



Western Michigan University
ScholarWorks at WMU

Dissertations

Graduate College

4-1994

The Effects of Utilitarianism on Performance Appraisal and Selected Output Measures in a Weberian Bureaucracy

Donald Wayne Wright
Western Michigan University

Follow this and additional works at: <https://scholarworks.wmich.edu/dissertations>



Part of the Social Policy Commons, and the Social Welfare Commons

Recommended Citation

Wright, Donald Wayne, "The Effects of Utilitarianism on Performance Appraisal and Selected Output Measures in a Weberian Bureaucracy" (1994). *Dissertations*. 1848.

<https://scholarworks.wmich.edu/dissertations/1848>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



THE EFFECTS OF UTILITARIANISM ON PERFORMANCE APPRAISAL AND
SELECTED OUTPUT MEASURES IN A WEBERIAN BUREAUCRACY

by

Donald Wayne Wright

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Public Administration
School of Public Affairs and Administration

Western Michigan University
Kalamazoo, Michigan
April 1994

THE EFFECTS OF UTILITARIANISM ON PERFORMANCE APPRAISAL AND
SELECTED OUTPUT MEASURES IN A WEBERIAN BUREAUCRACY

Donald Wayne Wright, D.P.A.

Western Michigan University, 1994

The research explored whether or not a relationship existed between the performance appraisal criteria used by the Social Security Administration and actual performance between 1984 and 1990. The issue was whether or not either of the diverse sets of performance appraisal criteria was more effective in improving performance than the other and whether or not base pay should be tied to the results of the performance appraisal system.

The methodology for this study was both quantitative and qualitative. The qualitative technique of content analysis was done on the sets of performance appraisal criteria to assess the extent to which each set had the research attributes of positive, negative, quantitative and qualitative referents. The quantitative analysis involved using regression analysis and a t -test to determine whether any significant improvement occurred within either set, including both sets together or between the sets. Processing times for Retirement, Survivor's and Disability claims were the operational data analyzed for improvement.

The Chi-square analysis of the research attributes indicated a significant difference between the two sets of criteria. It was found that the second set of criteria was more negative and quantitative than the first set. The quantitative analysis of the eight possible relationships between actual performance and criteria set revealed only one relationship improved while seven did not.

The conclusion was that neither set was associated with improved performance and that base pay should not be tied to the system until better measurement criteria can be emplaced. The conclusions also called for further research on the relationship between the Civil Service Reform Act of 1978 and performance standards government wide in terms of the research attributes used in this study.

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.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9426892

**The effects of utilitarianism on performance appraisal and
selected output measures in a Weberian bureaucracy**

Wright, Donald Wayne, D.P.A.

Western Michigan University, 1994

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

ACKNOWLEDGMENTS

The accomplishment of this degree must be shared by many. My wife Debra deserves the lion's share of the credit. The importance of her emotional, functional and faithful support can not be overstated. I would also like to thank my parents, Donald E. Wright, and Josephine and Harry Swank. I would also like to acknowledge Martin Kolm who frequently served as my devil's advocate and referee in the completion of this work. Each of them contributed by believing in my ability and by allowing me the time necessary to complete this dissertation. The patience, love and understanding of each was significant.

I would also like to thank Dr. Ralph C. Chandler, Dr. Kathleen P. Reding, and Dr. William P. Browne. Dr. Chandler's guidance on the theoretical aspect of this research was very important to understanding the true nature of the problem.

I would also like to thank Darlene Bell and Shelly Starkloff of the Social Security Administration. Without their assistance, the completion of this research would not have been possible.

Donald Wayne Wright

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
PREFACE	1
CHAPTER	
I. INTRODUCTION	2
Legal Aspects	2
Operational Aspects	3
Compassion	3
The Legal Basis	4
Operational Basis	11
Historical Perspective	12
Operational Changes	13
Compassion Aspects	26
II. REVIEW OF THE LITERATURE	32
Overview	32
Weberian Literature	32
Utilitarian Democracy	35
Victor Thompson	38
General Performance Appraisal Literature	39
Objective Criteria Literature	44
Social Security Literature	51
III. STATEMENT OF SCOPE	54
Overview	54
Locus of Research	55

Table of Contents—continued

CHAPTER	
Focus of Research	56
Summary of Scope	58
IV. ISSUE STATEMENT	59
Overview	59
Operational Issues	59
Purposive Issues	60
V. METHODOLOGY	63
Overview	63
General Description	63
Operational Definitions	65
Operational Definition - RSI Processing Time .	65
Operational Definition - DI Processing Time ..	65
Operational Definition - 1981 Performance Criteria	66
Operational Definition - 1986 Performance Criteria	67
Axiomatic Assumptions	67
Research and Null Hypotheses	68
Research Techniques	71
Content Analysis	71
Regression Analysis	73
t-test	74
Actions and Instructions	75
General	75

Table of Contents--continued

CHAPTER

Qualitative Data Gathering	75
Quantitative Data Gathering	76
Data Analysis - General	77
Qualitative Analysis	78
Quantitative Analysis	80
Limitations - General	81
SSI Processing Time Limitation	81
Payment Accuracy Limitation	81
Unavailability of Data Prior to 1984	82
Unavailability of Performance Ratings	82
VI. FINDINGS	84
Content Analysis	84
Quantitative Analysis	89
Statistical Independence Testing	90
Hypothesis Testing	91
VII. CONCLUSIONS AND RECOMMENDATIONS	98
Overview	98
Conclusions	98
Quantitative Analysis Conclusions	98
Qualitative Analysis Conclusions	102
Recommendations	105
VIII. SUMMARY	115

Table of Contents--continued

APPENDICES

A. Exclusion From Performance Appraisal	119
B. Summary Sheet for Performance Appraisal	122
C. Example of Completed Appraisal	125
D. Executive Organization Chart	128
E. Research Attributes	130
F. Position Description Groupings for Content Analysis	149
G. Graphic Representation of Content Analysis Results	152
H. Comparison for Claims Taking Research Attributes to Overall Attributes - By Set	155
I. Raw Data for Monthly Processing Times and Claims Counts	158
J. Trend Lines for Raw Data by Month	161
K. Chi-Square Conversion and Calculation	166
BIBLIOGRAPHY	173

PREFACE

"The federal civil service is the triumph of technique over purpose" - Wallace Sayre. In the many decades since this profound statement, this insidious "triumph" has marched from the arena of selecting staff for federal agencies to becoming the operant theory and driving force within federal agencies. No more egregious example of this tragedy camouflaged as "triumph" exists than the performance appraisal system of the Social Security Administration.

Machiavellian malignancy, guised as admirable objectivity, drives the system forward with little or no regard to the effects of the system either inside or outside the agency. This system is living testament to Weber's feared "iron cage of rules."

With Social Security ready to tie base pay to performance, no more vital challenge exists for the Social Security Administration than to remove the performance appraisal system from Weber's "iron cage of rules" and re-establish the dominance of purpose over technique.

CHAPTER I

INTRODUCTION

Legal Aspects

This introduction serves three fundamental purposes. First it will provide the legislative basis for two basic legal aspects of the current workings of the Social Security Administration. It is the interaction of these two aspects which provides the substantive basis for the analysis conducted in this research.

The first legislative aspect deals with establishing the legitimacy for taking and adjudicating retirement, survivor's and disability claims. The length of time which these claims require to take and adjudicate becomes the first legal basis for the research.

The second legal basis to be established is the basic requirement for mandated performance appraisal in the federal government in general. This will be accomplished by discussing the authorizing legislation and the roles of those agencies charged with the development and then the implementation of the legislation. It will also include very basic analysis of the law itself. This will establish the second legal basis for the research.

Operational Aspects

Secondly, the introduction will summarize the operational aspects of performance appraisal in the Social Security Administration (SSA). The actual operation and outcomes of the system will be discussed to provide a working knowledge of the system. This portion of the introduction will provide the operational context for the study.

Compassion

Finally, the issue of compassion in administration will be raised as it relates to the precepts contained in Without Sympathy or Enthusiasm by Victor Thompson. Thompson argues that compassion is unnecessary and uncalled for in public administration. While this position has remained academically unrivaled to the present time, it has one operational challenge which serves as a progenitor for this study. In 1990, Gwendolyn King, the Commissioner of the Social Security Administration, made a direct call for compassionate service. This portion of the introduction provides the basic theoretical perspective necessary to address the concept of compassion in this research.

The introduction of these precepts will establish the relationship between the legislative intent, theoretical underpinnings and finally the operational aspects of the performance appraisal system. By doing so, it will clearly

bring to bear the importance of compassion in the present administration in both the legislative and operational realms as it relates to performance appraisal. Understanding these legal, theoretical and operational premises and concepts should provide a useful framework in which this research can be understood.

The Legal Basis

The historic signing of the Social Security Act in 1935 established a broad basis for the federal government to be active in the business of social insurance. A small, relatively simple act at inception, it has grown vigorously over the years to include many types of benefits and many services. Two of the distinct and varied types of social insurance responsibilities serve as an operational and analytical basis for this research. These two aspects are Retirement and Survivor's Insurance (RSI) claims and also Disability Insurance (DI) claims. Each of these types of benefits has evolved over time and a brief history will be provided here to ensure that the scope of each type of benefit is understood.

Retirement and survivor's benefits were the first type of benefit paid by the system in 1937. Disability benefits became part of the Act in 1957. The United States Code Annotated, Chapter 7, Subchapter 2 contains the basic legislative authorization for SSA to accept and adjudicate

old age, survivor's as well as disability claims.

At the present time, a wage earner must have earned forty quarters of coverage at some time throughout their working life to receive these benefits. This method of becoming insured applies to old-age (retirement) and survivor's benefits. There is an additional requirement for disability benefits, that twenty of the forty quarters of coverage must be earned immediately prior to the onset of the disability.

These benefits are paid to insured individuals as well as their spouse(s), children, adult disabled children, some parents, or grandchildren. It is the acceptance and then the adjudication of the combinations of these claims which are the analytical focus of this research.

Initially, both the wage earner and the spouse had to be age sixty-five or older to receive these benefits. In 1956, the age at which women could become entitled to retirement benefits was reduced to age sixty-two. A similar change was made for men in 1961. For those years covered in this research, the earliest age at which retirement benefits could be paid was sixty-two for both males and females.

Thus, no change occurred between 1981 and 1990 which alters the definitions of the retirement claims taken and adjudicated. This ensures that during the period of study, there was no change in claim definition which could affect

this research. This study has no measurement artifact.

Survivors claims became payable in 1940. Initially, this included only a widow of a deceased worker and that worker's child who was under age eighteen. Beginning in 1950, the divorced wife of a deceased worker could receive Survivor's benefits if she was caring for a child of the deceased wage earner. In 1965, the age limit for widow's benefits was lowered to age sixty-two. In 1968, disabled widow's became eligible for benefits as early as age fifty. Effective with 1973, non-disabled widows could receive survivor's benefits as early as age sixty. It should be noted that no changes in claim definition occurred in 1981 or later. Thus, the definition of a survivor's claims is consistent throughout the entire time frame of the present research.

Disability benefits became payable in 1957 to anyone who was disabled and over the age of fifty. Effective with 1960, disability benefits could be paid to a person of any age who met the disability insured status requirement. It should be noted that no substantial change took place after 1981 which would alter the analysis performed in this research. While there were provisional changes in benefit computation rates and offset provisions, the taking and adjudicating of disability claims was not significantly affected.

Having established the legal basis for the types of

claims which are used in this research, the legal basis for the second legal component of the research, performance appraisal, will now be provided.

Performance appraisals in the Social Security Administration are not a matter of choice. Like every other federal agency, the Administration is required by the Civil Service Reform Act of 1978 to complete performance appraisals for all employees on a fixed and regular basis which utilize objective criteria. The Carter Administration intended to make the bureaucracy accountable for its performance. Accountability was to be determined via performance appraisals.

Public Law 95-454 amends Section 203 Chapter 43 of Title 5 of the United States Code regarding performance appraisal. The law clearly includes the field office employees of the Social Security Administration under Section 4301.2 which states that an "employee" means "an individual employed in or under an agency." The Act then lists seven exceptions to the definition, none of which apply to service providing domestic field office personnel of the Social Security Administration. See Appendix A for the complete list of these exceptions.

Having mandated performance appraisals, the legislators then chose to define "unacceptable performance" rather than acceptable performance. This choice gave the legislation both negative and default characteristics. Chapter

43, subsection 4301.3 determines that unacceptable performance is "performance of an employee which fails to meet established performance standards in one or more critical elements of such employee's position." This negative definition becomes the focus of the system. It is vital to note both the negative and default aspects of the legislation since they are later operationalized.

In a study by Pavett (1983), non-negative or positive feedback was found to be very important as it relates to the individual's ability to respond to the results of a performance appraisal. By choosing to focus on the negative, Congress may have prevented effective implementation from the outset.

The Act then provides the guidelines for the establishment of performance appraisal systems. The Act states that the system must be periodic in nature and that the employees should be allowed to participate in establishing the standards by which they will be judged.

Lovrich (1981) researched the effects of employee participation on the performance appraisal process in the State of Washington and found the effects to be quite positive when very broad participation is allowed. Similar results are derived from a study by Robson and Nelson (1991) conducted in Utah and another study by Steel (1983). Goodson and McGee (1991) found that employee participation in goal setting, which is analogous to criteria establish-

ment, was a vital component to the effective use of performance appraisal.

SSA did not allow this broad input. Instead, federal employee union representatives were allowed to represent employees in the establishment process and employees themselves were not involved. Perhaps many later difficulties could have been averted by taking this most basic step.

Subsection 4302.a.3 then provides a list of actions which can be taken based on the rating given under the system. As per this subsection, the agencies are required to "use the results of performance appraisals as a basis for training, rewarding, reassigning, promoting, reducing in grade, retaining and removing employees." Thus, the legislation ties these outcomes to the use of the system. In essence, Congress has concerned itself with the origination and ultimate use of a system without allotting itself any responsibility in the development, or most critically, the application of it.

The balance of this subchapter is spent defining the negative actions which can result from unacceptable performance. This is reflected in the subchapter headings such as: "Suspension For 14 Day Or Less"; "Removal, Suspension For More Than 14 Days"; "Reduction In Grade Or Pay Or Furlough For 30 Days Or Less." It seems clear that the focus of this section of the Act is negative. A major argument of this inquiry is that this negative inception

has been virused into the operational stages of the Social Security Administration's Performance Appraisal Plan and thus severely limits the effectiveness of the system.

The Act then assigns jurisdiction for monitoring the development and use of the performance appraisal process to the Office of Personnel Management (OPM). This executive agency is charged with prescribing guidelines which agencies must use to establish performance standards, implementation time frames and the frequency with which reviews are conducted.

These guidelines constitute the technical assistance referred to in subsection 4304.a. of the Act. OPM is also to exercise oversight of the development of performance appraisal systems by agencies and "periodically report its findings to the Office [of the President] and to Congress" (5 USC 4304.2).

The legislation also authorizes OPM to initiate corrective action with an agency should the agency developed system not meet the statutory requirements of the legislative intent. Finally, the Act then allows OPM to "prescribe regulations to carry out the purpose of this subchapter" (5 USC 4304.b.1). This is the legal basis from which the operational aspects of SSA's current performance appraisals were derived.

In summary, the legislative focus of performance appraisal development was negative and default oriented.

Congress has authorized OPM to oversee the efforts of an agency to develop and implement a performance appraisal system which meets the legislative intent. The contention in this research is that the system developed by SSA exhibits both default and negative traits operationally.

Operational Aspects

This portion of the introduction serves a dual purpose. The primary intent is to provide historic information about SSA's development of a performance appraisal system as required by Congress. A broad overview of the system's development and significant changes will be given.

The second purpose of this section is to provide a description of the operation of the performance appraisal system from 1981 through 1990. This will include information about the actions taken by supervisors to complete the reviews, time frames for completion, communicative activities between supervisor and employee as well as the limitations of the system.

The historical perspective and operational description should provide an adequate framework for understanding the uses of these sets of performance appraisal criteria. It will also facilitate an understanding of the resulting limitations.

Historical Perspective

Stemming from a negative legislative genesis, the performance appraisal system of the Social Security Administration is essentially a search for negatives. Any errors found during performance appraisals serve as a basis upon which the Administration can effectively discriminate between employees in regards to the several types of personnel outcomes listed previously. Given the negative language of the mandating act, this use of the system is hardly surprising.

It is interesting that in 1879 the Pension Office of the United States Government evaluated and retained employees based on the number of errors made in one year (Ingle, 1982)! It was soon dropped as a means of retention. This error focused effort is the de facto operandi in SSA at present. Perhaps this government fascination with error is a hundred year cycle.

Social Security developed appraisal criteria and submitted them to OPM for approval in 1979. These standards were accepted and became the operational basis for the performance appraisal system. The standards developed in 1979 were actually implemented in 1981. These standards remained in place through fiscal 1985.

A major change took place in 1985. Under the Reagan administration, a new, more objective, set of criteria was

emplaced. The new standards became more closely aligned with quantification. This was borne out by the content analysis portion of this research.

The retrenched standards developed during the Reagan administration ignored the historical difficulty of quantifying work which requires thought and judgement (Dresang, 1992). These criteria were implemented in October, 1986.

Operational Changes

The single significant operational change which took place is still in effect today. At the system's inception, supervisors were required to announce in advance what week the appraisal would be done for each employee. This changed in 1985. Supervisors are no longer required to inform an employee when an audit will be conducted and they do not do so.

Also, the review is now continuous in that no single week for the review is specified as was done before. It can be done at any time throughout the year without notice at the discretion of the reviewing official.

Thus, the supervisor can conduct one employee's review at a time when the work load is very high and another employee's review may be completed at a time when the work load is low. This is hardly a true basis for comparison. Yet, that is one of the Congressionally stated purposes of the system; comparison of employees for any of the various

outcomes mentioned in the enabling legislation.

The use of the second set of performance criteria was intended to make the process more objective and thus more acceptable to the employees. The major operational change was ostensibly to soften the impact of the process and reduce stress on the employees since the review process would not be limited to one week and therefore not as intense. These changes were to have produced a more positive system.

In 1990, Commissioner Gwendolyn King convened a work group to assess the performance appraisal system. Data was collected from a cross-section of the work force agency wide via self-developed questionnaire. The result was to produce a report entitled Guidelines for the Development and Application of SSA Performance Standards. A synopsis of the operational outcomes found by this report follows.

Three specific negative operational outcomes were being generated by the performance appraisal system. The National Performance Standards Work Group report to the Commissioner states these adverse outcomes very clearly.

The most basic operational finding indicated that some types of work were not being completed at all or suffered lengthy delays in favor of those types of work which counted for numeric credit in performance appraisals. Types of work, such as post-entitlement earnings activity, benefit verifications, explanation of the Medicare pro-

visions, etc, would routinely be left undone so that more quantitative work duties could be maximized.

This tendency skewed the workflow and allowed large backlogs of these cases to develop. The Administration is still plagued by this phenomenon to the extent that it conducts quarterly pending work load audits to monitor it. Backlogs have continued unabated. It should be borne in mind that only those claims which have been taken and adjudicated are part of this study. The analysis of this research is not affected by the elongation of claims in the adjudication queue.

The second major operational problem resulting from the performance appraisal system is that work completed frequently does not count for appraisal purposes. For example, a service representative may spend considerable time making phone calls and performing calculations to resolve an overpayment issue. However, he or she would receive no credit for this since he or she is not allowed to effectuate the final resolution of the overpayment. Another example is that a claims representative may spend a great deal of time calculating a complex widow's benefit yet not receive credit for this time spent since it counts as a single claim, just as though there were only one possible benefit rate to pay. These are examples of work which must be completed which does not count for appraisal purposes.

The third negative resultant consists of six operational "barriers". Each barrier presents a specific problem to SSA and will be treated individually here.

The first barrier was that too much work was being required with too few resources. Given the downsizing the agency has experienced, this was not unexpected. Unfortunately, as long as SSA remains an on-budget agency, this condition is likely to persist. Thus, politics is found in administration: What would Woodrow Wilson say?

The second barrier is the failure to credit or consider all work. Note that this is different than the first finding. This deficiency indicates that not all work is considered valuable, even though it must be done. For example, jobs now contain components such as data input which had previously been lower grade work. No credit has been given for assuming and performing these functions in addition to those already found in the job description.

The inadequacies of the performance appraisal system itself were identified as the third barrier. This essentially refers to such things as duties not listed, standards not fair or objective, expectations not clearly defined, and the appraisal system not being perceived as equitable.

Supervisory deficiencies in using the system have been identified as a significant barrier also. This fourth barrier encompasses such concerns as non-supportive manage-

ment, favoritism, poor management practices and inadequate or incompetent supervision. These concerns have to be addressed regardless of the type of system used. Cameron (1989) concluded that the most basic necessity for effective performance appraisals is a "supervisor's fair, impartial and sincere desire to help the employee grow and advance." This is the antithesis of what has occurred in SSA.

Supervisory personnel were given virtually no training in performance appraisals prior to implementation of the system. Martin and Bartol (1982) identify seven vital areas in which supervisors must be trained to effectively operate a performance appraisal system. These include: the proper use of the system; the organizational aspects of the system; how to set meaningful performance standards; avoid measurement errors; expected types of feedback and how to deal with it; training on how to coach an employee; and finally how to conduct the appraisal interview. Virtually none of these steps were taken by SSA.

Norton and Dunne (1983) also assign great importance to the function of the supervisor. They feel that the rater must use differentiating achievement, rather than critical incidents (like errors) to be the basis for truly effective evaluation. In a study by Vance, Winne and Wright (1983) it was found that variance in rating was significantly more attributable to the rater than to the

ratee.

The fifth category of barriers involves too much emphasis on production. This barrier includes such basic concerns as: too much emphasis on numerics; questionable priorities and lack of concern for quality. Thus, simple production became the focus with little concern for what was being produced. The underemphasis on quality is borne out by the content analysis of this research.

The final category of barriers is a litany of negative outcomes from the first five. The sixth barrier includes: lack of training; low morale; low job satisfaction; poor procedures; increased stress and no job satisfaction. The work group recognized these factors as being caused by the performance appraisal system. It should be noted that this research does not seek to establish cause and effect.

Taken together, these six barriers constitute a major obstacle for SSA to overcome in developing an effective performance appraisal system. These six barriers must be resolved in any system which SSA designs. Indeed, Wallace Sayre's lament that technique is superior to purpose seems evident in this set of findings.

These findings were reported to Commissioner King who then took two major actions. The first was to abolish the use of numerics in performance appraisal. She then charged the work group with the task of rectifying as many of these deficiencies as possible within the constraints of downsiz-

ing, budget limitations, and the hostile political climate which permeated the 1980's in the federal civil service. The result was the production of a Guide for Developing Performance Plans.

Since the abolition of the 1985 standards in 1990, a relatively short time has passed and numerous sets of performance criteria have been used. These criteria have been a continuous source of bargaining disagreement between the Administration and the American Federation of Government Employees (AFGE) which represents the employees of the Social Security Administration. Due to the brief and unstable nature of these criteria, no post 1990 data will be used in this research.

Now, a general description of the appraisal process will be provided. Each position within SSA has a fixed number of Generic Job Tasks (GJT's) which the employee will perform. These include such things as "makes timely re-entitlement decisions" and "provides accurate information to the public and others." These GJT's are comparable with those criteria described as "Appraisal-By-Objectives" in Monitoring the Use of Appraisal by Objectives in Iowa: Research Note by Dennis Daley (1990). SSA's system is similar to the Iowa system reviewed in this article and may well suffer from the same anemia; performance criteria which cannot be measured. Daley's study concluded that only one percent of the criteria could be measured.

Saltzstein (1983) also concludes that the most basic difficulty with performance appraisals is unclear goals or goals which cannot be measured.

This type of objective system, even with employee input, may lead to "wolves in supervisor's clothing" Edwards (1983). Edwards suggests that organizations overcome these "wolves" by having multiple primary raters. This is unlikely in SSA for two reasons. First, the extreme specialization of the programs and second, the lack of personnel to do it. Thus, it is unlikely that these suggestions will change the current procedures in SSA.

These GJT's become the building blocks for the performance appraisal by a single supervisor. The work of the subordinate is viewed in light of these specific tasks. This process is analogous to the type of system argued for in Behaviorally Anchored Rating Scales for Patrol Officer Performance Appraisal: Development and Evaluation by Bradley and Pursley (1987) and later by Millar (1991). At first thought to be very objective (as SSA's GJT's are thought to be), Behaviorally Anchored Rating Scales (BARS) are later found to be highly dependent on the variable of organizational commitment as perceived by the rater (Ross, 1983). SSA's system may suffer from this same phenomenon.

Stemming from the tone of the enabling legislation, the system focuses on the errors of each employee as it relates to the specific GJT's of that job description. The

primary function of the supervisor while conducting the appraisal is the unit review. The timing and thoroughness of this review is left to the discretion of the reviewing official.

Since the intent of the supervisor is to make comparisons for " training, rewarding, reassigning, promoting, re-reducing in grade, retaining and removing" employees, they are searching for errors which can be used as a basis to rank employees against one another should any of the outcomes listed above become necessary or possible.

The supervisor has several sources of information available to him or her while conducting the continual unit reviews. The results of these reviews become the basis for the year end performance appraisal. It is this rating which is used comparatively by the agency in effecting any of the outcomes provided for in the Civil Service Reform Act of 1978.

The primary means of information gathering is the desk audit. This process involves the supervisor being at liberty to review any and all materials in, on or around the employee's work station at any time without providing notice that he or she intends to do so. The employee is not usually present while this is done. It is not necessarily done at the same time, nor with any concomitant thoroughness, for each employee, even by the same supervisor. This fact alone could render the process far

more subjective than it is intended to be.

Folders, call back messages, teletypes, and any other SSA related material are subject to review. These items are checked for timeliness, recordation accuracy, payment accuracy, documentation accuracy, correctness in priority, timely follow ups, etc. Virtually any point of view the supervisor wishes to take is allowed in the completion of the desk audit.

A second vital source of information is the interview audit. This involves the supervisor actually being present while the employee conducts an interview with a member of the public. The supervisor is introduced and his or her purpose is stated to the interviewee. The supervisor is then allowed to be present throughout the interview and to take notes concerning both the nature of the interview and the conduct of the employee.

Another frequently used technique for information gathering is to use the mentor approach. This involves the supervisor turning to the most experienced co-worker in any given operational unit and asking that employee if he or she has noticed any deficiency on the part of the employee being evaluated. While this is not formally allowed, it does occur and often this input is incorporated into the unit review as though it were observed by the supervisor.

While other means of obtaining information do exist, these comprise the most basic and frequently used methods.

The results of these inquiries are then grouped by Generic Job Task. Cases which contain detected errors are then categorized by type of error within the appropriate Generic Job Task.

Then, without knowing the total number of similar cases processed, a rating is given for each critical and non-critical GJT based on the observed errors. There may have been one error in two cases or one error in 2,000 cases, but the one error found becomes the basis for the rating regardless of the total number of similar cases processed.

Several pertinent factors are not accounted for in the summary of cases. For example, the varying level of difficulty of each individual case is not taken into account, nor is it possible for the reviewing official to know the total number of cases processed since literally thousands are done each appraisal year by each employee.

Therefore, no ratio between cases processed and errors found can be determined. This is how the default nature of the enabling legislation is operationalized. Since no true percentage of cases handled can be determined, the judgment of performance defaults to the number of errors found irrespective of any other consideration. This factor significantly limits the objectivity that is possible through this appraisal process.

A rating is then assigned for each GJT and recorded

on a summary sheet. See Appendix B for an example of the summary sheet. Once the rating is recorded on the summary sheet, weighting is done since not all GJT's have the same weight. For example, critical GJT's have a heavier weight attached to their ratings than do non-critical elements. After each GJT has been weighted, they are summed.

This sum total of the employee's weighted ratings is then divided by the total number of weighted points available on all GJT's in the position description. Appendix C contains an example of this process and a more detailed explanation. This final number is the overall rating for that employee for that year. Unit reviews are held twice a year. This frequency of review was suggested by Ronald W. Clement (1986) in The Performance Appraisal Interview: What, When and How. It must be noted however that only a yearly final rating is given for outcome purposes.

The possible ratings are supposed to be one through five. A rating of one is unsatisfactory. Two is minimally satisfactory. Three is fully satisfactory. Four is excellent and the rating given to level five is outstanding. Each employee receives only one of these as a final rating regardless of how well or how poorly he or she did on any single GJT. It should be noted that while all five levels are defined, only levels one, two and three were provided to field offices for usage. This is a significant factor during the qualitative portion of this research.

In a recent case, the Federal Labor Relations Authority ruled that performance appraisal systems can be legally challenged even when they have been approved by OPM. The agency brought before the review board had defined only one standard, fully successful, for assignment by supervisors. This was done with OPM concurrence. This example is provided here to demonstrate that each agency has discretion in implementing the performance appraisal system. SSA's use is not atypical of the implementation strategies agencies have used to pervert the performance appraisal system.

The "objective" nature of the appraisal is lost in this system. The final rating given does not reflect the number of cases handled, the errors as a percentage of cases or any other objective measurement. Instead, the supervisor gives a rating of one through five based on the audits conducted during the unit review with no evidence whatsoever that this audit is representative of that one employee's performance as a whole.

Yet, the rating is tied to percentages. In the given example in Appendix C, the percentage for level three is between sixty and eighty.

This does not imply that only sixty to eighty percent of the cases for the rating period are handled accurately. Rather, it reflects what percentage of available rating points the supervisor saw fit to award the employee. Meas-

ure anything you want: when it does not relate directly to the final rating scheme, it is not truly an objective appraisal. There are two negative results of this limitation.

First, the lack of objectivity undermines the support of the performance appraisal system. The work group determined that neither subordinates nor supervisors found the system to be fair and/or thorough.

The second negative outcome generated by the system is that its nature does not allow the strengths of individual employees to be recognized. This is true for two reasons. Primarily, it focuses on the negative. The second reason involves its reductionist nature.

Determining a single rating outcome gives no credit to those particular abilities that any given employee may have. Commissioner King dealt with this issue in the preface to the Guide for Developing Performance Plans.

Compassion

From the introduction of The Guide for Developing Performance Plans comes the basis for the pursuit of this study and a response to Without Sympathy or Enthusiasm by Victor Thompson. Thompson argues that compassionate public service is impossible due to the effects of complexity and modernity on large organizations (such as the Social Security Administration).

His viewpoint leaves practitioners of the New Public Administration no longer twisting in the wind on the issue of compassion since he decrees that individual treatment in a bureaucracy is neither possible nor desired. In his view there is no room for compassion in administration. Imagine his consternation at this:

More than a year ago, the Commissioner suspended the use of numeric standards in performance plans for SSA employees. That action was taken because the emphasis on selected numeric goals was adversely affecting completion of all work SSA must perform to serve the public. (emphasis in original).

The wide spread reliance on a few selected, easy to measure standards and goals was, in fact, resulting in a breakdown of teamwork and counterproductive competition. Moreover, there was a failure to recognize the many **individual** activities which contribute to the **compassionate**, quality service the public deserves and had come to expect from SSA. Recognizing the need to develop a better, more balanced approach to performance management, the Commissioner charged a workgroup (consisting of representatives from all SSA components, field and headquarters and management and union) with examining the issues and developing an approach to performance management that would ensure that performance plans were broadly based enough to reflect all that is expected of employees and all that the public deserves in terms of service. (Guide for Developing Performance Plans, 1990, p. 2)

Perhaps no clearer knell could have been sounded for compassionate administration than this.

Social Security, with its "womb to tomb" service orientation is probably the federal government's most public contact oriented agency. It has a mission which clearly involves many situations in which compassion may be con-

sidered appropriate perhaps even necessary since there are frequent daily interactions with retirees, widows, orphans, disabled and terminally ill individuals, though Victor Thompson may disagree.

Thompson views compassion in terms of outputs. He equates the expression of compassion with bending or breaking the rules to assist someone who would not otherwise receive assistance. Thompson also equates compassionate administration with answering a child's cry for assistance from its parent. In effect, the agency stands in loco parentis. This view of compassionate administration is far removed from what Gwendolyn King's work group called for in the Guide for Developing Performance Plans cited above.

The call for compassionate administration yields no intent to alter the outcome or output of the interaction. There is no rule breaking or rule bending proposed. Claimants make no child-like cry for outcome altering compassion as Thompson asserts.

Claimants simply desire an understanding of the system. Perhaps the most frequently found forms of compassion are explanation and education about the system to which the claimant is, or is trying to become, a part. If any child-like quality exists, it is the need to understand, not the need to be administratively coddled as Thompson suggests.

Thompson also asserts that some compassion is worse

than no compassion at all; that some compassion makes the individual an object of pity; that he or she is felt sorry for; that it may make him or her feel not worthy of inclusion into the system since the outcome does not change. Machiavelli would be proud (Skinner and Price, 1990). The Prince's often quoted dictum that it is better to appear kind and actually be cruel apparently obtains in Thompson's experience.

He asserts that expression of compassion is really a cruel act, only appearing kind, since the outcome and output will not be changed. He overlooks the possibility that the act of compassion itself may have value to the participants of bureaucratic interactions.

This *usus loquendi* portrays compassion as a part of the process which may remain in varying amounts at the interaction terminus, even though the output and outcome are not changed by its presence. This non-positivist view is what Gwendolyn King's work group called for in the Guide for Developing Performance Plans cited above. The basis for Thompson's viewpoint is utilitarianism. This theoretical precept has as its goal doing the greatest good for the greatest number. Thompson asserts that this is the reason for which bureaucracy exists; bureaucracy is the most efficient means of doing the greatest good for the greatest number. This orientation severely limits the agency's ability to see the individual subordinate as a

stakeholder (Mitroff, 1989). It seems clear from the description of the performance appraisal system that SSA operates from this utilitarian point of view since all employees are given a single final rating regardless of their individual abilities.

Utilitarianism, operationalized as the greatest good for the greatest number, is not appropriate when the well-being of an individual is at stake. Compassionate administration will not tolerate this version of democracy. The basis for democracy is the belief in the value of each individual. Any system which diminishes the value of the individual by portending to do the greatest good for the greatest number diminishes each individual's value by doing so and ultimately undermines itself. This position is supported by Ingle (1982) when he identifies each "individual's levels of commitment, communication, cooperation, human interaction, trust and understanding of one's own expectations and those of others" (p. 329) as the only workable starting point for good performance evaluation.

Perhaps this provides a basis for understanding the reason for the negative outcomes of the performance appraisal system of the Social Security Administration. The system has focused on the negative, aggregate nature of the stakeholders (both internal and external), (Mitroff, 1989) and found itself wanting and failing in respect to those who are affected by it. This viewpoint is supported

by the research of Ralph Borras (1987) in The Delicate Fulcrum Point. His most vital conclusion is that all stakeholders be part of the performance appraisal process. As Gabris (1983) concludes, management must remember that performance appraisal is as much a "people" concern as it is a results or structural concern.

The performance appraisal system of the Social Security Administration affects the individual natures of both the employee and the claimant. Tying pay to performance is tenuous at best Gabris (1983) and McNish (1986). Gaetner and Gaetner (1985) find no support for the concept of pay for performance plans, yet Social Security presently approaches the implementation of one.

Hyde (1988) in The New Environment for Compensation and Performance Evaluation in the Public Sector concludes that the real effects of the performance appraisal system are felt by those who receive the organizational services, the public. Thus, both the employee and the public stand affected by the nature and results of the performance appraisal system.

This is the fundamental reason for which no greater challenge exists for the Social Security Administration than to reestablish the dominance of purpose over technique in the realm of performance appraisal.

CHAPTER II

REVIEW OF THE LITERATURE

Overview

This literature review consists of five specific parts. The first part seeks to establish the Weberian nature of Social Security Administration in general and of the performance appraisal system in particular. The second part regards the Utilitarian nature of the performance appraisal system. It seeks to establish the link between the present system and the Utilitarian tenet of doing the greatest good for the greatest number. The third section deals specifically with Victor Thompson's Without Sympathy or Enthusiasm. It seeks to firmly establish the link between his treatise and Utilitarian doctrine. The fourth section of this literature review concerns performance appraisal literature in general while the final section deals with writings specific to the federal performance appraisal system in general and SSA's performance appraisal system in particular.

Weberian Literature

The Social Security Administration is a classically

Weberian bureaucracy. There is a definite focus on rules, structures, authority, hierarchy, career and specialization (Shafritz and Witbeck, 1978). Weber's six characteristics of bureaucracy are met within the organization in general and within the scope of performance appraisal in specific.

Weber's first characteristic requires fixed jurisdictional areas governed by rules. This is evident in SSA's job descriptions since they are specific in jurisdiction. This includes official responsibility for conducting performance appraisals.

Characteristic two is the use of graded levels of authority with clear lines of subordination. The supervisor is clearly the superior in the performance appraisal process. The claims representative, service representative or clerical is clearly the subordinate.

The third characteristic of bureaucracy is management by written documents. The numerous written guidelines and policies for developing and implementing SSA's performance appraisals fulfill this attribute of a classic Weberian bureaucracy.

The fourth characteristic of bureaucracy is training. Each claims representative has been through basic training and many subsequent training sessions regarding legal changes, technical changes and operational changes. Yearly, each subordinate receives training on the Generic Job Tasks which they are expected to perform.

Weber's fifth characteristic of a bureaucracy is that the job must demand the full working capacity of the official. There is no doubt that this requirement is met in SSA today since backlogs exist as a matter of routine across the spectrum of jobs in SSA.

Finally, Weber states that officials must have technical knowledge to perform their duties. SSA's personnel manuals alone are several thousand pages long. There are specific technical instructions regarding the design, implementation and use of performance appraisal. Having met these conditions, SSA should be considered a Weberian bureaucracy.

Sayles (1987) identifies six flaws of large organizations which are related to the Weberian principles just stated. Being one of the federal government's largest agencies, SSA suffers from all of these conditions which include: multiple layers of management; difficulty in communication maintenance; many interdependent units with separate goals which must be coordinated; an emphasis on rigid plans; use of narrow, rigid job classification and descriptors; and more power games or status ploys which interfere with goal achievement.

It is safe to characterize SSA as a classically Weberian bureaucracy which exhibits all of the traits just presented. It is in this context that the performance appraisal system must be understood. To be clear, this re-

search does not seek to indict the Weberian nature of the bureaucracy, only to establish it as a basis for understanding the subsequent analysis.

Utilitarian Democracy

John Stuart Mill typifies the concept of utilitarian democracy. In his book On Liberty, Mill clearly believes that democracy exists to do the greatest good for the greatest number. The performance appraisal system of the Social Security Administration operates on this basis. It does so by treating everyone exactly the same, regardless of the individual's strengths or weaknesses as an employee. This treatment, in theory, is to benefit the greatest number of employees by being equitable to all, thus doing the greatest good for the greatest number.

Further study of Utilitarianism reveals a basic contradiction which is incorporated into the performance appraisal system of the Social Security Administration. This contradiction involves the incompatibility of egoism and utilitarianism.

Fundamental to the design of the performance appraisal system as ordered by Congress is the concept of self since it seeks to rank individuals against one another. By this ranking attribute, the Utilitarian concept of egoistic hedonism becomes a factor in the performance appraisal system. The system forces the individual employee to focus on

his or her own well being since ranking against other employees is the primary use of the system. In this way, Congress forced egoistic hedonism to be an unavoidable component in the appraisal process.

However, since the system seeks to do the greatest good for the greatest number (irrespective of self), it also has a utilitarian aspect. This aspect arises from the notion that the majority of the employees will benefit from the standardization of the system for purposes of retention, promotion, demotion, and so on.

John Plamenatz in The English Utilitarians says that no man could attain his own greatest happiness without all men doing the same. Even if this were possible, egoism and utilitarianism can not simultaneously exist. Whether the interests of a man are natural or contrived, there is no possible accord between the two doctrines. It seems that any system attempting to reconcile these is doomed to failure. Yet, that is exactly what the performance appraisal process as designed by Congress seeks to accomplish.

Perhaps Jeremy Bentham in Introduction to the Principles of Morals of Legislation prognostigatively reveals the difficulty with SSA's performance appraisal system. In the work just cited, he initially contends that pleasure and pain can be quantified. He says that any two minutes of pleasure felt by one man is the same as any one

minute of pleasure felt by two men.

He constructed a system of filicific calculus to ensure that the greatest good is done for the greatest number in terms of pleasure. He attempts to quantify that which is not readily quantifiable. He eventually admits that this may not be possible.

SSA's performance appraisal system can be understood as a twentieth century attempt to institute filicific calculus in a Weberian bureaucracy. It attempts to quantify the many facets of an individual's behavior within a large, complex organization, which are inherently difficult to quantify, into a single numeric product of performance.

Saltzstein (1983) views the complexity of individual human behavior as a reason for which performance appraisal has never been effective. Saltzstein and Bentham come to agreement on this egoistic aspect of individuality and realize that further quantification is likely useless.

It seems clear that egoism and utilitarianism are at odds in operational terms. In Without Sympathy or Enthusiasm, Victor Thompson has seemingly solved this difficulty by declaring individualism (egoism) inappropriate in Weberian bureaucracies (utilitarianism). If one must be chosen, egoism or utilitarianism, perhaps SSA has selected the wrong concept for focus in relation to its performance appraisal system. Perhaps to reestablish the dominance of purpose over technique will require a change

of focus from utilitarianism to egoism in relation to performance appraisals.

Victor Thompson

Thompson envisions modern bureaucracies as modern day organizations in which personal and compassionate service has no place. His basis for this is that utilitarianism will not allow it. Any time spent on compassionate or personal acts would necessarily detract from the proper utilization of the system, thus ensuring that the greatest good is not done for the greatest number.

Thompson asserts that any exchange in a bureaucratic setting constitutes theft. He regards taxes as public theft and the expenditure of a public official's time on compassionate interaction as a means of theft. Again he argues that this is due to the utilitarian requirements of the system for efficient processing.

It will be argued in this research that this very viewpoint, utilitarianism, is the most basic and most debilitating trait of the performance appraisal system of the Social Security Administration since the Congressional purposes for which it was designed are basically incompatible in terms of egoistic hedonism and utilitarianism. As Bentham realized, these concepts are not reconcilable.

Thompson could not be located for the purposes of an elite interview on the subject of administrative compas-

sion. His former employer, Florida State University, was contacted to determine his whereabouts. He is no longer teaching and no address could be obtained for him. Thus, other steps were taken to attempt to resolve this issue.

The Public Affairs and Information System (PAIS) was accessed to search for further literature regarding compassion and public service. This proved to be in vain. No listings were found which included both concepts. Subsequently, the Psychological Literature Abstracts were key word searched and yielded no listings regarding compassion and public service.

The author hopes that this research brings again to the fore the concept of compassionate administration. It seeks to fill a void in the literature which has existed since Without Sympathy or Enthusiasm. It appears that Thompson's work may have been the last word on the notion of compassionate administration, until now.

General Performance Appraisal Literature

In Merit Pay and Performance Evaluation: Enhancing Employee Motivation in Municipal Government Michael Abels (1989) identifies three components essential to the successful operation of a performance appraisal system in relationship to the merit pay plan for the City of Defiance, Ohio.

He argues that objective standards must be taken di-

rectly from the employees specific job duties. He also feels that mutually designed, quantifiable goals must be established and that there must be a minimum of two evaluations per year. Abels states that performance evaluation could be a means of reaching the intrinsic motivation of the staff. If this occurs then scores improve as the participants of the process gain an understanding of each others expectations and goals. These goals must be established and individual performance geared to them.

A 1989 study by Sims, Veres and Heninger also identify the individual as the key to developing effective performance appraisals. Indeed, they call for the necessity of training of appraisers to facilitate the cooperation of the individual.

Abels also maintains that two essential problems develop even when these ideal conditions are met. First, Abels fears that grade inflation will occur. He argues that this occurs especially in government since anything less than above average ratings means that mediocrity is being accepted; a situation which no government manager wishes to face.

The second major problem identified is that of rater inconsistency. The same performance may be rated differently by different supervisors. This occurs, perhaps unwittingly, but has a detrimental effect on the system nonetheless. If performance evaluations were truly objec-

tive, there would be no need for ratings. However, since there is no true objective standard in place, rating is a necessary component of the process according to Abels.

In a similar assessment of what constitutes effective performance standards, Brown (1987) suggests that effective standards are: realistic, specific, measurable (both in quantity and quality), consistent with the agency's goals, challenging, dynamic and understandable. This set is simply a more detailed description of Abel's desired characteristics with the exception that Abels does not recognize the issue of quality. While Cascio (1989) believes that effective criteria are: relevant, sensitive to individuals, reliable, acceptable and practical.

Mossholder, Giles and Wesolowski (1991) call for a code of ethics for performance appraisals in general. Their motivation for this action is that they believe that a code of ethics would enhance the level of trust which is essential for employees to have regarding the appraisal system. The major objective of the code of ethics is the protection of the employee's right to privacy.

Ammons and Rodriguez (1986) conducted a study of the managerial performance appraisal systems of 122 cities across the United States. The inquiry tried to assess the types and styles of performance appraisal being used by these cities and found great variance among them. The systems were analyzed according to formality of appraisal,

appraisal objectives, techniques used in the appraisal process, time allocation, appraisal frequency and the general satisfaction with the appraisal process. Other factors included organization size and form of government.

Their conclusion was that performance evaluation is inconsistently applied at the managerial levels of these cities at best. Also, there is a great tendency toward relying on ratings as opposed to objective criteria. Many systems are trait oriented and highly subjective at best. Ammons and Rodriguez argue that these weaknesses make them unreliable as a basis for administrative decision-making. It is exactly these administrative decision-making functions for which the Social Security Administration relies on its performance appraisal system.

The author's final conclusion is that many cities reported a lack of commitment from those called upon to evaluate others. They conclude that without objectivity and commitment, performance appraisal will be of very limited benefit. These findings are later substantiated in Nonmanagerial Performance Appraisal Practices in Large American Cities (England and Parle, 1987).

Danny Balfour (1992) conducted a study in the State of Florida to determine whether or not the amount of time invested in the performance appraisal system made any difference in perceived effectiveness of the system. Hans and Tyler (1980), cited by Balfour, identified performance ap-

praisal as the weakest facet of personnel administration.

Six major agencies of the State of Florida participated in this study. All six agencies granted elite interviews as defined by Nachmias and Nachmias (1992) with their personnel directors and career employees. However, only four of the six participated in the survey. The interviews with personnel directors served to reveal the level of time committed to the performance appraisal system. The surveys served to provide the perceptions of the system as it had been implemented.

The agencies were then divided into two groups, high investment agencies and low investment agencies. According to the criteria established in the research, SSA would be a high investment agency. Several types of statistical analyses were performed which lead to the following conclusions.

The most basic finding of the study was that no significant difference in supervisory effectiveness existed between the two groups. High investment agencies emphasized both the administrative and counseling aspects of performance appraisal. This dual approach was not successful in improving supervisory effectiveness. The basic lesson here is that these two goals are conflicting and cannot be reconciled, even by the high investment of time and money by the agency. This is clearly a case of the techniques being superior to the purpose.

The final conclusion of the study indicates that less should be expected from the performance appraisal system regardless of the level of agency investment. Balfour contends that performance appraisal is more attuned to administrative decision-making than it is to counseling. He recommends that these functions be separated and that other programs be put in place to counsel employees.

Hyde (1982) came to a similar recommendation regarding the separation of functions as did Zemke (1991). A study by Hopkins, Lovrich, Shaffer & Yale (1980) concludes that employees generally distrust the entire process regardless of intent.

Thus, at present, performance appraisal serves two masters with equal ineffectiveness. This final conclusion is compatible with those resulting from Robert W. Clement's research in Performance Appraisal in the Public Sector: Truth or Consequences (1984).

Objective Criteria Literature

Leonard Berger (1983) in The Promise of Criterion-Referenced Performance Appraisal, (CRPA) traces the chronological development of performance appraisals. This leads him to the recent development of CRPA's.

He promotes the objective nature of CRPA's stating that it is vital for CRPA's to be based on the job in its entirety. Also, it is important that employee participa-

tion be encouraged in the process whenever possible. He does note that, while objective criteria tend to be more accepted, they are often used comparatively and therefore promote competition which is counterproductive.

Dennis M. Daley (1990) in The Civil Service Reform Act and Performance Appraisal: A Research Note on Federal Employee Perceptions found that the objective criteria established for federal employees in 1981 were not effective. In fact, federal employees perceptions of objectiveness in evaluation declined after the objective standards were implemented.

Brown (1982) in Performance Appraisal: A Policy Implementation Analysis researches the implementation of performance appraisals under the Civil Service Reform Act. He takes the perspective of the Office of Personnel Management (OPM) for which he works. He describes five conditions which must be present for the successful implementation of performance appraisals. He then states that OPM has created these conditions, except for continuing legislative support. However, he finds that implementation problems exist within the agencies as they create a system to meet OPM guidelines.

He is only guardedly optimistic about successful implementation of the Act's requirements even under the five ideal conditions he elucidates. He states that there will be a great temptation to "pull the plant up by the roots to

see why it doesn't grow faster" (p.71). After twelve years of dismal failure, perhaps the uprooting time is here.

Nalbandian (1981) argues that objectivity can be attained, but at a very high cost. He argues that objectifying performance appraisal will distort the human dynamics at work in the organization. He maintains that if an organization is willing to accept this cost, four elements are necessary for success. These elements include trust by all persons involved, acceptance of the supervisors role in performing the appraisal, sensitivity to the human element and effective supervisory training. Even under these conditions, Nalbandian is not sure that objectivity is worth the cost.

In the performance appraisal literature in general, the issue of whether or not the system actually affects performance is rarely addressed. One piece of research which attempts to determine whether or not performance is actually affected by the performance appraisal system is by Marion E. Haynes (1973).

In Do Appraisal Reviews Improve Performance, Haynes determines that two individual characteristics are essential to the result of any performance appraisal system. These are the individual's level of aspiration and self-image. He finds that performance appraisals are a threat factor on the issue of self-image and a demotivator on aspiration unless the individual has internalized the or-

ganization's current goals. His research takes place in a large, private sector organization and deals with performance measured over time by consistent criteria. Haynes' results indicate that the performance appraisal system was ineffective in improving performance or behavior.

Another perspective on this issue is offered by Wholey (1991). He states that the critical step in the future of performance appraisals is to determine whether or not performance is actually affected. He is a former Secretary of Health and Human Services. He suggests that several, small pilot studies be conducted to assess this issue. This research constitutes a retrospective attempt at making this assessment.

Three dramatic departures from the standard literature are worthy of mention in this review. All deal with performance appraisal in relation to the total organization, although each takes a different approach.

The first departure is Evaluate or Not: That is Not the Question in which Sumek (1988) deduces that performance appraisal on the local government level is here to stay. His insight is that most performance appraisal is retrospective when it should be prospective. SSA's appraisal system is retrospective to the exclusion of prospectivity. Perhaps employees would better accept performance appraisal if it focused positively on the employee's future rather than negatively on the employee's past.

Another dramatic departure from the literature is provided by Ronald W. Clement (1987). In Performance Appraisal: Nonverbal Influence on the Rating Process, Clement argues that one shortcoming of performance appraisal is its reliance on the verbal nature of communication only. He suggests that there are three vital areas in which nonverbal factors are important.

One group is composed of the employee's social and leadership skill, including such things as the attributed causes of performance and the employee's similarity to the supervisor.

A second area of nonverbal concern is the features of the supervisor. This group includes the supervisor's level of knowledge about the employee, the experience as a supervisor and the relationship with the employee.

The final nonverbal area is the rating context. This includes the timing and purpose of the rating as well as the organizational culture in which the rating is given.

He contends that these areas do play a vital, though unrecognized, part in the appraisal process. He stresses the need for all parties involved to come to terms with these nonverbal factors. He argues that doing so will improve the system and make performance appraisals more useful.

Though this line of research is interesting and no doubt plays a part in appraisal, it is unlikely to be

valuable. This is so since the rater cannot change the employee or the context in which the appraisal is given and is unlikely to change his or her own features to accommodate any performance appraisal system.

The final departure from the usual literature perspective is Performance Appraisal in the Z Organization by Grover Starling (1982). Starling reviews the Japanese methods of performance appraisal and notices five basic differences from our most frequently used systems.

First, there is an acceptance of the subjective by the Japanese that is not present in either the public or private sector in the United States. If performance appraisal is subjective then perhaps we must come to terms with that facet of it before anything else can be accomplished.

The second attribute of the Japanese system is the recognition of the value of interdependence. It is stressed that teamwork is most important to goal accomplishment. Meanwhile, our systems in general, and SSA's in particular, pit employee against employee and decrease the probability of teamwork.

There is often no formal appraisal given for long periods of time in the Japanese system. This third attribute facilitates long term evaluation which is more comprehensive than the short term, less comprehensive counterparts generally found in American organizations. This amounts to the organization doing more work and doing less

evaluation.

The fourth attribute of the Japanese system is the informal, continuous and fluid flow of performance feedback. It is not formatted in the rigid, written, fixed time frame format that most American appraisals are. This reduces the stress between rater and ratee. This means that the performance appraisal system must be understood in terms of the entire organization, not as a system unto itself. This is consistent with the findings of Gilbert (1982).

The final difference between the American and Japanese systems is the decoupling of the appraisal system from the compensation system. The Japanese feel that this is necessary to reduce the threatening nature which performance appraisal has been documented to invoke in many American organizations.

In all, these attributes may hold some long term value for American organizations in the years ahead. However, Starling is not optimistic at this point in time. He feels that American organizations and governments are headed in a more explicit, more job related (not goal related) direction.

SSA's direction is unquestionably concomitant with the general trends found by Starling. Having indicated the direction in which it seems to be headed, perhaps now is the time to address SSA's system in particular.

Social Security Literature

In 1993, the General Accounting Office conducted interviews with personnel specialists from many agencies throughout the federal government including the Department of Health and Human Services. The basic conclusion of the study was that the performance appraisal system as it currently exists is not effective (Federal Times, 1993).

Roberta Peters, said "I believe that the current system is an impediment to federal employee productivity" (Federal Times, 1993, p.12). Abramson and Schmidt (1983) link the public perception of credibility to perceived service. They maintain that as long as civil servants are perceived as ineffective, performance appraisal will be criticized as being ineffective. He does not directly establish this linkage however.

Larry Slagle, Department of Agriculture Personnel Director believes that change is inevitable and that every single employee has a stake in it (Federal Times, 1993).

The conclusion of the GAO was that the performance appraisal system should be divorced from awards and promotions. Agencies need separate methods for determining who receives awards and promotions. They determined that the performance appraisal system should be used to enhance all career opportunities but that it was seen as a threatening experience which can not facilitate this outcome.

The Federal Times (1990) conducted a large survey of employee attitudes in the Social Security Administration about performance appraisal. The result of the study was that only eighteen percent of the employees felt that the system was fair. The survey also determined that there was a general distrust of the intended uses of the system. Employees were not convinced that the system was intended to help them perform better. Instead they felt that it was a threatening system.

In the Report of the National Performance Standards Workgroup, many possible reasons for this dissatisfaction were uncovered. The causes ranged from supervisory incompetence and the nature of the performance appraisal system itself to excessive quantification and lack of objectivity in standards.

The Haynes research aside, the majority of the current literature reveals the negative and normative outcomes of the system. However, any decision about any specific system, such as SSA's, should assess the effect which it has on actual performance.

Weberian bureaucracies may be able to defend the use of these types of systems, which are generally perceived as unfair, threatening and ineffective for the purposes of employee career enhancement and counseling, if they actually are effective in utilitarian terms. For this to be true, performance would have to improve. The purpose of this

research is to determine whether or not performance has improved thus validating the utilitarian aspects of the system.

This research serves to create this link between the performance appraisal system and actual performance in the Social Security Administration. This link is similar to the one which Haynes considered twenty years ago in the private sector. This research seeks to gain an understanding of whether or not there is a link between improved performance and the appraisal system within the Social Security Administration.

CHAPTER III

STATEMENT OF SCOPE

Overview

Three factors lead to the narrow scope of this research. First, it is limited by agency. Then, within the agency, specific components have been chosen for study. Within these components, certain programs were targeted for analysis via the performance appraisal criteria that were utilized between 1981 and 1990. Northrup and Perry (1985) find that the type of service, (human or physical), location, (headquarters or field) and agency focus (process or product, etc.) are three important factors in assessing appraisal systems. This research deals with human service, field operations and a definite agency focus on claims processing. The rationale for these selections will be presented here.

The Social Security Administration is the general scope for this research. Part of the Department of Health and Human Services, Social Security operates approximately 1300 domestic field offices, one central office, ten regional offices, seven payment centers, three data operations centers, numerous offices of hearings and appeals as

well as other types of specialized offices. See Appendix D for an executive organization chart of the Social Security Administration. The Administration also operates offices in various foreign countries through the Office of International Operations.

Locus of Research

This research is specifically limited in scope to the Social Security Administration's domestic field offices. The basic reason for excluding foreign operations is that the employees in those offices are not subject to the same performance appraisal system outlined in the introduction which applies to domestic employees. The field offices were chosen as the loci for this research since the vast majority of the employees who are subject to performance appraisals work in these locations.

Another factor contributing to the selection of this scope is the public service orientation of the domestic field offices. No other offices within the domestic scope of SSA, such as central or regional offices, have comparable public contact responsibilities to that of field offices. This significantly changes the performance appraisal criteria for field office employees when contrasted with other domestic SSA components.

Focus of Research

More specifically, the Title II entitlement programs of the Social Security Administration have been selected as the foci for this inquiry. This includes the Retirement/Survivors (RSI) and Disability (DI) programs. The Title XVI program, commonly called Supplemental Security Income (SSI) have been excluded. The Title XVIII program, commonly called Medicare, was also excluded. There are numerous reasons for these exclusions.

The basic rationale for excluding the Title XVI program is a data limitation which will be discussed at length in the methodology section of this proposal. However, other rationale exists for exclusion of this program.

First, SSI is an eligibility program, not an entitlement program. The basic difference is that Title XVI claims require much more work initially as well as frequent and regular follow ups which are not required for Title II programs. Thus, the nature of the work itself may make the SSI program an acceptable exclusion from the scope of this study. Also, it is funded from general revenues, not earmarked contributions like the entitlement programs which are the focus of this research. This clearly differentiates the two programs. Finally, the record keeping system for SSI claims is not automated and is less reliable than the Title II data. Given these limitations, the exclusion

of SSI claims seems acceptable.

The Title XVIII program was excluded because the Social Security Administration does not process the claims for Medicare. The processing of these claims is completed by a carrier or intermediary under the auspices of the Health Care Financing Administration (HCFA). Social Security field office personnel simply interpret the Explanation of Medicare Benefits statements, a miracle in its own right, after the claims have been paid by the carrier or intermediary. No useable measure of this activity exists.

The domestic field offices of the Social Security Administration have an operational and technical homogeneity which makes them an acceptable scope for study. Though offices vary greatly in size, all domestic field offices take the claims which are the foci of this research.

Additionally, all offices are tied to one central mainframe computer. The processing and evidentiary requirements are the same for every field office. The technical aspects of taking RSI and DI claims are the same regardless of office size or location. The claims counts, which are the essential component of the data used in this study, are automated. This makes them highly reliable and comparable.

Summary of Scope

In summary, the scope of the research is limited to Retirement and Survivor's Insurance (RSI) and Disability insurance (DI) claims processed by domestics field offices of the Social Security Administration. All RSI and DI claims taken between 1981 and 1990 are subject to analysis in this research. The issue statement which follows makes clear the reasons for which this scope is adequate and appropriate for this inquiry.

CHAPTER IV

ISSUE STATEMENT

Overview

The specific issue under investigation is the relationship between two sets of performance appraisal criteria and two specific output measures over a ten year time span. However, there are many factors which require explanation to ensure a clear understanding of this issue. Some of these factors are operational while others are purposive.

Operational Issues

Operational issues exist on several levels. It must be clearly understood that this research focuses only on work which has been accomplished. This means that the accumulation of workload backlogs is not a significant factor in the analysis of this data. It is recognized that these backlogs exist, but their existence does not affect the work which has been completed. Only that work which has been completed is included in the data for this study. Thus, the backlogs are not an intervening variable in the research.

Another operational issue is the downsizing of the

Social Security Administration. SSA, like every other federal agency, lost personnel during the years covered in this study. However, this factor does not directly affect the work measured by the RSI and DI processing times. This is true since the effect of downsizing would be the creation of backlogs. The non-mitigating nature of backlogs has just been presented.

A final operational issue is the distribution of claims. The distribution of RSI and DI claims did not affect processing times during the years covered in this study. It is raised as an issue here since it might have affected the results of the study. To the extent possible, the methodology used in this research has determined that claims distribution is not an issue to be operationalized and taken into account during the hypothesis testing portion of this research. The quantitative methods ensure that this conclusion is valid.

Purposive Issues

There are three basic purposive issues. These issues are negativity, objectivity and quantification. Each of these issues separately, and in some cases perhaps jointly, affected the performance appraisal system which is the subject of this study.

The negativity issue in this study arises in two forms. First, the enabling legislation is negative in

tone. Second, it is operationalized by the methods employed in utilizing the performance appraisal system regardless of the levels of quantification and objectivity. It may be that this system can not be perceived as a positive force within the agency, and thus become an effective means of improving performance, until this issue is resolved. It is hoped that this study will raise this issue to a level at which it can no longer be ignored.

The second purposive issue is objectivity. The enabling legislation clearly mandates objective performance appraisals. Yet, the ratings given to each employee who is covered under this system are inherently subjective since they are ultimately the opinion of the supervisor about the employee's performance. This is not truly objective since even the most basic numeric comparisons, like percentage of errors or frequency of errors can not be made. Yet fixed percentages are assigned as indicators of performance. For this system to be effective, a means of becoming more objective may be needed.

The last purposive issue is quantification. During the years covered in this study, it became the single most vital determinant of work flow. After a long courtship, quantification was married to supposed objectivity by the 1986 standards. This union was annulled by Commissioner's decree in 1990. While quantification does have a place in performance appraisal, it may not be the most effective

primary focus of the process.

An equitable and positive focus of the performance appraisal system should drive the system toward the agency objectives of timely and accurate payments as well as the desired levels of compassionate individual service. The accomplishment of this would certainly remove the system from the "iron cage of rules" and establish the dominance of purpose over technique.

CHAPTER V

METHODOLOGY

Overview

The methodology section of this proposal begins with a general description of the research design. This is followed by operational definitions for each variable under study. Next, an detailed set of axiomatic assumptions will be stated. The assumptions are followed by a section which makes explicit the hypotheses which drive the methodology. Then, an explanation of the research techniques to be used to test the hypotheses will be provided. Subsequently, a set of instructions necessary to complete the research is detailed. Finally, data limitations will be discussed.

General Description

This research is exploratory in nature. It attempts to establish whether or not performance improved over the time during which the 1981 and 1986 performance appraisal criteria were in use. This study utilizes secondary data which has been collected consistently over a substantial portion of the time frame of this study.

This data was collected specifically to report per-

formance and therefore does not have the limitation of having been collected for another purpose as is sometimes a limitation when using secondary data (Nachmias and Nachmias, 1992).

The investigation is narrow in scope and longitudinal in orientation. The narrow scope of the study is advantageous in that there are few intervening variables. The longitudinal orientation is important since this allows the program to be analyzed over a long period of time for variations which may have occurred at any time after implementation. These factors enhance the reliability of the study since it has a definite, consistent focus over a long period of time.

This research utilizes both quantitative and qualitative research techniques in the framework of axiomatic theory. First, the axioms are expressed as assumptions which are made explicit. Then the theorems are stated as hypotheses which are to be tested.

The qualitative portion of this research employs content analysis. Content analysis uses nominal data frequencies as a means for establishing any differences which may exist between the sets of performance appraisal criteria. The specifics of this procedure will be provided later in the methodology.

Parametric statistical techniques, utilizing interval level data, are the quantitative means of testing the hypo-

theses of this inquiry. The specific parametric techniques used include the t-test and regression analysis.

Operational Definitions

The variables under study in these hypotheses, Retirement and Survivor's Insurance (RSI) processing time and Disability Insurance (DI) processing time, have precise operational definitions. Each set of performance appraisal criteria will be also be operationally defined.

Operational Definition - RSI Processing Time

RSI processing time is defined as the true cumulative time required to complete a Title II Retirement or Survivor's claim. It begins with the date the application is taken or the protective filing form is completed. It ends with the day that the computer system processes the claim for payment even though an actual check may not be due for several months. These dates are automatically captured by the computer system. RSI processing time is measured in days expressed as decimal (i.e. a claim can take 3.6 days) and is considered interval level data.

Operational Definition - Disability Processing Time

DI processing time is the average cumulative time required to complete a Title II Disability claim. It begins with the date the claim is taken or protective

filing is completed. It omits the days that the claim spends en route to and is being processed by the Disability Determination Service (DDS). The exclusion of this time is necessary since the field office has no control over the processing time used by the DDS to reach a decision regarding whether or not the claimant is disabled. This exclusion is uniform throughout the agency. These days are automatically excluded from the processing time count by the computer. The excluded days begin with the day the claim is forwarded to the Disability Determination Service and a computer notice is generated by the claims representative in the field office. The excluded days end with the day the claim is receipted in via computer input back into the field office. DI processing time is measured in days expressed as a decimal (i.e. a claim can take 14.7 days) and is considered interval level data.

Operational Definition - 1981 Performance Criteria

The 1981 set of performance appraisal criteria is operationally defined as those regulations which are used by supervisors to determine a performance rating for an employee as specifically stated in the Performance Standards Handbook - 1981. These criteria were in place through fiscal 1985.

Operational Definition - 1986 Performance Criteria

The 1986 set of performance appraisal criteria is defined operationally as those regulations which are used by supervisors to determine a performance rating for an employee as specifically stated in the Performance Standards Handbook - 1986. These were in use until Commissioner King's decree to cease using them in 1990. These definitions are the delimiters for both the quantitative and qualitative portions of this research. Though inherently simple, they are accurate and reliable. RSI and DI processing times are completely automated with virtually no chance for human intervention or error. The sets of performance appraisal criteria are distinctly different. Written years apart by different work groups, they serve to define the agency's very diverse attempts to implement the legislative mandate for performance appraisal.

Axiomatic Assumptions

This section makes explicit the axiomatic assumptions of this research. According to Nachmias and Nachmias (1992), Axiomatic Theory consists of two parts, theorems and axioms. Axioms are the theoretical, untestable assumptions which are assumed to be true for the theory to be valid. The axioms appropriate for this research are expressed in the following assumptions.

1. If a set of performance appraisal criteria is effective then significant improvement will be observed within the time period that the criteria are in effect.

2. If both sets of performance appraisal criteria are effective then significant improvement will be observed across the time periods that the criteria are in effect.

3. It is assumed that the political environment was consistent across the entire time period and did not affect one set of performance measures more than the other.

4. It is assumed that the downsizing of SSA did not significantly affect the processing times of those RSI and DI claims which were processed. Downsizing would result in increased backlogs, not changes in processing time.

5. It is assumed that the effects of the performance appraisal criteria on the Title II claims taking process are similar to the effects on the clerical, service and Title XVI claims taking processes. This relates to the issue of the generalizability of the research.

Research and Null Hypotheses

Theorems are those propositions deduced from axioms which are able to be empirically tested. In this proposal, the theorems have been expressed as H₁ through H₈. These eight hypotheses are exhaustive of the operational relationships possible between the variables. Each hypothesis and its corresponding null will be provided.

The testing of these hypotheses should allow any relationships between the sets of performance appraisal criteria and actual performance between 1981 and 1990 to be ascertained. The decision criteria for acceptance or rejection of the null will be provided in the data analysis portion of this proposal.

1. H₁ The 1981 set of performance appraisal criteria was not associated with significant improvement in RSI processing time.

H_{1.1} The 1981 set of performance appraisal criteria was associated with significant improvement in RSI processing time.

2. H₂ The 1981 set of performance appraisal criteria was not associated with significant improvement in DI processing time.

H_{2.1} The 1981 set of performance appraisal criteria was associated with significant improvement in DI processing time.

3. H₃ The 1986 set of performance appraisal criteria was not associated with significant improvement in RSI processing time.

H_{3.1} The 1986 set of performance appraisal criteria was associated with significant improvement in RSI processing time.

4. H₄ The 1986 set of performance appraisal criteria was not associated with significant improvement in DI processing time.

cessing time.

H.4 The 1986 set of performance appraisal criteria was associated with significant improvement in DI processing time.

5. H. The 1981 and 1986 sets of performance appraisal criteria together were not associated with significant improvement in RSI processing time.

H.5 The 1981 and 1986 sets of performance appraisal criteria together were associated with significant improvement in RSI processing time.

6. H. The 1981 and 1986 sets of performance appraisal criteria together were not associated with significant improvement in DI processing time.

H.6 The 1981 and 1986 sets of performance appraisal criteria together were associated with significant improvement in DI processing time.

7. H. There was no significant difference in RSI processing time between the 1981 and 1986 sets of performance appraisal criteria.

H.7 There was a significant difference in RSI processing time between the 1981 and 1986 sets of performance appraisal criteria.

8. H. This was no significant difference in DI processing time between the 1981 and 1986 sets of performance appraisal criteria.

H.8 These was a significant difference in DI pro-

cessing time between the 1981 and 1986 sets of performance appraisal criteria.

Research Techniques

Several research techniques are employed in this study. Each technique will be discussed briefly along with the purpose it serves in this research. The techniques used include content analysis, regression analysis, and a t-test.

Content Analysis

Content analysis is a technique for making inferences by systematically and objectively identifying the specific characteristics of written or verbal messages (Krippendorff, 1978). It requires the consistent application of explicit rules for interpreting the meaning of communications.

Content analysis utilizes sampling units, context units and recording units. This research seeks to reveal two aspects of the respective sets of performance appraisal criteria. One is the quantitative versus qualitative aspect of the sets. The other is the positive versus negative aspect of the criteria.

The sampling units establish a general frame of reference in which the context and recording units are found. For purposes of this research, each Generic Job Task (GJT) constitutes a sampling unit.

A context unit is a sentence or descriptor within the sampling unit. A sentence is any grammatically complete thought in the GJT. A descriptor is any label, level or specific requirement of performance in the GJT. Each sentence or descriptor may contain one or more recording units.

The recording units are those specific words phrases or numerals which are indicative of the quantitative, qualitative, positive and negative nature of the performance appraisal criteria. See appendix E for a complete list of words and phrases which are considered positive, negative, quantitative and qualitative.

It should be noted that many words contained in the sets of performance appraisal criteria are neutral and are not included in any of the four sets mentioned above. No comprehensive list of these thousands of words has been compiled. For example, articles, pronouns, and propositions have been excluded from the four analytical groups. Words such as communicates, provides, obtains, files, responds, absorbs, applies, etc. have been judged neutral. This is appropriate since they speak to none of the four issues under study. They are simple recognition skills which are not inherently quantitative, qualitative, positive or negative.

Content analysis, under the guidelines provided above, will serve two functions. Initially, the quantitative ver-

sus qualitative and positive versus negative attributes (herein after called research attributes) of all GJT's will be determined. This allows a systematic assessment of each set of performance appraisal criteria as a whole for purposes of comparison.

Following the assessment of the GJT's as a whole, the research attributes for only those GJT's which affect taking RSI and DI claims will be analyzed. So doing will allow the comparison of those RSI and DI GJT's to each set as a whole. This comparison will assess the level of similarity of the RSI and DI GJT's to each set as a whole.

By accomplishing these two functions, the direct comparability of the RSI and DI GJT's to the Generic Job Tasks as a whole can be determined. This should enhance the internal validity of the study.

Regression Analysis

Regression analysis was used to determine whether or not any significant improvement occurred within the scope of each set of performance appraisal criteria. This was accomplished by analyzing the monthly processing times in an effort to determine whether or not any significant improvement can be measured.

The interval level data provided by the processing times is the basic measurement of performance used to determine whether or not any improvement which may have oc-

curred as analyzed by the regression procedure. Regression analysis is the method for testing the intra-criteria trends which are expressed in hypotheses H₁ through H₄. RSI and DI times are subjected to this analysis separately.

It is necessary to identify those variables which may intervene in the relationship between the performance appraisal criteria and processing times. One possible intervening variable is claims distribution. Regression analysis was used to establish that this was a non-intervening variable.

t-test

The inter-criteria analysis requires the use of the t-test which is a parametric statistic requiring interval level data as well. This analytical technique compares the means of two groups to determine if any significant difference between those groups exists. The specific test invoked in this research incorporates the concept of unequal variances. This is necessary since the variances differed between the 1981 and 1986 sets of processing times for both RSI and DI claims. In this research, RSI and DI processing times are separately subjected to the t-test procedure to determine if any significant difference in performance exists between the sets of criteria as measured by the interval level data of processing time.

Thus the 1981 and 1986 RSI processing times will be

compared for any significant difference. The 1981 and 1986 DI processing times will be compared for a significant difference as well. This technique provides the means for testing hypotheses H₁ and H₂. The hypothesis testing portion of this inquiry will be complete at this point.

Actions and Instructions

Both quantitative and qualitative data was necessary to complete the research. This section provides a comprehensive list of actions and instructions which have been taken to complete the project. These sections are general and representative of the steps overall. There are specific facts which are not presented here.

Qualitative Data Gathering

A literature review was completed from several perspectives to ensure that the subject of performance appraisal was thoroughly and fairly treated as it relates to this research. Several specific topics, including, Weberian bureaucracies, and utilitarian principles were researched and related to the literature on performance appraisal in general, objective appraisals and most specifically, the Social Security Administration's appraisal system. Additionally, the issue of compassion in administration was researched.

The 1981 set of performance appraisal criteria was

available from the Regional Personnel office. This is the only copy which could be located.

The 1986 set of performance appraisal criteria was provided by the union of American Federation of Government Employees. Having these two sets of appraisal criteria completed the qualitative data requirements for the research. The quantitative data was much more difficult to obtain.

Quantitative Data Gathering

Data collection was initiated in May of 1992. Several requests were denied on the basis that the requested data did not exist. Subsequent Freedom of Information Act requests were then filed beginning in March 1993. The information used in this research was finally delivered on December 22, 1993. This process devolved into a bureaucratic nightmare not soon to be forgotten.

The original intent was to use six processing measures over ten years to assess the effectiveness of the performance appraisal system. However, it was determined that several data limitations existed on the part of Social Security Administration.

The foremost limitation was that no valid Supplemental Security Income (SSI) data was available on a fixed monthly basis. Of the six intended measures, this eliminated two, SSI processing time and SSI payment accuracy.

The second data limitation involved RSI and DI payment accuracy. The Administration changed computer mainframes after 1990. No prior payment accuracy data has been converted to the new system. Paper records could not be located. Thus, RSI and DI payment accuracy could not be obtained for the years included in this analysis.

As a final result, only two of the six originally intended measures could be included in this study. RSI and DI processing times are included since they are the only viable measures available for the time period of the study. Even this data was not available in its entirety. Data was only available from April 1984 through December 1990. These months constitute the actual time frame for this research.

This litany of data collection limitations is noted not in an attempt to criticize SSA, but to make clear the initial research intent to use all viable measures and to demonstrate that the researcher was not biased in selecting RSI and DI processing times as a means of assessing the ability of the performance appraisal system to affect performance.

Data Analysis - General

The data analysis proceeded at different rates. The qualitative analysis was completed before the quantitative data was obtained. Though sequential in the following

narrative, several steps were concurrently completed.

Qualitative Analysis

The initial step in data analysis was the completion of the content analysis of the two sets of performance appraisal criteria. Sampling units, context units and recording units were selected and used consistently throughout both sets of criteria. The Generic Job Tasks were selected as sampling units. The sentences and descriptors were chosen as context units and then individual words, phrases or numbers were selected as recording units.

The next step in the content analysis was to review the Generic Job Tasks for all words, phrases, or descriptors which were indicative of the research attributes of quantity, quality, positive and negative. A comprehensive list of these words for each attribute can be found in appendix E.

It must be stressed that once a word was considered to represent one of the research attributes, it was counted every time it appeared in either set of performance appraisal criteria. This is necessary to insure the consistent application of the principles of content analysis.

The next step was to label the GJT's by position type and criteria set for later analysis. For example, Set 1, Service Representative, consists of all those GJT's which

applied to the Service Representative position from 1986 to 1990.

After this was done, each GJT for each position was marked as follows: Quantitative attributes marked in orange; Qualitative attributes marked in pink; positive attributes marked in green and negative attributes marked in yellow. A box of each color was then made at the edge of each page and the total number of each attribute was included in the box for easy reference and verification.

The position types were then collapsed into three categories for data analysis. These categories include claims representative, service representative and clerical. This grouping corresponds to the functional structure found in SSA's domestic field offices.

This was necessary since the position types found in the first set of performance appraisal criteria did not exactly correspond to those found in the second. For example, the Data Review Technician position was distinct from the service representative position in the first set. This was true even though the pay level, technical knowledge and responsibilities were essentially the same. In the second set, the positions were listed together as a combined position description. For a complete list of how the various position types were grouped, please review the material in Appendix F.

Once the grouping was complete, the data was entered

into a raw data file for the purposes of computing descriptive statistics based on the nominal data generated via this content analysis. All data was entered in a three digit format. Thus, if fifty-three qualitative attributes were observed 053 was entered. Likewise, if one hundred and five quantitative attributes were found 105 was entered. The results of this analysis will be discussed in the Findings chapter.

Quantitative Analysis

The quantitative analysis required two basic steps. The first was that the RSI and DI overall counts and RSI and DI processing times be entered into a raw data file. Overall counts were entered in a six digit format using lead zeros when necessary. RSI and DI processing times were entered in a three digit format. No lead zeros were necessary in this data set. The first step was to determine whether or not the claims distribution correlated significantly with the respective processing times. They did not. Had they done so, this relationship would have been operationalized via a factoring procedure and taken into account in the regression and t-test analyses which followed. The specific results of the claims distribution by processing time regression will be discussed in the findings chapter as will the results of the result of the hypothesis testing via regression.

Limitations - General

There are four basic areas of limitations to this research. The first involves the lack of SSI processing time data. The second limitation is the lack of payment accuracy data for either program on a monthly basis. The next limitation concerns the unavailability of RSI and DI processing time prior to 1984. The final limitation involves the lack of actual performance rating for individuals.

SSI Processing Time Limitation

The most basic limitation of this research is the lack of monthly processing time data for the different types of Supplemental Security Income (SSI) claims. The initial intent of this research was to include these processing times since they may have a different relationship to the performance appraisal criteria than that of the Title II programs. This limitation reduces the overall generalizability of the study but does not directly affect the analysis of the programs which are the focus of this research.

Payment Accuracy Limitation

The second data limitation is that no monthly data was available on payment accuracy. It could be argued that the performance appraisal criteria might affect the accuracy of

processing claims. Since it was not kept on a monthly basis during the years of this inquiry and was only available in the yearly format beginning with 1991, this study will contain no payment accuracy data.

Unavailability of Data Prior to 1984

The next limitation concerns the unavailability of processing time data prior to 1984. The original intent of this research was to use all years from the inception of mandated performance appraisals, 1981, through 1990 when the process was drastically changed by then Commissioner King.

However, a search of the records indicates that no processing time data is available prior to April 1984. While this is a limitation, it may be acceptable on the grounds that the implementation of a vast and new program like the performance appraisal system would have required a few years to achieve.

Unavailability of Performance Ratings

The final limitation of this study involves the unavailability of actual performance ratings to analyze with the processing time data. This would have allowed actual triangulation of the processing times. The Social Security Administration would not release this information citing the Privacy Act as a reason.

While these limitations do relate to the proposed research, it is unlikely that any of these limitations interfere with the relationship between the RSI and DI processing times and the performance appraisal criteria.

CHAPTER VI

Findings

This chapter presents the findings which can be substantiated from the data analysis. The findings deal with both the qualitative and quantitative aspects of the research. These findings will be used in the final section of this work to provide conclusions and recommendations.

Content Analysis

The results of the content analysis are summarized both statistically and graphically. The statistical output will be summarized here. For graphics see Appendix G.

The position types for each set of performance appraisal criteria were collapsed into three categories; claims, service and clerical. The frequency for each research attribute, quantitative, qualitative, positive and negative was totalled for each position type.

This resulted in a three (position type) by four (research attribute) cell table of nominal data suitable for Chi-square analysis. The frequencies for the 1981 performance appraisal criteria were converted to percentages and used as expected values in the Chi-square calculation.

The 1986 set of performance appraisal criteria were also converted to percentages for use in Chi-square analysis.

sis. Having standardized the performance appraisal criteria sets, Chi-square was computed using the first set as expected values and the second set as observed values. This would reveal whether or not any statistically significant difference existed between the two sets in terms of the research attributes. Chi-square was computed to be 111.89. A precise step by step progression for this calculation is given in Appendix K.

There were six degrees of freedom for this table. At the .05 level of significance, the table value for comparison to the Chi-square is 12.592. Since the value of Chi-square exceeds the table value, it can be concluded that a significant difference does exist between the first and second sets of performance appraisal criteria based on the research attributes of quantity, quality, positive and negative references. See Appendix H for the actual frequency counts, converted percentages and cell residuals resulting from the calculation.

The graphic display of this difference takes the form of stacked graphs. The darker shaded area in each of the graphs represents the negative and quantitative aspects of the sets of performance appraisal criteria respectively while the lighter shading represents the positive and qualitative aspects respectively.

The 1981 set of performance appraisal criteria had the following research attributes. The total number of attri-

butes in the set was 694. Of this total, 343 were quantitative while 220 were qualitative. Eighty-two were negative while the remaining forty-nine were positive.

In only quantitative versus qualitative terms, the 1981 set of performance appraisal criteria was more quantitative than qualitative. The set was 61% quantitative and 39% qualitative. Thus the set was explicitly more quantitative than qualitative.

In terms of positive versus negative research attributes, the set was more negative than positive. The set was 63% negative and 37% positive. This indicates that the set was explicitly more negative than positive. Although the percentages differed, the 1986 set exhibited the same basic configuration.

The 1986 set revealed a total of 5,109 research attributes. Of this total, 3,183 were quantitative while 1,057 were qualitative. The total number of quantitative and qualitative research attributes was 4,240. In percentage terms, the set was 75% quantitative and 25% qualitative. This indicates that this set was predominantly more quantitative than qualitative.

In terms of positive versus negative research attributes, the set had a total of 869 found. Of this total 772 were negative and 97 were positive. Expressed in percentages, the set was 89% negative while being only 11% positive. This set of research attributes, as was the 1981

set, was more negative than positive.

In summary, both sets of performance appraisal criteria were more negative than positive as operationally defined in this research. Both sets of performance appraisal criteria were also more quantitative than qualitative. It is vital to this research that it be noted that although both sets were more quantitative and negative, the difference between them was statistically significant.

The next step in the content analysis was to compare each set of claims taking GJT's to its respective set of GJT's as a whole. This procedure allowed analysis of the representativeness of the claims taking GJT's to the set of GJT's as a whole.

The results of these comparisons demonstrated that a good, general representativeness existed for each set of criteria. Appendix H summarizes the actual counts and percentages for each set of claims GJT's as compared to the entire set of GJT's, a brief written summary is provided here.

The 1986 criteria as a whole was 75% quantitative and 25% qualitative. The claims taking GJT's were 75% quantitative and 25% qualitative. The 1986 criteria as a whole was 89% negative and 11% positive. The claims taking GJT's were 83% negative and 17% positive.

The 1981 criteria as a whole was 61% quantitative and 39% qualitative. The claims taking GJT's were 65% quanti-

tative and 35% qualitative. The 1981 criteria as a whole was 61% negative and 39% positive. Meanwhile, the claims taking GJT's were 55% negative and 45% positive.

The content of each set of claims taking GJT's is very generalizable to its respective set as a whole with the quantitative and qualitative aspects of the 1986 set being an exact match. The greatest variance is found in the 1981 set on the issues of positive and negative with a 7% divergence. In all, the representativeness of the claims taking GJT's to their respective sets as a whole supports the generalizability of the results of the content analysis of the Title II data to the organization as a whole.

This operation completes the qualitative analysis portion of the research. In summary, the two sets of performance appraisal criteria are significantly different when considering the research attributes of quantity, quality, positive and negative references.

This conclusion is important to the hypothesis testing portion of the quantitative analysis which follows. If there had been no difference between the sets, then quantitative testing of performance for each time period, as required by hypotheses one through four, would have been unnecessary.

Also, the claims taking GJT's are representative of their respective sets of performance appraisal criteria as a whole. This is important to the generalizability of the

study. This premise is contained in assumption five as provided in the methodology section of this research.

Quantitative Analysis

The quantitative analysis required several sequential steps for completion. After obtaining the RSI and DI performance data, regression analysis was used to determine statistical independence of the processing times from the claims load. Any trends which developed within or between the sets of performance appraisal criteria would be revealed during the regression procedure. Then, t-tests were used to determine any statistically significant differences which may exist between the two sets of performance data.

There are two important aspects about the quantitative analysis in this research that should be borne in mind upon reading. One deals with the exploratory intent of the study and the other deals with the fact that this is a population study.

The first important aspect is that this is an exploratory study. This is not an explanatory study and although it utilizes interval level data, no cause and effect is inferred. This study seeks to determine whether or not a relationship exists, not determine the cause and effect of the observed results.

The second important factor to bear in mind is that this is a population study. All RSI and DI claims during

the analysis period are included in the study. This means that no sampling has been done. Thus, the quantitative techniques generate parameters, not statistics. There are no confidence intervals given in this analysis for this reason. Confidence intervals are involved when population means must be estimated from sample means. That is not necessary in this study. Bearing these in mind, the chosen quantitative techniques revealed the following parameters.

Statistical Independence Testing

The initial step in the quantitative analysis was to determine whether or not the processing times for RSI and DI claims were dependent on the number of claims taken on a monthly basis. If the RSI overall count (total number of claims taken in any one month) correlated with the RSI mean time for processing at .65 or above, then they would be considered dependent and would have to be factored into any further analysis. The same procedure was undertaken for DI overall count and DI mean time.

The RSI overall counts were entered in a raw data file in a six digit format and the RSI mean time was entered in a three digit format.

The RSI overall count and RSI mean times correlated at only .08799. This indicates that the monthly RSI mean time was not significantly correlated to the total number of RSI claims taken on a monthly basis. Thus, the RSI mean time

would be considered to be statistically independent for the analysis which was subsequently undertaken.

The DI overall counts were entered in a raw data file in a six digit format and the DI mean time was entered in a three digit format. The DI overall count and DI mean times correlated at only at .08320. This indicates that monthly DI mean time was not significantly correlated to the total number of DI claims taken on a monthly basis. Thus, the DI mean time would be considered statistically independent for the analysis which was then performed.

Hypothesis Testing

There were eight specific hypotheses to be tested in this research. These hypotheses were exhaustive of the relationships possible in the data. Each of the eight will be discussed separately here and a statement of acceptance or rejection will be made.

For those hypotheses related to the 1981 set of performance appraisal criteria only, the first thirty months of data were used. Thus, for intra-criteria analysis for 1981 hypotheses $N=30$. For the intra-criteria analysis of the 1986 set, forty-eight months of data are used. For those hypotheses involving only the 1986 criteria, the $N=48$. For those hypotheses which involve both sets of criteria, the $N=78$.

It should be noted that the t -test was a one-tailed

test since the critical region of acceptance or rejection of the hypothesis is on the right or positive side of the distribution. The cumulative results of all of these sets of hypotheses will be discussed in the conclusions and recommendations chapter. All raw data is found in Appendix I. While the trend lines for these raw data sets are found in Appendix J.

The first hypothesis stated that the 1981 set of performance appraisal criteria was associated with no significant improvement in RSI processing time. Only the processing times for those months subject to the first set of performance appraisal criteria were used to test this hypothesis. This was accomplished by utilizing the "select if" procedure of SPSS and selecting PACS (performance appraisal criteria set) one. These cases were then subjected to regression analysis.

The regression procedure for hypothesis one resulted in a correlation coefficient of .15866. This indicated that the regression line was very weak. To determine whether or not any significance was found, a t -test for the regression line was completed. The t -test yield was a -.850. This result was not significant at the .05 level of significance.

The hypothesis that no significant improvement in RSI processing time was associated with the first set of performance appraisal criteria is accepted. The null hypo-

thesis that significant improvement did occur is rejected.

The second hypothesis stated that the 1981 set of performance appraisal criteria was associated with no significant improvement in DI processing time. Only the DI processing times for those months subject to the first set of performance appraisal criteria were used to test this hypothesis. The regression analysis for hypothesis two resulted in a correlation coefficient of .41127. This is a moderate association at best. To determine whether or not any significance was found, a t -test was completed for the regression line. The t -test yield was a -2.387. This t -test value was significant at the .05 level of significance, but in a negative direction. Thus, the second hypothesis that no significant improvement in DI processing time was associated with the first set of performance appraisal criteria is accepted. The null hypothesis that significant improvement did occur is rejected.

The third hypothesis stated that the 1986 set of performance appraisal criteria was associated with no significant improvement in RSI processing time. Only those months subject to the second set of performance appraisal criteria were used to test this hypothesis.

The regression analysis for hypothesis three resulted in a correlation coefficient of .23719. This is an weak correlation. To determine if any significance was associated with this regression line, a t -test was completed. The

t -test yield was -1.656. This result was not significant at the .05 level of significance. Thus, the third hypothesis is accepted. The null hypothesis that significant RSI processing time improvement is associated with the second set of performance appraisal criteria is rejected.

The fourth hypothesis stated that no significant improvement in DI processing time is associated with the second set of performance appraisal criteria. Only those DI processing times for months subject to the second set of performance appraisal criteria were used to test this hypothesis.

The regression analysis for hypothesis four resulted in a correlation coefficient of .03685. This indicates that virtually no linear relationship exists. To determine if any significance was associated with this regression procedure, a t -test was completed. The result was a -.250 which was not significant at the .05 level of significance. Thus the fourth hypothesis is accepted. The null hypothesis that the second set of performance appraisal criteria was associated with significant improvement in DI processing time is rejected.

The fifth hypothesis stated that both sets of performance appraisal criteria together were not associated with significant improvement in RSI processing time. All months were used to test this hypothesis.

The regression analysis produced a correlation coef-

ficient of .20067. This indicates an extremely weak linear relationship. To determine if any significance was associated with relationship a t-test was completed. The yield was a -1.786 which was not significant at the .05 level of significance.

The fifth hypothesis which holds that no significant improvement in RSI processing time is associated with both sets of performance appraisal criteria is accepted. The null hypothesis that significant improvement in RSI time is associated with these two sets of performance appraisal criteria is rejected.

The sixth hypothesis held that both sets of performance appraisal criteria together produced no significant improvement in DI processing time. The regression produced a correlation coefficient of .19633. This indicates an extremely weak association. To determine whether or not any significance is associated with this regression, a t-test was performed. This yielded a -1.746 which was not significant at the .05 level of significance.

Thus, the sixth hypothesis that no significant improvement in DI processing time is associated with both sets of performance appraisal criteria is accepted. The null hypothesis that significant improvement in DI processing time was associated with both sets of performance appraisal criteria is rejected.

The seventh hypothesis was that there was no signifi-

cant difference in RSI processing time between the two sets of performance appraisal criteria. The RSI processing times associated with the first set were compared to the RSI processing times associated with the second set. This comparison was made via t-test. Using the separate variance estimates, the t-test value was 5.86. This was significant at the .05 level. This indicates that significant difference did occur in RSI processing times between the sets. The hypothesis is rejected and the null is accepted. In the Conclusions and Recommendations chapter, a possible explanation for this significant difference will be given.

The final hypothesis was that no significant difference in DI processing times exists between the sets of performance appraisal criteria. The DI processing times associated with the months in the first set were compared via t-test to the DI processing times for those months in the second set. Using separate variance estimates, the t-test value was -1.52. This was not significant at the .05 level. This indicates that there was no significant difference in processing times between the sets of performance appraisal criteria. The hypothesis is accepted. The null hypothesis that there was a significant difference in DI processing times between the two sets of performance appraisal criteria is rejected. This completes the hypothesis testing portion of the research.

In summary, seven of the eight hypotheses were

accepted. The intent of this study was to determine association. By using interval level data these associations have been tested. A summary will be provided in the Conclusions and Recommendations chapter.

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

Overview

This section will provide a set of conclusions and recommendations to the Social Security Administration in particular and more generally to Congress which initiated this system. The conclusions are based on both the qualitative and quantitative aspects of the research. The conclusions will serve as a basis for the recommendations.

Conclusions

The conclusions offered as a result of this research are both quantitative and qualitative in nature. While some are derived specifically from the hypothesis testing, others are the result of the qualitative analysis or from general evidence provided throughout the research.

Quantitative Analysis Conclusions

These conclusions deal specifically with the results of the hypothesis testing portion of this research. The results of the Chi-square analysis are discussed later in the qualitative analysis conclusions part of this chapter.

Neither set of performance appraisal criteria were

associated with significant improvement in processing times for Disability Insurance claims separately. There was no improvement in DI processing time within either the first or second set of criteria.

Thus, the conclusion here is that neither the first set of performance appraisal criteria, which was somewhat more positive and qualitative, nor the second, which was more quantitative (and supposedly more objective as a result) and more negative, was associated with any improvement in performance as measured by DI processing times from 1984 through 1990.

Neither set of performance appraisal criteria separately was associated with significant improvement in processing times for Retirement and Survivor's claims. There was no improvement in RSI processing time within either set of performance appraisal criteria.

Thus, the conclusion here is that, neither set, regardless of the configuration of positive, negative, quantitative or qualitative attributes, was associated with improved performance as measured by RSI claims processing time.

The sets of performance appraisal criteria taken together were not associated with improved performance in either RSI or DI claims. It must be remembered that the taking of these types of claims constitutes the vast majority of the work performed by the Social Security Admini-

stration. The conclusion here is that the majority of the work completed by SSA, as measured by the performance appraisal system, has not improved even though very different approaches to performance appraisal have been tried. Perhaps no greater empirical referent for the triumph of technique over purpose exists than this.

There was no significant difference between the two sets of criteria as they are associated with DI claims taking and adjudicating. The conclusion reached here is that one set was not more effective than the other. While there were definite and demonstrable differences between the two sets, neither was associated with improved performance between the two.

There was a significant difference in RSI processing times between the two sets. However, other factors could have contributed to this associative difference. Since this is simply an exploratory study, no reason for significant difference will be given, however, an additional factor will be provided for any subsequent research which may be conducted along these lines.

During the time period of this study, the Field Office Systems Enhancement (FOSE) project was implemented. This involved the purchase, installation and use of several thousand computer terminals. The Terminal Acquisition Project (TAP) portion of FOSE had these terminals in place and working by 1990. It is possible that this dramatic in-

crease in microcomputing capability could account for the significant difference in RSI processing times between the two sets.

As this study does not seek to establish cause and effect, this development is raised only as a possible factor which may have influenced the rejection of this hypothesis and which should be taken into account if similar concerns are raised in the future.

The purchase of these computers occurred over all of the years of the study. As these computers were operationalized, the processing times should have improved. If these computer purchases are factored into the RSI processing times at the rate of .0164 per month in a cumulative manner, the t -test for RSI processing time is a -13.31 which is significant at the .05 significance level. It is important to realize that this is a negative result. This is not to imply that any negative result would have been produced, only to show that the factor of microcomputing may have had an effect on the outcome.

It should be noted that even without taking into account the advantages of microprocessing, seven of the hypotheses which stated that no significant improvement occurred, were accepted. Finally, when the advantages of microcomputing were taken into account, the eighth hypothesis, stating that there was no improvement in RSI processing time between the two sets, would also have been true.

Qualitative Analysis Conclusions

The qualitative analysis sought to explore two aspects of SSA's performance appraisal system. The first was to determine whether or not any difference existed between the 1981 and 1986 sets of performance appraisal criteria. The second aspect was to determine if any differences found were significant.

There were significant differences between the two sets of performance appraisal criteria. The 1986 set was significantly more negative and quantitative than its 1981 predecessor. The method for determining this was to use the qualitative research technique of content analysis on each set of performance appraisal criteria.

The results of the process were then analyzed via the Chi-square statistic for significance. There was a significant difference between the two sets of criteria on the issues of positive, negative, quantitative and qualitative attributes.

This makes clear that significantly different approaches to performance appraisal have been attempted by SSA to improve performance. The conclusion here is that neither set, both of which were more negative than positive and more quantitative than qualitative, are associated with any significant improvement in performance.

It must be remembered that the Office of Personnel

Management approved both of these sets of performance criteria. This is an indication that the attempts of the Social Security Administration to establish effective performance appraisal criteria were within the guidelines established by Congress. This is the rationale for those recommendations of this study which are directed toward Congress.

It is further concluded that SSA's two attempts to establish productive performance appraisal criteria are similar to the tone and intent of the authorizing legislation. The default and negative aspects of the system have been revealed throughout the research.

Stemming from the tone of the enabling legislation, Congress virtually assured the failure of the agency developed system by focusing on the negative and allowing default judgements instead of objectivity. Congress is also responsible for the multiple and conflicting uses for which the system must be used. It is unlikely that the individual employee who is being evaluated for purposes of comparison against others will perceive the system as an appropriate counseling mechanism as Congress intended.

The agency bears responsibility for the failure of the system to be associated with improved performance. It developed a system which does not directly relate to the work actually completed. Also, SSA is responsible for not taking into account the types of work and the facets of the

work while constructing appraisal criteria. It is also the operational responsibility of Social Security that the system has a default focus. This is an agency choice, within the Congressionally established guidelines, plain and simple.

It was also the Social Security Administration's choice to focus on the quantitative aspects of the work. This had several debilitating effects on the workflow. A review of these was given in the introduction to this dissertation. This focus on quantity was removed from the performance appraisal system in 1990 in an effort to reinstate the concept of quality. The struggle continues today. The conclusion in this regard is that quantity has continuously been emphasized over quality between 1984 and 1990.

Thus, in the final summary, both the quantitative and qualitative analysis demonstrates that performance was not positively associated with the performance appraisal system as it existed between 1984 and 1990. It certainly seems that the technique of performance appraisal does not fulfill the most basic purpose for which it was intended. Perhaps this is the most important reason for reestablishing the dominance of purpose over technique. It is hoped that the following recommendations can facilitate the reestablishment.

Recommendations

There are five basic recommendations as a result of this research. Some are directed to the Social Security Administration while others are directed toward Congress. Some are operational in nature while others are research oriented. It is the author's fervent hope that these five basic recommendations, taken as a whole, may someday produce a performance appraisal system that facilitates improved performance and improved stakeholder satisfaction.

The most basic recommendation of this research is that Congress must not tie base pay to performance until an adequate means of assessing both the quantity and quality of performance for each individual exists. A review of the shortcomings of the system as it relates to measurement is appropriate to substantiate this recommendation.

The current performance appraisal system fails to measure several important aspects of the work. First, it does not measure all of the work completed by any individual employee. Thus, the employee could do required work but not receive credit for it. This is the most basic operational deficiency which should be remedied before Congress considers tying base pay to performance.

An additional operational aspect which must be corrected before Congress attempts to tie base pay to performance is that not all of the work completed counts for ap-

praisal purposes. The systems should be made as holistic as possible. Specifically, this should include the component of quality. This would require a drastic restructuring of the job descriptions and Generic Job Tasks. This step should be taken in conjunction with making sure that all work required receives credit. Until these shortcomings can be remedied, Congress should take no step to tie base pay to performance.

The second recommendation is that SSA's performance appraisals should be used in a more individual manner. While it is understood that the comparative aspect can not be eliminated, it is suggested that the focus become more balanced between the two. This recommendation is specifically provided to SSA since they have a great deal of latitude in changing the system as it currently operates. The present system has two congressionally specified uses. The utilitarian use is to compare for promotion, retention and so on. Meanwhile, the more egoistic use is to counsel the employee who is not performing well.

To this point, the predominant use of the system has been the comparison for promotion aspect. This is essentially the utilitarian aspect of the system. It seeks to do the greatest good for the greatest number by supposedly making objective comparisons of performance possible for everyone. Even if complete objectivity were possible, Nalbandian's research indicated that the cost in human

terms may be too high. As Gabris reminds us, performance appraisal is as much as human endeavor as it is a technical or systems endeavor.

It seems likely that Jeremy Bentham or John Stuart Mill would support the present system on utilitarian grounds. They would contend, as does Thompson, that individualism, or more properly egoistic hedonism, has no place in public administration. The utilitarian nature of the present system seems an overt attempt to revive, in an organizational context, the concept of filicific calculus.

It is unlikely that this attempt at filicific calculus will be any more successful than was Bentham's. This position is taken since the quantification of work in SSA has proven as difficult to quantify as were pain and pleasure for Bentham. This type of utilitarian thinking ignores the concept of quality and individualism.

Perhaps the system could be better accepted if the stakeholder aspects of it were emphasized. This would mean that each appraisal would have the individual employee's interests as a basis. First, the focus should be shifted from errors to positives whenever possible. It should be commonplace for the supervisor to have adequate means for assessing the actual percentage of rating points earned. If this could be done, with the supervisor in the counseling role instead of the utilitarian role, the system may improve the performance of the stakeholders, both

internal and external.

Secondly, the appraisal process should be future oriented. The argument made here is that the past can not be altered and that any sincere attempt to improve performance must take place in the future. To this end, subordinates and supervisors should agree on goals to be achieved by the subordinate which will improve his or her performance and allow the eventual rating to improve as performance does.

If the internal stakeholder (employee) performs better, then the external stakeholder (client) is the one who truly is the recipient of the change. Gwendolyn King made a direct call for compassionate service to the public. Perhaps this service must start with compassion expressed internally in the manners suggested here.

Thompson is revolted at the thought of compassionate public service. He believes that it is not possible or wanted in the bureaucratic setting. However, compassion when operationalized as recognition and reward of the individual's contributions to performance should not be unwanted nor unwelcome in the public service. Perhaps once compassion is attained internally, it can be extremely useful externally.

Having a supportive supervisor with an adequate measurement system who is interested in having improved performance for the sake of all stakeholders is possible.

This type of system would require much work, but the basics are available under the current system. The supervisors could engage more in the counseling role, it is in their job description. The process could become future oriented with the past as a guide concerning what type of support is needed. Finally, the limitations and abilities of each individual should be taken into account. It could be argued that they employer accepted these limitations and abilities upon hiring the individual and that the performance appraisal process is an appropriate process for recognizing this commitment.

Many of the recommendations given here could possibly facilitate this change in orientation to compassionate service. It is hoped that others will further define and make useful the concept of compassion in the public service. If this occurs then at least one objective of the research will have been attained.

An essential element of this is the focus on the employee's future performance, not the past mistakes. It should be part of the review process to set definite, measurable objectives to be accomplished by the next appraisal. Note that this is necessarily an individual undertaking since each employee will have differing strengths and weaknesses. Perhaps by concentrating positively on the future, individuals will perform better so that the agency can perform better overall.

The third recommendation is that the performance appraisal system should be decoupled from the promotion system. This is addressed to Congress at whose mandate the system is used in this way. At the present time, it is those employees who maximize their work credits who get promoted. It is recognized throughout the agency that the last two performance appraisals are the most significant determinant of promotion.

Anyone playing the game can appear to do substantially more work than the others with whom they are being compared for purposes of promotions. Therefore, this system virtually ensures that it is the best player of the game, not the best performer, who gets promoted. It is exactly this use of the system which may influence an agency's subordinates to pervert the system as the Commissioner's report concludes SSA employees have done.

There is no specific recommendation at this time concerning an appropriate replacement for the performance appraisal in the promotion process. That determination is out of the scope of this inquiry. Such a determination would require an analysis of the components used in promotion decisions and other analytical steps before a suitable replacement could be recommended. Perhaps if the present system were to be revamped, it could still have a place in promotions. However, the known shortcomings at the present time, make this usage of the system inappropriate.

In summary, the three operational recommendations that pay not be tied to performance, that the system become more compassionate and future oriented and that it be decoupled from the promotion system may produce a system which is more useful. These three operational recommendations may help reestablish the dominance of purpose over technique.

In addition to the operational recommendations just given, several recommendations for research are provided here. It is hoped that completion of these research initiatives could further enhance the system.

The most basic research needed should assess the enabling legislation in positive, negative, quantitative and qualitative terms. There are two specific reasons for recommending this agenda of research.

Most importantly, Congressional intent may be inferred from this research. The entire Civil Service Reform Act of 1978 should be reviewed in an attempt to determine what the true purposes of the Act were. The legislation should be analyzed on its own merits (no pun intended) to determine what true messages may be contained in the wording of the document.

Perhaps there is a semiotic relationship which will be revealed in completing this research. This type of relationship analyzes the surface meaning to assess the deeper meaning of any written or symbolic communication. If there is to be the reestablishment of purpose over

technique, then perhaps this piece of proposed research is a good genesis for a new system to ensure it.

The second research recommendation is dependent on the completion of the first. Once the Civil Service Reform Act has been analyzed for intent traits, then the performance appraisal criteria from all federal agencies could be gathered for analysis.

This research prospect has several interesting possibilities. One basic step would be to establish the intent traits from the enabling legislation and compare each agency's performance appraisal criteria to them. This would allow a broad analysis of what the individual agencies have done as a result of the Act. This would in effect become a usage distribution analysis. In addition, it could be used to develop profiles of usage based on other factors.

One such factor might be departments, such as Health and Human Services or the Department of the Interior. This could be done to determine how much the structure and function of the agency is a determinant of the effectiveness of performance appraisals.

Another factor which may be used to analyze usage distribution is agency goals or mission. The type of agency and the agency goals could be tested for goodness of fit with the attributes of the performance appraisal system.

The second use for this type of analysis would be to assess the actual performance each agency has experienced while using the various criteria and methods. This would provide a basis for assessing whether or not certain types of criteria were associated with improved performance regardless of the type of structure and function.

It may also be wise for Congress to do this since it now appears that performance appraisal systems as a whole can now be legally challenged. If it were determined that certain requirements of the Generic Job Tasks were routinely not met by a group or class of employees, then the federal government may be liable to pay damages to those who were denied or refused promotions based on the flawed system.

A similar fate befell the PACE exam which was used in the early 1980's as a basic assessment of ability. It was determined that certain classes of prospective employees were unfairly treated by the test. It is possible that the performance appraisal system could suffer a similar fate as this type of system comes under closer scrutiny as Congress readies the effort to tie base pay to performance.

In summary, these five recommendations are very broad and general. This is true since the system to which they apply is very broad and general. Some are directed toward the Social Security Administration while others are directed toward the United States Congress. Still others are

directed to future researchers in the area of performance appraisal.

The recommendations take in virtually every aspect of the performance appraisal system of the Social Security Administration. From the nature of the enabling legislation to the possible outcomes associated with the system, all of it needs to be researched, reviewed and revamped. The recommendations, taken as a group, suggest that perhaps the time has come to "pull the plant up by the roots to see why it doesn't grow faster" as Brown (1982) feared. To further the analogy, based on all of the problems cited in the Commissioner's Report and the absence of association with any type of improvement in performance, perhaps the plant has been discovered to be a weed which Congress should uproot.

CHAPTER VIII

SUMMARY

This research has produced results and recommendations which should be helpful to future researchers in the area of performance appraisal in general and to those who sow and reap the harvest of the performance appraisal system in the Social Security Administration.

The performance appraisal system of the Social Security Administration is nurtured by default decisions about employee ratings which are derived from an essentially negative Congressionally created climate. The operational focus of the system is rooted in a soil which is quantitative almost to the exclusion of quality. It is a bottom line driven decision making process which is akin to a modern day attempt at organizational filicific calculus.

The system's primary use is utilitarian in nature since it is most frequently used for comparison purposes almost to the exclusion of counseling purposes. This amounts to weeding out those that don't measure up to the system's subjective and unproductive standards at the present time regardless of their future potential. This limits the compassion which can be expressed to stakeholders in the form of individuality. This is true whether the stakeholder is internal or external.

The system has not been associated with any improvement in performance of either RSI or DI claims at any time since 1984 when the microcomputing improvement in the claims system is taken into account. Thus, even technology does not help the plant grow as it presently exists.

It would seem clear after this analysis that operationally, productively, and normatively, this system has not been associated with any type of positive outcome. Perhaps the time to uproot the plant has arrived after thirteen years of non-positive outcomes. Whether the system remains planted or is uprooted to plant another seed, as Congress would surely do, a famous quote from Thomas Paine comes to mind.

The harder the conflict, the more glorious the triumph. What we attain too cheap, we esteem too lightly: it is dearness only that gives everything its value. I love the man that can smile in trouble, that can gather strength from distress and grow brave by reflection. 'Tis the business of little minds to shrink; but he whose heart is firm, and whose conscience approves of his conduct, will pursue his principles unto death. (Reader's Digest, February, 1993, p. 141).

Perhaps what has gone before in the realm of performance appraisal has been attained too cheaply. Perhaps only the surface meaning of performance has been utilized. Perhaps the system has been taken too lightly, arranged with only compliance and not dearness at heart.

It certainly appears that performance appraisal will

be an organizational conflict for perpetuity. As we face the current troubles of performance criteria which can not be accurately measured, work which must be done and not credited and so on, it must be remembered that the opportunity for glory exists.

The strength gathered in the face of this distress must result from a focus on research and commitment to purpose, not technique. There must be continued research efforts concerning performance appraisal for the sake of individuality and compassion within organizations. This must be done while remembering that it is the purpose, not the technique, which will lead to triumph in the end.

Perhaps Wallace Sayre was ahead of his time with his famous quote that the federal personnel system is the triumph of technique over purpose. It seems more appropriate today than it did then. In Vice-President Gore's attempted reinvention of government, one of the cornerstones of the program is the elimination of the twenty thousand page personnel manual. Wallace Sayre would likely approve of this development.

While Victor Thompson may consider compassion inappropriate in the administration of a bureaucracy, he must now look at where the lack of compassion expressed as individual treatment of stakeholders, both internal and external, has lead us. If knowledge, understanding and fairness are forms of compassion, then perhaps they can

displace the dispassionate forms which presently drive the system at present, rules, ratings and comparison.

Perhaps this research can plant a new seed. While there must be a utilitarian element to any system developed, perhaps it does not have to be the overwhelming strain in the hybrid that develops. While there do need to be rules to maintain the system, perhaps they can be more flexible and thereby more compassionate so as to encourage growth. While there must be some level of uniformity of treatment, perhaps the individual strengths of the plants can be recognized and balanced with a positive bottom line in the future as the goal for all. When the plant grows in this way, then will be found the triumph of purpose over technique.

Appendix A
Exclusions From Performance Appraisal

These seven specific exclusions from the requirement for performance appraisal are given in the Civil Service Reform Act of 1978. The exclusions are provided here to ensure the understanding that the domestic field office personnel of the Social Security administration are not included within their scope.

The specific exclusions are:

1. An employee outside the United States who is paid in accordance with local native prevailing wage rates for the area in which employed.
2. An individual in the foreign service of the United States.
3. A physician, dentist, nurse or other employee in the Department of Medicine and Surgery, Veterans' Administration whose pay is fixed under chapter 73 of title 38.
4. An administrative law judge appointed under section 3105 of this title.
5. An individual in the Senior Executive Service
6. An individual appointed by the President
7. An individual occupying a position not in the competitive service excluded from coverage of this subchapter by regulations of the Office of Personnel Management.

There are no domestic field office personnel covered under any of these seven exceptions. Administrative law

judges work in the Office of Hearings and Appeals, not in field offices. No other exclusion pertains to the Social Security Administration directly.

Appendix B
Summary Sheet for Performance Appraisals

PART III—RATING OF RECORD

A. NAME OF EMPLOYEE (LAST, FIRST, MI.)

* B. RATING LEVEL ASSIGNED
(Check one only)

- ☐ LEVEL 1—UNACCEPTABLE
☐ LEVEL 2—MARGINALLY SUCCESSFUL
☐ LEVEL 3—FULLY SUCCESSFUL
☐ LEVEL 4—EXCELLENT
☐ LEVEL 5—OUTSTANDING

* C. CHECK ONE AND FILL-IN

DATE(S)

1. ☐ Sept 30 rating for appraisal year ending

mo	day	yr
09	30	--
2. ☐ Extended appraisal period ending

mo	day	yr
--	--	--

For appraisal year ending

mo	day	yr
09	30	--
3. ☐ Special rating dated

mo	day	yr
--	--	--

(For within-grade increase)

* NOTE: Above entries must be repeated in Section III F on reverse side (for data entry purposes).

D. COMPLETE THIS PART IF RATING DERIVED BY CONSOLIDATING TWO OR MORE INTERIM (OR WIG) RATINGS

1. Show date(s) of the interim (or w/g) rating(s) recorded on separate forms SSA-1303 →

MONTH	DAY	YEAR

MONTH	DAY	YEAR
2. Describe method used to consolidate

E. RATING OF RECORD CERTIFICATIONS

1. Appraising Official—Signature

2. Reviewing Official—Signature

3. Awards Manager—Signature
when necessary or check
if approval documented
elsewhere ☐

Date

Date

Date

4. Employee verification
of receipt

Signature

Date

NOTE: A FULLY SUCCESSFUL RATING IS REQUIRED FOR A
WITHIN-GRADE INCREASE

PART II—Continued

Use this space for more documentation or comments, if necessary

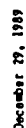
Appendix C
Example of Completed Appraisal

SSA EMPLOYEE PERFORMANCE MANAGEMENT SYSTEM PERFORMANCE RATING (See PMS Chapters 430-1 and 430-3)						PLEASE TYPE OR PRINT CLEARLY IN BLACK INK																																					
PART I - EMPLOYEE IDENTIFICATION																																											
A. NAME OF EMPLOYEE (LAST, FIRST, MI.) WRIGHT DONALD W					B. SSN 376-70-6249																																						
PART II - INTERIM RATING																																											
A. POSITION IDENTIFICATION																																											
1. Position Title SERVICE REPRESENTATIVE					2. Position Number 03C3170		3. Series and Grade GS 962 08																																				
4. Organization (Office, Division, Branch, etc.) 3 MT PLEASANT MI DIST OFC																																											
B. GJT RATINGS																																											
FROM (date plan issued) → 10 01 92				TO (ending date) → 09 30 93																																							
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>1. GJT Number</th> <th>2. Critical or Non-Critical</th> <th>3. Weight (1-5)</th> <th>4. Rating Assigned (1-5) Appraising Official</th> <th>5. Reviewing Official</th> <th>6. Weighted Score (Wt. X Rtg.)</th> </tr> </thead> <tbody> <tr><td>7</td><td>N</td><td>2</td><td>3</td><td></td><td>6</td></tr> <tr><td>14</td><td>N</td><td>2</td><td>3</td><td></td><td>6</td></tr> <tr><td>71</td><td>C</td><td>4</td><td>3</td><td></td><td>12</td></tr> <tr><td>117</td><td>C</td><td>4</td><td>3</td><td></td><td>12</td></tr> <tr><td colspan="6">TOTAL WEIGHT → 12 TOTAL WEIGHTED SCORE → 36</td></tr> </tbody> </table>				1. GJT Number	2. Critical or Non-Critical	3. Weight (1-5)	4. Rating Assigned (1-5) Appraising Official	5. Reviewing Official	6. Weighted Score (Wt. X Rtg.)	7	N	2	3		6	14	N	2	3		6	71	C	4	3		12	117	C	4	3		12	TOTAL WEIGHT → 12 TOTAL WEIGHTED SCORE → 36						C. INTERIM SUMMARY RATING COMPUTATION 1. Preliminary Computations a. Enter ees total weighted score for all GJTs from B7 36 b. Enter the total weight of all GJTs on which ee was rated from B6 12 c. Multiply a x b to compute ees highest possible total weighted score 60 d. a - b = This is the ees percentage score used to determine the summary rating in Block 2 below 60 2. Summary Rating Level (Check one box) <input type="checkbox"/> 1-UNACCEPTABLE- Rated level 1 on any critical GJT <input type="checkbox"/> 2-MARGINALLY SUCCESSFUL- Rated level 2 on any critical GJT or percentage score of less than 60% <input checked="" type="checkbox"/> 3-FULLY SUCCESSFUL- Percentage score of at least 60% but less than 80% <input type="checkbox"/> 4-EXCELLEN- Percentage score of at least 80% but less than 95% <input type="checkbox"/> 5-OUTSTANDING- Percentage score of 95% or higher			
1. GJT Number	2. Critical or Non-Critical	3. Weight (1-5)	4. Rating Assigned (1-5) Appraising Official	5. Reviewing Official	6. Weighted Score (Wt. X Rtg.)																																						
7	N	2	3		6																																						
14	N	2	3		6																																						
71	C	4	3		12																																						
117	C	4	3		12																																						
TOTAL WEIGHT → 12 TOTAL WEIGHTED SCORE → 36																																											
B. Use this space to document any GJT Rating of 1, 2, or 5 (Continue On Reverse if Needed)																																											
D. INTERIM RATING CERTIFICATIONS (Not Required if Rating of Record Completed Concurrently)																																											
1. Appraising Official-Signature				Date		2. Reviewing Official-Signature																																					
3. Performance Award Manager-Signature				Date		4. Employee Verification of Receipt-Signature																																					
						Date 01/08/93																																					
E. EMPLOYEE COMMENTS (Continue on Reverse if Needed)																																											
PART III-FROM REVERSE SIDE-CONTINUED																																											
F. AFTER COMPLETING PARTS III.B AND III.C., ENTER THE FOLLOWING (For Data Entry Purposes)																																											
1. Rating Level From Part III.B (1-5) 3		2. Number Next to Box Checked in III.C. (1-3) 1		3. Earliest Date Entered in III.C.		MO DAY YR 09 30 93																																					

Form SSA-1303TR (4-89)

PART III-RATING OF RECORD														
A NAME OF EMPLOYEE (LAST, FIRST, MI) WRIGHT DONALD W														
* B RATING LEVEL ASSIGNED (Check one only) <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"> <input type="checkbox"/> LEVEL 1-UNACCEPTABLE <input type="checkbox"/> LEVEL 2-MARGINALLY SUCCESSFUL <input checked="" type="checkbox"/> LEVEL 3-FULLY SUCCESSFUL <input type="checkbox"/> LEVEL 4-EXCELLENT <input type="checkbox"/> LEVEL 5-OUTSTANDING </div>	* C CHECK ONE AND FILL-IN DATE(S) <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 5px;"> <tr> <td style="width: 70%;">1. <input checked="" type="checkbox"/> Sept. 30 rating for appraisal year ending</td> <td style="width: 30%;">09 30 93</td> </tr> <tr> <td>2. <input type="checkbox"/> Extended appraisal period ending For appraisal year ending</td> <td>09 30 </td> </tr> <tr> <td>3. <input type="checkbox"/> Special rating dated (For within-grade increase)</td> <td>09 30 </td> </tr> </table>		1. <input checked="" type="checkbox"/> Sept. 30 rating for appraisal year ending	09 30 93	2. <input type="checkbox"/> Extended appraisal period ending For appraisal year ending	09 30	3. <input type="checkbox"/> Special rating dated (For within-grade increase)	09 30						
1. <input checked="" type="checkbox"/> Sept. 30 rating for appraisal year ending	09 30 93													
2. <input type="checkbox"/> Extended appraisal period ending For appraisal year ending	09 30													
3. <input type="checkbox"/> Special rating dated (For within-grade increase)	09 30													
* NOTE: Above entries must be repeated in Section III-F on reverse side (for data entry purposes)														
D COMPLETE THIS PART IF RATING DERIVED BY CONSOLIDATING TWO OR MORE INTERIM (OR WIGII) RATINGS														
1. Show date(s) of the interim (or WIGII) ratings recorded on separate forms SSA-1303 <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 5px;"> <tr> <td style="width: 25%;">MONTH</td> <td style="width: 10%;">DAY</td> <td style="width: 10%;">YEAR</td> <td style="width: 25%;">MONTH</td> <td style="width: 10%;">DAY</td> <td style="width: 10%;">YEAR</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>			MONTH	DAY	YEAR	MONTH	DAY	YEAR						
MONTH	DAY	YEAR	MONTH	DAY	YEAR									
2. Describe method used to consolidate														
E. RATING OF RECORD CERTIFICATIONS														
1. Appraising Official-Signature <div style="text-align: center; margin-top: 10px;"><i>A. Ghadiri</i></div> Date <u>10-7-93</u>	2. Reviewing Official-Signature <div style="text-align: center; margin-top: 10px;"><i>Thomas Edwards</i></div> Date <u>10/8/93</u>	3. Awards Manager-Signature when necessary or check <input type="checkbox"/> if approval documented elsewhere Date <u>10/8/93</u>												
4. Employee verification of record: <div style="text-align: center; margin-top: 10px;"><i>Donna Wright</i></div>		Date <u>10/8/93</u>												
NOTE: A FULLY SUCCESSFUL RATING IS REQUIRED FOR A WITHIN-GRADE INCREASE														
PART II-Continued														
Use this space for more documentation or comments, if necessary.														

Appendix D
Executive organization chart



Appendix E
Research Attributes

The material in this appendix identifies the specific research attributes which were utilized in the content analysis portion of this research. The following pages will group the attributes by category and then provide the specific research attribute along with the rationale for its inclusion in that particular category.

Quantitative Research Attributes

1. Digits - meaning any numeric representation of actual performance, such as level 3 performance.

Rationale: This research attribute is included since it makes the employee view any performance in terms of a quantitative outcome, regardless of the activity performed.

2. Acceptable Speed - meaning that the performance of the employee is being judged in ratio terms on some factor of job performance.

Rationale: It is the rapidity of motion or task completion which is quantified by this research attribute.

3. Accurate(ly) - means that some facet of the work is completed with a lower ratio of time or resources required.

Rationale: This research attribute involves many aspects of the work, such as typing, data input, form completion, evidentiary acceptance, follow-ups, etc. which are judged by how frequently corrections must be made. This is the quantification of errors.

4. Ahead of Schedule - means that some quantity of time was saved.

Rationale: This research attribute is a means of quantifying both work and time in a single concept.

5. Complete - means that any activity performed was done

fully.

Rationale: This research attribute quantifies activity. It bring to the employee the perception of quantity in that they may do work which is not complete and therefore requires more effort. It is the quantification of effort as it relates to many tasks which the employee must perform on an on going basis.

6. Current - means that all possible actions have been taken.

Rationale: This research attribute quantifies activity. There is an organizational focus on being current which is expressed in performance appraisals in terms such as "your work is 70% current".

7. Effectively - means that the action taken required no further activity and therefore no further time or resources.

Rationale: This quantifies both time and resources since either could be required if inefficiency on the part of the employee is found.

8. Efficiently - means that the resources necessary to achieve some outcome was kept to a minimum.

Rationale: This research attribute quantifies the resources required to perform any of the many tasks which SSA employees must complete.

9. Higher degree of...- means that some expectation of

percentage exists as it relates to the performance of a specific task.

Rationale: This research attribute quantifies activity. It is a ratio description of an activity such as "you required a higher degree of training." It raises the quantitative aspect of performance by treating it as a ratio.

10. Minimal - means that some basic level of achievement was observed.

Rationale: This research attribute quantifies effort or resources. For example, "the employee does minimal review of claims" or "the employee is minimally satisfactory."

11. Minimize - means to require as little as possible of time or resources.

Rationale: This research attribute quantifies the time or resources necessary to complete a task. It specifically makes the employee aware of the organizational need for expediency, such as "the employee should minimize delays in processing."

12. Promptly - means that every action should be taken as soon as possible.

Rationale: It quantifies the issue of time.

An example may be that "the employee promptly greets all visitors" without acknowledging that this may well

be impossible.

13. Timely - means that actions are taken at the correct time.

Rationale: This research attribute quantifies the accomplishment of activities. There are certain actions which can only be taken at certain times. When this is done, then the action is considered to be timely. It is expressed in performance appraisal as a percentage, such as "the employee took 90% of the actions on the case timely."

14. Statistical - means that the performance is reported by the employee for purposes of quantification directly.

Rationale: Each week the employee must prepare a report on the cases cleared that week within certain categories. These statistical reports become the basis for promotion, pay increase and award decisions. This research attribute quantifies performance in specific categories which are used for many quantitative purposes.

Qualitative Research Attributes

1. Adequate - means that the action taken was sufficient to accomplish the organizational goal or outcome.

Rationale: This research attribute concerns the quality of the work done, such as "the interviewer was adequate in his or her explanation."

2. Appropriate - means that the action taken was properly taken given the organizational situation.

Rationale: This research attribute concerns the usefulness of the action, such as "the employee took the appropriate steps to clear the claim."

3. Complex - means that the situation requires a more qualified effort to resolve.

Rationale: This research attribute concern the individual nature of each case presented, such as "the interview was complex with many issues to be resolved."

4. Courteous - describes the quality of the employee's conduct regardless of the quantity of actions.

Rationale: This research attribute deals with the interpersonal conduct of the employee, such as "treats visitors in a courteous manner."

5. Dignity - describes the respectfulness of attitude which the employee exhibits toward others.

Rationale: This research attribute deals with the the employee's attitude toward others, such as "the

employee respected the dignity of the caller."

6. Ease of - means that some task was made easier by the employee's actions.

Rationale: This research attribute deals with the qualitative nature of the work. While there is no quantification of the work, such as "more work is accomplished," this attribute deals with the fact that whatever work is accomplished is accomplished easier, such as "the employee facilitated the ease of retrieval."

7. Non-Routine/Unusual - means that the situation required unusual handling.

Rationale: This research attribute deals with the extenuating circumstances which arise in organizational interactions, such as "the claimant's deafness caused the interview to be non-routine."

8. Proper - means that the activity or method chosen met the requirements for the situation.

Rationale: This research attribute is qualitative in that it there is agreement between the needed activity and the task accomplished, such as "the employee took the proper action to complete the task."

9. Respect - means that the employee acknowledges the values of others.

Rationale: This research attribute is qualitative in that it refers to the qualitative ability of the

employee to respect others, such as "the employee respected the limitations of the client."

10. Skillfully - means that more than basic competence is exhibited by the employee in interpersonal terms.

Rationale: This research attribute is qualitative in that it deals with the quality of interpersonal interactions without regard to the number conducted.

11. Sufficient - means that the action taken is acceptable for the situation.

Rationale: This research attribute is qualitative in that the action taken has been adjudged as adequate, for example, "the employee did sufficient work to justify the use of overtime."

12. Tactfully - means that the employee was sensitive to the situation.

Rationale: This is qualitative in that it assesses the ability of the employee to be considerate of those with whom he or she must deal, such as "the employee dealt with the issue tactfully."

13. Tailored/Understandable/Easiest to understand/Clear - means that some aspect of performance was adjusted to the ability of another.

Rationale: This is qualitative in that the employee must be able to adjust his or her level of speech to the individual, such as "he tailored his speech to the claimant's level of understanding."

14. High Quality - means that some aspect of the work is recognized as being above average in quality terms. Rationale: This is qualitative since it can be used regarding an employee's performance without regard to quantity or frequency, such as "the employee did a high quality interview."

15. Unruly - means that a difficult situation was encountered with a difficult to deal with client. Rationale: This research attribute is qualitative in that it describes the quality of the interview regardless of quantity or frequency with which it occurs, such as "the employees controlled the unruly claimant for the time of the interview."

16. Realistic - means that the employee has an valid assessment of a situation. Rationale: This research attribute is qualitative in that it applies regardless of the quantity or frequency of the occurrence, such as "the employee sets and maintains realistic priorities."

17. Comprehensive - means that the claimant was treated in a holistic manner. Rationale: This is qualitative in that it describes a situation, such as "the employee treated the matter comprehensively." Note that there is no element of quantity.

18. Clear - means that either written or oral communication

was understood.

Rationale: Clarity in qualitative in that it is accomplished when one party understands another, regardless of the quantities of time and resources required.

Positive Research Attributes

1. Correct(ly) means that an assigned task was completed as directed.

Rationale: This research attribute is positive in that any task done correctly would reflect positively on the employees, such as, "the employee correctly accepted the evidence."

2. Independent(ly)/On Own Initiative/Takes Initiative/Needs no Assistance/Without Referral to Supervisor - are positive in the they indicate the employee is self motivated.

Rationale: This research attribute is positive in that it indicates that the employee can act on his or her own without the assistance of higher graded employees or supervision, such as "employee clears claims without...".

3. High level of Knowledge/Extensive Knowledge are positive in that they indicate the accumulation of useful job knowledge. Notice that no specific quantitative amount is attached to the attribute, such as "the employee uses a high level of knowledge in dealing with claimants."

4. Positive Benefits/Is Recognized/Improves Understanding/Positive Image are positive in that it assesses the effects of the employee's efforts on the public, such as "the employee's appearance creates positive

benefits in the community."

5. Successfully Fulfills is positive in that it raises the issue of successful accomplishment of a specified task by the employee.

Rationale: This is positive in that the concept of the employee successfully accomplishing a task is positively evaluated in performance appraisal, such as, "the employee successfully fulfilled the duties."

6. Toward the journeyman position is positive in that it indicates the employee is progressing professionally.

Rationale: This progression is positive in that the employee is exhibiting attributes which the organization values, such as "employee made progress toward the journeyman position during the review period."

7. Learns quickly is positive in that it indicates the employee assimilates job related information rapidly.

Rationale: This research attribute is positive in that the organization values the quick assimilation of information since the rules and regulations change continuously. This would be expressed in performance appraisal as "the employee learns new material quickly."

8. Does not require reviews is positive in that it indicates that the employee can operate without the review process being used.

Rationale: This attribute is positive in that the organization does not have to expend resources for review by higher graded employees or supervisors if the employees can operate without review. This is expressed in performance appraisals such as "the employee does not require reviews to learn new material."

Negative Research Attributes

1. Delays/neglects/untimely/remiss - is negative in that it means some specific task is not completed as expected.

Rationale: This research attribute is negative in that any delay means that some organizational goal or expectation was not met at the time expected, such as "the employee delayed the processing of inputs."

2. Difficulty - is negative in that it indicates that the employee is unable to easily grasp the job requirements or perform some task.

Rationale: This research attribute is negative in that the employee is made aware that they are not performing a task or function to the satisfaction of the supervisor. It is used in performance appraisal such as, "the employee has difficulty absorbing new material."

3. Error - is negative in that it makes the employee aware of mistakes.

Rationale: This research attribute raises the issue of mistakes which force a negative image into the relationship between the supervisor and the employee. It is used on performance appraisal such as "the error resulted in..."

4. Fails - is negative in that it means that the employee is incapable of some required activity.

Rationale: This is negative in that the employee is

being told that the performance is not acceptable and¹⁴⁵ must be corrected. It is used in performance appraisal such as "the employee fails to meet established goals."

5. Inaccurate - is negative in that the issues of mistake or error in detail is raised.

Rationale: Any inaccuracy is an indication that the employee is not performing with an acceptable attention to detail and is used in performance appraisal such as, "the employees inputs are inaccurate."

6. Inadequate/less than expected - is negative in that the issue of incomplete or ineffectiveness in performance is found.

Rationale: This research attribute raises that issue of the employee not being able to perform up to established expectations.

7. Inappropriate - is negative in that some aspect of the employee's behavior was not acceptable.

Rationale: Being inappropriate indicates the failure of the employee to understand the required behavior or task decision. It is used in performance appraisal such as "the employee's interviewing style was inappropriate."

8. Incomplete/Insufficient are negative in that the employee did not finish some required task.

Rationale: When the employee's work is insufficient or incomplete, the issue of being able to perform all parts of the job is raised. It is used in performance appraisal such as "the employee's work is often incomplete at the time of input."

9. Ineffective - is negative in that it raises the issue that the employee's actions were not able to resolve the situation as intended.

Rationale: Being unable or perceived as unable to effectively resolve a situation or perform a task gives the employee and the supervisor a negative perception of the subordinate.

10. Misrouted - is negative in that the use of the term indicates that the employee does not understand the organization.

Rationale: This research attribute gives the employee and the supervisor the impression that the agency material was not forwarded properly due to lack of knowledge or oversight. It is used in performance appraisal such as "the employee misrouted the claim which resulted in a longer processing time for the claim."

11. Needs direction/ongoing direction/repeated direction/ requires assistance/requires direction/requires guidance/receives direction/close observation by the supervisor/supervisory intervention - are

negative in that they raise the issue of the employee not being able to operate on his or her own without help.

Rationale: These attributes are negative in that they raise the issue of the employee requiring additional agency resources to accomplish agency specified goals or tasks. It is used in performance appraisal such as "the employee requires ongoing direction to meet level 2 performance."

12. Retraining/Follow-up training required - is negative in that the employee is requiring additional resources from the organization.

Rationale: These attributes are negative in that the employee requires additional effort or resources from the organization which is a negative perception by the supervisor. It is used in performance appraisal such as "the employee requires retraining to meet expectations."

13. Tentative/ Uncertain - are negative in that a basic inability of the part of the employee is focused on by the supervisor who chooses these descriptors of behavior.

Rationale: These attributes indicate that the employee exhibits the inability to reach a decision or to reach a decision as quickly as desired by the organization.

14. Limited knowledge - is negative in that it implies that

the employee does not have enough technical awareness to perform his or her job.

Rationale: This attribute is negative since it gives the employee and the supervisor the impression that the employee has not attained the level of knowledge expected by the organization.

15. Rarely shows initiative is negative in that it implies that the employee is not self motivated.

Rationale: The employee is perceived as unable to solve a problem or accomplish a task independently. It is used in performance appraisals such as "the employee rarely shows initiative in solving routine problems."

Appendix F
Position Description Groupings for Content Analysis

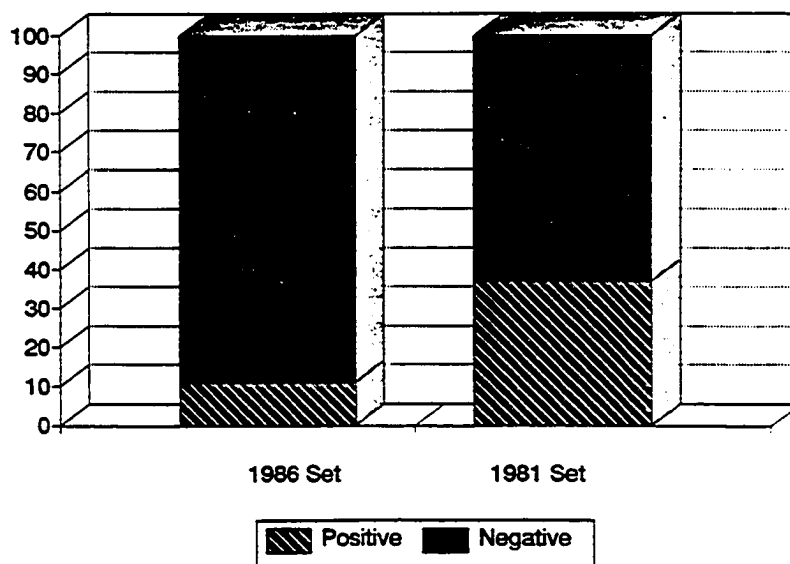
The specific positions descriptions were grouped as follows for purposes of analytical classification. These groupings correspond to the operational units in the field offices of the Social Security Administration.

Grouping	1981	1986
Clerical	Receptionist	Development Clerk (GS4)
	Data Transcriber	Development Clerk (GS3)
	Claims Clerk Typing	Development Clerk (GS2)
	Field Operations Clerk	
	Secretary Steno Typing	
	Secretary Steno Trainee	
Service	Data Review Technician	Service Rep (GS7)
	Data Review Trainee	Service Rep (GS6)
	(Tele) Service Rep	Service Rep (GS5)
	(Tele) Service Rep	Data Review Tech (GS7)
	Trainee	
	Service Rep/ Data	Data Review TEch (GS6)
	Review Tech	

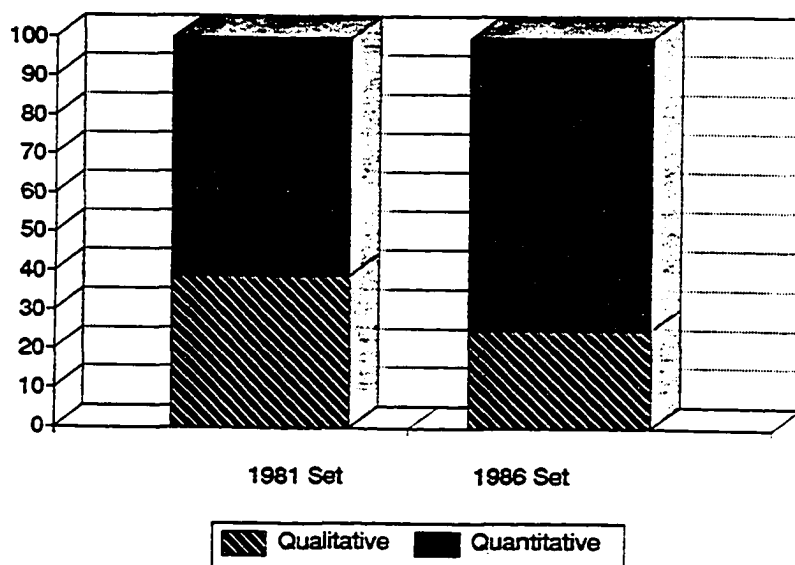
Claims	Lead Contact Rep	Field Rep
	Claims Rep	Claims Rep T2 (GS10)
	Claims Rep Trainee	Claims Rep T16 (GS10)
	Field Rep	Claims Rep T2 (GS9)
	Field Rep Trainee	Claims Rep T16 (GS9)
	Resident Rep	Claims Rep T2 (GS7)
	Operations Analyst	Claims Rep T16 (GS7)

Appendix G
Graphic Representation of Content Analysis Results

Comparison of 1981 and 1986 Sets



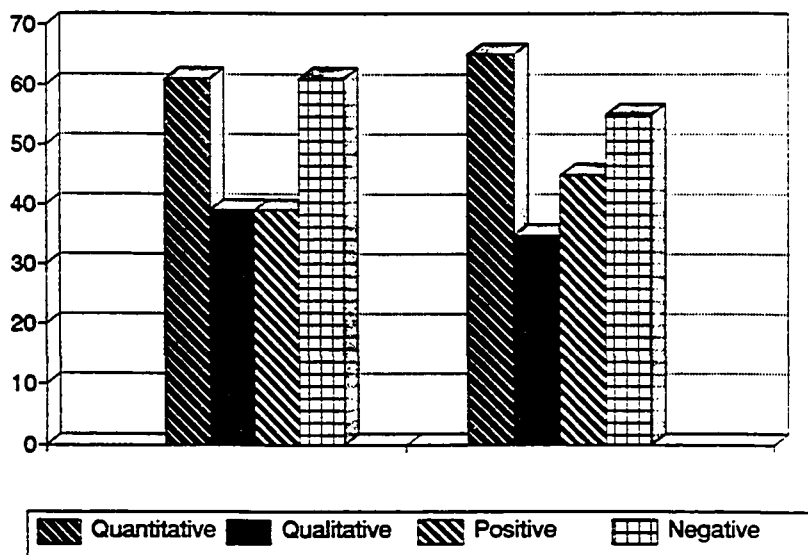
Comparison of 1981 and 1986 sets



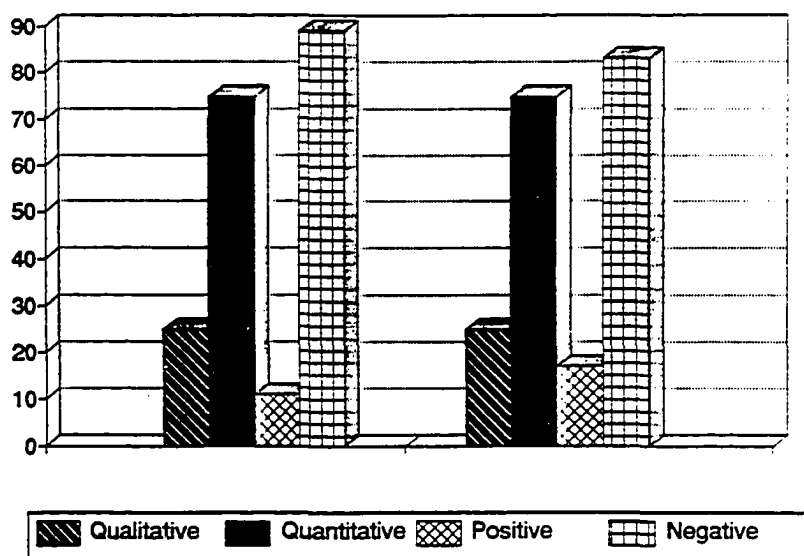
Appendix H

Comparison for Claims Taking Research Attributes to Overall Attributes - By Set

Comparison of All and Claims GJT's 1981



Comparison of All and Claims GJT's 1986



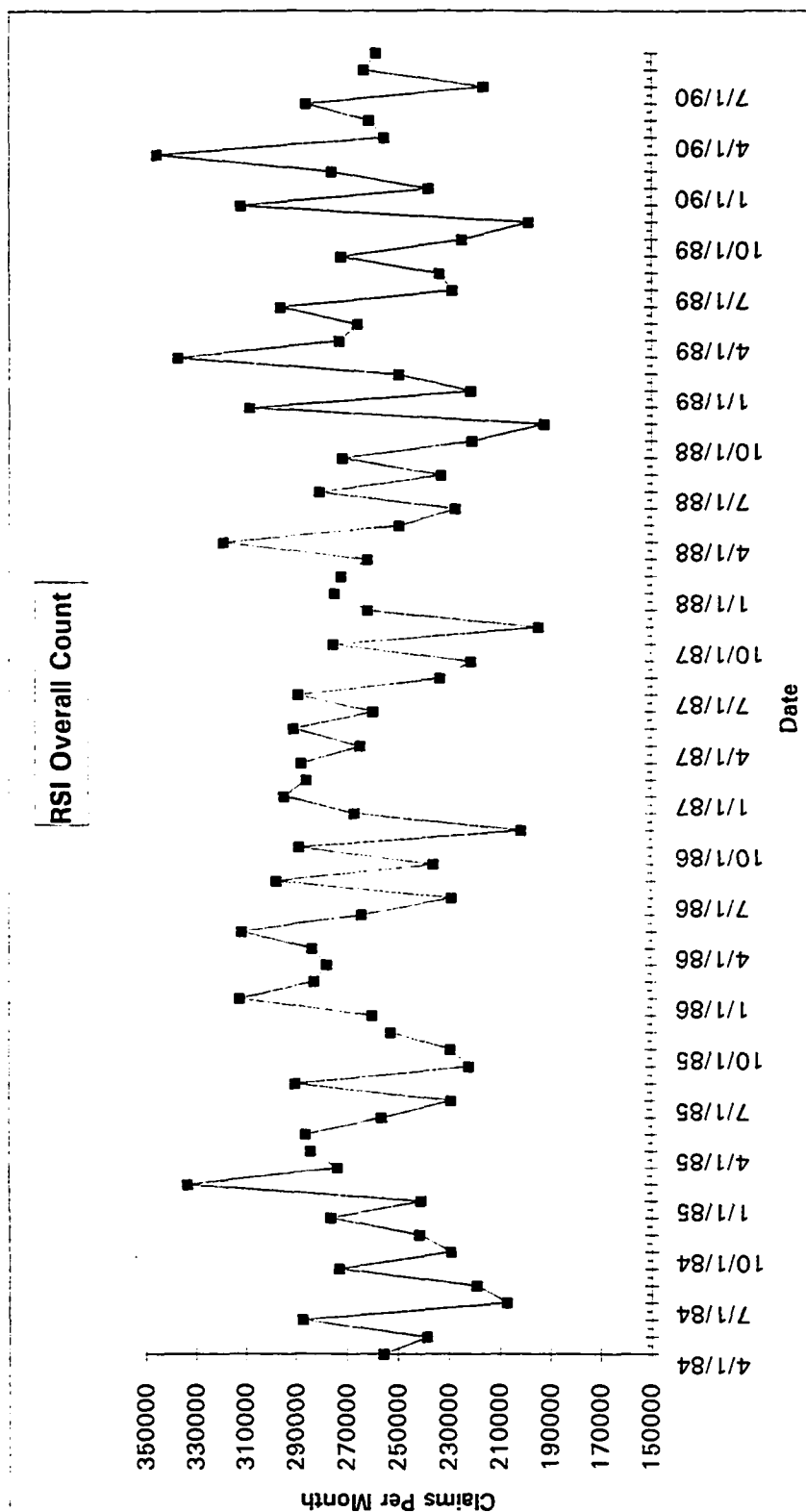
Appendix I

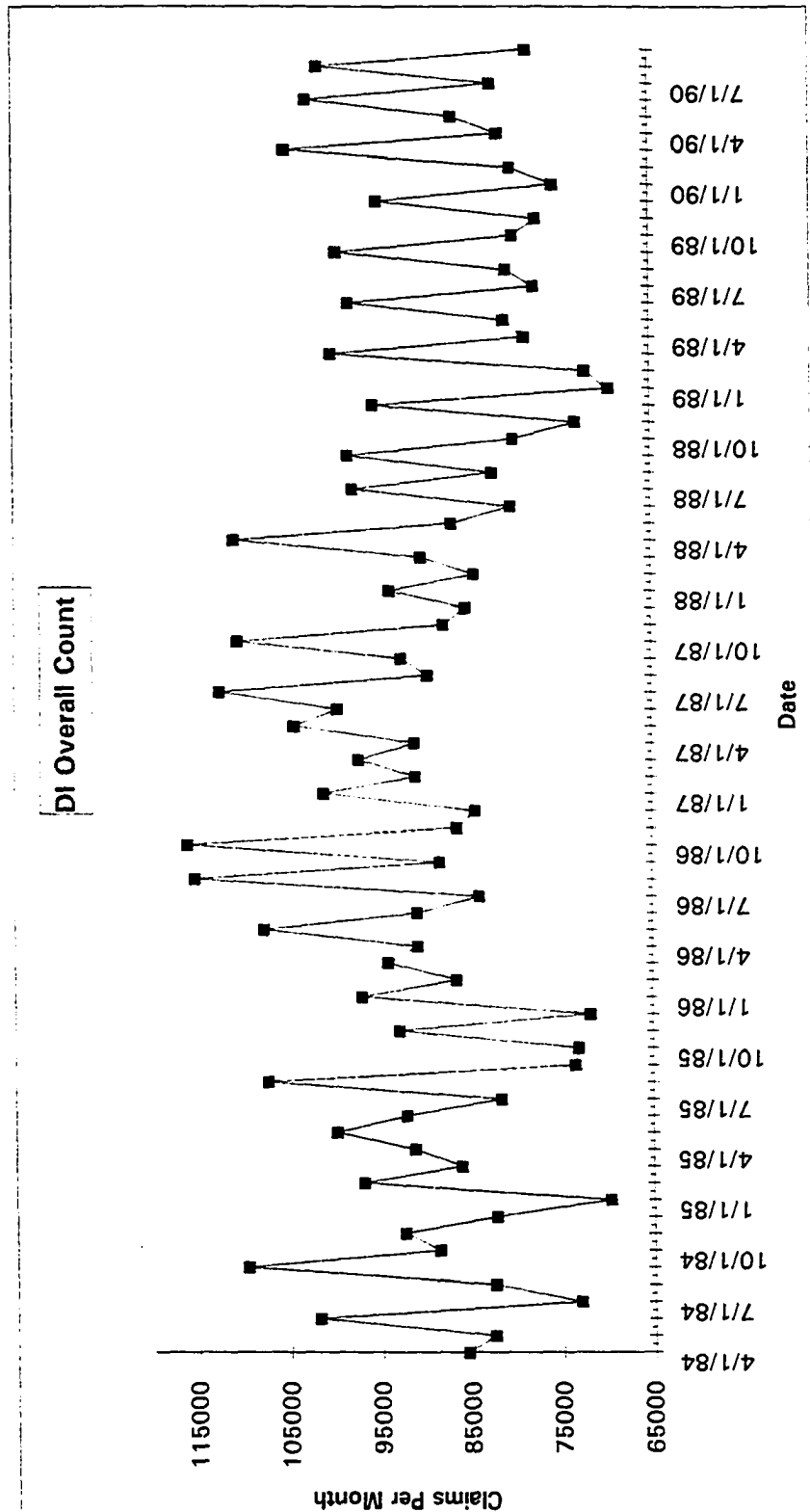
Raw Data for Monthly Processing Times and Claims Counts

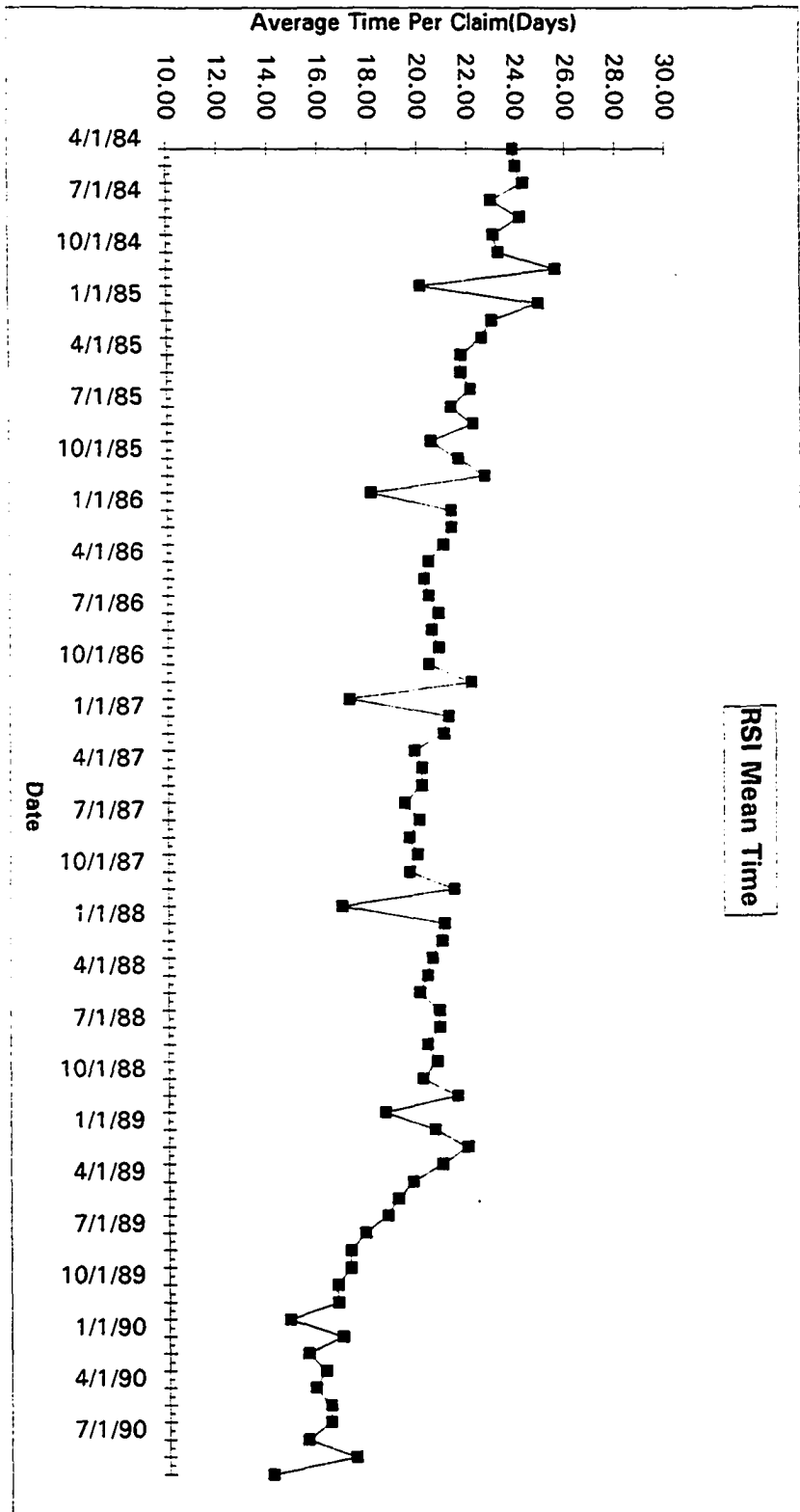
Date	RSI OC	RSI MT	DI OC`	DI MT
4/1/84	255381	23.90	85404	66.70
5/1/84	238527	24.00	82419	67.20
6/1/84	287190	24.30	101884	69.00
7/1/84	206997	23.00	72995	69.90
8/1/84	218940	24.20	82372	71.00
9/1/84	272864	23.10	109559	71.10
10/1/84	229127	23.30	88569	70.40
11/1/84	241387	25.60	92350	70.90
12/1/84	276377	20.10	82140	72.30
1/1/85	240772	24.90	69697	78.30
2/1/85	333291	23.00	96824	75.00
3/1/85	273271	22.60	85968	72.60
4/1/85	283999	21.70	91213	69.40
5/1/85	285976	21.70	99766	68.90
6/1/85	256238	22.10	92126	67.50
7/1/85	228960	21.30	81540	67.30
8/1/85	289982	22.20	107280	66.50
9/1/85	222034	20.50	73524	63.70
10/1/85	229112	21.60	73171	64.50
11/1/85	252624	22.70	92848	69.10
12/1/85	260060	18.10	71850	80.60
1/1/86	312383	21.30	96973	88.40
2/1/86	282416	21.30	86475	89.00
3/1/86	277644	21.00	94100	86.80
4/1/86	283438	20.40	90740	84.20
5/1/86	311577	20.20	107595	81.80
6/1/86	264042	20.40	90845	81.70
7/1/86	228816	20.80	83875	80.80
8/1/86	297650	20.50	115130	78.80
9/1/86	236012	20.80	88290	79.90
10/1/86	288533	20.40	115953	77.70
11/1/86	201521	22.10	86306	78.50
12/1/86	266956	17.20	84298	80.70
1/1/87	294444	21.20	101096	81.90
2/1/87	285549	21.00	90945	78.20
3/1/87	287567	19.80	97178	72.40
4/1/87	264543	20.10	91035	70.60
5/1/87	290786	20.10	104291	69.30
6/1/87	259603	19.40	99519	68.30
7/1/87	288983	20.00	112323	69.90
8/1/87	233466	19.60	89504	69.90
9/1/87	221146	19.90	92463	69.80
10/1/87	275235	19.60	110329	71.30
11/1/87	194470	21.40	87766	72.30
12/1/87	261698	16.90	85294	73.90
1/1/88	274412	21.00	93681	78.70
2/1/88	271969	20.90	84300	78.80
3/1/88	261813	20.50	90226	75.10

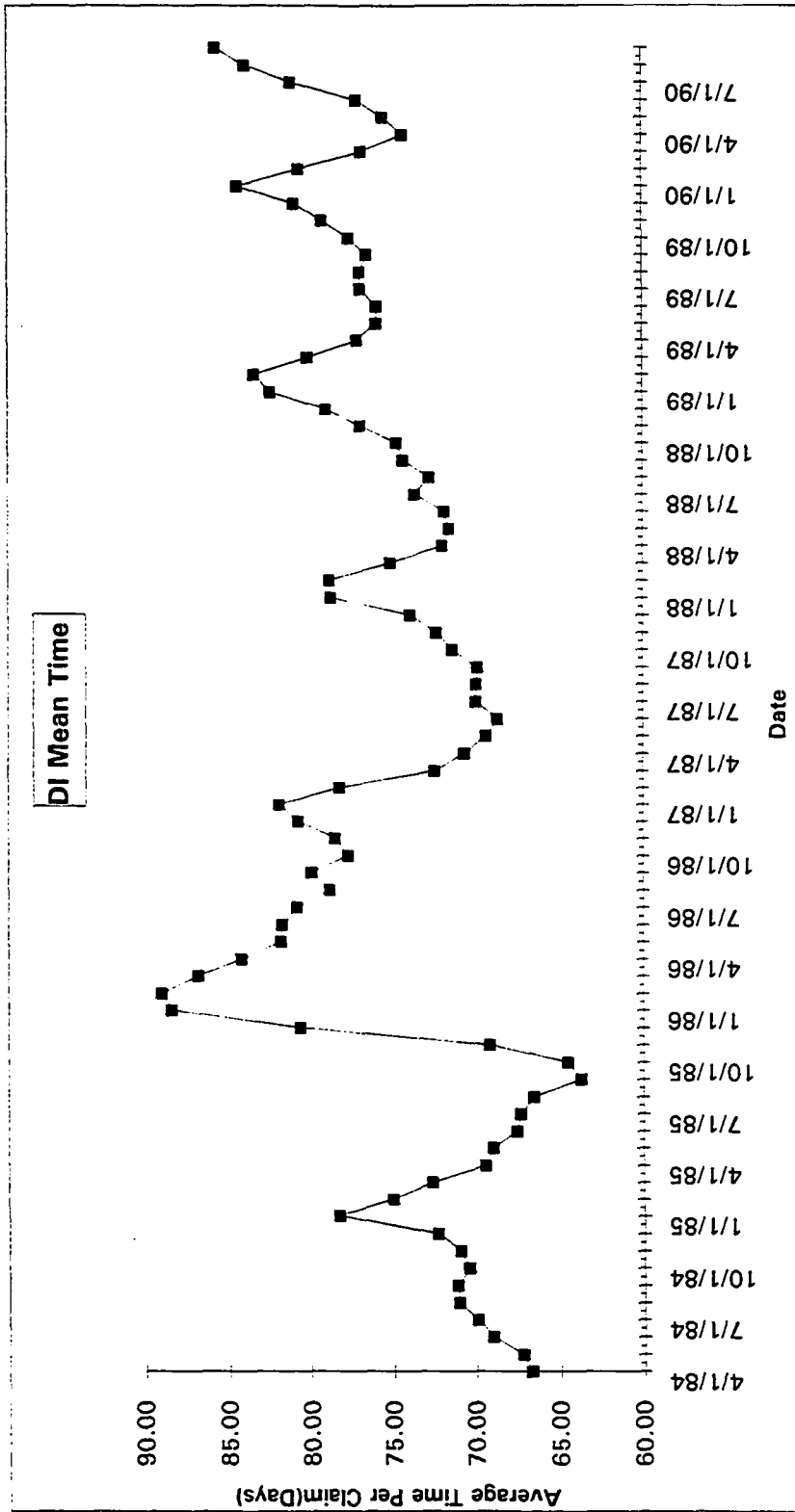
4/1/88	318854	20.30	110722	71.90
5/1/88	249213	20.00	86771	71.50
6/1/88	227257	20.80	80240	71.80
7/1/88	280410	20.80	97740	73.60
8/1/88	232884	20.30	82215	72.70
9/1/88	271349	20.70	98227	74.30
10/1/88	220622	20.10	79938	74.70
11/1/88	192270	21.50	73152	76.90
12/1/88	308514	18.60	95399	79.00
1/1/89	221112	20.60	69488	82.40
2/1/89	249574	21.90	72099	83.40
3/1/89	337062	20.90	100040	80.10
4/1/89	272905	19.70	78684	77.10
5/1/89	265428	19.10	80883	75.90
6/1/89	295880	18.70	98148	75.90
7/1/89	228556	17.80	77637	76.90
8/1/89	233580	17.20	80662	76.90
9/1/89	271988	17.20	99391	76.50
10/1/89	224626	16.70	79922	77.60
11/1/89	198628	16.70	77375	79.20
12/1/89	311738	14.80	94899	80.90
1/1/90	237926	16.90	75493	84.40
2/1/90	275860	15.50	80137	80.60
3/1/90	345540	16.20	104977	76.80
4/1/90	255211	15.80	81514	74.30
5/1/90	260993	16.40	86544	75.50
6/1/90	285758	16.40	102683	77.10
7/1/90	216452	15.50	82258	81.10
8/1/90	263020	17.40	101344	83.90
9/1/90	258399	14.10	78323	85.70

Appendix J
Trend Lines for Raw Data by Month









Appendix K
Chi-Square Conversion and Calculation

Raw Frequency Counts 1981 Data

	Clerical	Service	Claims
Quantitative	38	91	214
Qualitative	40	66	114
Positive	02	13	34
Negative	18	22	42

The total number of attributes in the matrix is 694. The process for determining the converted frequencies which are presented next is to take the attributes found in that cell and divide by 694. A minimum value of one was attributed to each cell due to rounding.

Converted Frequency Counts 1981 Data

	Clerical	Service	Claims
Quantitative	5	13	31
Qualitative	6	9	16
Positive	1	2	5
Negative	3	3	6

These figures are converted percentages from the raw frequency counts for the 1981 data. These converted percentages become the expected values in the Chi-Square calculation in the last page of this appendix. These figures will be located in the small, upper left hand boxes of each cell.

Raw Frequency Counts - 1986 Data

	Clerical	Service	Claims
Quantitative	1268	1235	754
Qualitative	410	432	215
Positive	30	31	36
Negative	375	284	170

The total number of attributes in the matrix is 5240. The process for determining the converted frequencies which are presented next is to take the attributes found in that cell and divide by 5240. A minimum number of one was attributed to each cell due to rounding.

Converted Frequency Counts - 1986 Data

	Clerical	Service	Claims
Quantitative	24	24	14
Qualitative	8	8	4
Positive	1	1	1
Negative	7	5	3

These figures are converted percentages from the raw frequency counts from the 1986 data. These converted percentages become the observed values in the Chi-square calculation in the last page of this appendix. These figures will be located in the larger area of each cell of the matrix found on the last page of this appendix.

Configuration of Expected and Observed Frequencies

	Clerical		Service		Claims	
Quantitative	5		13		31	
	24		24		14	
Qualitative	6		9		16	
	8		8		4	
Positive	1		2		5	
	1		1		1	
Negative	3		3		6	
	7		5		3	

$$\chi^2$$

The Computation for Chi-square for the matrix above is that the observed values (in the larger area of the cell) is subtracted from the expected value (in the smaller cell) with the difference being squared. The squared difference is then divided by the expected value.

For example, the bottom left hand cell of the matrix would be computed as: $3 - 7 = -4$ and then $4 \times 4 = 16$, and then $16/3 = 5.33$. In the final matrix on the next page, this formula was followed for every matrix.

Final Chi-Square Calculation

	Clerical	Service	Claims
Quantitative	72.2	9.31	9.32
Qualitative	.67	.11	9
Positive	0	.5	3.2
Negative	5.33	.75	1.5

The sum of the cell values is 111.89. This is a very large values for Chi-square. It indicates a significant difference between the expected values established by the first set of performance appraisal criteria and the second set which were used as observed values in the calculation.

It must be remembered that the sets were standardized by converting them both to percentages. If this had not been done, the resulting Chi-square figure would have been hundreds of times larger due simply to the vast difference in the total number of attributes found in each set (694 in 1981 and 5340 in 1986).

BIBLIOGRAPHY

- Abels, M. L. (1989). Merit Pay and Performance Evaluation: Enhancing Employee Motivation and Accountability in Municipal Government. Public Management, v 71, 11-12.
- Abramson, M. & Schmidt R. (Mar/Apr 1983). Politics and Performance. Public Administration Review, 275-280.
- Ammons, D. & Rodriguez, A. (Sep./Oct. 1986). Performance Appraisal Practices For Upper Management in City Governments. Public Administration Review, 460-466.
- Balfour, D. L. (1992). Impact of Agency Investment in the the Implementation of Performance Appraisal. Public Personnel Management, v 21, n 1, 1-15.
- Berger, L. (Summer 1983). The Promise of Criterion-Referenced Performance Appraisal. Review of Public Personnel Administration, v3, 21-32.
- Borras, R. (June 1987). The Delicate Fulcrum Point. Public Management 6, v 69, 23-24.
- Bradley, D. E., Pursley, R. D. (March 1987). Behaviorally Anchored Rating Scales for Patrol Officer Performance And Appraisal: Development and Evaluation. Journal of Police Science And Administration 14, 37-45.
- Brown, D. G. (Summer 1987). Development of Performance Standards: A Practical Guide. Public Personnel Management 16, 93-114.
- Brown, R. W. (Spring 1982). Performance appraisal: A Policy Implementation Analysis. Review of Public Personnel Administration, v 2, n 2, 69-85.
- Cameron, J. R. (Feb. 1989). Performance Evaluations Re-Evaluated. The Police Chief, 53-56.
- Cascio, W. F. (Winter 1982). Scientific, Legal and Operational Imperatives of Performance Appraisal Systems. Public Personnel Management, v 11, 367-374.
- Civil Service Reform Act of 1978, S 4301, 5 U.S.C.

- Civil Service Reform Act of 1978, S 4304, 5 U.S.C.
- Clement, R. W. (Spring 1987). Performance Appraisal: Non-Verbal Influences On The Rating Process. Review of Public Personnel Administration 7 (2), 14-25.
- Clement, R. W., & Aranda, E. K. (Fall 1984). Performance Appraisal in the Public Sector: Truth or Consequences. Review of Public Personnel Administration, v 5, n 1, 34-42.
- Clement, R. W., & Stevens, George E. (Spring 1986). The Performance Appraisal Interview: What, When and How. Review of Public Personnel Administration 6, 43-58.
- Daley, D. M. (1990). The Civil Service Reform Act and Performance Appraisal: A Research Note on Federal Employee Perceptions. Public Personnel Management, v 19, n 3, 245-251.
- Daley, D. M. (1991). Great Expectations, Or A Tale of Two Systems: Employee Attitudes Toward Graphic Rating Scales and MBO Based Performance Appraisal. Administration Quarterly, v 15, n 2, 188-209.
- Dresang, D. L. (1991). Public Personnel Management and Public Policy. (2nd ed.). New York: Longman.
- England, R. E., & Parle, W.M. (1987). Non-Managerial Performance Appraisal Practices in Large American Cities. Public Administration Review, v 47, 498-504.
- Federal Times (June, 1993) Do You Like How You Are Being Rated?
- Gabris, G. T. (Winter 1986). Recognizing Management Technique Dysfunctions: How Management Tools Often reate More Problems Than They Solve. Public Productivity Review, v 40, 3-18.
- Gabris, G. T. (Fall 1986). Can Merit Pay Systems Avoid Creating Discord Between Supervisors and Subordinates?: Another Uneasy Look at Performance Appraisal. Review of Public Personnel Administration, v 7, n 1, 70-89.
- Gaertner, G & K. (May 1985). Contingent Pay For Federal Managers. Administration and Society Performance, v 17, 1-7

- Gaertner, K. N. & Gaertner, G. H. (May 1985). Performance Contingent Pay for Federal Managers. Administration and Society, v 17, n 1, 7-19.
- Gilbert, G. R. (1982). Performance Appraisal and Organizational Practice: A Post Reform Review. Public Personnel Management Journal 11, 318-321.
- Goodson, J. R. & McGee, G. W. (1991). Enhancing Individual Perceptions of Objectivity in Performance Appraisal. Journal of Business Research, v 22, n 4, 293-303.
- Haynes, M. E. (Mar/Apr 1973). Do Appraisal Reviews Improve Performance. Public Personnel Management 2, 128-132.
- Hoffman, C. C., Barry, N. R., and Holden, L. M. (1991). A Comparison of Validation Criteria: Self vs. Supervisor Ratings. Personnel Psychology, v 44, n 3, 601-619.
- Hopkins, H. R., Lovrich N., Shaffer, P. & Yale, D. (May/Jun 1980). Do Public Servants Welcome Or Fear Merit Evaluations Of Their Performance? Public Administration Review, 214-221.
- Hyde, A. C. (Winter, 1988). The New Environment for Compensation and Performance Evaluation in the Public Sector. Public Personnel Management 17, 351-357.
- Hyde, A. C. (1982). Performance Appraisal in the Post Reform Era. Public Personnel Management Journal 11, 294-305.
- Ingle, H. T. (Winter 1982). Contemporary Issues in Federal Evaluation Policy: New Linkages Between Personnel and Program Assessment Processes. Public Personnel Management 11, 322-331.
- Kleiman, L. S., & Gordon, M. E. (Apr. 1986). An Examination of the Relationship Between Police Training Academy Performance and Job Performance. Journal of Police Science and Administration 14, 293-298.
- Lovrich, N. R, Jr., Hopkins, R. H., Shaffer, P. L. & Yale, D. A. (Summer 1981). Participative Performance Appraisal Effects Upon Job Satisfaction, Agency Climate and Work Values: Results of a Quasi-Experimental Study in Six State Agencies. Review of Public Personnel Administration 1, 51-73.

- Martin, D. C., & Bartol, K. M. (1986). Training the Raters: A Key to Effective Performance Appraisal. Public Personnel Management, v 15, n 2, 101-109.
- McNish, L. C. (Fall 1986). A Critical Review of Performance Appraisal at the Federal Level: The Experience of PHS. Review of Public Personnel Administration, v 7, n 1, 322-56.
- Millar, K. I. (1990). Performance Appraisal of Professional Social Workers. Administration in Social Work, v 14, n1, 65-85.
- Mitroff, I. Stakeholders of the Organizational Mind. Mossholder, Kevin W., Giles, William F. and Wesolowski, Mark A. (1991). Information Privacy and Performance Appraisal: An Examination of Employee Perceptions and Reaction. Journal of Business Ethics, v 10, n2, 151-156.
- Nalbandian, J. (May/Jun 1981). Performance Appraisal: If Only People Were Not Involved. Public Administration Review, 41, 392-396.
- Nalbandian, J. (May/Jun 1982). Performance Appraisal. Public Management (3), v 41, 392-396.
- Nigro, L. (Jan/Feb 1981). Attitudes Of Federal Employees Toward Performance Appraisal And Merit Pay. Public Administration Review, 84-86.
- Northrup, A. & Schmidt, R. (Mar/Apr 1985). Task Environment Approach to Organizational Assessment. Public Administration Review, 275-280.
- Pavett, C. M. (1983). Evaluation of the Impact of Feedback on Performance and Motivation. Human Relations, v 36, 641-654.
- Payne, T. (January 1994). Points to Ponder. Reader's Digest, 141.
- Plamenatz, J. (1966). The English Utilitarians. London: Compton Printing.
- Robson, R. E., & Nelson, Dalmas H. (1991). Supervisor and Non-Supervisor Agreement and Dissonance Regarding Performance Appraisal. Review of Public Personnel Administration, v 11, no's 1 & 2, 121-130.

- Rosinger, G., Myers, L. B., and Girard, W. (Spring, 1982). Development of Behaviorally Based Performance Appraisal Systems. Personnel Psychology 35, 75-88.
- Ross, J. D. (Summer, 1983). Commitment and Loyalty as a Determinant of Performance Ratings. Review of Public Personnel Administration, v 3, No 3 (Summer 1983): 105 - 116.
- Sayles, Leonard R. (November 1987). Organization Size, Effectiveness and Human Values. Administration and Society.
- Shafritz, Jay M. & Witbeck, Philip H. (1978). Classics of Organization Theory. Oak Park, Illinois: Moore Publishing Company.
- Sims, R. R., Veres, J. G. & Heninger, S. M. (Spring 1987). Training Appraisers: And Orientation Program for Improving Supervisory Performance Ratings. Public Personnel Management 16, 37-46.
- Skinner, Q. & Price, R. (eds.), (1988). Machiavelli: The Prince. New York: Cambridge Press.
- Social Security Administration. (1990a). Guide for Developing Performance Plans. Washington, D.C.: U.S. Government Printing Office.
- Social Security Administration. (1990b). Guidelines for the Development and Application of SSA Performance Standards. Washington, D.C.: U.S. Government Printing Office.
- Social Security Administration. (1990c). Report of the National Performance Standards Workgroup. Washington, D.C.: U.S. Government Printing Office. Some Employees Believe Appraisal System is a Disincentive. (1993, April). Federal Times. p. 1-18.
- Spitz, D. (ed.). (1975). On Liberty: John Stuart Mill. New York: W. W. Norton and Co. Inc.
- Starling, G. (1982) Performance Appraisal in the "Z" Organization. Public Personnel Management 11, 343-351.
- Steel, B. S. (Summer 1983). Implementing Performance Appraisal: Comparing Views Among Employees, Supervisors and Personnel Specialists. Review of Public Personnel Administration, v 3, n 3, 73-93.

- Sumek, L. J. (Feb 1988). Performance Evaluation: Evaluate or Not? That Is Not the Question. Public Management 70, 2-9.
- Thompson, V. A. (1975). Without Sympathy Or Enthusiasm. University of Alabama Press.
- Vance, R. J., Winnie, P. S., & Wright, E. S. (Autumn, 1983). A Longitudinal Examination of Rater and Ratee Effects in Performance Ratings. Personnel Psychology 36, 609-620.
- Wholey, J. S. (1991). Using Evaluation to Improve Program Performance. The Bureaucrat, v 20, 55-59.
- Zemke, R. (1991). Do Performance Appraisals Change Performance? Training, v 28, n 5, 34-39.