.0%	100.0%

A contingency table was prepared to determine if there was a significant relationship between gender and education. Cell counts were maximized by classifying education as dichotomous (having or not having at least a 4 year degree. The chi-square statistic was highly significant, $X^2(1,\underline{N}=170)=32.04$, $\underline{p}<.0001$. Thus, there was a strong significant interaction between gender and education. Subsequent analyses also needed to take into account the interaction between gender and education. Table 8 is the contingency table of cell counts. Females in the sample were significantly less educated than males in the sample.

Table 8

Contingency Table of Cell Counts Showing the Interaction Between Gender and Education

	Female	Male
No 4 year degree	34	7
4 year degree +	40	89

About a third (32.7%) of the respondents were in a managerial or supervisory position. Table 9 is a contingency table of gender and position. A chi-square analysis revealed a strong significant relationship between gender and position, $X^2(1,\underline{N}=162)=14.86$, $\underline{p}=.0001$. Thus, another gender interaction was uncovered in the sample. Males were significantly more likely than females to be in a managerial position.

Table 9

Contingency Table of Cell Counts Showing the Interaction Between Gender and Position

	Female	Male
Manager	11	42
Non manager	59	50

Over two thirds (69.0%) of the respondents had been evaluated using the 360-degree feedback performance appraisal system less than three times, and 31% had been evaluated three or more times. In contrast, about half (50.3%) had evaluated others using the 360-degree system fewer than three times, and 49.7% had evaluated others three or more times. Spearman's rank order correlation revealed a moderate positive relationship between the number of times they had been evaluated with the number of times they had been evaluated by others ($rs=.63$, $P<.001$).

Overall Responses to the Survey

Table 10 shows how the sample responded to the 20 items on the survey. The table is sorted by level of agreement. That is, respondents most strongly agreed with the items near the top of the table and most strongly disagreed with the items near the bottom of the table. The six items that respondents indicated the strongest agreement were:

I like the idea of my boss knowing how my peers and customers rate me.

Compared to the one rater (supervisor) rating systems, the 360-degree system is MORE fair.

The 360 system has given me a more honest opinion about my performance.

The "Customer" feedback I received from the 360-degree system is the MOST important information.

Those who provide me with feedback on the 360 really want me to succeed.

I have used the feedback I received from the 360-degree system to improve my job performance.

The six items that respondents indicated the strongest disagreement were:

Improvements in overall performance within my organization are directly linked to our use of the 360-degree feedback system.

Since my organization has started using the 360-degree system, overall organizational performance has improved.

Since we started using the 360-degree feedback system, employee moral has improved.

Regardless of how well others rate me, my overall rating will be what the boss wants it to be.

I would prefer we STOPPED using the 360-degree system.

The 360 System is a way for those who DON'T like me to get even.

An obvious paradox is that respondent's said they used the feedback from the 360-degree system to improve their own job performance. However, they did not feel that improvements in overall performance within their organization were directly linked to their use of the 360-degree feedback system.

Table 10

Responses to the 20 Items in Descending Order by the Mean Average Level of
Agreement.

	Mean & SD	Agree	Neutral	Disagree
I like the idea of my boss knowing how my peers and customers rate me.	3.97 0.88	136 75.6%	34 18.9%	10 5.6%
Compared to the one rater (supervisor) rating systems, the 360-degree system is MORE fair.	3.59 1.26	110 61.1%	28 15.6%	42 23.3%
The 360 system has given me a more honest opinion about my performance.	3.39 1.19	104 57.8%	33 18.3%	43 23.9%
The "Customer" feedback I received from the 360-degree system is the MOST important information.	3.38 1.14	84 47.5%	60 33.9%	33 18.6%
Those who provide me with feedback on the 360 really want me to succeed.	3.37 1.00	88 49.2%	64 35.8%	27 15.1%
I have used the feedback I received from the 360 degree system to improve my job performance.	3.33 1.13	103 57.2%	31 17.2%	46 25.6%
The feedback I have received from the 360 system has resulted in improving my performance.	3.22 1.02	89 49.4%	42 23.3%	49 27.2%
The 360-degree performance evaluation system is VERY fair for measuring my job performance.	3.20 1.22	89 49.7%	36 20.1%	54 30.2%
The questions used in the 360 reflect what management thinks is important.	3.18 1.08	81 45.5%	51 28.7%	46 25.8%
The "Peer" feedback I received from the 360-degree system is the MOST important information.	3.05 1.02	63 35.2%	70 39.1%	46 25.7%

(table continues)

	Mean & SD	Agree	Neutral	Disagree
The "Supervisor" feedback I received from the 360-degree system is the MOST important information.	3.01 1.11	59 33.0%	71 39.7%	49 27.4%
The 360-degree system is a VERY effective way to measure my work contribution.	2.97 1.14	71 39.4%	42 23.3%	67 37.2%
Using the 360 removes the "politics" using the rating system.	2.93 1.28	71 39.7%	33 18.4%	75 41.9%
The questions used in the 360 accurately reflect what is important for me to be effective in my job.	2.81 1.13	63 35.0%	41 22.8%	76 42.2%
The 360 System is a way for those who DON'T like me to get even.	2.74 1.23	55 30.6%	34 18.9%	91 50.6%
I would prefer we STOPPED using the 360 system.	2.61 1.33	40 22.3%	50 27.9%	89 49.7%
Regardless of how well others rate me, my overall rating will be what the boss wants it to be.	2.58 1.25	51 28.7%	22 12.4%	105 59.0%
Since we started using the 360-degree feedback system, employee morale has improved.	2.56 1.05	33 18.3%	57 31.7%	90 50.0%
Since my organization has started using the 360-degree system, overall organizational performance has improved.	2.54 1.01	26 14.4%	67 37.2%	87 48.3%
Improvements in overall performance within my organization are directly linked to our use of the 360-degree feedback system.	2.34 1.00	20 11.1%	62 34.4%	98 54.4%

Research Questions

There were three research questions for this study:

1. Does the 360-degree feedback appraisal <u>effectively</u> measure job performance?

2. Does the 360-degree feedback appraisal <u>fairly</u> measure job performance?

3. Does the 360-degree feedback appraisal <u>contribute to improving</u> organizational effectiveness?

Implicit in these research questions was the idea that the three constructs (effectively, fairly, and contribution) were unique, and that respondents would be able to distinguish between them in their evaluations. Since the survey was designed around these constructs, it was assumed that a factor analysis would confirm the existence of the three constructs. Four items (3, 14, 19, and 20) were designed to measure perceptions of the effectiveness of the 360-degree feedback program. Six items (1, 2, 12, 15, 17, and 18) were designed to measure perceptions about the fairness of the 360-degree feedback program, and six items (4, 8, 9, 10, 13, and 16) were designed to determine the degree to which the 360-degree feedback program contributes to improving organizational effectiveness.

A factor analysis was performed on the 16 items. A varimax rotation was used and the exit criteria was set so a factor would not be extracted unless it explained at least five percent of the total variance in the responses to the items. Surprisingly, only one factor emerged, and it explained 50.3% of the total variance. Either respondents were not able to distinguish among the three constructs, or their attitudes with respect to the three constructs were essentially the same. Thus, the factor seemed to represent "overall"

perceptions of the 360-degree system, and there was no support for the three-construct approach.

Table 11 shows the simple structure factor loadings and communalities for each of the 16 items. As expected, items 17 and 19 had negative factor loadings because of the negative phraseology (i.e., "The 360 System is a way for those who DON'T like me to get even" and "Regardless of how well others rate me, my overall rating will be what the boss wants it to be"). Most of the factor loadings were quite high (.7 or greater), indicating that these items had strong relationships with the factor.

Table 11

Varimax Simple Structure Factor Loadings and Communalities for the Initial 16
Variables Solution

Item	Factor Loadings	Communalities
1	0.835306	0.697736
2	0.792461	0.627994
3	0.860964	0.741260
4	0.696539	0.485167
8	0.788380	0.621543
9	0.728885	0.531274
10	0.712147	0.507153
12	0.673886	0.454122
13	0.785526	0.617052
14	0.720870	0.519654
15	0.752435	0.566158
16	0.311267	0.096887
17	-0.505861	0.255896
18	0.608753	0.370580
19	-0.460940	0.212466
20	0.861911	0.742890

Note: Underlining indicates low communalities.

As revealed by the table, the communalities were low for items 16, 17, 18, and 19. A communality is the sum of the squares of all the factor loadings for that variable. That is, it provides an indicator of how much of the variability in an item can be explained by the factor(s). Since only one factor was extracted, the communalities show the proportion of variability in each item that was explained by the factor.

In order to refine the factor, the four items with low communalities were excluded, and another factor analysis was performed excluding the four items. The proportion of explained variance increased from 50.3% to 60.0%. Table 12 shows the simple structure factor loadings and communalities for each of the 12 remaining items. All factor loadings and communalities were high indicating that the composite factor was good representative of the individual items.

Table 12

Varimax Simple Structure Factor Loadings and Communalities for the Final 12 Variables Solution

Item	Factor Loadings	Communalities
1	0.832910	0.693739
2	0.805333	0.648561
3	0.866030	0.750008
4	0.705968	0.498391
8	0.801342	0.642149
9	0.746077	0.556630
10	0.725244	0.525980
12	0.656268	0.430688
13	0.789925	0.623982
14	0.725089	0.525754
15	0.751313	0.564471
20	0.860124	0.739813

Cronbach's alpha coefficient was used to measure the reliability (internal consistency) of the items making up the factor. Cronbach's alpha has an upper limit of one and a lower limit of zero. It is considered to be a conservative estimate of the reliability among a group of items. Cronbach's alpha was .939, an extremely high value. Thus, it was concluded that the factor provided very high reliability. A factor score was constructed for each respondent by summing the products of the factor loadings and the standardized responses. Table 13 shows the Pearson's product-moment correlation coefficients among the items making up the factor with the factor itself. The high correlation coefficients confirm the cohesiveness of the factor.

Table 13

Correlation Matrix of the Final 12 Variable Solution and the Factor Scores

Factor	1	2	3	4	8	9	10	12	13	14	15
1	0.84										
2	0.80	0.71									
3	0.86	0.83	0.74								
4	0.70	0.49	0.53	0.56							
8	0.80	0.59	0.57	0.60	0.48						
9	0.74	0.49	0.53	0.56	0.50	0.76					
10	0.72	0.52	0.54	0.58	0.43	0.67	0.60				
12	0.66	0.49	0.43	0.52	0.45	0.47	0.34	0.45			
13	0.79	0.54	0.57	0.58	0.74	0.58	0.58	0.47	0.58		
14	0.73	0.58	0.52	0.61	0.42	0.51	0.56	0.43	0.47	0.56	
15	0.75	0.64	0.56	0.61	0.41	0.59	0.50	0.50	0.48	0.52	0.53
20	0.86	0.79	0.72	0.77	0.58	0.61	0.52	0.56	0.53	0.63	0.57	0.65

One of the goals of the study was to determine if there were significant relationships among the demographic characteristics of the sample and their perceptions of the 360-degree system. Stepwise multiple regression analysis was used to explore these relationships. The dependent variable was the respondents factor score, and the independent variables were their demographic characteristics.

The factor score (i.e., the dependent variable) was computed directly from the factor loadings. For all items in the factor, each subject's responses were first standardized (i.e., expressed in terms of standard deviations away from the mean). The standardized score was then multiplied by the factor loading for that item, and the sum of the products for all the items became the factor score for that individual. This resulted in factor scores that were positive for people that had favorable attitudes towards the 360-degree system, and negative factor scores for people with less favorable attitudes. Because the data was standardized, the resulting factor scores were balanced so that the sum of the positive factor scores was approximately equal to the sum of the negative factor scores. The mean factor score was .02 (SD=.72) and the range was from -16.7 to 15.9. The Kolmogorov-Smirnov statistic indicated that the distribution of factor scores was not significantly different from a normal distribution (KS=.81, p>.05).

All of the independent variables were nominal or ordinal data. Thus, dummy variables (dichotomous variables coded as one or zero) were created. In order to reduce the number of dummy variables, a median split was used on the ordinal variables to create a single dummy variable, rather than one for each category. In addition,

independent variables were also created for the gender-age, gender-education, and

gender-position interactions. The final list of independent variables was:

Male	1=male 0=female
Over 45 years	1=over 45 years 0=45 years or less
Four-year degree	1=has degree 0=does not have degree
Manager	1=manager 0=nonmanager
Male times over 45 years	(interaction)
Male times four-year degree	(interaction)
Male times manager	(interaction)
Evaluated 3+ times	1=yes 0=no
Evaluated others 3+ times	1=yes 0=no

One variable (over 45 years) was stepped into the regression model. The overall

F-ratio was significant, $F(1,157)=5.195$, $p=.024$. However, the r-squared statistic revealed

that only 3.2% or the variability in the factor score could be explained by age. While the

relationship between age and the factor score was significant, the relationship was very

weak, and the proportion of explained variance was too small to make it a useful model.

Table 14 shows the regression statistics for the model. Since the coefficient was positive,

older respondents (over 45) tended to be perceive the 360-degree system more favorably

than younger respondents (45 or less).

Table 14

Regression Coefficients Using Age to Predict Factor Scores

Variable	Coef.	Beta	F-Ratio	P	SE
Over 45 years old	2.6344	0.1790	5.1952	0.0240	1.1558
Constant	-1.3404		1.0279	0.3122	1.3221

A one-way analysis of variance (ANOVA) model was prepared to test whether or

not there were significant differences in the factor scores among the five organizations

that participated in the survey. The dependent variable was the factor score, and the factor

variable was the organization. There were five levels for the factor: (a) Department of

Energy, (b) Department of Commerce, (c) AMEC, (d) Corp of Engineers, and (e) NWM.

The eight respondents from other organizations were excluded from the analysis because

they did not represent a cohesive group. Table 15 shows that the results of the ANOVA

were highly significant, $F(4, 167)=9.797$, $p<.0001$.

Table 15

One-way ANOVA Showing That There Were Significant Differences in Perceptions
Among the Five Organizations

Source	DF	Sum of Squares	Mean Squares	F-Ratio	P
Organization	4	1722.858	430.714	9.797	<.0001
Error	167	7342.256	43.966		
Total	171	9065.113			

Since the F-ratio was significant, post-hoc testing was required. Least significant

difference t-tests were performed to compare all the organizations to each other in order

to determine which organizations were different from the others. Table 16 shows the

mean and standard deviation of the factor scores for all five organizations. A positive

mean indicates greater satisfaction with the 360-degree system and a negative mean

indicates less satisfaction.

Table 16

Means and Standard Deviations of the Factor Scores for the Five Organizations in the Sample

Organization	Mean	SD	N
Dept of Energy	-0.41	6.66	104
AMEC	7.95	5.45	16
NWM	5.50	2.59	6
Corp of Engineers	-0.83	5.42	15
Dept of Commerce	-4.03	7.91	31

The post hoc t-tests revealed that respondents from AMEC and NWM shared similar positive attitudes towards the 360-degree system. The Department of Energy and the Corp of Engineers employees both had slightly negative attitudes towards the 360-degree system, and the Department of Commerce employees stood alone with strongly negative attitudes relative to the employees of the other organizations. Table 17 shows the results of the post-hoc least significant difference t-tests.

Table 17

Post-hoc Least Significant Difference t-tests Comparing All Combinations of Organizations on Their Mean Factor Scores

Organization 1		Organization 2	t	df	p
Dept of Energy	&	AMEC	4.697	118	≤.0001
Dept of Energy	&	NWM	2.124	108	.0351
Dept of Energy	&	Corp or Engineers	0.229	117	.8189
Dept of Energy	&	Dept of Commerce	2.669	133	.0084
AMEC	&	NWM	0.771	20	.4415
AMEC	&	Corp of Engineers	3.685	29	.0003
AMEC	&	Dept of Commerce	5.871	45	≤.0001

(table continues)

Organization 1		Organization 2	t	df	p
NWM	&	Corp of Engineers	1.978	19	<u>.0496</u>
NWM	&	Dept of Commerce	3.224	34	<u>.0015</u>
Corp of Engineers	&	Dept of Commerce	1.535	44	.1267

Note: Underlining indicates a significant difference.

The fact that the employees from the Department of Commerce had such negative attitudes compared to the other respondents prompted a multiple regression analysis on just those respondents. The dependent variable was the factor score and the independent variables were the demographic information.

Two variables were stepped into the regression model and the F-ratio was highly significant, $F(2,22)=7.373$, $p=.0035$. Furthermore, the corrected r-squared (corrected for low n) revealed that 34.7% of the variability in factor scores could be explained by the regression model. Table 18 shows the regression coefficients. The beta weights for both significant variables were approximately equal, indicating that the predictive ability of the two variables were about the same. The positive coefficient for managers means that managers were more likely to have positive attitudes compared to non-managers. The negative coefficient for those who had more experience being evaluated indicates that the more they had been evaluated by the 360-degree system, the less they liked it.

Table 18

Regression Coefficients Using Managerial Status and Experience Being Evaluated by Others to Predict Factor Scores for Department of Commerce Employees

Variable	Coef.	Beta	F-Ratio	P	SE
Manager	8.0889	0.4289	6.7155	0.0167	3.1214
Evaluated 3+ times	-10.7298	-0.4329	6.8411	0.0158	4.1023
Constant	3.3460		0.4830	0.4943	4.8144

An additional stepwise multiple regression analysis was performed where their experience level being evaluated by others was not included as a potential independent variable. The idea was to determine if their experience evaluating others would enter into the model instead of their experience being evaluated by others. There was a strong rationale for this since Spearman's rank-order correlation coefficient between the two variables was strong (rs=.852, p<.0001). Surprisingly, their experience level in evaluating others was not significant. Managerial status still was entered into the model, and by itself, explained 18.1% of the variability in factor scores, $F(1,23)$=6.305, p=.0195. Apparently, being evaluated by others was associated with negative attitudes towards the 360-degree system, but the frequency that they had personally evaluated others did not significantly affect their attitudes.

Of particular interest was the fact that managerial status had such a strong significant effect on attitudes towards the 360-degree system for the Department of Commerce employees. Yet, it had not been a significant factor when the entire sample was evaluated. An additional stepwise multiple regression analysis was performed on all the respondents who were not from the Department of Commerce (N=149). The

dependent variable was the factor score and the independent variables were the demographic information. One variable was stepped into the model and the F-ratio was significant, $F_{(1,132)}=14.792$, $p=.0002$. The frequency that they had been evaluated by others was the only significant variable, and by itself, explained 10.1% of the variability in the factor scores. Managerial status was not significant. Furthermore, the coefficient was positive, indicating that those who had more experience being evaluated by others tended to have more positive attitudes towards the 360-degree system. This was in direct contrast to the results for the Department of Commerce employees. Table 19 shows the results of the regression.

Table 19

Regression Coefficients Using Experience Being Evaluated by Others to Predict Factor Scores for Respondents Who Were Not Employed by the Department of Commerce

Variable	Coef.	Beta	F-Ratio	P	SE
Evaluated 3+ times	5.4972	0.3174	14.7923	0.0002	1.4293
Constant	-0.3109		0.0550	0.8149	1.3256

CHAPTER 5

Summary, Conclusions, and Recommendations

Summary

The primary purpose of this study was to develop the solution to the management problem of selecting a measurement instrument (performance appraisal) that effectively and fairly measure individual contributions to organizational effectiveness. The question of how to measure job performance has long been a matter of interest, debate, and research. Additionally, the paradigm shift of identifying and meeting customer demands has become an essential element of competing in a global market. The ability to evaluate how well an organization or individual is meeting customer demands and expectations has become an area where effective and fair measurement is needed. After reviewing a number of appraisal systems the 360-degree feedback appraisal program was selected as an evaluations system that could possibly solve the problem.

The 360-degree feedback appraisal system is also known as the full-circle appraisal, multirater assessment, and group performance appraisal. It is an approach that gathers behavioral and performance observations from different external and internal sources to evaluate the performance of an individual or group. Usually, the rated individual is evaluated by a combination of supervisor, peers, subordinates, and customers. The results of the appraisal are compared against organizational strategies, values, and business objectives. Feedback is then provided to the rated employee. The

objective is to identify areas for both organizational and individual improvement

(Hoffman, 1995).

To develop a solution to the management need for a validated instrument that can

be used to effectively and fairly measure an individual's job performance, a study of the

360-degree system was conducted. The purpose of the study was to conduct a subjective

survey study to validate the use of the 360-degree feedback appraisal program as an

effective and fair measure of individual job performance as perceived by employees and

supervisors. A survey study was conducted of employees and supervisors whose job

performance was rated/evaluated (or who have rated/evaluated the job performance of

others) using the 360-degree system. The study sought to answer three research

questions:

1. What is the perceived effectiveness of the 360-degree system to measure

individual job performance?

2. What is the perceived fairness of the 360-degree system to measure individual

job performance?

3. How do employees and supervisors perceive the 360-degree system as a

contributor to improving organizational effectiveness.

A survey instrument was developed (see appendix B), and a pilot test was

conducted. Based on the recommendations developed by the pilot-test, the survey was

revised and then distributed to the U.S. Army Management Engineering College

(AMEC), U.S. Department of Commerce, U.S. Department of Energy—Golden Field

Office, U.S. Army Corps of Engineers--Rock Island District, and Norwest Mortgage, Inc. (NWM).

Conclusions

One hundred eighty surveys were returned, which represented a response rate of 60%. The overall results of the survey were that the respondents' perception of the effectiveness, fairness, and ability to improve "individual" performance of the 360-degree feedback appraisal program are slightly negative. However, the study found a number of significant demographic differences useful for profiling differences in these perceptions.

Implicit in the research questions was the idea that three constructs (effectively, fairly, and contribution) were unique, and that respondents would be able to distinguish between them in their evaluations. However, when a factor analysis was performed, it was found that respondents were not able to distinguish among the three constructs, or their attitudes were essentially the same. This finding was not surprising. It was considered logical that those who would perceive the 360-degree system as fair would also perceive it as effective, and a contributor to performance. Therefore, the only construct considered in making conclusions was respondents' overall scored perceptions, based on the total survey.

To better understand the results of the survey, a detailed statistical analysis of the demographic data was performed. Specifically, the demographic data was examined to determine if any interaction existed since this would have implications for subsequent

analysis. A strong significant interaction between gender and age was found – females in the sample were younger than the males in the sample who responded. There was also a strong significant interaction between gender and education – women in the sample were significantly less educated than males in the sample. Another strong significant interaction was found between the relationship between gender and position – males were significantly more likely than females to be in a managerial position.

While the above was found when the overall survey was analyzed, the nongovernment agency (NWM) was an exception. The demographic data for the six respondents from NWM were age 45 years or under (N=5), female (N=5), advanced education (N=4), supervisor (N=1). The one supervisor respondent from NWM was female. Finding there was a possible demographic difference between employees of government and nongovernment organizations was not expected or considered. Unfortunately, because only one nongovernmental agency was sampled, and its sample rate was only 6, no useful comparison could be made using the variable government and nongovernment.

Discovering that females in the overall sample tend to be younger, less educated, and in lower managerial positions than males was expected. The overwhelming majority of those surveyed were working in governmental agencies. Although major improve-ments in the percentages of female working in management positions have been made, many agencies continue to be male dominated. This finding was also supported by similar findings presented in the U.S. Department of Labor's Glass Ceiling Commission reports and the Center for Research on Women (Burbridge, 1994).

One of the goals of the study was to determine if there were significant relationships between the demographic characteristics of the sample and their perceptions of the 360-degree system. A weak positive relationship (r^2=.032) was found for older respondents (over 45). This is interpreted as meaning they perceived the 360-degree system as slightly more favorable than younger respondents. When coupled with the above relationships, it can be concluded that male respondents were also more likely to be older, more educated, in supervisory positions, and to perceive the 360-system as more positive.

A number of explanations for this finding are possible. First, as indicated in Table 17, being evaluated by others was associated with negative attitudes towards the 360-degree system, but the frequency respondents personally evaluated others did not significantly affect their attitude. Since those in the positive 360-degree system profile could be expected to be raters, as opposed to and more frequently than being rated, their perception would be more positive. Second, it is probable that this group of respondents in the positive profile (older, male, and management) were the decision makers, implementers, and champions for the 360-degree system. Third, the results of the 360-degree appraisal would be less likely to cause any adverse career or job impact to this group of employees. Fourth, this group might be less likely to criticize themselves (management) for implementing a system that was not successful. Fifth, the respondents in senior positions would be in a better position to evaluate the effectiveness of the 360-degree system, and therefore base their perceptions on known information or facts. They

could be expected to have access to information related to organizational and personnel effectiveness which could support their perceptions.

AMEC and NWM were exceptions to the general finding of a slightly negative perception of the 360-degree system. The positive perception of NWM was not considered in this conclusion as being significant. NWM presented an insufficient sample size of only six responses. Additionally, no analysis could be performed to determine if any significant relationship existed between the workforce demographic and the perceptions of governmental and nongovernmental employees.

As presented in chapter 4, Table 16, two organizations stood out on each extreme. These organizations were AMEC and the Department of Commerce. A comparison of responses between organizations revealed that respondents from AMEC had the most positive perceptions of the 360-degree system. Those respondents from the Department of Commerce had the most negative perceptions of the 360-degree system. When the demographics of the respondents of these two organizations were analyzed, it was found that AMEC fits the male, older, more educated, and supervisor profile previously identified more than those responding from the Department of Commerce. Table 20 shows the demographic comparison of between the two organizations.

Table 20

Demographic Comparison Between AMEC and Department of Commerce Respondents

Category	AMEC	Dept. of Commerce
Male	81.2%	46.6%
Age (Over 40)	81.2%	56.6%
Education (B.S.+)	87.5%	60.0%
Supervisor	31.2%	23.3%

Another explanation for AMEC's strong positive perception of the 360-degree system may be attributed to the fact that the organization had a vested interest in the success of the 360-degree system. AMEC provided consulting services to administer, process ratings, and develop feedback reports for the 360-degree performance management system. Generally, the employees of AMEC worked independently providing consulting and teaching services intended to maximize organizational effectiveness. One of the tools they advised using was the 360-degree system; and, when fully implemented, as they represented it, a positive perception could be expected. This conclusion was indicated by the written comments from AMEC. While the written comments received from the respondents tended to identify the 360-degree system as being too behavioral and not asking the right questions, these and the other comments centered around procedural and implementation issues.

As presented in chapter 4, Table 16, the respondents from both the Department of Energy and the Corp of Engineers shared slightly negative attitudes towards the 360-degree system. When reviewed collectively it was concluded that, except for AMEC and NWM, the remaining organizations had negative perceptions towards the 360-degree

system. The two organizations (AMEC and NWM) that perceived the 360-degree system as very positive should not be considered as significant. When these organizations were considered alone it was found that they did not represent the population. The sample sizes for NWM and AMEC were small, and AMEC should be considered suspect because of their demographic profile and potential bias toward the 360-degree system.

The very negative perception of the 360-degree system reported by respondents from the Department of Commerce required further review. Not only did this department stand alone with strongly negative attitudes relative to the other organizations, they were also in direct contrast in another area. When an analysis was performed to determine if experience with the 360-degree system (number of time rated or as a rater) was likely to affect perceptions, respondents from the Department of Commerce were in direct contrast to the other organizations. Managerial status and experience using the 360-degree system had a strong significant effect on attitudes. As shown in chapter 4, Table 18, the analysis indicated that the more they had been evaluated by the 360-degree system, the less they liked it.

In an attempt to explain why the responses from the Department of Commerce were so negative compared to the other respondents, a separate multiple regression analysis was performed on just those respondents. Additionally, several telephone conversations and a personal field visit was conducted with this organization. The purpose of these communications was to discover why their responses were so inconsistent with those of the other surveyed organizations. It was found that a number of

factors may have contributed to the very negative perception toward the 360-degree system and why the managerial status demographic variable was significant.

The general morale and attitude within the organization appeared very low. This was evidenced in the written comments and verified through discussions and the field visit. Contributing factors for the general poor morale and attitude toward the 360-degree system were that Commerce Secretary Mr. Ron Brown had recently been killed in a plane accident, the organization had just been re-organized, a number of supervisor positions had been re-aligned to nonsupervisory, it was rumored that Congress was planning to completely close the department, and employees were generally suspicious of change.

An underlying opposition to changing, especially the performance appraisal system which has considerable "personal" career impact, during the turbulent time listed above was considered as being significant. To fully understand this impact it is important to know that when a governmental agency reorganizes, closes, or conducts a Reduction In Force (RIF), the employees' performance rating score becomes an important factor for job survival. The change to a new, unknown, multirater system appears to have created a great deal of skepticism and fear. Some feared that this new system of scoring could jeopardize their ability to compete for jobs in outside agencies.

The significant difference in perceptions among supervisors discussed above may also be explained by the general opposition to change and concern over career impacts. Another explanation found was that because the department had been reorganized, many of those who reported themselves as supervisors were no longer in supervisory positions.

Therefore, their perceptions were that of an employee, and not as a supervisor. To fully understand the dynamics, cause and effect, and nature of this large organization was beyond the scope of this study. However, it presents an excellent opportunity for future study.

Another paradox identified was respondents saying they used the feedback from the 360-degree system to improve their own performance. However, they did not feel that it contributed to improvement in overall performance within their organization. This difference may be attributed to the design or the implementation of administering the 360-degree survey used by those organizations.

Many of the written comments returned (see appendix D) state that the 360-degree survey questions used to evaluate them did not measure their job performance. Respondents state the questions were too behavioral, lacked specific detailed performance measure, were more of a personality test rather than a performance measure, and the results from the 360-degree rating had not been used to create individual development plans to enhance performance. This was also evidenced by the question, "The questions used in the 360 accurately reflect what is important for me to be effective in my job." A significant number (42.2%) of the respondents disagreed with this question. This paradox presented another area where additional study would be useful.

The original problem this study attempted to solve was to find a validated instrument that can be used to effectively and fairly measure an individual's job performance. It was felt that the 360-degree system only partially meets that requirement. The slightly negative perception found (excluding the extremes discussed above) can be

explained as dissatisfaction in management's administration, implementation, and follow through. Respondent comments state the 360-degree appraisal did not ask the right questions, results were not tied to rewards, and people could pick their raters and therefore, manipulate scores. The system has the potential for being fair, effective and contributing if appraisals are designed correctly, and management commits to integrating them into the overall resource and personnel development process.

Recommendations and Implication for Social Change

The findings of this study make a significant contribution to the business and human resource communities and have the potential for creating social change. The contributions are presented as both theoretical and practical. The importance of this study is demonstrated first through the presentation of a personal evaluation system for quantitative analysis of user perceptions and second by the identification of significant demographic variances in those perceptions.

A cultural shift within the workforce continues to take place. Organizations continue to re-organize, re-engineer, re-size, and re-tool to meet changing demands. One of the results from these organizational changes are less bureaucratic and flatter organizational structures. The current structural trend appears to be the use of more work teams and ad hoc structures. These teams are formed as a way to meet the speed of customer and technological change. To effectively cope with this cultural shift, management must have a performance evaluation system that has more encompassing

criterion. This criterion should include employee acceptance of the system and a system that contributes to the performance of the individual, group, and organization.

The performance management school or perspective was concerned with accurately measuring performance. However, accuracy may not necessarily be the most effective measure of success. This study provides management with a picture of the diversity in perceptions of fairness and effectiveness between various demographic elements common to every organization. With this information management has the opportunity to make more informed evaluation and performance evaluation decisions. The results from this study should significantly contribute to development of a more effective performance measurement instruments that would create both organizational and positive long-term social change.

This study indicates several areas for future research, but the most promising areas are research that link to the work that has already been done by this study. Specifically, additional research is needed in the following areas

1. Research to identify and differentiate between the effective and noneffective 360-degree systems.

2. Research to identify which post evaluation processes are most effective.

3. Research to identify the most effective method for administering the 360-degree system (e.g., selecting raters, reducing bias, and standardized scoring).

4. Research to determine if the 360-degree system is perceived differently between government and nongovernment employees, and between different job categories.

5. Research to understand the paradox of why respondents reported they used the feedback from the 360-degree system to improve their own performance. However, they did not feel that it contributed to improvement in overall performance within their organization.

6. Research to understand the relationships and organizational nature identified within the Department of Commerce.

REFERENCES

Air University, Department of the Air Force. (1996). Air University sampling and surveying handbook. Maxwell AFB, AL: Author. [Online]. Available: www.au.af.mil/au/hq/selc/smplntro.htm [1997, May 19].

Babbie, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing.

Babbie, E. R. (1992). The practice of social research (6th. ed.). Belmont, CA: Wadsworth Publishing Company.

Balian, E. S. (1994). The graduate research guidebook (3rd. ed.). New York: University Press of America.

Bournellis, C. (1995). Peer evaluations: The impact of performance. Lan Times. [Online]. Available: www.wcmh.com/lantimes/95oct/510c125a.html [1997, Jan. 10].

Brotherton, P. (1996, May). Candid feedback spurs changes in culture. HR Magazine, pp. 47.

Burbridge, L. C. (1994). The glass ceiling in different sectors of the economy: Differences between government, non-profit, and for-profit organizationss. [Online]. Available:www.ilr.cornell.edu/library/e_archive/glassceiling/ 3/3front.html [1997, Dec. 6].

Covey, S. R. (1996). Conditions of empowerment. Principle-Centered Leadership. [Online]. Available: www.afpc.af.mil/quality/covey.html [1996, Dec. 31].

Dauphinais, G. W. (1996). Middle management's new roles. HR Focus, 73(10), 12-14.

Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Massachusetts Institute of Technology.

Department of the Army. (1996). Summary of 360 evaluation test survey results. Rock Island, IL: U.S. Army Engineer District, Rock Island.

Dixon, J. R. (1991). The new performance challenge: Measuring operations for world-class competition. Homewood, IL: Irwin.

Drucker, P. F. (1974). Management: Tasks, responsibilities, practices. New York: Harper and Row.

Feldman, J. M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, 127-148

Gerst, R. M. (1995). Assessing organizational performance. Quality Progress, 28(2), 85-88.

Gravetter, F. J., & Wallnau, L. B. (1992). Statistics for the behavioral sciences (3rd. ed.). New York: West Publishing Company.

Hoffman, R. (1995). Ten reasons you should be using 360-degree feedback. HR Magazine, 40(10), 82-85.

Ilgen, D. R., Barnes-Farrell, J. L. & McKellin, D. B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? Organizational Behavior and Human Decision Processes, 54, 321-368.

Landy, F. J., & Farr, J. L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

Lapin, L. L. (1973). Statistics for modern business decisions. New York: Harcourt Brace Jovanovich.

Levinson, H. (1976). Appraisal of what performance? Harvard Business Review, 54, 30-32.

Long, T. J., Convey, J. J., & Chwalek, A. R. (1985). Completing dissertations in the behavioral sciences and education. San Francisco, CA: Jossey-Bass Publishers.

Madsen, D. (1992). Successful dissertations and theses. San Francisco, CA: Jossey-Bass Publishers.

Mariotti, J. (1997, January). Tough bosses, easy bosses. Industry Week, 57.

Mayhew, D. (1997). Performance appraisal. [Online]. Available: http://www.doeren.com/do03007.htm [1997, Dec. 23].

McMillan, J. H., & Schumacher, S. (1993). Research in education: A conceptual introduction. New York: HarperCollins College Publishers.

Miaoulis, G., & Michener, R. D. (1976). An introduction to sampling. Dubuque, IA: Kendall & Hunt Publishing.

Milliman, J. F., Zawacki, R. A., Norman, C., Powell, L., & Kirksey, J. (1994). Companies evaluate employees from all perspectives. Personnel Journal, 73(11), 99-103.

Milliman, J. F., Zawacki, R. A., Schulz, B., Wiggins, S., & Norman, C. A. (1995). Customer service drives 360-degree goal setting. Personnel Journal, 74(6), 136-141.

Murphy, K. R., & Cleveland J. N. (1991). Performance appraisal: An organizational perspective. Boston: Allyn & Bacon.

Patterson, D. (1996). Rethinking performance appraisal. [Online]. Available: http://www.fryberger.com/news/html [1997, Dec. 23].

Patterson, T. F. (1987). Refining performance appraisal - A reliable and useful method. Journal of Extension, 25 (4). [Online]. Available: gopher://gopher.ext.vt.edu: 4070/00/joe/ 1987winter/a5 [1998 Jan. 6].

Peters, T. (1987). Thriving in chaos: Handbook for a management revolution. New York: Harper and Row.

Pierre DuBois and Associates Industrial Psychologists Inc. (1996). 360° feedback. [Online]. Available: www.apexdigital.com/dubois/360_eng.html [1997, Feb. 12].

Roberts, G. E. (1992, May-August). Linkages between performance appraisal system effectiveness and rater and ratee acceptance. Review of Public Personnel Administration, 19-41.

Rose, K. H. (1995). A performance measurement model. Quality Progress, 28(2), 63-66.

Rosnow, R. L., & Rosenthal, R. (1996). Beginning behavioral research (2nd. ed.). Englewood Cliffs, NJ: Prentice Hall.

Rossi, P. H., Wright, J. D., & Anderson, A. B. (Eds.). (1983). Handbook of survey research. New York: Academic Press.

Schneier, C. E., Beatty, R. W. & Baird, L. S. (1986, May). Creating a performance measurement system. Training and Development Journal, 74-79.

Sproull, N. L. (1995). Handbook of research methods: A guide for practitioners and students in social sciences (2nd. ed.). Metuchen, NJ: The Scarecrow Press.

Texaco. (1996). Texaco announces comprehensive plan to ensure fairness and economic opportunity for employees and business partners. Company Information. (News Release, Dec. 18). [Online]. Available: www.texaco.com/compinfo/pr12_18.html [1997, Jan. 13].

U. S. Department of Energy. (1996). 360-degree performance appraisal survey results: 1995/1996 Comparison. Las Vegas, NV: Nevada Operations Office.

Weaver, W. T. (1996, November). Linking performance reviews to productivity and quality. HR Magazine, 96(6).

Wood, R., & Marshall V. (1993). Performance appraisal: Practice, problems and issues. Paris, France: Organisation for Economic Co-operation and Develepoment (OECD).

Yin, R. K. (1994). Case study research: Design and methods (2nd. ed.). Thousand Oaks, CA: Sage Publications.

APPENDIX - A

GLOSSARY OF TERMS

360-Degree Feedback Appraisal System - Also known as the full-circle appraisal, multirater assessment, and group performance appraisal. It is an approach that gathers behavioral and performance observations from different external and internal sources for the purpose of evaluating an individual's performance.

Internal Customer (Sources of Evaluation) - Customers, such as individuals or organizational elements, that are part of the same larger organization from which the goods or services are received. For example an accounting sections may provide accounting services for other sections within their company (internal customer) as well as for other companies (external customers).

External Customer - Customers, such as individuals or organizational elements, that are not part of the same larger organization from which the goods or services are provided. (See Internal Customer above.)

Matrix Support Organization - Also referred to as the grid, project, or product management structure. An organizational structure that combines the functional and product forms of departmentation in the same organizational structure. For example a matrix organization could have an engineering department, that tends to be permanent and has a functional managers in charge of the engineering functions. At the same time the engineers in the department are assigned on a temporary basis to one or more product units or projects.

APPENDIX - B

SURVEY

360-DERGEE FEEDBACK PERFORMANCE SYSTEM
(EFFECTIVENESS AND FAIRNESS)

Sample Letter

TO: [Insert Name and Address]   Date:

      This survey is being conducted by the undersigned, a Walden University student, as part of a Ph.D. dissertation research project.. [Mr./Ms. Insert name, title of appropriate official] has granted the permission to survey members of your organization.

      I assure you that your responses to this survey will be kept confidential. No supervisors, employer or others members of your organization will be allowed to read your completed questionnaires. Neither Walden University nor I will allow the release of any data that could identify you by name or position. Only totaled survey statistical and demographic results will be published.

      If you have any questions, recommendations, or concerns related to this survey please feel free to contact me at the address or number listed below.

      Thank you for taking the time to complete this survey - your response is both needed and important!

Sincerely,

William B. Morgan
PO Box 1405
Jackson, NJ 08527
Phone/Fax: 1(908) 370-3430
Email: wmorgan@waldenu.edu

SURVEY

360-DEGREE FEEDBACK PERFORMANCE SYSTEM
(EFFECTIVENESS AND FAIRNESS)

_____

Date

_____

Survey Number

Consent Statement

You are invited to help in a research study of *The Perceived Effectiveness and Fairness of the 360 Degree Feedback Performance System* by answering questions and making comments. You were selected as a possible participant because your job performance has been rated/evaluated (or you have rated/evaluated the job performance of others) using the 360 degree feedback system.

This study is being conducted by: WALDEN UNIVERSITY, 155 5th Avenue South, Minneapolis, MN 55401.

Procedure: If you agree to take part in this study, we would ask you to take approximately thirty (30) minutes to do the following things:

1. Read the "Sample Instructions" on the next page.
2. Answer the six questions located on the "Demographic Page."
3. Answer the "360 Degree Feedback Performance Program Survey" questions.
4. Mail the completed survey in the attached self addressed, stamped envelope.

Risk and Benefits of Being in the Study:

Risk - Because of the procedures used for this survey to maintain confidentiality, no risk to participants are known or anticipated.

Benefits - Potential benefit of this study include: (a) permitting employees to express their perceptions of the fairness and effectiveness of the 360 degree performance system, (b) permitting management to assess employee perceptions of the 360 system.

Compensation: None.

Confidentiality: The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be kept in a locked file. Only the researchers will have access to the records.

Voluntary Nature of the Study: Your decision whether or not to participate will not affect your current or future relationship with the University or your employer. If you decide to participate, you are free to withdraw at any time without affecting those relationships.

Contact and Questions: The researchers conducting this study are William B. Morgan, Ph.D. Candidate, and Dr. Joseph Barbeau, Advisor, telephone (908) 370-3430.

Thank You For Participating!

Instructions

1. In the survey you are about to fill out, the five point rating scale is used. The questions ask you to MARK one of several numbers that appear next to an item. You should MARK the number that best describes your opinion using the following scale:

> 1 = Strongly Disagree
> 2 = Disagree
> 3 = Neutral
> 4 = Agree
> 5 = Strongly Agree.

For example:

> a. If you *Strongly Agree* that the 360 degree feedback evaluation is absolutely a FAIR system to evaluate YOUR job performance, then you would mark as follows:
> [1] [2] [3] [4] [5̸]

> b. If you *Strongly Disagree* that the 360 degree feedback evaluation is a FAIR system to evaluate YOUR job performance, then you would mark as follows:
> [1̸] [2] [3] [4] [5]

2. In marking your answers, please remember the following points:

> a. Answer ALL items, do not omit any.
> b. Mark only ONE number per question.

3. Your feedback is important, so you should consider each item carefully.

4. After completing all of the survey, please return it in the attached self addressed envelope.

> If the envelope has been misplaced, please return the survey to:
> William B. Morgan
> Survey Administrator
> P.O. BOX 1405
> Jackson, NJ 08527

Or, contact Mr. Morgan directly at Telephone/FAX: (908) 370-3430; Email: wmorgan@waldenu.edu and another envelope will be delivered to you.
Thank you!

Demographic Page

For statistical purposes only, please record: DATE: _____

1. Your age: Years

         [ ] Less than 25   [ ] 41 - 45
         [ ] 26 - 30   [ ] 46 - 50
         [ ] 31 - 35   [ ] 51 - 55
         [ ] 36 - 40   [ ] Over 55

2. Your gender: [ ] Female   [ ] Male

3. Highest level of education attained:
   (Please choose only one.)
   [ ] High School/GED
   [ ] College, no degree
   [ ] Associate Degree (A.A., A.S)
   [ ] Bachelor's Degree (B.A., B.S)
   [ ] Advanced Degree (M.A., M.S., Ph.D., M.D., etc.)

4. Your Current Job Category/Classification:
   (Please choose only one.)
   [ ] Executive
   [ ] Supervisor/Manager
   [ ] Staff (Non-Supervisor)
   [ ] Employee (Non-Supervisor)
   [ ] Other, specify: _____

5. How many times have YOU been evaluated using the 360 Degree Feedback Performance Appraisal System?

         [ ] 1 - 2   [ ] 5 -6
         [ ] 3 - 4   [ ] More than 6 times

6. How many times have you evaluated OTHERS using the 360 Degree Feedback Performance Appraisal System.? (Provided input another workers 360 Appraisal as a supervisor, peer, or customer.)

         [ ] 1 - 2   [ ] 7 - 8
         [ ] 3 - 4   [ ] 9 - 10
         [ ] 5 - 6   [ ] More than 10 evaluations

360 DEGREE FEEDBACK PERFORMANCE PROGRAM SURVEY

The 360 Degree Feedback Performance Program is also know as the *Full-circle appraisal, multirater assessment,* and *group performance appraisal.* This appraisal system represents a multisource appraisal from a full circle of people with whom you interact. Input for the final appraisal usually includes supervisors, peers, colleagues, subordinates, and customers. Please answer the following questions based on *your* perception of the effectiveness and fairness of this appraisal system.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. The 360 degree performance evaluation system is VERY fair for measuring my job performance.	[1]	[2]	[3]	[4]	[5]
2. Compared to the one rater (supervisor) rating systems, the 360 degree system is MORE fair.	[1]	[2]	[3]	[4]	[5]
3. The 360 degree system is a VERY effective way to measure my work contribution.	[1]	[2]	[3]	[4]	[5]
4. I have used the feedback I received from the 360 degree system to improve my job performance.	[1]	[2]	[3]	[4]	[5]
5. The "Customer" feedback I received from the 360 degree system is the MOST important information.	[1]	[2]	[3]	[4]	[5]
6. The "Peer" feedback I received from the 360 degree system is the MOST important information.	[1]	[2]	[3]	[4]	[5]
7. The "Supervisor" feedback I received from the 360 degree system is the MOST important information.	[1]	[2]	[3]	[4]	[5]
8. Since my organization has started using the 360 degree feedback system, overall organizational performance has improved.	[1]	[2]	[3]	[4]	[5]
9. Improvements in overall performance within my organization are directly linked to our use of the 360 degree feedback system.	[1]	[2]	[3]	[4]	[5]

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
10. Since we started using the 360 degree feedback system, employee morale has improved.	[1]	[2]	[3]	[4]	[5]
11. I would prefer we STOPPED using the 360 system.	[1]	[2]	[3]	[4]	[5]
12. Those who provide me with feedback on the 360 really want me to succeed.	[1]	[2]	[3]	[4]	[5]
13. The feedback I have received from the 360 system has resulted in improving my performance.	[1]	[2]	[3]	[4]	[5]
14. The questions used in the 360 accurately reflect what is important for me to be effective in my job.	[1]	[2]	[3]	[4]	[5]
15. Using the 360 removes the "politics" from the rating system.	[1]	[2]	[3]	[4]	[5]
16. The questions used in the 360 reflect what management thinks is important.	[1]	[2]	[3]	[4]	[5]
17. The 360 system is a way for those who DON'T like me to get even.	[1]	[2]	[3]	[4]	[5]
18. I like the idea of my boss knowing how my peers and customers rate me.	[1]	[2]	[3]	[4]	[5]
19. Regardless of how well others rate me, my overall rating will be what the boss wants it to be.	[1]	[2]	[3]	[4]	[5]
20. The 360 system has given me a more honest opinion about my performance.	[1]	[2]	[3]	[4]	[5]

Additional comments:

THANK YOU FOR YOUR COOPERATION!

Please return the survey in the attached self addressed, stamped envelope. If this envelope has been misplaced, please return the survey to:

William B. Morgan
Survey Administrator
P O. BOX 1405
Jackson, NJ 08527

Or, contact Mr. Morgan directly at Telephone/FAX: (908) 370-3430; Email: wmorgan@waldenu.edu and another envelope will be delivered to you.

APPENDIX - C

SURVEY COMMENTS

360 DERGEE FEEDBACK PERFORMANCE SYSTEM
(EFFECTIVENESS AND FAIRNESS)

SURVEY COMMENTS

Below are the written comments received from those responding to the "360 Degree Feedback Performance System" survey. The comments are separated by organization.

Norwest Mortgage, Inc.

I do not support using the 360° for performance appraisal. I do support it as a method of development.

Comment to questions 5,6,7, the most "import" data is not just from one group. I need to look at what is being asked and see who is (in) the best position to give feedback.

U.S. Army Corps of Engineers

Question in the 360 did not reflect my job or performance. It was a standardized list used to test the system.

Questions used were from a different organization with different goals, values, mission. Reason for response to 14 & 16. Good luck on you Ph.D research.

We haven't used it long enough to have an impact. One small sample group w/o large participation from all sectors.

Our current questions do not reflect individual job performance questions – that needs to be fixed.

Used 360 one time - Questions not really tailored to MVR - interesting to see what feedback results were - interesting to see variances between rater, senior rater and peers. Interesting conclusion in mine was senior rater gave lowest, peers highest, rater very close to peers. Conclusions I reached is rater knows what I am doing. Maybe if questions were different my conclusions of value would have been different.

U.S. Army Management Engineering College

I strongly believe in the 360^0 Perf. Appr. System! However, thru past experience, depending on who you select as your raters makes a huge difference in your rating-obviously. I've found women rating women cannot always be objective. Therefore, the next time I'm rated, it will be from only men who know about the work I do. Men are much more fair in this instance.

The 360 is too behavioral. I worked alone and others did not know what I was doing. The teaming aspect was missing from my effort. The key to the 360 is that management along with the rating power also delegate award Moines to the teams. My experience was that when the teams decided on monetary awards the key mgt. gripe was overcome.

I believe the greatest benefit of the 360 as employed by AMEC is the elimination of the very contentious supervisor-employee appraisal process. Even favorable appraisals under that process were often provided in an unpleasant interaction. AMEC totally eliminated supervisor's control of the process. I believe we underused the outputs, but it was still a very beneficial change.

The 360^0 survey does not ask the right questions.

The 360 system was one of several initiatives undertaken concurrently; their individual contributions to performance improvement is not measurable. There have been a few instances of subordinate ratings being lower than appropriate because of not seeing the big picture.

For the 360^0 to be effective - such things as cash awards, seniority, etc. must be removed from the system.

Our system was designed and implemented by the employees as a part of becoming "self-directed" teams. While this transition is not complete (and may never be), 360 was a very positive contribution to this process.

I have changed organizations since I was at AMEC. My new organization uses the old supervisor (one rater) system and I would much prefer to use the 360 (as both a supervisor and employee).

U.S. Department of Commerce

Comments are not considered so why comment.

Moral is at an all time low. The worst in the last 10 years.

The questions are critical, will never be perfect.

Only some (a very few) people use the 360^0 system <u>professionally.</u> The rest us all "N/A's" or BS. It really means nothing to most of us because <u>no one</u> on the team knows exactly what the others are doing (just general information and not specifics). Others use it to get back at people.

The questions used in the 360 don't really pertain to the quality of work I am doing.

Under our system I am anonymously rated on model leadership and peer behaviors and customer service attributes. These 3 anonymously ratings (from peers, subordinates and customers) count for 50% of my annual appraisal. My boss provides the other 50%.

Q1- Everyone is afraid to mark people low therefore we see grade/rating cluster in the upper 3rd on most of the rating elements. Also based on how much a person is liked. Q3- If comment for #1 were not the case I would rate this as a #4. Q5.- Not applicable used only for peer review. Q7-Unable to distinguish supervisor from my teammates. Q11- Its better than other options. Q12- Some...others just gripe! Q14- We revised them 4 mos. ago so they are better. Q15- Its just a different kind of politics! Q16- Revised by team members...not management. Q17- It can be, and has been upon occasion. Q19- Peer review only accounts for 25% of my rating. CMTS: I think you get a sense of what I am saying... If people would be more honest in the way they rate others while not "dumping" on those who they are not close too it would be better but this is a "cultural" issue not structure... <u>The 360^0 process and structure are great</u>...Its how people implement the process that is the problem!!

Q3- It <u>can</u> be depending on the willingness of participants to take the exercise seriously. Q5,6-They are equally important. My peers are my customers to a great extent since we work together and they depend on me. Q8-No an effective catalyst or large enough motivational factor to change behavior patterns. Perhaps unique to this setting. Any system that does not have the right input can't produce the right outcome. Q10-No direct link. However, I do believe employees have a more accurate sense of their performance level. Q11- Our system has primarily evaluated team behavioral aspects which was appropriate for the firs couple years after reorganizing. We need to revise the system to assess performance outcomes. Q12-Mixed motivations. Q13-I have made some conscious efforts when feedback has been specific. Q14-Somewhat – a "component" of effectiveness but does not encompass the total picture. Q15-Would any system be perfect? Q16-For our system, it reflects team perf. mgmt. theory and application. CMTS: It would have been interesting if you had compiled data on <u>how</u> folks are using 360^0 as it probably would impact or have some correlation as to how employees feel about the system. For example: Is feedback given orally face to face? Individually? In a group setting? Is feedback given anonymously? In paper format? Electronically? Is 360^0 used

for feedback purposes only? or, is it tied to a performance appraisal and given weight? Is it tied to the pay out of a performance award? Etc.

U.S. Department of Energy - Golden Field Office

Getting my ratings from my customers is far better than having the manager of the office play politics with the appraisals.

My closing comment, but hopefully your first impression: Although I can only assume the premise (thesis statement) for this effort, I truly believe this effort's completion really serves as the basis for what questions you should be asking (what factors also need to be considered.). Hopefully you'll consider doing a second more deeply meaningful survey to tackle these. And yes, the deeper and more complicated you get, the lesser the response rate to the survey.

This survey seems a little too simplistic, superficial and poorly designed for a Ph.D. candidate. There are courses on how to do surveys and I question if this even approaches a textbook example. The point is – How valid are the results of a poorly conceived survey? But in this case, there has probably been too much effort made to change it and you'll proceed on. Right?!

Generally speaking, I think the 360 system is a better system than the previous performance rating system. Based upon what I've seen however, I don't think it is working as effectively as it could. I think more care should be given to the relevance of the rater lists developed by staff. The results of the 360 for our office as a whole seem to indicate very little differentiation in performance of various staff – I don't believe this to be accurate. We need to have better understanding of why this is.

360 lacks specific detailed performance measurer – actual products.

MARS does not reward best performers. Prior system tied cash to outstandings.

The flat supervisor structures leaves supervisors w/20-25 employees. Even those actively involved in day-to-day activities can not know what is going on. They will tend to rate highly those who are better politicians. The 360 can give good feedback to supervisors even though persons are not to ask persons to rate who might be negative for rating (This would provide a truer picture but only if supervisor does not see it.). Only 8 of 15 did the rating; fewer made comments.

I disagree with the use of the word "performance" related to the 360. The 360 review was not advertised to be used as a job performance measure. Now it seems to be an easy way to eliminate any responsibility of management to measure performance of individuals.

The 360 degree feedback performance system, though not perfect, is a definite improvement from the earlier traditional performance evaluation system (mostly an employee/supervisor interaction), because it is based on both internal and external assessments by individuals and/or organizations familiar with the work of those being evaluated.

I got positive feedback on the value of the survey from "outside" customers. They appreciate being asked and ease of input to the format. Windows application would help improve ease of input though.

We have been told the 360 is not a performance evaluation tool for mgmt. – it is a behavioral tool to be used by the employee to improve overall performance. If evaluators are chosen by the employee and do not reflect a good crosscut of peers and customers, the eval. will be of little use to the employee. It depends on employee motives – receive a high score or receive valuable feedback for improvement. Unfortunately, the majority pick the former so little improvement is noted.

360 is only a tool. When you have a system of heavy seniority and tenure; where mediocrity and incompetence cannot be rooted out, a 360 is an aspirin for a corpse.

The 360 has its merits in rating an employee however if a customer or another employee give a rating that has less than 6 month working with an individual, this is not a fair rating. Sometime it happens this way. What can be done?

Questions and elements of the 360 should be more precisely written to eliminate the rater from guessing as to how to answer. Many negative answers are given because of unclear questions.

Due to budget cuts program funding has been eliminated and a program manager is going forward with work activities authorized by director (supervisor). What option(s) can be given for a performance evaluation? The 360 would not apply. There is very little contact with customers, state, etc. (To be addressed).

This year the 360 was used entirely to rate performance. Most of my evaluators (customers) only know a portion of my performance. A better process would be for the evaluation to be based 50% on 360 input, and 50% on management input.

System is good overall as a "system." However, there is no linkage of awards to ratings, the elements don't always reflect what one does, and management priorities/involvement is completely missing. Keep using system and focus on improvements.

The 360 DFPS is a good concept. However, the implementation of the system into our organization require improvement to increase fairness and accuracy of ones performance. my experience has been people use the system to get even with people they don't like. Also, if they are in a bad mood, they will take their frustration out on you, and that can affect your rating.

Although the 360 system provides a broader base for the evaluation of performance, the de-linking of the rating from awards ($) has provided a general disincentive to improve. People are motivated by $ and ratings and since there are now more politics involved in awards, morale has suffered.

Would like to see a study of appraisals done the old way as compared to the 360. If employees have improved dramatically in ratings using 360, then 360 is over stating employee performance. Try letting supervisors pick raters for their employees.

The 360 system looks at attitudes and behaviors, not performance against critical objectives. There may be a link between "behavior" and "performance" but it isn't always a strong one.

I think the 360 provides some good feedback, but tends to measure "attitude" and "approach" to the job as opposed to actual performance or results.

One is at the mercy of how a person rates – some rate high, some low. In comparison to others you could look poor but be better. Value of 360 is the written comments.

We have not tied the 360 performance rating with an individual development plan to enhance performance, get more training in weak areas, or to improve work effort. We haven't closed the loop.

We (the organization) don't do a good job of explaining what the different ranking mean.

My supervisor never took my evaluation seriously, and since I was a supervisor and took the evaluation of my employees very seriously I thought the system was bad. However, although my evaluation is better, I don't believe my ex-employees get the amount and quality of feedback as they did from me over my 18 years as a supervisor! This system protects employees from poor supervisors, but does not replace good supervisory feedback!!!

My direct supervisor has harbored a strong dislike for me since I begin working for him. His comments in my 360 were extremely negative – but countered by the wonderful statements made by my customers. The contradiction in the comments seem to indicate that a supervisor with a grudge can continue to treat someone negatively in spite of professional excellence in an individual. Also, workload and initiative aren't necessarily reflected. Someone who doesn't have a large workload can excel. There is no process to review/provide feedback for improving performance. Despite my excellent peer evaluation – my supervisor continues to ignore my abilities and has been insulting numerous times. I think that the supervisor should be rated by all his employees and staff rated by all staff to get better results. Also – the process doesn't take into account folks standards – a 5 given by one person is equivalent to a 3 by another – it's the professionalism and judgment by individuals who care about outcomes vs. "amiability."

This 360 rating system sounds great in theory. But DOE has structured the system such that it is a popularity contest. The comments are contradictory and oftentimes whimsical. I have seen nothing in the 360 system that actually measures performance.

I feel the number of raters (currently 12) should be reduced to no more than 6. I also feel the supervisor should not be allowed to decide who the raters should be or increase the number of raters the individual has chosen.

If you are allowed to use the people you choose then the system would work well. But when you are told who to use it's another story.

Staff needs to be reminded of the purpose of using the 360 performance. There is such a potential for growth that is not being realized. Mgt. needs to spend more time discussing results with employees. Employees need to stop trying to figure out who said what about them. This all diminishes the process.

The 360 can be beneficial if the system is optimized – i.e., the right factors are evaluated for a given position – generic factors or behavior do not give a total picture; -- without those, the process is interesting, but not terribly meaningful.

I don't believe that the 360 system is any better or worse. I think most people send them out to only those that will give them a favorable rating. I believe it takes additional time and work with little or no impact.

The questions used on our 360 evaluation are hard for our outside customers to answer because they do not address the issues or areas that are important to them. The outside customers have no knowledge of some of the areas being evaluated. I have found that my supervisor gives me more useful feedback than the survey does.

360 system is a good tool but should be used in conjunction with a mgt. rating system. Many raters selected by employees on 360 system don't really understand the total job and what needs to be considered overall in assessing performance.

Although I feel that the 360 reviews I have received reflect how I think I am performing; overall, I think the system is flawed. In the last evaluation, the average office scores reflect that everyone is working at an exceptional level (4 out of 5). This can't be accurate.

The problem with the 360 system is employees choose raters they know will give them positive feedback, so it becomes a popularity contest. I'd prefer a two-tiered approach where the "old" supervisor review system was coupled with the 360 approach. The 360 would be incorporated into the supervisor's review. Another problem is the questions asked are too generic, not specific to the job.

Use the 360 system in conjunction with a supervisor rating – present system jumble them all up.

Bosses look at 360 ratings, which isn't fair from the aspect that some select only their "cronies" as raters. More requirements for number and variety of raters should be made so all are rated by a larger field.

The 360 should be used in conjunction with a regular performance rating system – it is great for getting feedback from customers and peers – but only tells part of the story. It is also easy to "game" the system when used alone.

There is no standard of comparisons between raters. One rater has a different standard than the next. Management pays more attention to the numbers than the comments. Questions do not relate to all positions.

The way the 360 is implemented at my place of employment is more of a personality test – i.e. who likes me, who doesn't . It does not truly measure my job performance or work relationships. I believe a 360 can be valuable if implemented correctly. I don't believe my employer is capable of using this type of system to benefit the office – only certain individuals.

I believe 360 systems can be effective in providing "objective" performance assessments. However, our system has not been implemented effectively. Responses appear to have been based on perceptions, not facts. Also, the high average scores indicate that some offices used the system differently than others.

I believe that the 360 performance rating system can be an excellent tool for evaluating and improving performance if the system is used correctly. The 360 appraisals should be

given to the employee and not used as a supervisory appraisal. The employee can then use the 360 as a tool to improve their performance without having to worry that the comments on the 360 will be used against them.

The 360 has been used in our organization as a supervisory appraisal. The problem with this is that the 360 rating system only ask questions about an employee's personality, not job performance. The questions in the 360 don't ask about or measure performance. Also, employees have been forced to put evaluators who don't interact with the employees as raters. In addition, people were contacted by supervisors and told what to say in an employee's 360. Also, many people did not use the evaluations for providing constructive criticism, they were used to express their personal opinions of the employee. As a result, some employees only put their friends on their list of evaluators, not people they actually interacted and worked with. Since the 360 can be manipulated and controlled, we might as well go back to the single supervisory appraisal system.

On the 360 evaluation, a "0 to 10" scale would be better than the "0 to 5" scale. A "0 to 10" scale would allow the rater a little more choice and flexibility.

Q5,6,7- You can't tell whose feedback is whose! Since all ratings are confidential, how can we answer these 3 questions?

Q10- Poor design, since in this case employee morale has been affected by other significant external factors.

Q11- I believe the 360 system should be supplemented by a supervisor's rating system, or should be used as input in supervisory rating. It is too subjective and raters too diverse to be of much benefit.

Q14- First round-no; most recent round-yes.

Q15- Obvious built in "good ol' boy" system if you pick your raters, how many people knowingly pick people they have had difficulties/problems with. Yet some people had some or all of their raters picked for them.

Q19-I am not certain what is meant by this statement so I did not respond.

<u>Other</u>

Overall – I really like 360 feedback but many "bugs" have to be worked out before it is a fair and accurate reflection of job performance.

None of the 360 feedback systems in which I've been involved have been used to determine job performance (i.e. in accomplishing responsibilities of the job), but in evaluating behaviors, skills or work styles. 360 was not used in annual performance appraisal.

Q20-If raters have a true understanding of my job and my performance.

CURRICULUM VITAE

WILLIAM B. MORGAN, CFP, Ph.D.
PO Box 1405, Jackson, NJ 08527
Phone/FAX: (732) 370-3430
Email: wmorgan@waldenu.edu

EDUCATION: Ph.D. Walden University
M. S. Strayer College
B. S.   University of Maryland

DISSERTATION   "Validation of the 360-Degree Appraisal Program as an Effective
TITLE: and Fair Measure of Individual Performance"

PROFESSIONAL EXAMINATIONS AND DESIGNATIONS:
Certified Financial Planner (CFP)
Securities Exams and Licenses: Principal - Series 24, General Securities - Series 7,
Series 63, Series 65

PROFESSIONAL EXPERIENCE:
President, Logistic Management Sciences, Inc., Jackson, NJ 1998-Present.
Developer of quality logistics, training, management, fielding, and consulting services.

Independent Contractor & Consultant to COBRO Corporation, Eatontown, NJ 1995 - Present. Researched and developed three "Performance Specifications and Standards" courses, the "Cross-Functional Training Course" and the "Team Leader Management Course." Primary consultant and course developer for management, organizational and logistical courses.

Administrative Vice President and Registered Principal, Integrated Financial Planning Services (Broker/Dealer) 1994 - Present. Responsible for insuring Broker/Dealer regulatory compliance with both federal and state securities registration and licensing agencies throughout the United States.

Owner, Morgan Financial Services 1991 - Present. Provider of management, tax, and financial consulting; training course development; and, independent business and management consulting.

TEACHING EXPERIENCE:
- Adjunct Professor, Community College of Philadelphia (1998-Present).
- Senior Instructor for Management, Team and Leadership Training for COBRO Corp. (1995-Present).
- Business, Financial, and Tax Planning Lecturer (1985 - Present)
- US Government (Dept. of Army) Contracted Workshop Facilitator (1993-Present).
- Senior Instructor Exec. Tax Service and H&R Block Tax Courses (1989 - 1995)
- Instructor/Lecturer, U.S. Army Signal Schools Assoc. Master Instructor (5 yrs); Past Instructor at Central Texas College, and City College of Chicago.

IMAGE EVALUATION
TEST TARGET (QA-3)

150mm

6"
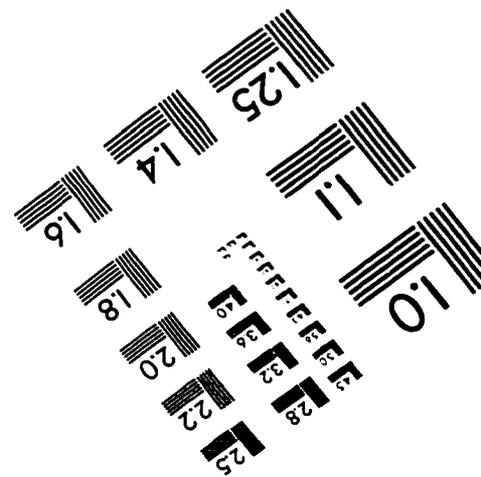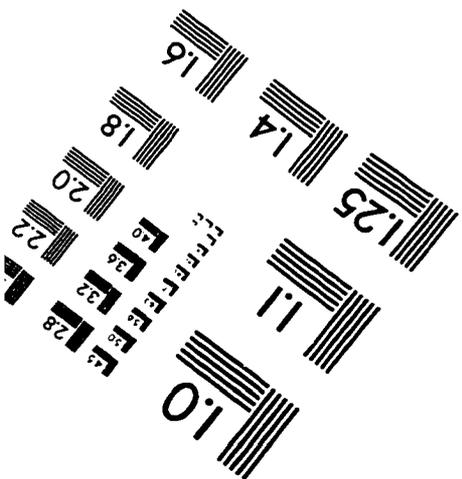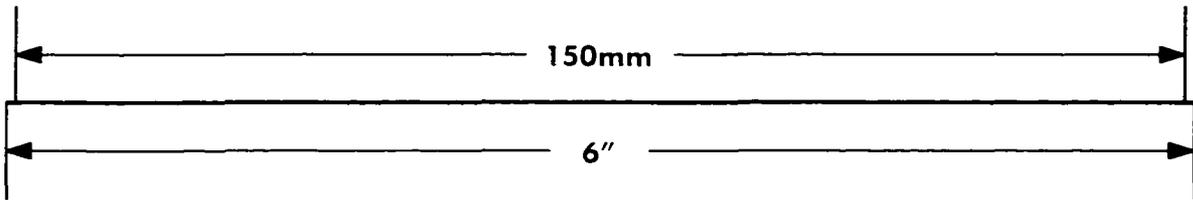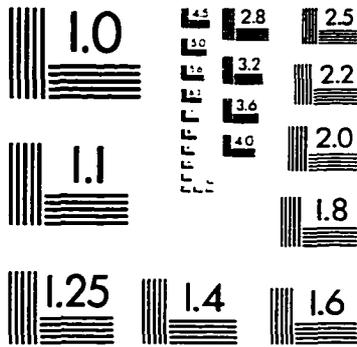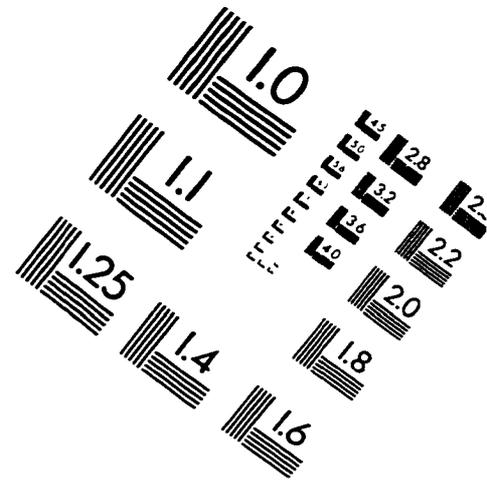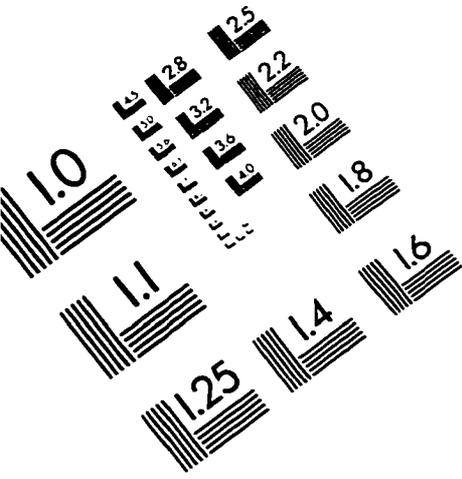
APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved