THE MEASUREMENT AND MODIFICATION OF ABSENCE BEHAVIOR AND OTHER EMPLOYEE TIME USES DURING IMPLEMENTATION OF AN ABSENTEEISM CONTROL INFORMATION SYSTEM: A TIME-SERIES STUDY

By

Ian Alexander Miners

A DISSERTATION

Submitted to:
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

School of Labor and Industrial Relations

ABSTRACT

THE MEASUREMENT AND MODIFICATION OF ABSENCE BEHAVIOR AND OTHER EMPLOYEE TIME USES DURING IMPLEMENTATION OF AN ABSENTEEISM CONTROL INFORMATION SYSTEM: A TIME-SERIES STUDY

By

Ian Alexander Miners

A substantive and empirical reconceptualization of the absence research paradigm, this project addresses limitations of the conventional paradigm reported in prior research by developing a detailed system for time-series analysis. To integrate concerns of conceptual consistency and measurement raised in the literature (Chadwick-Jones, Nicholson and Brown, 1983; Hammer and Landau, 1981), absence, attendance and time scheduled off have been reconceptualized in the present work as alternate forms of employee time-use. These have been operationally defined as a set of 13 interdependent moving values calculated daily across the duration of an absence control program. A time-series split-half reliability test was used to determine the reliability of these measures. Both cross-sectional T-tests and a time-series correlation technique were applied to investigate program impact and to test substantive hypotheses about absence behavior. For both the T-tests and correlations, hypotheses of two kinds were examined. In each

case, the first kind focused on testing the relative merit of the procedures used to examine the second or substantive kinds of hypotheses. Since the absence literature postulates significant moderations of absence behavior by annually recurring seasonal changes or holidays, substantial attention was also devoted to distinguishing the extent of such intra-annual shifts so that they could be considered in the assessment of program impact. Major substantive hypotheses dealt with whether employee time-use changed from short-term to long-term absence and from illegitimate to legitimate forms of absence.

The results generally showed the measures to be an order of magnitude more reliable than measures used in prior research (Breaugh, 1981; Hammer and Landau, 1981), although under small sample conditions the measures became less reliable. The absence control program providing the multi-year database of this research did clearly reduce short-term non-legitimate absence significantly and a co-occurrent increase in time worked on the job for each employee also took place. There also occurred a powerful legitimacy effect, a shift of employee time-use from non-legitimate absence into long-term absence which was still legitimate after program installation.

On a methodological level, the evidence brought to bear on procedural hypotheses clearly supported the use of group level time-series analysis, and in particular supported the time series correlation technique employed here. The evidence also supported the contention that daily cross-sections of individual absence data are distributed highly non-normally. Despite this finding, T-tests comparing daily cross-sections of individual data over time still showed results that are consistent with the results of group-level time series analyses. The comparison of daily cross sections has obvious practical merit since it allows estimation of costs and benefits. But for inferential tests to analyze the significance of trends in time-use behavior, conventional cross-sectional procedures are at best tenuous and where possible these should probably be set aside in favor of the time-series procedures.

On this theoretical level, a strong substitution of long-term legitimate absence for short-term non-legitimate absence ($r = 0.90$) occurred across all four different time intervals examined during the span of the program. The evidence strongly supported the legitimacy hypothesis. When the effect of legitimacy was controlled for however, no evidence of a group-level progression toward longer-term absence was apparent.

With regard to program impact, the absence control project itself was, by design, a managerial and technical system installed for reporting and guiding the control of employee absenteeism. It was a contractual, OD and research success for the consultants, and an administrative success for the DOT. However, the impact of the program was somewhat diluted by substitution of long-term legitimate absence for short-term non-legitimate absence, and by an externally imposed increase in legitimate absence due to 80 hours of paid leave granted by the City Government in concession bargaining. In short, the consultants and the DOT succeeded beyond the goal that had been set. But overall absence changed very little since the city authorities gave back part of the time saved, and the DOT long-term absence control policies were not designed to prevent much of it from shifting into long-term legitimate absence.

DEDICATION

To Dad, who taught me to think in models and to build the models into things that work ...

To Mom, who taught me to love books and to wonder about human behavior ...

To Mike, who taught me his trade and gave me a clear path to learn ...

Those gifts were well given. This one is for all of us.

ACKNOWLEDGEMENTS

A great many people have poured their efforts into this work. Lionel Little of The Department of Transport in Detroit and Tony Juliano of Moore and Juliano Associates were most helpful in providing access to the research data.

Dr. Michael Moore of the School of Labor and Industrial Relations at Michigan State University guided the design and installation of the absence control system. It was Mike's insight into the need for precise detail in day-to-day data collection that created the detailed raw data analyzed here. And it was his support and guidance that nurtured the dissertation research process.

Dr. Tom Patten has helped guide my progress from the beginning of the Ph.D. program. Both he and Dr. Dan Ilgen have contributed their time and critical insights to the structure of this document and to refining its content. They have been very helpful about providing quick and sincere feedback and the dissertation is a much better product because of their help.

Dr. Larry Foster and Dr. Mary Zalesny also contributed some helpful suggestions and special thanks are due to them for their encouragement and support.

There are several special people in the Computer Services Department at The University of Toledo, Toledo, Ohio, whose support was invaluable. Wendy Fraker especially, and Donna, Gregg and Lance -- thank you for the personal attention you gave to the awkward problems this research posed.

My friends and family have helped when they could and have tried to understand and share this work. There were many times when their faith in me made a great difference in the small, cold hours of the morning.

Jan Tabernik, my editor, secretary, and friend, has typed every draft at every stage of this dissertation. Often she had to find a way to do the work at the last minute or at night and without any compromise of quality. Thank you, Jan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

Current work in personnel research has a greater need than ever before to find and show strong relationships between personnel practices and concrete dollar consequences. At first glance, absenteeism is one tangible consequence where costs and the control of them should be easily measured and managed. Apparently this has not been the case. The field still does not have an effective response to the problem. In 1982 a rough estimate (an adaptation of Mirvis and Lawler, 1977) was updated using the 1981 CPI and national absence statistics, placing overall absenteeism costs at $53.6 billion, while payroll related costs alone were about $23.6 billion. This burden alone is a compelling reason for reducing absence if possible. And if an effective system existed with which to monitor and analyze changes brought about through absence control efforts, then a systematic knowledge could be developed about what works best to control absence. But absence research reported to date has not demonstrated an effective means for program evaluation and absenteeism management.

Meanwhile, there exists a growing emphasis in the literature (Breaugh, 1981; Hammer & Landau, 1981; Steers and

Rhodes, 1978) on reconceptualization of absenteeism (Chadwick-Jones, Nicholson & Brown, 1983) and on refinements in measurement. Several new opportunities have emerged for using innovations in data preparation (Stevens, 1951; Hamblin, 1974; Tukey, 1977) and analysis to enhance this work (Glass, Wilson & Gottman, 1975; Montgomery & Johnson, 1976; Bloomfield, 1976; Thrall & Engleman, 1981; Liu, 1981).

The research reported below is an attempt to establish and apply a stronger paradigm for absence research and constitutes a significant departure from the conventional approach to absence measurement. Both the conceptual basis and the operational definition of the measured variables are new.

An Introduction to the DOT Absence Control System

During 1978 and 1979 a comprehensive system for identifying and reporting absenteeism and for implementing control practices to reduce non-legitimate short-term absence was installed at the Department of Transport in Detroit, Michigan. The project was implemented in the stages described below. Further detail is provided in Chapter III.

(a) Top management support was secured for the project and extensive supervisory input into the determination of various forms of employee time-use was encouraged. Supervisory input was facilitated using a hierarchically sequenced series of structured problem-solving groups and management training workshops. Data for July 1, 1978 to June 30, 1979 were collected with the payroll data for use as a pre-program baseline period in later analyses. During the preprogram installation period, nominal groups were used to focus employee, supervisor, and manager input regarding causes of absenteeism, solutions to the "absenteeism problem," and incentives for attendance. These meetings were part of both a diagnostic and programmatic effort to raise awareness of absenteeism at the DOT. It was begun roughly nine months before the computer aspects of the program were implemented.

(b) A Real-time information system was created to computerize the flow of information about employee time uses so that timely and adequately detailed reports could be provided to supervisors. These reports were to become the crucial information base required for consistent supervisory implementation of absence control policy.

(c) The absence control program itself was implemented on a real and high-impact basis, coming on-line in stages during May 1980. In those last 30 days of the installation period (June 1980), absence control steps were actually implemented. Employees and supervisors were shown how the system worked and were given a chance to learn what impact it would have on them.

(d) On July 1st of 1980, an additional component of the absence control program went into effect. Each employee was given a "bank" of casual leave, to be used at his or her discretion, as paid, legitimate, short-term absence. This bank of casual leave, plus holidays and annual vacation, were to be renewed each year in accord with the collective bargaining agreements with the several unions involved. In effect, the absence control system thereby provided an annual renewal for each employee of legitimate discretionary absence days on July 1 of each year.

(e) As employees worked and the system ran, many employees prematurely used up their annual bank of casual leave and then began to accumulate incidents of non-legitimate absence. At that point, the corrective discipline system began to operate. Each successive incident resulted in a report to the supervisor advising (1) what step in the discipline process had been reached, and (2) what specific action must now be taken.

In addition to disciplinary action, an incentive bonus of about $100.00 was awarded to each employee whose absence level was still low enough to qualify at the end of the first year of the program.

After two years of operation, the program was still running strongly and enjoyed continued supervisory support right up to the top management level. At that time the data for this present research were copied and released for analysis.

Objectives

The central project in this dissertation has been designed to evaluate the impact of an absence control program on employee absence behavior. Four objectives had to be addressed in the process in order to build a clear chain of research procedure linking the raw data into a focused body of evidence that bears on each substantive hypothesis.

Procedural Objectives:

1. Reliable and sensitive measures for the study of absence had to be created and tested empirically, both in terms of their reliability and for overall sensitivity in use to both long-term changes and intra-annual changes.
2. A controversy in the literature regarding the study of absence frequency data in preference to absence severity (or duration) was examined using the derived measures. It was hypothesized that when an appropriate test of reliability was used, severity measures would not be less reliable than frequency measures.

Substantive Objectives:

3. Several substantive concerns about hypothesized shifts in absence behavior from one form to another in the presence of absence control programs were investigated. This was done so that any such shifts could either be distinguished from expected overall long-term forms of program impact or identified as unintended but significant program related changes.

a. If absence behavior shifted from short-term to long-term absences, showing progressively more severe forms of withdrawal over time (regardless of whether the short-term absences involved were legitimate (allowed) absences, this would imply that the program induced a progression effect at a group level.

b. If absence behavior shifted from patterns of time-use that were made non-legitimate by the absence control program to patterns that were still legitimate, this would indicate a legitimacy effect. Depending on what new forms of absence appeared and to what extent, a legitimacy effect might indicate success or partial failure.

c. Due to the annual renewal of legitimate short-term absence days for all employees, these effects were expected to be restimulated in an annual cycle. If they accurred, there would then have been evidence for annual cycles of "seasonal" shifts as well as for long-term shifts that span the entire post-installation period. Both have importance for theory testing and for program evaluation and absence control policy. Also, the presence of such short-term shifts would have meant that the overall impact of the program could only be determined when the procedure controlled for renewal effects.

4. The overall impact of the absence-control program which provided the data for this research was evaluated. The consultants had specifically contracted to reduce short-term non-legitimate absence. Concerns noted above implied that this might have had unintended consequences. The program was therefore evaluated both in terms of the extent to which reduced short-term non-legitimate absence occurred, and in terms of what major concomitant changes also took place.

Distinctive Features

Six features distinguish this dissertation project from prior work.

1. Absenteeism was conceptually and operationally defined here as one form of employee time use behavior. Since the same rationale applies to any other form of employee time use ... such as time worked on the job, or time scheduled off ... this permitted an "internally comparable" set of measures for all forms of employee time use to be created.
2. Due to the internal comparability of the measures and the unusually high quality and detail of the unprocessed raw data, it was also possible to design procedures that recategorize all the daily values into their correct time-use types as these changed with duration of a given time-use behavior over time. For example, an absence

might at first be recorded as a short-term absence without leave. If it then persisted for more than five working days, it became a long-term absence. Measurement integrity requires that short-term absences which accumulate day-by-day to become long-term absences must then be recategorized from the first day of the occurrence onward. The logic and sequencing of sorting procedures employed in all such recategorizations done here is described in Chapter III.

3. Once the consistent and corrected classification of employee time-uses over time was provided for, a new kind of time-series measure for both frequencies and durations of the various different employee time uses could then be derived from the cleaned raw daily data. Both 100-day moving values (short-term) and annual moving values were calculated for each type of time use on each and every day of the program, for each employee, and for each selected subpopulation of employees. These derived daily estimates, based on data across either the past 100 days (sensitive to changes within the year) or across the preceding 365 day period (only sensitive to changes from year to year), are, by construction, proportionately less radically affected by short-term fluctuations than single-day or weekly based measures, and provide more stable estimates of long-term variation in the group time use patterns than the raw daily frequency

or severity values would, without being insensitive to daily fluctuation.

4. The derived statistics that were the focus of all substantive inquiry in this research project are primarily group statistics. In no case did any substantive analysis rely directly upon estimates of individual time use behavior except for the cross-sectional normality tests and T-tests. The research was unusual but not unique in this since others have used group measures of absence behavior (Chadwick-Jones, Nicholson and Brown, 1982). This group focus was central to two further innovations that are pivotal in subsequent applications of inferential statistics.
   a. By calculating a daily estimate of the group mean value for each time-use measure, a problem characteristic of individual absence data distributions could be substantially reduced by construction. Absence research which depends directly upon the absence behavior of individual employees has been characteristically undermined by asymmetric data distributions that are both upwardly skewed and abruptly truncated at the lower end (Hammer and Landau, 1981). Under other circumstances, the asymmetry could be removed by simple transformation procedures, but the severe truncation makes this procedure inap-

propriate for individual absence data. Fortunately, the group means rarely equal zero and so they do not suffer from a truncation problem. Also, the group and not the individual is the appropriate research focus for both theory development and support of absenteeism control policy decisions.

b. The daily group means also provide a theoretically far more appropriate set of data values than individual values for estimating the reliability of the measures over time. Traditional absenteeism research has attempted to test the over-time stability of absenteeism data by repeatedly comparing two sets of absenteeism data values, each taken from the same set of individuals, but at two points in time. That procedure cannot show the over-time stability of absence measurement procedures because it is designed so as to discover only the stability of individual absence behavior between times ... although, if individual absence behavior itself were not variable, there would never be any absenteeism to study.

c. The dissertation used time series correlation of the daily group-means for one randomly selected half of the values for each measure, with the daily group means for the other half over specifically selected periods within the program, as the basis for its major tests of substantive hypotheses.

5. The distributional normality of the individual values used in the construction of T-tests was examined using the Kolmogorov Normality test statistic "D". As a consequence, the relative appropriateness of using T-test statistics can be shown.
6. The use of both a 100-day base period and a 365-day base period for calculating derived time-use values allows both a good sensitivity to short-term changes (using the 100-day base), and an exclusive sensitivity to long-term inter-annual changes using the 365-day base. (Seasonal effects, holidays and intra-annual short-term effects that repeat yearly are simply excluded by construction.)

<u>Dissertation Structure</u>

The main portion of this dissertation is comprised of six chapters. Chapter I introduces the project and describes the structure and content of the dissertation. Chapter II reviews the literature and relates this to specific hypotheses that were tested on absenteeism measurement and theories of absence behavior. A summary of the central issues, both in measurement and theory of absence behavior is presented on page 27. In Chapter III, the procedures used for data collection in the present research are briefly described, sample demographics are summarized, and the logic and procedures for creating derived measures of absence behavior and other time-use patterns are explained. (See Figure 5 on page 48.) On page 47 these measures are summarized in a comprehensive list. Chapter IV briefly describes the research methodology and presents the hypotheses that were tested in each step of the dissertation project. On page 57, Figure 6 outlines the sequence of steps included in data analysis and hypothesis testing. Chapter V reports the results of data analysis in both tabular and graphical form, with comments on the purpose and implications of each report. Chapter VI presents a brief summary and a discussion of the findings and the conclusions that follow and comments on directions for further research on employee time uses.

The appendices to this report present greater detail on both the results and many of the methodological and procedural problems resolved in the present work. For reader convenience, a brief summary of these steps has been included in the dissertation itself, with references provided to the more detailed description included in the appendices.

CHAPTER II

Literature Review and Explanation of the Hypotheses

General Introduction to the Literature Review

A discussion of the literature relevant to research on absence measurement and theory is introduced below and has been presented in greater detail throughout this Chapter along with accompanying hypotheses. In essence, two underlying themes pervade this literature.

1. How can absenteeism be reliably and precisely measured? In recent years, the measurement problem has gained more and more attention. It is the subject of the first step in each phase of hypothesis testing for the present research. Prior to the use of T-tests, the underlying distributions were examined for normality. And prior to the use of time-series correlations, the respective time-series measures were examined for reliability.
2. How does absenteeism behavior actually change over time from one kind of time-use to another? (This theoretical question was first raised by Hill and Trist (1952, 1953, 1962) in work that represents forms of absenteeism as different solutions to an employee's alienation from work, an hypothetical need to withdraw from the job.)

A list of these themes, including several questions raised by the absenteeism literature and addressed by this dissertation, has been included toward the end of this Chapter on page 41. In the work to establish a viable paradigm for absence research, the problems of theoretical conceptualization and operational definition have been extensively intertwined, perhaps slowing the evolution of both. These issues have been the focus for specific treatment in Chapter IV through Chapter VI, and in the appendices. However, a theory-based phrasing of all hypotheses tested in the present work has been prepared in the terms of their respective theoretical foundation. Throughout the review of literature presented below, those theory-based articulations of research hypotheses have been presented. These appear in the order that their respective value to this research was established within the present literature review. The cross-references which accompany these theory-based hypotheses refer to the related but operationally defined hypotheses articulated in Chapter IV.

A conceptual consistency in integrating these theory-based hypotheses with the procedural hypotheses had to be established in terms of the actual research setting. To support this link within the present research methodology, hypotheses concerning the concept of "annual renewal" of the employees' casual leave days have also been presented here. To establish a common ground for later discussion of the literature, the disciplinary practices in employment relations and the various measures of workforce absenteeism that are already established as conventions within the field have been summarized below.

Practices in Employment Relations Which Relate to Absence Management and Control

The discipline and labor arbitration practices in labor relations require that rules be enforced consistently, across all unionized employees, and without exceptions (Elkouri and Elkouri, 1974). A rule must apply and be applied to all employees to the same extent and in the same way. Furthermore, under Title VII of the Civil Rights Act of 1964 and related court decisions, the employer may not discriminate in employment against protected minority employees on an idiosyncratic and arbitrary basis. Just cause and fair and equitable policy and practice must be demonstrated. The application of discipline to control absence is an action which directly affects an employee's career opportunities, and it falls under the jurisdiction of laws protecting the rights of workers in employment. Accordingly, it is necessary that an absence control program implement control practices in a way that affects all employees in the same way and to the same extent. That is, the rules must be consistently applied in all cases.

Yet, consistent application of the rules to all absence occurrences, allowing corrective discipline procedures to be applied, requires that the company have a full and detailed knowledge at all times of each employee's absenteeism record. Prior to discipline, a supervisor must know exactly how many legitimate absence days the employee may still claim before corrective discipline begins, and which stage of corrective discipline the employee has reached, and what action must be taken the next time the employee violates the rules for absence control. In any company of even moderate size, or where employees do not always have the same supervisor, this detailed knowledge will not be available unless a systematic effort is made to provide it.

Furthermore, since an absence control and discipline policy is vulnerable and can be undermined by allowing individual exceptions, it must be constructed to exclude such exceptions and to provide a workable and reasonable policy for all employees as a group. As a result, the proper focus of program design and evaluation is at the level of the group, while those policies must be continuously implemented by uniform enforcement based upon detailed and current information at the level of the individual.

Conventional Measures of Absence

Absence has been operationally defined in various ways in the literature. Throughout these manifestations, three forms adopted by the U.S. Bureau of Labor Statistics (B.L.S.) have achieved general acceptance.

1. The absence rate (Hedges, 1977; Miner, 1977; Taylor, 1981) is a percentage based on the duration of absence over time. The typical form is (absence rate) = (number of hours lost to absence per month) divided by (average number of employees multiplied by number of hours usually worked), all multiplied by 100.
2. The incidence rate (Hedges, 1977; Taylor, 1981) also is a percentage, but this one indicates overall frequency of employee absence in any given unit of time. The typical form is (incidence rate) = (number of workers absent) divided by (total employed), multiplied by (100). The time period used as a base is usually one week, although 100-days, 100-weeks, or one year have also been used.
3. The severity rate is also a duration-based percentage rate, but this one is a percentage of time absent in a given period for the absent employee group alone. The term "severity" itself has been established in use by the absence literature to refer to severity or duration of absence as a form of withdrawal. Increased severity means longer duration of absence. The typical form is (severity rate) = (average number of hours lost by absent workers) divided by (average number of hours scheduled for work by absent workers) multiplied by (100).

These three standard B.L.S. measures of absence have been adopted and extended by Chadwick-Jones, Nicholson & Brown (1982) in a recent reconceptualization of research on absenteeism. Their extensions are of two kinds. The less pervasive innovation was the presentation of several measures custom-tailored to address specific substantive issues. These are as follows.

(a) The Blue-Monday Index (BMI)

$$\underset{\text{(per week)}}{\text{BMI}} = \frac{\text{Difference Between Total Absence Levels on Monday and Friday}}{\text{Number of Employees in Sample}} \times 100$$

(b) The Worst-Day Index (WDI)

$$\underset{\text{(per week)}}{\text{WDI}} = \frac{\text{Difference Between Total Absence Levels on Best and Worst Days}}{\text{Number of Employees in Sample}} \times 100$$

(c) Short-Term Index (STI)

$$\underset{\text{(per week)}}{\text{STI}} = \frac{\text{Total Number of One- and Two-Day Absences Starting in the Week}}{\text{Number of Employees in Sample}} \times 100$$

(d) Percentage of "Cleans" of Total (CLEANS)

$$\underset{\text{(per week)}}{\text{CLEANS}} = \frac{\text{Number of Employees With No One- or Two-Day Absences Per Year}}{\text{Number of Employees in Sample}} \times 100$$

(e) Mean Length of Absence in Days (MLA)

$$\text{MLA} = \frac{\text{Days Lost Due to Absence}}{\text{Frequency of Absence Incidents}}$$

(f) Labor Turnover Rate (LTR)

$$LTR = \frac{\text{Number of Employees Leaving Per Year}}{\text{Number of Employees on the Payroll}} \times 100$$

(g) Accident Absence Rate (AAR)

AAR = Number of Accidents at Work That Caused a 3 Day or More Absence, Per 100 Employees Per 100 Weeks

Although these seven measures are intriguing, nonetheless, the second innovation in measurement presented by Chadwick-Jones, Nicholson & Brown has far more sweeping implications. In this case, the researchers varied not the substantive focus of measurement, but the time interval across which the observed behavior was to be aggregated. Three time intervals were used:

a. rate per week

b. rate per year

c. rate per 100 weeks

The resulting derived data allowed analysis of absenteeism patterns by comparison of over-time plots based on weekly rate levels. This focus on Time-Series analysis is a fundamental requirement for research that asks questions about over-time shifts in the patterns of absence behavior. The use of several different spans of time as the base period for measurement also allows an "adjustment of focus" for the research to discover events that may be either broader in scope or greater in detail than a single "focus" can illuminate. In the present research, the time-series measurement thrust has been expanded in two ways. First, the precision of measurement was further refined from weekly to even shorter, daily and hourly intervals, and second, the time spans across which observations were to be aggregated were constructed to use moving values based on both 100-day totals and 365-day totals. The resulting database is suitable for addressing the substantive and methodological issues posed in the current research. It is also, by design, sensitive to daily changes which should enable it to reveal even the weekly patterns that the custom-tailored measures presented by Chadwick-Jones, Nicholson & Brown were derived to discover. Details are presented in Chapter III and further explained in Appendix B.

Finally, the work of Glasser (1970) suggests that some measure of "time on the job in between absences" would be very helpful in distinguishing time-use by the non-absence-prone employee group from that of the absence-prone group. It would also help clarify where the employees' work or absence time is being shifted to due to program induced changes or substitution of one form of absence for another. A discussion of theories concerning how and why absence behaviors change over time follows the comments on subgroup differences below.

Methodological Issues

The mainstream operational definition and conceptualization of absence research has predominantly been an intra-individual one (Hill and Trist, 1953; Steers and Rhodes, 1978, 1980; Nicholson and Goodge, 1976; Breaugh, 1981; Hammer and Landau, 1981). Both the design of the measures and sampling procedures, and the use of cross-sectional tests of over-time reliability reflect this stance. From this intra-individual focus stem a number of problems for the traditional research paradigm. Since those problems and the testing of means to overcome or aovid them constitute a central theme in the present work, literature and hypotheses concerning both theory and methodology have been reviewed.

Various researchers (Latham & Pursell, 1975; Steers & Rhodes, 1978 and 1980; Breaugh, 1981; Hammer & Landau, 1981; Terborg, Less, Smith, Davis & Turbin, 1982; Chadwick-Jones, Nicholson & Brown, 1982; and numerous others) have noted the superior over-time test-retest reliability of absence frequency over absence severity data. Severity is a term used to describe the length of absence "spells" while frequency means how often these episodes occur. In brief, the correlations previously reported between sets of absence frequency data for a group of individuals at time 1 and again later at time 2 are almost without exception considerably higher than over-time correlations between datasets for the same employees that are measures of absence duration. In one response to this instability, Latham & Pursell (1975, 1977) have proposed a change in focus from absence to attendance. Apparently, attendance statistics are more reliable (less variable) than duration of absence, and absence statistics are often confounded in data collection by sick leave abuse, etc., so that the real reasons for absence may not be made evident from the customary self reports. This proposal has been challenged by Ilgen (1977), who notes that studying attendance will not tell us about the details of absence behavior and that, in any case, the customary test-retest reliability statistic is at best rarely (and then only under carefully controlled assumptions) an appropriate choice for a measure of absence behavior in which many researchers have reported considerable variation over time (Glasser, 1970; Garrison & Muchinsky, 1977; Nicholson, Jackson & Holmes, 1978; Breaugh, 1981; Hammer & Landau, 1981; Markham, Dansereau & Alutto, 1982). Responding to the controversy in a slightly different approach, Smulders (1980) cites statistics on absence behavior in the Netherlands and notes that using values for frequency over severity would cause a gross underestimation of the seriousness of the absence problem and is, in any case, theoretically inappropriate where what is of theoretical interest is the absence duration.

Another area of absence research (Cheloha & Farr, 1980; Nicholson, Brown & Chadwick-Jones, 1977; Nicholson & Goodge, 1976; Parasuraman, 1982; Mobley, 1982) argues for the identification of specific subcultural groups of employees who tend to be more prone to absence or to attendance, respectively. Nicholson, Brown & Chadwick-Jones (1977) specifically promote the construct of "absence-proneness". This group basis for sampling is fundamental to their more recent research on absence (Chadwick-Jones, Nicholson & Brown, 1982).

There is also a body of work accumulating which specifically examines the statistical properties of absence-prone groups (Glasser, 1970; Behrend & Pocock, 1976; Garrison & Muchinsky, 1977). Glasser applied stochastic techniques from epidemiology to assign employees to absence-prone and non-absence-prone groups for further analysis of their behavior. His work strongly supports the selection of such subcultural samples and also advances another absenteeism measure or index ... "scheduled time worked between absences". Behrend & Pocock (1976) have also noted clear trends in absence-proneness over time and problems of skewness in the data distribution.

Another body of work which addresses absence-proneness (Garrison & Muchinsky, 1977) strongly supports the construct of absence-proneness but points out that membership in the absence-prone group tends to change over time and may be a relatively short-lived condition at the level of the individual employee. (A substantial change in group composition over a 21 month period was reported.) This suggests that group composition should be re-established for each time period across which there may be a substantial shift in membership.

Finally, work by Hammer & Landau (1981) notes that in addition to serious unreliability in the severity measures, the absenteeism data distributions also suffer from characteristic problems of truncation at the lower end of the distribution (a large cluster of very low absence employees), and from severe positive skewness due to the inclusion of a relatively few atypically large extreme values. This extreme value problem has been noted by others. Ilgen & Hollenbeck (1977) noted the low reliability of absence severity data and censored (threw out) the extremely long sick leave absences to improve the symmetry of the distribution and reliability of the measure. It can also be argued that both the truncation concern and the extreme value problem are matters that stem from including in the sample of absence-prone employees, other employees: (a) who don't belong because they are not absent enough to belong ... to be members of the absence-prone group, or (b) who don't belong because they are not present at work enough to truly influence or be influenced by membership in the absence-prone group. In short, there are reasons of both theoretical consistency and methodological rigor for separating these types of extreme values from the absence-prone group. Unfortunately, as the research of Garrison and Muchinsky (1977) indicates, such a selection process must be redone on a very frequent basis if the research is not to artificially exclude newly absent-prone workers or include others who have, in fact, "reformed".

In the balance, it is conceptually reasonable to define absence as a condition of limited participation in employment, characterized by neither consistent attendance nor extreme non-attendance, which are respectively the lower and upper limits on the range of employee time-use as it relates to absenteeism.

Hammer & Landau (1981) have noted procedures for transforming data to reduce non-normality in data distributions due to the inevitable presence of extreme values in the dataset. But they recommended instead the use of Tobin's Probit analysis procedure, due to the severity of typical truncation problems. The data transformation procedures mentioned by Hammer and Landau are now customary tools (Tukey, 1977) for improving the symmetry of distribution in datasets used in the natural sciences. If the problem of truncation at the lower end due to directly relying on data from non-absence prone employees could be overcome, such methods could and possibly should be applied to absence data. The group-focus which is basic to the time-series analyses in this present work conveniently circumvents the truncation problem by construction, except in the T-test data which remain cross sections of individuals' data and not group data. It also operates to drastically reduce the effect of extreme values on the group level derived measures used for time-series correlation. Consequently, testing the effect of such transformation has been left for later work.

Yet, to analyze only data for absent employees would involve rejection of the information about related and alternative time-uses contained in the data for the excluded extreme sub-groups. Also, the analysis would represent only the respective subgroup, not the workforce which is the unit of theoretical and practical interest. An alternative procedure was employed to evaluate change in absence behavior in the present research. The data for all employees were categorically sorted and aggregated across the sample group on a daily basis as component elements of the group-level values for each of the respective time-use behaviors instead of being analyzed as behavior from specific employees. Then, each type of absence-related behavior was investigated in its own right as a distinct category of group-level time-use behavior. The identification of absence-prone employees (Garrison and Muchinsky, 1977), which was likely to be made difficult by shifts of particular employees over time between the absence-prone subculture and the committed non-absence-prone subculture was thereby made unnecessary. (A very convenient side-effect since the primary purpose of implementing policies for absence measurement and control was to induce just such shifts.) The operational definitions and procedures for deriving those group-level measures have been described briefly in Chapter III and are presented in detail in Appendix B.

Given the concerns raised in the above literature with respect to non-normality of the intra-individual and cross-sectional absenteesim data distributions, any research procedure which relies upon distribution-based inferential hypotheses tests (such as T-tests) must first examine the extent to which the use of those tests may be relied upon. For practical reasons, such as estimating the costs of absenteeism or the benefits attributable to an absence control program, some sort of direct comparison between daily cross-sections appears to offer a very helpful first approximation. The use of both annual and quarterly (100-day) moving values as derived data in the present research was in part chosen specifically to simplify interpretations of such cross-sectional comparisons. These comparisons were implemented in the present work and so were the associated T-tests. The following hypothesis was used to examine the merit of those T-tests for statistical inference.

Hypothesis:

> The distribution of employee time-use as absenteeism will, in general, depart from normality so severely that the confidence levels and related statistics associated with T-test comparisons between the means of daily cross-sections are, at best, questionable and inaccurate. (See Chapter IV, Hypothesis 1).

The second major methodological issue raised in the above absenteeism literature concerns the apparent unreliability of traditional absence measures. This appears to stem from a number of sources and has stimulated a controversy over whether absence severity (or duration) has any merit as a measure at all, or whether it should in fact be abandoned in favor of frequency-based measures. Mathematically this frequency-severity controversy is related to the problem of data transformation. Frequency is, by construction, a power transformation of severity -- a transformation of the numerical value for severity to the zero power. The practice of using power transformation to adjust data distributions toward symmetry (and often thereby increasing the apparent strength of observed relationships) is common in the natural sciences, but a transformation to the zero power <u>always</u> reduces the transformed value to 1, and this never allows reconstruction of either the raw data or the subsequent summary statistics by transformation back toward the

original state. Use of frequency over severity is the least conservative power transformation that could be selected, removing along with some measurement error most of the information in the raw data. The stronger reliabilities which have traditionally been observed for absence frequencies as opposed to severities are partially indicative of whatever remains of the information in time-use severities after transformation to the zero power has occurred. Since most of the observed variation tends to get removed along with that process of transformation, the correlations between two cross-sections of frequency data ought to be higher by construction.

From a procedural standpoint, another problem also undermines the use of cross-sectional over-time comparisons to test reliability. Variation both from measurement or sampling error and from real changes in the absence behavior over time is grouped together in the respective cross sections. Any differences noted between the two cross-sections must therefore be a combined effect of both measurement or sampling error and change in the observed behavior. And, these two contributing component effects are undistinguishably combined with each other in the cross-sectional data so that any real changes observed over time will operate directly to reduce the apparent reliability -- the correlation between measures of respective individuals' absence behaviors over time. Once again, the apparently low reliability statistic is to be expected -- an artifact of the analytical procedure rather than the measure of absence itself.

To circumvent this problem, the present procedure has employed a time-series approach both to measurement itself and to analysis, instead of only using the traditional comparisons of two daily cross-sections. As a result, the reliability of the measures over-time may be ascertained using split-half correlations between two time-series of group means. These means represent the daily values of time-use behavior for randomly selected halves of a sample of workers. The reliability statistics generated can be affected by sampling error and by measurement error, but will not be undermined by the inclusion of observed "real" changes along with measurement error in the reliability data. Details on the time-series correlation procedure are presented in Chapter IV and in Appendix C.

Given the construction of the time-series measures used here, the concerns raised above about reliability could then be addressed through specific tests of the new time-series measures. This was done by testing the following two hypotheses.

Hypotheses:

The time-series correlations between the two sub-sample time-series of group means for both severity measures and frequency measures will be significant and positive and approach r = 1.0. (See Chapter IV, Hypothesis 10.)

and,

Neither frequency nor severity measures demonstrated a substantially greater time-series reliability than the other. Both were reliable measures. (See Chapter IV, Hypothesis 10-a.)

Variation Due to Demographic Subgroup Differences

In a given workforce, a number of systematic variations in absence behavior attributable to population demographics of employees can be expected to occur. Mowday and Spencer (1981) noted differences in employee absence behavior associated with type of work and scope of job occupation. Steers and Rhodes (1978, 1980) offer a broad-based, theoretical model including a very wide selection of plausible demographic influences. Chadwick-Jones, Nicholson and Brown (1982) suggest that subsample groups selected on the basis of industry, sex, age, type of occupation, and pay practices may all be expected to differ in absence behavior. Further work by Hill and Trist (1955) suggests that length of service, individual accident rate, and time of shift worked, as well as a number of other issues may also determine employee subgroups that are characteristically different in absence behavior.

Demographic variations within absence behavior are of interest both because these would allow the effectiveness of measures to be examined under varying subcultural conditions and because analysis within these subgroups, exclusive of other employees, should result in less "error" of measurement. Nonetheless, comparative analysis by demographic subgroups has been deferred for detailed attention in later work. The random selection of a sample to represent <u>all</u> employees in the DOT workforce presumably precluded any undue and excessive influence from over-representation of demographic subgroups. A comparison of sample demographics with the workforce demographics has been presented in Chapter III, Figure 2.

Preprogram Shifts in Employee Time Use

Since the absence control program included a heavy component of management development and supervisor training prior to the start-up of the system, some anticipatory changes in absence levels were expected to occur prior to full enforcement of the control program. To check for and discover the general direction of such preprogram changes, the time-use data for the first 30 days of the baseline year were compared with the data for 30 days directly preceding full start-up of the program in its first complete year. What was really expected was a reduction of short-term discretionary absence, but since backlash effects due to employees taking advantage of a last chance to "get away with" absenteeism was another possible outcome, the following hypothesis was tested -- as a null hypothesis.

Null Hypothesis:

HO: There were no significant changes in any employee time-uses prior to program installation. In particular, absenteeism levels will have remained steady. (See Chapter IV, Hypothesis 2.)

or,

H1: Significant preprogram changes in absenteeism occurred before full installation. Since the two time periods involved were both in the early summer and prior to full program implementation, seasonal changes and other cyclical effects were not expected to be substantiated. More specific details have been presented in Chapter IV.

Patterns of Cyclical Change in Absence Behavior

Various studies on absenteeism have noted the need for longitudinal research (Markham, Dansereau & Alutto, 1982a, 1982b; Nicholson, Jackson & Howes, 1978; Breaugh, 1981; Hammer & Landau, 1981). To date, most such work has been conducted by using multiple cross-sectional comparisons and graphical plots of change over time. These procedures were extended by Chadwick-Jones, Nicholson & Brown (1982) in detailed week-by-week graphical comparisons which reveal definite cyclical and systematic time-series patterns of change over time.

The investigation of these phenomena, using precise and reliable day-by-day measures in a systematic time-series analysis is a necessary step in studying absence. There has been evidence presented by many prior researchers which did lead to an expectation of seasonal shifts and holiday effects in the present research, and which suggested a recurrence of such behavior in annual cycles. Where the research methodology depends on comparing daily cross-sections (as in the T-tests here), these cyclical changes could easily be incorrectly incorporated into estimates of overall change. So, it was appropriate to examine the data for various types of seasonal shifts prior to further analysis.

Furthermore, the absence control program itself included an annual renewal of the number of casual leave and vacation days available to each employee, and this happened every year on July 1. As a result, the occurrence of cyclical effects was all the more likely to follow a pattern of annual renewal. Also, the seasonal and holiday effects noted above seemed very likely to be compounded by other short-term intra-annual fluctuations due to the unusually strong observed substitution effects discussed below.

Three general types of hypotheses about cyclical change were tested using T-test procedures between carefully selected points in time. But, because of the empirical conditions under which raw data were originally generated, the operational hypotheses presented in Chapter IV had to test for some of the above effects as combination effects, could only test for purely seasonal effects within the baseline year, and could test for annual renewal effects exclusively across periods within the post installation years. In short, the "theoretical hypotheses" below don't necessarily correspond directly to "operational hypotheses" as tested (see Chapter IV). In some cases, combined effects of two or more different cyclical shifts had to be made the focus of inquiry. Specific details of the circumstances surrounding the careful selection of daily cross-sections to be T-tested have been explained in Chapter IV. The following three theoretical hypotheses about cyclical changes were investigated, not to find all instances of occurrence, but to get some guidance for refining the interpretation of later overall changes attributable to the program itself.

Hypotheses: (See Chapter IV):

> Absenteeism levels will show seasonal differences, increasing in midwinter and midsummer, but decreasing in early spring and fall (Hypothesis 3, Hypotheses 5 and 8).

And,

> Absenteesim will tend to have increased at holiday periods in late December and early July due to a recurring holiday effect (Hypothesis 4, Hypotheses 5 and 7).

And,

> Due to program installation and a legitimacy effect (discussed below), post installation levels of short-term non-legitimate absence will have been significantly higher prior to annual renewal rather than after it; and short-term legitimate absence will have increased after annual renewal (Hypothesis 8).

Theories of Unintentionally Induced Over-Time Change in Absence Behavior: Internal Substitution and Progression Effects

Theories about internal shifts or substitutions between different forms of absence (an internal "loophole" through which employees could very possibly find "safe" ways to continue their absenteeism and which would dilute the desired impact of an absence control program) have often been advanced in the literature (Hill & Trist, 1953, 1955, 1962; Chadwick-Jones, Nicholson & Brown, 1982; Nicholson, 1976; Braun, 1978). Two of these together, the legitimacy hypothesis and the progression hypothesis, posit shifts of absence behavior from illegitimate short-term to increasingly legitimate longer term absence behavior under the impact of an absence control policy. Substitution of one form of withdrawal for another has been discussed in terms of: accidents and sick leave substituting for regular work (Hill & Trist, 1953); absence plus overtime work for regular work (Allen, 1981); absence instead of other forms of withdrawal (Mobley, 1982); absence prior to turnover; and tardiness or sabotage or drug abuse or turnover as substitutes to absence (Parasuraman, 1982).

Progression:

Early theories (Hill and Trist, 1953; Herzberg, Mausner, Snyderman & Capwell, 1957) posit a fairly simple progression of individual employee absence toward increasingly severe withdrawal from work. The focus in those studies was on the progression of individual employees into increasingly severe forms of withdrawal and not on a group-wide progression effect. As a consequence, the time-series procedures in the present research could only have discovered support for the progression effect if the entire sample showed a general tendency toward increased withdrawal from work. Since that would have forced the DOT to increase their workforce, this was not likely. But, so that it would be possible to check for such a group-level progression effect, the following hypothesis was tested.

Hypothesis:

> A pattern of progression from short-term legitimate absence into long-term legitimate absence will have occurred at the group level. A negative and significant correlation between these two forms of absence will be found when the two series of daily means are correlated (see Chapter IV, Hypothesis 11).

Legitimacy:

More recent work (Parasuraman, 1982; Hammer & Landau, 1981; and Nicholson, Jackson & Howes, 1978) posits a progression in the pattern of substitution from uncertified and non-legitimate but less serious absence to certified and legitimate absence, and to longer term and more serious forms of withdrawal including turnover. The substitution effects would tend to occur in the same direction as the progression effects under the implementation of absence control. A shift of employee time-use toward more severe but legitimate forms of absence from short-term non-legitimate absence could very likely occur. In the absence of specific practices to prevent these substitutions, such internal shifts are a highly plausible side effect of absence control. Given daily precision of measurement and a high level of reliability in the measures, it was possible to distinguish here between these two patterns in a few selected specific comparisons, either on the basis of timing of the behavioral shifts between the various measures across two cross-sections (the T-tests to compare means), or by using a time-series correlation technique. Procedures to do this have been briefly discussed in Chapter IV with a more detailed methodological note in Appendix C. The results are presented in Chapter V.

The legitimacy effects were one very plausible overall result of the absence control program, an alternative to working. The T-test procedure offered a simple but workable "estimate" of program impact which could be used to discover overall legitimacy effects under specific circumstances as detailed in Chapter IV. The following hypothesis was therefore tested.

Hypothesis:

> After program installation, significant year-to-year changes between corresponding dates will have occurred and, in particular, non-legitimate absence will have decreased while legitimate forms of absence will have increased (Hypothesis 6 and Hypothesis 7).

The time-series correlation procedure was specifically tailored to allow analysis of shifts within time periods between different forms of time-use. A discussion of the method is presented in Chapter IV and supplemented with further detail in Appendix C. By means of this procedure the following hypothesis was tested.

Hypothesis:

After program installation the level of non-legitimate short-term absence will have tended to decrease and the levels of long-term and short-term legitimate absence will have tended to increase over time. As a result, the time-series of daily mean levels for legitimate and non-legitimate absence behavior will be negatively and significantly correlated (Hypotheses 12 and 13).

Program Evaluation: An Overall Long-Term Assessment of Program Impact

The absence control program was specifically intended to stimulate and control a reduction in short-term non-legitimate absence. It was not contractually specified but was generally presumed that this would produce a corresponding increase in time worked. Despite the possibility of substitution effects subverting this unspecified but intended impact of the program, both T-tests using the cross-sectional data and time-series correlations using the time-series of daily mean time-use levels were implemented.

The T-tests were designed to discover long-term, year-to-year changes between the baseline year and the post-installation period by comparing mean time-use levels on corresponding dates. The following hypothesis was T-tested.

Hypothesis:

> Short-term non-legitimate absence in the baseline year will have decreased significantly by the end of the first year after installation, and time worked will have increased in a corresponding amount. The same patterns of change will also have become evident in comparison of the baseline year and the second year after full installation (Hypotheses 6 and 9).

To test for the same relationship as a continuing pair of interdependent trends over time, the following hypothesis was tested using the time-series correlation procedure.

Hypothesis:

> After program installation, time-series correlations between short-term non-legitimate absence and time worked will be significant and negative (Hypothesis 14).

It should be noted with regard to the various hypotheses above that the T-tests used to compare means of two daily cross-sections establish only the direction and size of effects brought about by a trend over time and provide an estimate of significance for those effects. The time-series correlations, on the other hand, only establish the strength, direction, and significance of the relationship that produced the effects, and say very little about their magnitude. These two procedures complement each other. To the extent that both help establish the direction of change, there is some redundancy.

A Summary of Theoretical and Measurement Issues

From the above discussion, a number of key points may be drawn to summarize the issues raised from analysis of the absenteesim literature itself.

1. There are internal trade-offs between absenteeism and other forms of employee time-use, such as time worked between absences and scheduled time off. Spending time in any one of these activities excludes doing the others.
2. Time-use behaviors such as absenteeism undergo patterns of continuous change over time so that careful study of them requires a time-series approach instead of a comparison of cross-sections.
3. There are short-term weekly shifts and longer term cyclical and seasonal shifts in the patterns of time-use behavior that must be identified and separated from inter-annual, long-term changes in order to assess the impact of absenteeism control.
4. Some forms of absence are more legitimate than others and employees tend to be absent for a variety of different reasons. There may be a causal relationship between the legitimacy of an absence type and its occurrence.

5. The occurrence of absence may also tend to vary somewhat depending on demographic characteristics of the workforce, notably: type of occupation, sex, age, industry, compensation practices, time of shift worked, length of service, and individual accident rate.
6. Absence control policy decisions require data that describe time-use for the group. Absenteeism discipline policies must apply and be applied to all employees equally if enforcement is to be equitable and legally defensible.
7. For comparison with absence figures across samples from different work forces, absence rates have typically been designed as overall summary statistics, reflecting various patterns of group behavior independent of the numbers of employees involved.
8. Duration of the absence, and frequency of occurrence are both popular objects of measurement for employee time-use behavior. There exists a lively controversy as to which is a better measure. Duration is more conceptually attractive. Frequency has been recommended as a preferred alternative because it has a higher test-retest reliability correlation.
9. Data from absence behavior tends to distribute itself abnormally. There is a truncated lower end due to the inclusion of values for employees who are never or almost never absent. There is a strong positive skew

to the distribution which could be easily and conveniently corrected if the truncation problem were not present.

10. Each employee has a characteristic short-term tendency to be absent-prone or non-absent-prone. This tendency changes a great deal over time so that accurately determining membership in the absence-prone or non-absence-prone group would require re-establishing the group composition daily.

A set of time-use measures which is responsive to these insights suggested by prior research has been developed for the present work and described in detail in Chapter III and in Appendix B. The derived measures used in the dissertation are listed and annotated in Figure 14 and Figure 15. These data were used to test operational hypotheses based on the theory-based hypotheses presented above, to examine in turn the respective theories about absence behavior. The methodology and operational hypotheses have been presented at length in Chapter IV, with detailed methodological notes in Appendix C. Results and interpretations are summarized in Chapter V.

CHAPTER III

Data Collection, Sample Demographics, and the Derived Measures

In order to address the conceptual and measurement issues raised by absence research, a very particular kind of database was required. The details of access and data collection for such a sample have described below, followed by a description of the sample demographics. Perhaps the most complex refinement in the present work, and a central element of the research program, was the conceptualization and creation of the various derived time-series measures of time-use by the workforce. In the latter part of Chapter III is a detailed presentation of the conceptual framework and procedures which generate the derived time-use measures.

Access to the Research Site and Data Collection Procedures

During 1977 and 1978, the Department of Transport (DOT) in the City of Detroit contracted with Moore & Juliano, Inc. of Lansing, Michigan (the consultants) to have them create and install a system for absenteeism measurement, reporting, and control. The data for the present research were generated as an integral part of that system. With the cooperation of both the Moore & Juliano firm and the DOT, access to data tapes containing the entire database was made possible. The purpose of the present inquiry includes an evaluation of that project. Detailed access was therefore provided to gather information about how the control system was installed, how the system for data collection was developed, how the data collection itself was administered, and how the program was implemented. A brief description follows.

The Consulting Intervention:

From the beginning, top management in the DOT was strongly committed to the development and implementation of the absenteeism program. Prior to full-scale program implementation, most supervisory personnel in the DOT were given training in the use of managerial techniques for performance planning, planning-group communication, and group problem-solving. From this base, the managers chose absenteeism as the problem they wished to address. Ultimately, the contract to install an absence reduction program was established. In the process of working together to diagnose absenteeism problems through interviews, nominal groups, and historical research, a detailed, guided exploration and categorization of their absenteeism problems occurred. From this base, through a hierarchical system of structured problem-solving workshops, the DOT Personnel Department and the consultants determined a detailed set of types of employee time-use (or absence) for all DOT employees. This typology was to be used later in coding the daily reasons for time-use of all employees on the DOT payroll, every day of the program as part of the reporting and control system and as part of the research effort to discover the extent of the problem as it evolved over time.

In many cases the data collected were already in existence for payroll purposes so that care and accuracy in data collection were mandated by law. For other types of time-use, particularly in the detailed absence categories, a technical classification decision might occasionally be required from the DOT about how to assign each specific day's worktime where some non-routine employee activity was involved. It was important to avoid forcing supervisors into a judgement-call position where they would be responsible for an on-the-spot determination of how to interpret the rules regarding an employee's time-use. The earlier training and problem-solving program had created a communication channel for propagating the necessary message. Supervisors were trained to recognize and make only the routine judgments already anticipated and defined in the context of the absence control program. They were to defer the non-routine decisions to the personnel department. Detailed technical training of the personnel staff and payroll clerks was provided. There, judgments could be made in a way consistent with the overall absence control policy and contractual commitments with the various affected unions.

Indeed, the absence-control program was actually worked out with these unions before implementation, although management never solicited public support due to the political nature of unions. All unions accepted the program without grievances. The personnel department centralized the absence control policy decisions under the control of an absenteeism program coordinator who was directly responsible to the supervisor of the payroll and personnel departments, and who was trained in policy, labor relations, data entry, and related aspects of policy implementation.

In addition to pre-program training and centralized administration of policy-making decisions, the program provided weekly computer reports on departmental absenteeism directly to the department supervisors. Other reports of a monthly, quarterly, and yearly nature were provided to the personnel department, line managers, and top executives. Each supervisor was required to implement counseling and disciplinary procedures such as warnings and disciplinary days off as soon as these became necessary due to evidence of excessive, non-legitimate absenteeism for any given employee. Steps in the policy were based on the historical research study conducted by the consultants which determined average absence rates (days absent per year or month) and inactivity rates (percentages of the workforce absent per year or month). These are listed by organizational unit in Figure 1 below, along with a tabular summary of absence control procedure. Steps in the policy occurred when the employee hit various percentage levels (i.e., Step Three at

50%, Step Four at 100%, Step Five at 150%, termination at 200%). A continuous track record was maintained for each employee and the record of violations was wiped clean every year on a rolling one-year basis. Each July 1, employees received a fresh bank of casual leave and vacation days as provided for by collective bargaining agreements. Based on each employee's track record, both disciplinary measures and an attendance bonus were awarded.

The overall effect of all these procedures was an unusually consistent and uniform implementation of both data collection and absence control. Accordingly, the data from July 1980 onward contain little or no contamination due to supervisors parlaying discipline decisions into political and personal favors, a problem present in the archived baseline data. These issues were structured out of the supervisor/subordinate relationship and out of the supervisor's role in every way it could be done by utilization of time reporting systems and specially trained data entry personnel. Careful pre-planning to prepare for data collection and training of the supervisors was vital. The vast bulk of time-use data were recorded directly from the payroll cards ... particularly for the baseline year (July 1, 1978 to June 30, 1979).

Unit Total	Historical Averages During the Baseline Year: Average Inactivity Rate %	Average Non-Legitimate Rate (Days per Employee per year)
D.O.T. Total	4.3	10.18
D.O.T. Excluding Trans.Division	4.08	9.58
Transportation Division	4.17	9.8
TEO	4.17	9.8
Rolling Stock (Maintenance)	4.5	10.57
Mechanics	4.7	11.04
Plant Maintenance	3.16	7.44
Personnel	4.25	9.98
Administration	1.17	2.75
Audit	1.42	3.33
Other	.92	2.15

Attendance Control Step:	1	2	3	4	5	6
Occurrence Trigger:	*Casual Leave Exhausted	First Non-Legitimate Absence	5	10	15	20
Hours Trigger:			40	80	120	160
Step Action Required:	Casual Leave Caution	Step 2 Notice	Written Warning	Second Written Warning	Suspension & Mandatory Attendance Adjustment Period (2 week min.)	Thirty Day Suspension Pending Discharge
Percent of DOT Historical Average:			50%	100%	150%	200%
Action Taken by:	Supervisor	Super-visor & Atten-dance Coordin. at loca-tion	Super-visor Atten. Coord.	Attend. Control Board & Attend. Control Mgr.	Attendance Control Manager (Personnel Division)	Attend. Control Manager (Personnel Division)

(Union representation is appropriate for steps 2-6 since these call for discipline steps.)

Moore and Juliano, Incorporated
1980

Figure 1: The City of Detroit Department of Transportation Attendance Control Operations Guide

The reasons for unscheduled absence were limited to a fixed set of codes and each event was recorded by the employee's supervisor if the situation was routine and clearly required no judgement calls. Otherwise these decisions were referred to the program coordinator, primarily to preserve consistency for enforcement of the related disciplinary measures wherever use of discipline became necessary.

Despite all the care given to proper coding of the time-use data, two kinds of irregularities did occur in the records of certain demographic data taken from the payroll cards. Some of the demographic data that were collected were miscoded such that those categories of data were rendered unusable. Also, some of the demographics, including both salary and hourly wage data, were stored on the new data tapes in a machine-specific compressed format and are not easily accessible outside of the DOT system. Fortunately, the portion of demographic data which was coded in usable form is complete enough to provide an excellent sampling frame for the present research. More details on the demographics are presented below.

The same type of irregularity in the data coding affected certain categories of the time-use data itself, especially in the baseline period (1978-1979). Again the effect was consistent. But, although its presence could be identified, this type of irregularity could not always be pinned down to one unique reason for occurrance. So, in the program which generates the derived dataset, a procedure has been included to recode those ambiguous data as "missing values."

Preparing the Time-Use Database

A brief summary of procedures for recoding the raw data and deriving the time-use indices has been presented later in this Chapter, and a detailed explanation is included in Appendix B. In essence, there was an extreme abundance of raw data made available for the present work ... some ten million characters of data stored in a datastructure which implicitly made use of the order of storage to keep track of the date to which each data value referred. These data had to be cleaned to correct any systematic miscoding, and then had to be reorganized as follows.

The demographic data were separated out from the time-use data and edited to eliminate redundancy, and to find (and where possible, correct) any errors or omissions. The data on absence type and absence severity and date (time-use data) were also split out into a separate database. In the process, the date information was used to calculate a relative date for each day's data, starting with day 1 on January 1, 1978. At this time, the social security number for each employee was replaced with a four digit case label to preserve confidentiality, and this was included as a sorting field in both the demographic data file and the file on relative date, absence type and severity. Only those demographics to be used in further research were included. Furthermore, the new case label, relative date, absence (or time-use) type, and absence severity were all written out to a separate tape file.

This new file is unpacked, restructured so it can be used with more flexibility in the data analysis. It contains all the raw data and its respective relative date information, and it is bigger than the original, about 30 million characters in size instead of 10 million, even without the demographic data. Yet this absence data was still in very raw form. Much sorting and recategorizing, as well as the calculation of annual moving values for each measure on each time-use index, each day, had yet to be done. The rationale and procedures for creating derived data structures have been explained briefly following the descriptions of sampling procedures and workforce demographics below. For convenience of comparison, the sample demographics have been presented together with the workforce demographics in Figure 2. The sheer size of the detailed database for the entire workforcce and the amount of computer memory and processing time required to sort and prepare each case for the derived database made it necessary to analylze a representative sample of the cases rather than the entire workforce dataset. As a result, the following sampling procedure was implemented.

The Sampling Procedure

Selection of a sample had to satisfy a number of requirements. (1) The sample had to be randomly selected by a procedure which gave every one of the 2336 cases an equal chance of being selected into either half of the split sample. (2) The two split halves had to be randomly selected at the same time such that both included about the same number of cases and the total number of cases desired was about 160 (a rough estimate of the maximum number of cases the computer could effectively process). (3) To maximize the amount of independent information brought from the overall workforce data and selected into the sample, no case was considered for inclusion more than once. Sampling was done without replacement. (4) For convenience in passing the resulting file of selected cases to the program which executed the next step in analysis, the data had to be kept as a fortran data file. The simplest tool to do this turned out to be a small, custom-designed Fortran program.

The sampling program actually worked as follows. The computer read in three values given it by the fortran program: a randomly chosen large odd number, a value for the number of random choices the computer was to make (2336), and a sampling fraction (160/2336 in the present instance), which the computer was to use in deciding to include or reject a case from the sample. The program then used a fortran subroutine (called GGUBS) from the International Mathematical and Scientific Library (a bank of standard computer programs that the Fortran language users can include in their own software to solve special but commonly encountered problems). The subroutine which was used in this case took the three values given to it and generated 2336 numbers between 0 and 1.0, drawn from a uniform random distribution. Then the computer read in the first case of data from the demographic data file, compared the first random number (between 0 and 1.0) with the sampling fraction (160/2336 to get about 160 cases out of the entire 2336 cases), and if the random number was less than or equal to the sampling fraction, the computer included that first case in the sample. If the random number was greater than the value of the sampling fraction, the case was then omitted from the sample and the computer tested the second random number against the sampling fraction to decide whether to include or omit the second case. Then it tested the third random number and decided about the third case, etc. At the end of the process a total of 164 cases had been selected for inclusion. The four extra cases (164 rather than only 160 as planned) were presumably generated as a consequence of chance variations in the values generated by the random number generator.

By this procedure the 164 case sample was selected. At the same time, the overall sample was split into two subsets. In effect, the first case to be included was written onto a file for the first subset, then a second case was randomly selected from the cases that had not yet been considered for inclusion and this second case was written into the second subset. Then a third case was selected from the remaining cases and written into the first subset, etc. In each case selection, no case was considered for inclusion which had already had a chance to be included in either half (or subset) of the sample. After all 2336 cases had been randomly selected and included in each subset, and since no case could be present in both subsets, the entire sample was composed of 164 unique randomly selected cases. The representativity of this 164 case sample has been examined by comparing demographics for those cases with demographics for the 2336 cases in the entire DOT workforce. Results are summarized in Figure 2.

<u>DOT Workforce Demographics</u>

The Department of Transport is a predominantly male, black workforce composed of about 60% bus drivers, 20% mechanics, and 20% distributed across various other occupations. The total number of employees is variable, but over the course of the present research, 2336 employees participated. A detailed summary of demographic characteristics was prepared for both the overall workforce and the sample drawn to represent the entire workforce. Those details are presented in Appendix A. Both have been summarized below in tabular form to allow convenient comparison of the sample with the workforce characteristics.

Characteristic	Value for All Employees	Sample Value
Number of Employees	2336.0	164.0
Median Seniority Year	1974	1974
Modal Seniority Year	1979	1979
Median Age in 1982	37	38
Modal Age in 1982	32	32
% Transportation Equipment Operators	57	55
% of Single Employees	83.0	83.5
% of Employees with no Dependents	68	73
% of Workforce by Race and Sex		
% Black	79.5	78.0
% White	18.9	22.0
% Male	84.8	87.2

Figure 2: DOT Employees' Demographic Characteristics, For All Employees and for the Research Sample

As Figure 2 indicates, demographic characteristics for the sample are generally quite similar to those for the workforce on the characteristics listed. Furthermore, as the differences between the median and modal values for age and seniority year indicate, both the sample and the overall workforce are going through changes in composition, of the same direction and magnitude.

Procedures to Generate the Derived Time-Use Indices (See Appendix B for more detail.)

The raw absence data contains daily values for the case number, relative date, absence type, and "absence" severity in hours. The first two values are assigned to the second two and identify to which employee and on what day those last two data values relate. The meanings of these absence data values are dependent on several things.

1. Of primary importance is the reason for absence (or worked time) which was originally coded for the specific employee on a given day.

2. Second, the duration of the occurrence of the time-use over time may persist until it forces a recoding of the entire occurrence or at least of part of it. Indeed, the duration of an absence episode may initially be understood to be short and yet later may turn out to last for weeks, or even months. Reassignment of such time-use occurrences to corrected categories from their onset is the first of two main functions served by the Fortran program which generated the derived database.

3. Third in importance but critical to the integrity of the derived data are a number of systematic coding errors which had to be removed from the data before they could be interpreted and used.

The logical structure of relations between absence reasons has been used to create a program that converts each primitive "absence reason" (Figure 3) and duration coded in the raw data into derived measures (Figure 5) for one or more of the 13 time-use indices. Figure 4 is a flowchart of the recategorization procedures. Data values listed in Figure 3 were resorted according to a set of recategorization rules to determine the values of derived data items listed in Figure 5. The recategorization rules used have been explained following the flowchart.

Read the flowchart (Figure 4) from right to left, following the arrows. Note that where the condition associated with a given arrow in the flowchart was satisfied, the data values in the raw data categories (on the right) will have been processed as contributing elements and added into the associated higher-level "resultant" category of time use (on the left). "Resultant" or derived data categories are each marked with bold-faced alphabetic characters in the respective box of the flowchart, and these alphabetic characters correspond to the alphabetic labels associated with each of the derived time-use categories in Figure 5. Similarly, there are bold-faced numeric labels associated with each raw absence or time-use category listed in the flowchart and these numeric labels correspond to the numeric labels for each Raw Absence Data Type listed in Figure 3.

In the original data, coding procedures describing absence, time worked or time scheduled off for DOT employees used 45 reasons to categorize these various time uses. These reasons are listed on the following page in Figure 3. The actual language used to label these reasons is the official DOT terminology for categories of time use.

<u>Rules for Recategorization of the 45 Raw Absence Reasons Over Time (Refer to Figure 4)</u>

The 45 categories of time-use listed as raw absence reasons in Figure 3 above were combined into the 13 derived categories listed in Figure 5 as follows.

1. Any employee time that was recorded as laid off, resigned, discharged for cause, retired, deceased, or transferred, where the employee was not rehired in 365 days or less, remained categorized as termination time. (Box M in Figure 4.)

2. Termination of an employee may be reversed later. Depending on the reasons for reversal and the time between termination and rehire, this termination time had to be recategorized into a specific resultant category. A routine, non-punitive layoff for five days or less became scheduled time off (Box D). If it lasted more than five days but less than 365 days, it became a long-term legitimate absence (Box J). If the employee voluntarily quit and was rehired in five days or less, this was recategorized as

Type 0: Time Worked on the Job

Code	Definition
1	Created for generating the derived database as a code for Day Present and Worked.
2	Off Day - Worked (employee was asked to work on a regularly scheduled day off instead of some other work day to correct a work scheduling problem for the DOT)

Type 1: Scheduled or Contractually Provided Absenteeism

Code	Definition
3	Regularly Scheduled Day Off
4	Off Day - Trade
5	Off Day - Adjustment
6	Casual Leave Day
7	Casual Leave Time
8	Vacation
9	Swing Holiday
10	Holiday
11	Death in Family
12	Occupational Injury
13	School/Training (DOT Related)
14	Leave of Absence
15	Military Leave
16	Jury Duty
17	Union Business
18	Conventions (DOT approved)
19	Laid Off
20	Civil Service Business
21	Suspended
22	Not Scheduled
23	Court Time (Witness or DOT related)

Figure 3: Summary of the Raw Absence Data Types as Recorded in the Department of Transport Absence Management Database or as Modified and then Used to Generate the Derived Data

Continued on next page

Figure 3 Continued

Type 2: Long-Term/Non-Recurring Absenteeism

Code	Definition
24	Long-Term Disability
25	Sick Employee (6 days or more)

(These codes may be used to handle catastrophic * illnesses approved by the Personnel Division without the 6 day restriction.)

Code	Definition
26	Resigned
27	Discharged
28	Retired
29	Deceased
30	Transferred
31	Long-Term with Injury
32	S & A with Injury

Type 3: Excessive/Recurring Absenteeism

Code	Definition
33	Sick Employee
34	Sick Family Member
35	Death - Not Family
36	Wedding
37	Moving
38	Off with Permission
39	Absent without Permission
40	Single Miss (TEO's only) **
41	Double Miss (TEO's only) **
42	Strike
43	Tardy
44	Other
45	Unexcused Absence

Notes:

1. According to the DOT Attendance Control Policy, employee absenteeism was grouped into three meta-categories, with time worked on the job included, there are now four:
   Type 0 - Time Worked on the Job
   Type 1 - Scheduled/Contractually Provided Absenteeism
   Type 2 - Long-Term/Non-Recurring Absenteeism
   Type 3 - Excessive/Recurring Absenteeism
2. These types are reworked as shown in Figure 4 to create the various levels and categories of derived data.

* Illnesses such as cardiac arrest, kidney failure, serious injury, etc.

** "TEO's" are Transportation Equipment Operators (bus drivers).

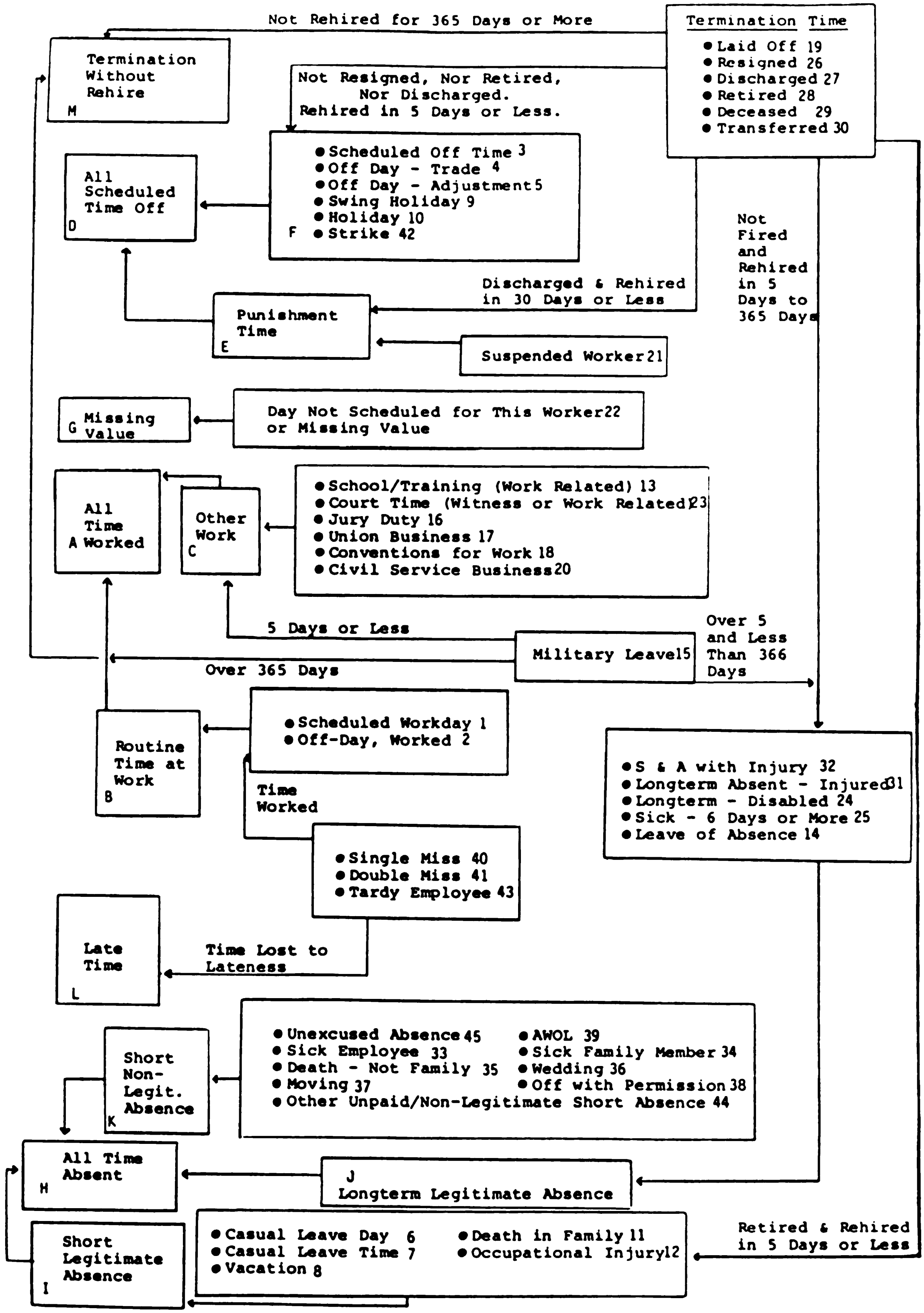


Figure 4: Logical Relations Among Different Uses of Employee Time

a short-term legitimate absence (Box I). Termination without rehire for 365 days or more remained a termination and the time use was assigned accordingly (Box M). Discharge for cause, followed by a rehire in 30 days or less was reassigned as punishment time (Box E), and other disciplinary suspensions were also put into this category. Time for a worker recorded as deceased who later turned up on the record in any category other than a form of termination was reassigned as missing data (Box G).

3. Employee time that was originally scheduled as scheduled off-time, an off-day traded from some other employee or for some other adjustment, or a swing holiday, public holiday, or strike, all were categorized as routine time off (Box F).

4. Punishment time and routine time off both contributed to become "all-scheduled-time-off" (Box D).

5. Any day not scheduled for the worker and any missing values were both reallocated as missing values (Box G).

6. Military leave for five days or less, school or training related to work, time spent in court as a witness or otherwise related to work, jury duty, union business, conventions for work, and civil service business were all recategorized as "other work" (Box C).

7. Military leave for 365 days or less was recategorized as long-term legitimate absence (Box J). Over 365 days, it became termination time (Box M).

8. Scheduled work days and off days that were worked, along with time worked on days when the employee arrived late, were all recategorized as routine time at work (Box B).

9. Both routine time at work and other work became components of all time worked (Box A).

10. Any time consumed by employees in arriving late was categorized as late time (Box L). Late days in most cases had to be split into late time and routine time worked (Box B).

11. In the absence of casual leave, instances of: unexplained absences, sick employees, a death outside of the immediate family, time off to move, absence without leave, time off to care for a sick family member or to attend a wedding, unpaid time off with permission, or other brief unpaid absences, all were categorized as non-legitimate short-term absence (Box K).

12. Casual leave days, casual leave time (part of a casual leave day split up), death in the family, time off due to occupational injury, and vacation time all became short-term legitimate absence (Box I).

13. Long-term employee time off due to sickness or accident or injury, long-term dis}iled employees, or a leave of absence for more than five days all were clustered under long-term legitimate absence (Box J).

14. Short-term non-legitimate absence, long-term legitimate absence, and short-term legitimate absence were components of all time absent (Box H).

These 13 different types of time-use categories stand in respect to each other in very specific interdependent relationships. There is some internal hierarchical clustering among the different time uses, and because the use of time is an exclusive behavior, to choose any one fundamental time-use behavior precludes simultaneously emitting any other. As a result, the flow of time-use from one category into any other or combination of others can be examined quite conveniently.

Following the paths and rules outlined in the flowchart (Figure 4), the raw absence reasons in Figure 3 have been recategorized to create the values for the 13 derived time-use measures below in Figure 5.

For a more detailed description of the procedures used to recategorize the raw data and to generate the derived time-use values needed to test hypotheses in the present research, refer to Appendix B.

Three Different Measures of Time Use: Frequency, Severity and Mean Severity per Occurrance

After recategorization of the time uses into their corrected categories, three types of moving values were calculated as measures for each time-use index, each day, for each employee's data: (1) "frequency" of occurrence for the employees on this time-use across the preceding specified

Time Use Category/Index	
A	All time worked on or off the job
B	Time worked on the job
C	Time worked off the job
D	All scheduled time off
E	Time off due to punishment
F	Routine normal time scheduled off or strike
G	Missing value or unscheduled day
H	Time absent for all reasons
I	Short-term legitimate absence
J	Long-term legitimate absence
K	Short-term non-legitimate absence (except late)
L	Late time
M	Unended termination time (no rehire)

Figure 5: Derived Time-Use Measures

length of time; (2) total duration or the sum of all "severity" values for the employee on this type of time-use across the preceding specified length of time; and (3) "mean severity" per occurrance, which was calculated as a ratio of the severity over the frequency for the respective employee on the specific day. Two different lengths of time were used as base periods for calculating the moving values. 365-day moving values were used to isolate the long-term inter-annual effects because these would exclude intra-annual cycles by construction. 100-day moving values were used to examine intra-annual patterns of time-use because these were: (a) short enough not to cancel out annual cycles such as renewal or seasonal effects, (b) long enough to provide a relatively stable base for the moving value, and (c) divisible by 100 so that an estimate of the corresponding daily time-use could be made at a glance. Moving values were also used instead of single-day values (as in much conventional research on absence) in order to make the means a convenient moving average suitable to smooth out sporadic highs and lows to be expected due to the variability of individual absence behavior. When these values were then averaged across their respective sample groups, they were to provide a larger aggregate or base of group behavior for study. And, from this base, a more stable and reliable estimate of the mean daily values on each time-use index was then calculated for the respective sample group.

These three types of moving values, for both the 100-day base and the 365-day base, and across all 13 categories of time-use, became the substantive basis of the research reported in Chapter IV and Chapter V. A more detailed description of the procedure and the measures themselves is presented in Appendix B.

CHAPTER IV
Methodology and Operational Hypotheses

Introduction:

This introduction is a brief survey of the framework of procedures and tests implemented here. Further detail on each aspect follows. The methodology and hypotheses described below are specifically designed to address questions of program impact and to investigate theories of absence behavior by analysis of group level measures. The steps in this analysis and related hypotheses follow exactly the sequence presented in Figure 6. Then the results generated by implementing these procedures also follow the same sequence and these have been presented and discussed in that order in Chapter V.

Because of the problems of non-normality typically encountered in data distributions like the ones used here in the T-tests, and because of the controversy over the use of severity versus frequency measures, it was first necessary to test procedural hypotheses concerning normality for the T-test data and reliability for the time-series measures, and to establish the relative merit of the present program evaluation paradigm. Then a selection of tests had to be executed to check for spurious short-term shifts such as seasonal differences. In addition to shedding light on some non-central substantive issues, that information was

Cross Sectional Analyses

A

Normality Tests

1. The 30-day datasets used to assess pre-program impact.
2. The 100-day short-term T-test data.
3. The 365-day long-term T-test data.

B

T-Tests

1. On the 30-day datasets testing for pre-program overall shifts.
2. On the 100-day dataset testing for pre-program patterns:
   a. seasonal effects
   b. holiday effects
   c. combined seasonal & holiday effects
3. On the 100-day dataset testing for change during the program.
   a. year to year change within the program excluding seasonal and holiday effects and intra-annual program induced effects
   b. holiday effects but not seasonal or annual renewal effects
   c. seasonal effects and annual renewal, but compensating for holidays
4. On the 100-day dataset to examine overall inter-annual shifts due to the program.
5. On the 365-day dataset to determine overall program impact on the annual time-use indices exclusive of seasonal effects and intra-annual substitution effects and holidays.

Time-Series Correlations

C

Split-Half Time Series Reliability Correlations

1. On the 100-day dataset
   a. during the baseline year, after the 100-day moving values were developed from day 281 to 546.
   b. during the post-installation period from day 982 onward when the 100-day moving values were again fully developed.
2. On the 365-day dataset from day 1247 onward, the span during which moving values of the long-term annual measures were fully developed.

D

Time-Series Correlational Tests of Substantive Hypotheses

1. On the 100-day dataset to see the strength of correlations indicating short-run:
   a. progression effects
   b. legitimacy effects
   c. combined progression & legitimacy
   d. Shifts toward or away from work or absence.

   * These will include co-occurrent intra-annual shifts.
2. On the long-term 365-day dataset to see the strength of long-term inter-annual effects, exclusive of annually repeating seasonality, holidays and renewal effects.
   a. progression
   b. substitution
   c. combined progression & substitution
   d. shifts into or away from work or absence

Figure 6: A Categorical Summary of Tests and Analyses

needed to modify the interpretation of subsequent substitution hypotheses and the evaluation of the program. A description of the normality testing procedure and the related rationale is presented here, with further details in Appendix C. Also in Appendix C is a methodological note on the procedures for time-series correlations. These were used to calculate split-half reliabilities for each of the time-use measures and to test substantive hypotheses concerning progression and legitimacy effects and shifts from absence toward more time worked.

The rest of this Chapter presents four clusters of methodology and related hypotheses. First, the data distributions for the T-test data had to be examined. Prior research has shown such distributions to be generally non-normal. In order that the relative merit of using the present data for T-tests could be assessed, the degree of non-normality in these data cross-sections required investigation. This investigation includes the steps outlined in Box A of Figure 6.

Second, the rationale and structure of the T-tests themselves were presented with related hypotheses (Box B of Figure 6). Three sets of T-tests were implemented. (A) Using the data for 30-day periods at the start of the baseline year and at the end of the pre-program installation interval, a set of general tests were done to see whether anything changed between those two time periods, prior to the absence control program itself (see Figure 6, Box B, Item B(1). (B) Similarly, the 100-day data from the base-line year were used to find out whether any seasonal or holiday effects were evident in the data since these data would have, by construction, revealed such changes, but would not be contaminated by later program induced effects. Like the pre-program overall changes, seasonal influences could have moderated any changes induced by the program and contaminated the T-tests for program impact. If no such changes were revealed, then changes later induced by the program as annual renewal of casual leave days on each July 1 would be interpretable as effects more likely due to the program than such unintended effects. The procedures involved are Steps B-2a, B-2b, and B-2c in Box B of Figure 6. A third, fourth and fifth set of T-tests were implemented to test substantive hypotheses designed to examine the data for impacts of the program. These are Steps B-3a, B-3b, B-3c, B-4 and B-5, outlined in Figure 6.

Third, due to the controversial unreliability of absence measures reported in prior research, it was necessary to establish the reliability of measures designed for the present work. The procedures used for this have been outlined in Box C of Figure 6, and a detailed methodological note on the time-series split-half correlation procedure has been presented in Appendix C.

Fourth, a similar time-series correlation procedure was used to test substantive hypotheses of both short-term and long-term change. These tests have also been outlined in Figure 6, Box D.

In all the following detailed presentation of methodology and hypotheses, the sequence of procedures and problems outlined in Figure 6 has been adhered to. In Chapter 5, the same sequence of steps has also been followed for presentation of the results.

Rearranging the Case Data by Relative Date

Once the thirteen derived time-use measures were created, the new database had to be sorted ... rearranged for group-level analysis so that all values for the 39 time-use indices that pertained to a given day were grouped together regardless of which employee they previously referred to. This was done for both the 100-day measures and the annual (365-day) measures, for the 82 cases in each random half of the sample representing all DOT employees, and the data files were then merged for these two randomly selected halves to make a third overall 164-case sample. Each of these three sorted files was written out onto a new tape file for use in later analyses for both the 100-day measures and the annual measures. In the discussion which follows, the term intra-annual effects means "cyclical and annually recurring effects based on a 100-day period," while inter-annual effects means analysis" means "based on a 365-day base period and relatively unaffected by intra-annual cycles." The letter labels A, B, C and D refer to specific clusters of steps outlined in Figure 6.

Procedures in A: Normality Tests of the Data Distributions

As a precursor to interpreting the T-tests that were to check for seasonality, substitution and preprogram change effects, or to interpreting the T-tests for overall program impact, the normality of the daily cross-sections of data had to be assessed. Using the SAS "univariate" procedure, the datasets for both the 100-day measures and the annual measures were analyzed to determine the normality of the data distributions on the dates chosen to be the focus of T-test calculations. The Kolmogorov D-statistic was used as well as more traditional statistics describing the empirical data functions. See Appendix C for details. This step was applied only to the 164 cases of data generated on selected days for the entire sample from all employees. The following procedural hypothesis was tested.

HYPOTHESIS 1:

> Asymmetry and non-normality in the data distributions as shown by the Empirical Data Function (EDF) statistics will be severe enough that the applicability of normality-based inferential statistics must be questioned.

Selected results of the normality tests are summarized in Chapter V, Table 1.

Procedures in B: T-tests of the Cross-Sectional Comparisons

Introductory explanation: For exploratory purposes and to assess consistency with established practice and past findings in absenteeism research, T-tests of the significance of differences in means over time were to be done, together with an interpretation of the EDF statistics and D-test results. These cross-sectional comparisons were done for five substantive reasons:

1. To spot-check for possible pre-program shifts in time-use attributable to anticipation of the absence control program by comparing the 30-day data cross sections;
2. To examine the patterns of time-use over time and test the significance of predicted pre-installation seasonal and holiday effects;
3. To test on the 100-day measures the substitution hypotheses that various specific shifts in employee time-use did occur from one form of absence to another during the period of the absence control program itself;
4. To test for significance of program-induced change using the 100-day dataset; and
5. On the 365-day dataset, to assess overall program impact on the annual time-use indices, exclusive of annually repeating patterns.

B-1. Pre-program Incidental Shifts Using Measures Based on a 30-Day Period:

Since there was only one year of baseline data, the derived data do not really constitute a fully annual set of moving values until the last day of the baseline year, nor do they constitute a legitimate and fully developed part of the 100-day measures until day 281. Prior to that point, the cumulative values of the time-use indices indicate both (1) observed patterns of time-use and (2) an incidental increase in these indices due to the fact that accumulation of annual values for the behavior had not yet had a full year to occur. On the other hand, essentially the same condition held true for the first year of data collected under the program, and there was one month of data collected between June 1 and June 30, 1980, which was immediately prior to full enforcement but after the start-up of the computerized absence-control program itself. It was quite possible that some very real shifts in time-use behavior might have occurred due to pre-program training and planning activities and due to employees anticipating the program. At the same time, seasonal and/or holiday-induced shifts were not expected between June and July prior to the start-up of the absence-control program. It appeared reasonable to compare time-use indices from 30 days into the baseline year (Day 211) with time-use indices after 30 days of data collection, just before the program started with a fresh bank of casual leave for each employee (Day 912).

Time-use indices for July 30, 1978 were accordingly compared with time-use indices for June 30, 1980, to spot-check for pre-program changes. Because of the brevity of the span (30 days) over which these data were being collected, and some possibility of seasonal contamination (due to a 1-month difference in timing), the results would have been be somewhat suspect, even if EDF statistics indicated normality in the distributions. Nonetheless, the following hypothesis was tested.

HYPOTHESIS 2:

H-0: Significant changes will not be observed in the mean levels of employee time-use between the end of the first month at 30 days into the beginning of the baseline year (July 30) and the end of the pre-program period (June 30, 1980).

The findings with respect to preprogram change were interpreted to determine the kind and extent of anticipatory socialization effects involving any of the 13 categories of time-use. In order to distinguish later findings on subsequent tests from anticipatory preprogram shifts, the extent of preprogram change had to be assessed. Results for Hypothesis 2 have been presented in Chapter V, Table 2.

B-2. T-tests of Pre-Program Changes on the 100-Day Measures:

To distinguish program impact and other shifts from incidental cyclical changes it was needful to assess the extent of holiday related and seasonal differences. This was done as follows.

B-2a. Seasonal effects:

"Seasonal" shifts in employee time-use were likely to co-occur with hypothesized progression and substitution effects during the post-installation period. Accordingly, the simplest way to examine the data for seasonal changes was to focus on pre-program data, and for the same reasons described above (duration of the observation period being less than one year), this meant that only the 100-day measures could effectively be used. The most dramatic evidence of seasonal shifts seemed likely to surface by comparing time-usage at the end of the winter months with time-usage at the end of the summer months, exclusive of data for the 4th of July holiday period and the Christmas season. Accordingly, the following hypothesis was tested for each type of time-use measured.

HYPOTHESIS 3:

There will be significant differences in the mean of the 100-day time-use indices between April 14, 1979, and October 20, 1979. Specifically, there occurred significant seasonal change in mean levels of employee time-use across the 6-month seasonal interval.

The 100-days preceding these dates do not co-occur with major holidays which would tend to have an erratic short-term effect, except for Labor Day and Easter respectively, balancing each other's effect. Results for Hypothesis 3 have been presented in Chapter V, Table 3.

B-2b. Holiday effects:

An hypothesis was also tested that there would be a "holiday effect" evident in comparison of the data including the Christmas season with data for the period immediately following. Time-use indices for April 14, 1979 were compared with those for January 5, 1979, on the theory that by early January any increase in the moving values due to absences related to Christmas holidays would become evident in the indices, and these would have been smoothed into the 100-day moving values by April 14. The hypothesis was as follows:

HYPOTHESIS 4:

Prior to program installation and absence control, there will be significant differences in employee time-use behavior between April 14, 1979, and January 5, 1979, a significant holiday effect.

Results for Hypothesis 4 have been presented in Chapter V, Table 4.

B-2c. Combined seasonal and holiday effects:

To check for significant changes in short-term time-use patterns between dates when the combined seasonal and holiday conditions were quite different, a comparison was made between values of the 100-day measures near the end of the baseline year and the values of the measures just after the Christmas season, on January 5, 1979.

HYPOTHESIS 5:

> There was a significant pre-program change in the 100-day time-use indices between late spring/early summer (June 15, 1979) and late fall/early winter (January 5, 1979). A combined holiday effect and seasonal shift will have occurred in the 100-day pre-program time-use measures.

Results for Hypothesis 5 have been presented in Chapter V, Table 6.

B-3. T-tests of Post-Installation Changes:

Within the span of the program (after installation) a number of changes were expected. These were examined in three sets of hypothesis tests as follows.

B-3a. Inter-annual change:

Differences between identical dates in 1981 and 1982 would, by construction, be primarily indicative of year-to-year change rather than cyclical variation. Some such long-term impact was expected. Accordingly, differences between the measures were T-tested across the following intervals.

-- March 10, 1981 - March 10, 1982

-- September 8, 1980 - September 8, 1981

-- October 20, 1980 - October 20, 1981

-- April 23, 1981 - April 23, 1982

HYPOTHESIS 6:

Significant year-to-year changes will have occurred in the 100-day measures of short-term absence and so will significant changes in the 100-day measures of long-term absence between comparable dates in 1980, 1981 and 1982. The 100-day measures of both short-term absence and long-term absence will show that inter-annual shifts occurred in levels of time-use as absence across the post-installation duration of the program. In particular: legitimate and long-term absences will have increased, non-legitimate and short-term absence will have decreased, and there will have been a significant decrease in the overall absence index, plus an increase in the levels of time-worked.

Results for Hypothesis 6 have been summarized in Chapter V, Table 6 and Table 7, for frequency and severity measures respectively.

B-3b. Holiday effects:

Changes will occur in levels of absence and other time uses in association with major holidays across periods when seasonal and annual renewal effects are not expected to have much effect. This was T-tested by comparing values on the 100-day measures for early winter (which includes Christmas) with those of late winter, and by comparing values for early summer (which includes July 4th) with those for late summer/ early fall. Accordingly, differences between the measures were T-tested across the following intervals.

-- March 10, 1981 - April 23, 1981

-- March 10, 1981 - April 23, 1982

-- September 8, 1980 - October 20, 1980

-- September 8, 1981 - October 20, 1981

HYPOTHESIS 7:

Holidays will have a significant impact on time-use behaviors. Levels of time-use at work were lower and levels of time-use in absence will have increased in association with holidays.

Results for Hypothesis 7 have been summarized in Chapter V, Table 8.

B-3c. Seasonal effects and annual renewal:

Because of the timing of annual renewal, seasonal effects and annual renewal should have both operated in the same direction, inducing high levels in short-term absence and lower levels of time worked. On the assumption that the Christmas and July 4th holiday periods would both have effects of about the same magnitude and direction, differences across carefully selected dates would show primarily this combined seasonal and renewal effect when similar holidays were set in comparison between winter and summer. To test for combined seasonal and renewal effects, differences in the 100-day time-use indices across the following intervals were tested.

-- March 10, 1981 - September 8, 1980

-- March 10, 1982 - September 8, 1981

-- April 23, 1981 - October 20, 1980

-- April 23, 1982 - October 20, 1980

HYPOTHESIS 8:

> There will have been significant changes in the time-use indices due to combinations of annual renewal, holiday, and seasonal shifts.

Results for Hypothesis 8 have been summarized in Chapter V, Table 9.

B-4 & B-5. Inter-Annual Shifts due to Program Changes:

T-tests were also implemented on both the 100-day and the 365-day measures to determine overall program impact on time-use indices where they could not, by construction, be influenced by annually repeating renewal, holiday, or seasonal effects. For the 100-day measures, this meant selecting identical dates for analysis from year to year. Because the 365-day base, by construction, automatically disincludes a day's worth of data 365 days before every time it includes one more day, any such annual cycles would have had no effect on the 365-day indices once those indices were fully developed and included a full 365 days of time-use behavior. Only inter-annual long-term change should remain evident in those 365-day indices ... the long-term program impact. The inter-annual shifts due to substitution and progression can be better detected using time-series correlations. But, overall change in the means under the impact of the program should be made evident by comparing levels of the 365-day indices from year-end to year-end. This was done by comparing the following dates.

June 30, 1981 - June 30, 1979

June 30, 1982 - June 30, 1979

And, for program impact continuing after the first program year,

June 30, 1981 - June 30, 1982.

HYPOTHESIS 9:

> There were significant overall reductions in non-legitimate absence and increases in legitimate absence and time-worked as a result of the program.

An unexpected result, due to the workers being granted 80 extra hours (2 weeks) of paid time off, made the 100-day results somewhat misleading. As a result, only the 365-day results have been reported in detail here. The 365-day results for Hypothesis 9 have been summarized in Chapter V, Table 10. A short note and discussion of the 100-day results also appears there. A graphical summary of the T-test results in presented in Chapter V, Figure 7.

Explanatory Comment on Long-Term Substitution Effects

Overall shifts from short-term to long-term absence behavior will tend to occur over time across the intervals from the baseline period to the first year of the program, and from the first to second year of the program, and from the second to the third year of the program. Because there was a practice in force of annually renewing for each employee the bank of casual leave days (which could be taken at will and which were to cover all short-term legitimate absence), there would presumably be an annual recycling of the substitution effect. The practice was to renew the bank of casual leave days on July 1 each year. Accordingly, differences in time-use patterns about that date should have provided evidence of any intra-annual substitution/progression effects that occurred. These, however, would be excluded by construction from differences in the 365-day indices, and evidence for substitution evident in the

365-day data would, therefore, be indicative of a long-term and inter-annual overall trend, not a cyclical intra-annual effect.

A Note on Calculation of the Means and Selection of the Split-Halves

In order to test the reliability of the time-use measures using the split-half correlation technique, two comparable sets of the sample data had to be compared by correlating the time-series of their means. That, in its turn, required for each index that two equivalent samples representing the workforce be selected and that their two comparable sets of means be calculated. This was done as follows.

For both the 365-day data and the 100-day data, each of the three subsets of data (164 randomly selected cases from the DOT database, and the two halves of that dataset ... split out in samples of 82 cases each by selecting every second case) were then subjected to the SAS "MEANS" procedure. More explicitly, that means a total of six datasets were analyzed: 164 cases of annual data and two subsets of 82 cases of annual data; 164 cases of 100-day data and two subsets of 82 cases of 100-day data. The means for all 39 time-use indices for each dataset each day of the program were written out day by day into a separate tape file for each set of data. These means provided the basis for all further analysis beyond the level of the T-tests, analysis based upon time-series comparisons.

C. Time Series Estimates for Split-Half Reliability in the Measures.

C-1. To test reliability of the measures for both the 100-day and the 365-day datasets, the series of daily mean values for each time-use index was compared with its counterpart in the other split-half by time-series correlation of the mean daily level of the index for all employees in the one half with that for the other half ... over all days of the baseline period or program for which the means actually represented the fully developed state of the measure. Since the relationship this procedure is designed to test for is one of near-identity where any differences between the halves would be artifacts of measurement or sampling error and not of substantive differences, these correlations should ideally approach $r = 1$. (See Appendix C for a note on the time-series correlation procedure.) This generated three sets of correlations for the 39 different time-use measures. The three periods were: (a) For the 100-day data, during the baseline period, from day 281 to day 546; (b) for the 100-day data, during the post-installation period, from day 982 to the end; and (c) for the 365-day data, during the post-installation period, from day 1246 to the end.

HYPOTHESIS 10:

> The correlation between the means on the two split halves will be significant and will approach one for each of these three sets of reliability correlations.

Results for Hypothesis 10 have been summarized in Chapter V, Table 11. Note: A test of the frequency-severity controversy could also be done by inspection of the overall set of time-series correlations. This leads to Hypothesis 10-a.

HYPOTHESIS 10-a:

> The time-series reliability estimates of daily mean values for both frequency and severity indices will both be highly reliable when these are tested by correlation with their split-half equivalents over time, but neither will have uniformly demonstrated a significantly higher reliability.

Results are presented in Chapter V, Table 11. Under special conditions, a number of the reliability values for especially small samples of time-use behavior could be anomalously low. The list below summarizes various conditions which could bring about such anomalously low reliabilities for the time-series correlation procedure. See Appendix C for a full explanation.

Conditions Conducive to Anomalously Low Reliability in the Time-Use Measures

1. The sample size can be too small, including too few employees.
2. Very few employees may participate in the respective time use.
3. The time-use itself may rarely occur, even for those who did choose it.
4. The time-use may be so carefully regulated or may otherwise have occurred with such consistency that the amount of observed variation underlying the split-half reliability estimate was too tiny.
5. The amount of time-use in one split half or the other may be constant over time causing a standard deviation of zero and making calculation of a correlation impossible.
6. The subdivision of overall categories of measurement into more detailed ones (e.g., three detailed absence categories are components of All Time Absent) may introduce too many distinctions for the amount of behavior in the sample and so cause the subsample of behavior assigned to one or more of the detailed time uses to become too small for effective study.
7. The length of the time-series may be too short to effectively establish reliability for the series.

8. The length of the observation period for collecting each day's values on each measure may be too short to smooth out the sporadic daily fluctuations.

D: Time-Series Correlations to Test the Substitution Hypotheses and Investigate the Impact of the Absence Control Program Over Time

Both the 100-day measures and the 365-day measures had potential for examining substitution effects induced by the program. Five specific subseries were created using appropriate portions of the time-series. From these two types of measures the new datasets were constructed by selecting out the first 99 days or 364 days respectively, for the time segments when part of the apparent variation was there merely due to the construction of the measures and the accumulation of daily values into incompletely developed 100-day or 365-day measures. To rely on such data would have incorrectly inflated the time-series correlations. The following datasets were then subjected to the SAS Proc Corr procedure to generate rectangular matrices of time-series correlations and associated significance levels.

On the 100-day dataset:

From day 281 to day 546 ===> baseline period

From day 982 to day 1277 ===> first program year ... after first 100 days of data collection

From day 1278 to the end (1680) ===> last 14 months

From day 982 to day 1680 ===> entire post installation period ... after first 100 days of data collection

And, on the 365-day dataset:

From day 1246 to day 1680 ===> the last 14 months

The 100-day datasets would show effects due to whatever happened to be occurring ... seasonal, holiday, annual renewal, substitution, or long-term change effects ... with only a limited capability to distinguish annual cycles from other co-occurrent patterns. However, the 365-day dataset is relatively insensitive to annual cycles by construction and could, therefore, be used to examine the purely inter-annual, non-cyclical effects. The use of these respective sets of data is described below.

D-1. Correlational Tests on the 100-Day Dataset:

These tests were done to investigate four general trends in time-use behavior over time.

D-1a. Group-level progression effects:

The first trend to be examined was a hypothetical progression from less severe to more severe forms of absence and withdrawal from work, for all employees in the workforce in general. This progression effect was investigated by correlation of the group-level time-series measures for short-term legitimate and long-term legitimate absence. A significant negative correlation would have supported the hypothesis that an inverse relationship between short and long-term absence existed and that a progression effect occurred.

HYPOTHESIS 11:

> The time-series correlations between the short-term legitimate and long-term legitimate absence indices in the 100-day dataset within the span of each year will show that significant and negative correlations occurred. Short-term legitimate absences were replaced by long-term legitimate absence each year.

Results for Hypothesis 11 have been summarized in Chapter V, Table 12.

D-1b. Legitimacy effects:

The argument has been made that under the impetus of an absence control program, employees will shift the form of their absence instead of eliminating absence from their behavior. Non-legitimate absence under such conditions will tend to be replaced with legitimate absence. To test for this effect, the following hypothesis was tested.

HYPOTHESIS 12:

> Time-series correlations between non-legitimate and legitimate short-term absence in the 100-day dataset over all four datasets will be negative and significant, indicating a legitimacy effect on the short-term absence.

Results for Hypothesis 12 have been summarized in Chapter V, Table 12.

D-1c. Combined progression and legitimacy effects:

If both these effects were present and working together, then the combined effect should have been stronger than either progression or legitimacy when the two operated in tandem. To check for this combined effect, the following hypothesis was tested.

HYPOTHESIS 13:

> Time-series correlations between non-legitimate short-term absence and legitimate long-term absence will be significant and negative, indicating a combined progression and legitimacy effect.

Results for Hypothesis 13 have been summarized in Chapter V, Table 12.

D-1d. Shifts toward work and away from non-legitimate short-term absence:

The intended result of the program was to induce a shift directly from short-term non-legitimate absence into work itself rather than various other time-uses. To verify this impact of the absence control program, the following hypothesis was tested.

HYPOTHESIS 14:

> Time-series correlations between time worked on the job and the short-term non-legitimate absence measures will be negative and significant.

Results for Hypothesis 14 have been summarized in Chapter V, Table 12.

Each of the above hypotheses (11, 12, 13 and 14) was tested with respect to each of the four time periods (see page 90). across which correlations among the 100-day measures had been calculated.

D-2. Time-Series Correlations Between Measures in the 365-Day Dataset:

These correlations were calculated for data from day 1246 onward to examine the strength of long-term inter-annual effects, exclusive of annually repeating seasonality, holidays, and renewal effects. As was done for the 100-day data, the 365-day measures were analyzed for four patterns of change: (a) progression, (b) legitimacy, (c) combined progression and legitimacy, and (d) shifts between work and absence. Results are presented in Table 12.

For all of the above hypotheses tests, tables which summarize the results have been explained, presented and then discussed in Chapter V. The sequence of presentation in Chapter V for results of the procedures described above has been described in Figure 6. Step-by-step, it follows the same general order of presentation used here in Chapter IV to outline the methodology and operational hypotheses.

CHAPTER V
Results and Discussion

Introduction:

The results presented here in Chapter V are organized to parallel the discussion of methodology presented in Chapter IV. Sections A, B, C, and D in the methodology have described the procedures which produced corresponding Sections A, B, C, and D in these results. See Chapter IV, Figure 6 for an outline of these procedures. Results for all hypotheses which rely upon a specific analytical procedure have been presented together, either in a single table or in a cluster of consecutive tables or time-series graphs. A selective summary of the results relevant to each hypothesis has been reported here. A more detailed presentation of the results can be found in Appendix D. For reader convenience, a brief descriptive guideline for interpreting the various results and summary statistics presented in the tables precedes each table or cluster of tables. Accompanying each respective cluster of tables, the results have been discussed and interpreted relative to the related hypotheses presented in Chapter IV. The extent of support for each hypothesis was assessed and examined briefly in the light of other co-occurrent findings.

Results of the (A) normality tests and statistics describing distributional form have been presented first; followed by (B) the T-tests; (C) the time-series split-half reliability correlations; and (D) the time-series correlational tests of substitution effects.

Note: For completeness, the database was constructed to allow for the occurrence of termination time. But for all results presented here, no findings on time-use indices for termination time have been reported because in the randomly selected sample of 164 cases, which is the basis of the present report, no occurrences of termination time were reported in the raw data. Consequently, that time-use category is neither needed nor available to help account for the impact of the program on the system of employee time uses.

A: Normality and Empirical Data Function Statistics for the T-Test Data

The selected EDF statistics described below have been tabulated here to illustrate the distributional forms typical of data used in the T-test analyses. Since there was an overwhelming repetitiveness in these EDF statistics, only statistics for All-Time-Absent have been reported here, and only for the dates selected for T-tests of year-to-year overall change. A more detailed set of tables in Appendix D allows comparison of the absence distributions with those for all-time-worked and for all-time scheduled off. EDF statistics for the 30-day, 100-day and 365-day datasets have been presented in Table 1 below, following a brief explanation of the meaning and usage of each type of statistic in the table.

Below is an annotated list describing EDF statistics in Table 1 below.

Relative Date:

Refers to the relative date to which the daily cross-section of T-test data pertain. Day 182 means July 1, 1978, day 183 means July 2, 1978, etc.

N:

Refers to the number of cases for which data existed on the respective day. Variation in N is indicative of the extent of turnover across the time duration of this study. Of the 164 cases in the sample, 126 were there on day 546, others hired on later, and still later a few employees apparently quit.

Mean and Med:

Refer to the group mean value of the selected time-use index on a particular day, or to the group median value, respectively. The units of measurement involved are: frequency of absence episodes per time period, total severity of absence episodes added across the time period, and mean severity (which is a ratio calculated each day as severity divided by frequency). The size and direction of differences between mean and median are crude but useful indicators of the direction and magnitude for departures from normality. Where the mean is greater than the median, the distribution will generally be positively skewed.

Std. Dev.:

Refers to the standard deviation for data in a particular daily cross-section.

Skew:

Refers to the skewness of the data distribution, ideally, in a normal distribution, skew = 0. Any negative skewness in time-use data will tend to be rather small because such a result is limited by minimum values at zero. On the other hand, positive skewness in time-use data can get quite large, depending on the number and size of unusually large outliers above the mean.

Kurtosis:

The "peakedness" of the distribution relative to a normal distribution, ideally = 0. Negative values mean a flat distribution, positive values mean a high modal concentration.

Table 1: Normality and EDF Statistics for Selected 30-Day, 100-Day and 365-Day Data Distributions Used for Cross-Sectional Tests of All-Time-Absent

Index	Relative Date	N	Mean	Med.	Std. Dev.	Skew	Kur-tosis	Prob. D	> D
30-Day	211	129	1.64	2.0	1.14	0.61	0.68	0.19	<0.01
Frequency	912	151	1.19	1.0	1.05	0.90	0.98	0.23	<0.01
30-Day	211	129	103.30	120.0	58.12	-0.69	-0.84	0.15	<0.01
Severity	912	151	125.58	152.0	61.59	-1.10	-0.02	0.19	<0.01
30-Day Mean	211	129	67.25	56.0	53.31	0.65	-0.60	0.16	<0.01
Sev./Occur.	912	151	73.77	74.5	65.06	0.38	-1.14	0.17	<0.01
100-Day	546	126	3.44	3.0	2.09	1.04	1.64	0.17	<0.01
Frequency	1277	151	1.74	2.0	1.51	1.07	2.55	0.17	<0.01
	1642	141	3.77	3.0	2.44	0.51	0.08	0.14	<0.01
100-Day	546	126	419.30	477.0	156.33	-1.59	1.53	0.25	<0.01
Severity	1277	151	281.83	336.0	151.65	-0.96	-0.44	0.20	<0.01
	1642	141	412.06	504.0	197.96	-1.13	-0.05	0.23	<0.01
100-Day Mean	546	126	144.95	124.0	108.28	1.94	5.45	0.16	<0.01
Sev./Occur.	1272	151	134.43	114.3	127.11	1.03	0.49	0.15	<0.01
	1642	141	130.59	103.2	122.12	2.12	5.53	0.17	<0.01
Annual	546	126	11.17	11.0	5.44	0.53	0.35	0.12	<0.01
Frequency	1277	151	9.34	9.0	5.29	0.23	-0.23	0.08	<0.04
	1642	141	12.21	12.0	6.80	0.39	0.27	0.06	<0.15
Annual	546	126	1383.33	1594.0	517.67	-0.95	-0.17	0.16	<0.01
Severity	1277	151	1322.63	1583.0	591.38	-1.34	0.31	0.27	<0.01
	1642	141	1490.70	1727.0	591.42	-1.42	1.10	0.20	<0.01
Annual	546	126	151.20	128.0	108.38	2.68	9.79	0.19	<0.01
Sev./Occur.	1272	151	155.07	135.4	106.95	1.59	4.46	0.11	<0.01
	1642	141	147.33	122.1	127.49	3.30	17.74	0.13	<0.01

D:

Refers to the Kolmogorov "D" statistic described in Chapter IV. The D statistic varies between 0 and 1, with lower values indicating less normal distribution. In a normal distribution, D = 1.0.

Prob > D:

Means "the probability of this distribution producing a higher D statistic." The D and Prob > D statistics occur as pairs.

Discussion of Results for the Normality and EDF Statistics

In general, the D statistics are all less than 0.5 and most are less than 0.4, indicating that the samples were definitely non-normal, while the associated probabilities of the underlying distribution producing a higher D statistic were without exception less than 0.01 (D was highly significant). Skew and kurtosis were generally positive and a clear relationship exists between the sizes of the mean and standard deviation. This set of distributions is definitely non-normal and positively skewed. The relatively large standard deviations in the values of 30-day severity per occurrence (compared to the respective means) are a consequence of compounding two sources of variation in measurement. The severity per occurrence (mean severity) was calculated as a function of 30-day severity divided by 30-day frequency.

The T-tested data distributions are clearly not normal. Hypothesis 1 is definitely supported. Changes in the means remain helpful as indices of the absolute levels of time-use, but estimates of confidence levels associated with those results are not at all accurate. These results are entirely consistent with prior findings about cross-sections of data in the absenteeism literature. On the other hand, these findings with respect to the individual values in daily cross-sections relate to the T-tests only. The data distributions for time-series of the group mean values that were used in later correlational analyses are an entirely different matter, and as means based upon the daily cross-sections, those will be less affected by extreme values and should be distributed more normally.

B: Results for the T-Tests of Cross-Sectional Comparisons

T-tests were implemented on differences between the means of 17 different, time-specific daily datasets to examine five sets of hypotheses (see page 57 and Figure 6). An explanation of the reasons for choosing these particular dates has been presented in Chapter IV. Results of the normality and EDF analyses for these data distributions indicate very serious non-normality, just as prior research would imply. Nonetheless, for consistency with prior research, the T-test results are presented below. In most cases the variances of the two distributions were different enough (F-tested) that the T-test for distributions with unequal variance was used. This decision was made separately for each T-statistic, however, since the F-test values and probabilities were all available.

For the T-tests, the table or tables of selected statistics associated with each hypothesis have been presented, one hypothesis at a time, followed immediately by a brief interpretation of the results and a discussion of support for the hypothesis. More detailed T-test results are included in Appendix D. Values presented in the T-test tables include:

- **Var/Index:** the variable and time-use index tested
- **Relative Date 1:** the first of the two dates for which the T-test was done
- **Mean 1:** the mean value of the index variable on Date 1
- **Std.Dev. 1:** the standard deviation for the data distribution at Date 1
- **Relative Date 2:** the second of the two relative dates for which the T-test was done
- **Mean 2:** the mean value of the index variable on Date 2
- **Std.Dev. 2:** the standard deviation for the data distribution at Date 2
- **F-Prob:** the probability of equal variance in the two data distributions

- **T:** the T-test statistic, in most cases calculated for unequal variance
- **T-Prob:** the probability that the difference between the means occurred by chance
- Units of measurement were as follows: (a) frequency was reported as number of episodes per period of observation; (b) severity was reported as number of hours across all episodes in the period of observation; (c) mean severity was reported as the daily ratio of severity over frequency (or mean number of hours per time-use episode).

Note: The number of missing values reported here and elsewhere in this research is a direct result of excluding every questionable raw value from further analysis and reassigning them as missing values. Missing values, as a result, do reduce sample size and may absorb undocumented changes in other time-use indices but are not otherwise included in calculation of the other time-use indices.

B-1: Discussion of T-Tests on Overall Pre-Program Shifts (See Table 2)

Contrary to the null statement in Hypothesis 2, some highly significant shifts in time-use behavior were apparently recorded between the start of the baseline year and the month prior to program commencement. Anticipatory effects due to pre-program training apparently did have an impact as the discussion below indicates.

Across the pre-program interval, between July 30, 1978, and June 30, 1980, the severity and frequency indices for all time worked both changed by about 25%, but in reverse direction, indicating that more time was worked with fewer interruptions. Mean "severity" (or duration) of time worked was up from about 67 to 73 hours per occurrence. The same was true for routine time worked which is the largest single contributing element in all time worked. Significance levels of P = 0.002 were reported for changes in both of these indices. At the same time, scheduled time-off dropped to 69 hours from 104 hours per worker, with a frequency that did not change much at all, and a drop in mean severity from 52 hours per occurrence to 23 hours. Absence levels increased from 31 hours per employee in 30 days to 43 hours per employee. An improvement in quality of data recorded is evident, between the baseline period where the data came from prior payroll records, and the later period

Table 2: T-Test Results for the Hypothesis that Pre-Program Shifts in Time-Use Behaviors Occurred - 30-Day Data

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	211			912					
Frequency		1.64	1.14		1.19	1.05	0.37	3.47	0.00
Severity		103.30	58.12		125.58	61.59	0.50	-3.11	0.00
Mean Sev.		67.25	53.31		73.77	65.06	0.02	-0.92	0.35
Routine Work	211			912					
Frequency		1.65	1.16		1.15	1.04	0.19	3.80	0.00
Severity		103.09	58.07		125.11	61.52	0.50	-3.01	0.00
Mean Sev.		67.60	54.01		72.91	65.16	0.03	-0.75	0.45
All Scheduled Off	211			912					
Frequency		4.31	2.11		4.49	1.49	0.00	-0.77	0.43
Severity		103.60	60.25		69.48	41.83	0.00	5.14	0.00
Mean Sev.		52.02	78.11		23.24	47.85	0.00	3.64	0.00
Missing Values	211			912					
Frequency		0.26	0.50		0.09	0.29	0.00	3.24	0.00
Severity		43.53	86.20		10.65	46.34	0.00	3.88	0.00
Mean Sev.		40.50	82.19		10.65	46.34	0.00	3.66	0.00
Routine Off	211			912					
Frequency		4.10	2.45		4.39	1.61	0.00	-1.15	0.25
Severity		59.26	34.29		56.23	21.71	0.00	0.87	0.38
Mean Sev.		12.02	6.26		11.84	4.46	0.00	0.26	0.79
All Time Absent	211			912					
Frequency		1.23	1.07		1.08	0.96	0.18	1.25	0.21
Severity		31.73	45.96		43.78	61.31	0.00	-1.88	0.06
Mean Sev.		22.50	41.03		36.59	60.21	0.00	-2.32	0.02
Sht. Leg. Absence	211			912					
Frequency		0.89	0.97		0.66	0.86	0.17	2.07	0.03
Severity		22.36	35.17		20.30	33.51	0.57	0.50	0.61
Mean Sev.		16.94	31.33		17.86	32.23	0.74	-0.24	0.80
Long Leg. Absence	211			912					
Frequency		0.02	0.12		0.11	0.32	0.00	-3.46	0.00
Severity		3.72	29.77		18.75	58.13	0.00	-2.78	0.00
Mean Sev.		3.72	29.77		18.75	58.13	0.00	-2.78	0.00
Sht. Non-Leg. Ab.	211			912					
Frequency		0.38	0.66		0.36	0.70	0.58	0.27	0.78
Severity		5.64	13.63		4.72	10.91	0.01	0.62	0.53
Mean Sev.		4.20	9.14		3.47	7.86	0.08	0.71	0.47

NOTE: Sample sizes for the daily cross-sections ranged from 145 to 164 case where observers were trained to record time-uses with greater accuracy. The mean number of hours of missing values dropped from 43 hours to 10 hours per employee per month.

These changes were probably due to the extensive participation of employees and supervisors in group meetings where absenteeism was discussed. In this case, a positive "Hawthorne Effect" seems to have been induced. There were significant increases in both time worked (P = .002) and time absent (P = .06), with a corresponding significant decrease in time scheduled off (P = .0001). It is also notable that **frequency** indices and **severity** indices do not always vary, either in the same directions or by corresponding amounts. Indeed, they may in fact have some tendency to do the reverse for these 30-day measures, a possibility that will be investigated and discussed later, using the time-series correlations. While such a tendency would imply a need to stop using frequency and severity as alternative but parallel measures for the same "absence," the relationship is quite congruent with theories of substitution and progression effects.

T-test results indicate that prior to full program installation, substantial anticipatory shifts in relevant patterns of time-use behavior had already occurred. This finding indicates that the results of later T-tests of program impact which incorporate these anticipatory shifts must be interpreted as the result of both the pre-program activities and the absence control program itself.

It should be emphasized once again that these pre-program data, although randomly selected, come from less than a 7% sample of the workforce, and for the assessment of pre-program shifts, these were collected in each of the two cross sections across only 30 days of observation. As such, this was a relatively small sample of time-use behavior. D-statistics have indicated a serious non-normality in the respective distributions. Accordingly, although difference in the means has practical significance and is, in itself, a helpful statistic, the T-test probability statistics reported here are at best tenuous. It should also be noted that for time-use reported in Other Work, Punishment Time, and Late Time, the index values are based upon scarce behavior from a much smaller sample of employees, so that results for those time-use variables are even more questionable. For this reason, the results for those indices have not been reported here, although they are included in Appendix D. Note also the extremely low levels of long-term legitimate absence reported prior to the program installation.

A Note on Internal Consistency and Relations Among Different Time Uses

There is a deliberate but noteworthy internal consistency that should be present, by construction, among the changes of severity values especially, one of which demonstrates the integral construction of the *set* of measures. Consider the following example. There is a general "fund" of eight hours available in each employee's day, seven days a week. For the 30-day values reported in Figure 2, this fund must total to 8 hours times 30 days, or 240 hours. Since the combined measures *must* account for how all those hours get used, an internal total of 240 hours per employee, (8 hours per day for 30 days), provides a common base or account that all measures of time-use basically draw against. Because the moving values include rounding errors and are calculated against a variable number of employee "cases" due to turnover, the total of (all-time-worked) plus (all-time-scheduled-off) plus (all-time-absent) will tend to vary to some extent above and below this target level. But, a rough equivalence should exist and indeed it does, a verification and testimonial to the integrity of the measures (total = 239 hours instead of 240). This same type of internal relationship holds for the 100-day and 365-day measures also, although the "fund" of hours involved is greater.

B-2: T-Tests for Pre-Program Intra-Annual Change

B-2a: Discussion of T-Test Results on Pre-Program Seasonal Shifts

During the baseline year, on the 100-day data, there was no evidence of any significant seasonal change in absence behavior itself whatsoever. This is important in that prior to the program, highly significant changes occurred in both the amount of time worked (up from 342 hours per employee to 409 hours, T-Prob = 0.003), and in the amount of time scheduled off (down from 355 hours per employee to 287 hours, T-Prob = 0.002), and in that during the following year the 30-day data have indicated a shift in absence toward lower short-term absence and more time worked. This suggests that the shifts in absence behavior itself did not begin until after the baseline year when the installation period had started. The pre-program seasonal differences above presumably reflect summer vacation time. It is interesting to note that short-term non-legitimate absence became a much more variable time-use behavior in the winter months than it was during the summer. Furthermore, there was a very large decrease (from 174 to 90 hours) in the amount of time accounted for by missing values across the period (see Table 3 below).

Some seasonal shifts in non-absence time-uses were occurring prior to program installation, and Hypothesis 3 is supported for time worked and for time scheduled off

Table 3: T-Test Results on Hypotheses of General Seasonal Shifts in the Pre-Program 100-Day Measures: October 20, 1978 to April 14, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	294			469					
Frequency		3.65	2.44		3.23	2.08	0.07	1.48	0.13
Severity		342.92	192.61		409.95	165.98	0.09	-3.00	0.00
Mean Sev.		103.83	102.19		156.40	136.01	0.00	-3.51	0.00
Routine Work	294			469					
Frequency		3.70	2.54		3.14	1.94	0.00	1.98	0.04
Severity		340.27	191.79		409.51	166.05	0.10	-3.10	0.00
Mean Sev.		101.53	96.80		157.66	135.75	0.00	-3.82	0.00
All Scheduled Off	294			469					
Frequency		12.58	6.49		13.65	5.29	0.02	-1.45	0.14
Severity		355.48	203.25		287.87	139.84	0.00	3.11	0.00
Mean Sev.		98.66	206.04		48.59	138.99	0.00	2.29	0.02
Missing Values	294			469					
Frequency		0.67	1.23		0.40	0.76	0.00	0.21	0.03
Severity		175.69	287.29		90.23	190.11	0.00	2.81	0.00
Mean Sev.		115.12	225.31		71.03	170.35	0.00	1.77	0.07
Routine Off	294			469					
Frequency		12.12	6.99		13.45	5.46	0.01	-1.71	0.08
Severity		177.74	4.66		194.97	66.93	0.00	-1.69	0.09
Mean Sev.		13.92	4.86		14.15	3.70	0.00	-0.42	0.67
All Time Absent	294			469					
Frequency		3.32	2.28		2.87	2.16	0.52	1.63	0.10
Severity		106.61	126.05		107.23	154.33	0.02	-0.04	0.97
Mean Sev.		41.64	83.67		56.70	140.03	0.00	-1.05	0.29
Sht. Leg. Absence	294			469					
Frequency		2.51	1.94		1.77	1.64	0.06	3.33	0.00
Severity		67.45	83.39		52.60	97.88	0.07	1.31	0.19
Mean Sev.		30.99	70.53		33.28	91.39	0.00	0.22	0.82
Long Leg. Absence	294			469					
Frequency		0.03	0.21		0.02	0.12	0.00	0.71	0.47
Severity		9.61	79.78		8.56	75.57	0.54	0.11	0.91
Mean Sev.		6.51	51.63		8.56	75.57	0.00	-0.25	0.79
Sht. Non-Leg. Ab.	294			469					
Frequency		1.09	1.55		1.22	1.36	0.15	-0.68	0.49
Severity		29.55	64.91		46.08	109.13	0.00	-1.48	0.14
Mean Sev.		19.48	56.43		28.21	93.21	0.08	-0.91	0.36

(changes at least partly controlled by the employer), but not supported at all for absence behavior which occurs at the discretion of the employee. This result is clearly different from conventional findings in absence research. It also means that effects upon absenteeism induced by the program itself may be free from spurious and large cyclical effects.

However, once again note that the EDF statistics show little evidence of normality in the distributions compared here, and that these T-test statistics should be considered with caution, especially since missing values accounted for about 40% as much employee time as does routine time worked during the winter of the baseline year prior to the entire intervention. Although the data about time worked or time paid and absent had to be accurately recorded because these were originally legally audited payroll data, no such constraint was placed upon other unpaid categories or upon the detailed record of reasons for absence, time off, etc., until the installation of the program began. To minimize contamination where error could have otherwise been introduced by miscoded data, these were identified and corrected, or where correction wasn't possible, converted to missing data.

B-2b: Discussion of the T-Tests on Pre-Program Holiday Shifts Observed in the 100-day Measures

The 100-day moving values for the measures were aggregated across a short enough period to be sensitive to holiday and seasonal effects which would be smoothed out of the 365-day data. Nonetheless, the Christmas Holidays had no effect at all on long-term legitimate absence, nor any significant effect on any of the absence measures except the frequency of short-term legitimate absence, which decreased, as expected, from the holiday to non-holiday period from 2.30 to 1.77, T-Prob = 0.01. It appears very likely that the co-occurrent non-significant decrease in the severity index for short-term legitimate absence of five hours per employee was the result of many employees taking a legitimate personal leave day during the Christmas period. Except for this minor fluctuation, the holiday had little effect on absence per se prior to the program itself (see Table 4 below).

However, as in the 30-day cross-sections, significant shifts did occur in time worked and in time scheduled off. Time worked per 100 days went up from 354 hours per employee to 409 hours, T-Prob = 0.01, and time scheduled off went down after the holiday period by a corresponding amount from 355 hours to 287 hours per employee (T-Prob = 0.002). Apparently both the employees and the employer cooperated to arrange legitimate time off during the holiday period.

Table 4: T-Test Results of Holiday Shifts in the 100-Day Pre-Program Time-Use Measures: January 5, 1979 to April 14, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	370			469					
Frequency		3.45	2.30		3.23	2.08	0.24	0.79	0.42
Severity		354.70	201.76		409.95	165.98	0.03	-2.40	0.01
Mean Sev.		113.94	107.86		156.40	136.01	0.01	-2.78	0.00
Routine Work	370			469					
Frequency		3.38	2.19		3.14	1.94	0.17	0.93	0.35
Severity		353.93	201.73		409.51	166.05	0.03	-2.42	0.01
Mean Sev.		115.53	111.97		157.66	135.75	0.03	-2.72	0.00
All Scheduled Off	370			469					
Frequency		12.63	6.68		13.65	5.29	0.01	-1.35	0.17
Severity		355.79	203.72		287.87	139.84	0.00	3.12	0.00
Mean Sev.		109.67	228.24		48.59	138.99	0.00	2.60	0.01
Missing Values	370			469					
Frequency		0.71	1.35		0.40	0.76	0.00	2.32	0.02
Severity		174.57	292.18		90.23	190.11	0.00	2.75	0.00
Mean Sev.		112.70	234.37		71.03	170.35	0.00	1.63	0.10
Routine Off	370			469					
Frequency		12.26	7.15		13.45	5.46	0.00	-1.50	0.13
Severity		177.67	100.00		194.97	66.93	0.00	-1.44	0.15
Mean Sev.		12.99	5.80		14.15	3.70	0.00	-1.91	0.05
All Time Absent	370			469					
Frequency		3.12	2.20		2.87	2.16	0.80	0.94	0.34
Severity		95.71	136.53		107.23	154.33	0.17	-0.64	0.52
Mean Sev.		48.81	127.11		56.70	140.03	0.27	-0.47	0.63
Sht. Leg. Absence	370			469					
Frequency		2.30	1.74		1.77	1.64	0.48	2.54	0.01
Severity		57.79	97.35		52.60	97.88	0.95	0.43	0.66
Mean Sev.		29.92	89.02		33.28	91.39	0.77	0.30	0.76
Long Leg. Absence	370			469					
Frequency		0.02	0.12		0.02	0.12	1.00	0.00	1.00
Severity		6.33	71.14		8.56	75.57	0.49	-0.24	0.80
Mean Sev.		6.33	71.14		8.56	75.57	0.49	-0.24	0.80
Sht. Non-Leg. Ab.	370			469					
Frequency		1.07	1.56		1.22	1.36	0.12	-0.81	0.42
Severity		31.59	85.68		46.08	109.13	0.01	-1.19	0.23
Mean Sev.		18.65	66.03		28.21	93.21	0.00	-0.95	0.34

Hypothesis 4 is supported for shifts in time-use that were at least partly subject to control by the employer, but not for absence in general. Contamination of the program evaluation by a pre-existing Christmas holiday effect would not appear to have been a problem, although this may not generalize from Christmas holidays to the summer July holiday period.

B-2c: Discussion of the T-Tests on Combined Seasonal and Holiday Effects in the Pre-Program 100-Day Measures

Between Christmas 1978 and June 15, 1979, there was very little change in absence behavior. Once again there was an increase in time worked, from 354 hours per employee to 423 hours, T-Prob = 0.003. There was also a significant decrease in time scheduled off, from 355 to 255 hours per employee, T-Prob = 0.0001; and a very large decrease in time accounted for by missing values, 174 hours to 48 hours, T-Prob = 0.0001. Again, the change in time absent was non-significant except for the frequency of short-term legitimate absence (see Table 5).

Hypothesis 5 is therefore supported to the extent that combined seasonal and holiday shifts are evident in time use behavior that is at least partly under the control of the employer. However, no significant evidence exists that a significant Christmas and winter to late spring holiday/seasonal effect on discretionary non-legitimate absence occurred within the baseline year and before the program went into effect.

Table 5: T-Test Results for the Combined Seasonal and Holiday Effects on the 100-Day Pre-Program Measures: January 5, 1979 to June 15, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	370			531					
Frequency		3.45	2.30		3.40	2.20	0.60	0.17	0.86
Severity		354.70	201.76		423.01	159.02	0.01	-3.02	0.00
Mean Sev.		113.94	107.86		150.06	123.72	0.12	-2.50	0.01
Routine Work	370			531					
Frequency		3.38	2.19		3.24	2.03	0.40	0.53	0.59
Severity		353.93	201.73		422.51	158.70	0.01	-3.03	0.00
Mean Sev.		115.53	111.97		157.04	124.26	0.24	-2.82	0.00
All Scheduled Off	370			531					
Frequency		12.64	6.69		13.65	4.51	0.00	-1.43	0.15
Severity		355.79	203.72		255.44	104.66	0.00	4.98	0.00
Mean Sev.		109.67	228.24		28.59	76.95	0.00	3.82	0.00
Missing Values	370			531					
Frequency		0.71	1.35		0.23	0.54	0.00	3.75	0.00
Severity		174.57	292.18		48.25	131.87	0.00	4.47	0.00
Mean Sev.		112.70	234.37		39.94	112.14	0.00	3.18	0.00
Routine Off	370			531					
Frequency		12.26	7.15		13.57	4.61	0.00	-1.74	0.08
Severity		177.67	100.00		204.84	53.32	0.00	-2.52	0.01
Mean Sev.		12.99	5.80		15.15	3.03	0.00	-3.76	0.00
All Time Absent	370			531					
Frequency		3.12	2.20		2.85	2.07	0.48	1.02	0.30
Severity		95.71	136.53		126.46	174.34	0.01	-1.58	0.11
Mean Sev.		48.81	127.11		71.03	154.34	0.03	-1.26	0.20
Sht. Leg. Absence	370			531					
Frequency		2.30	1.74		1.63	1.69	0.73	3.15	0.01
Severity		57.79	97.35		61.29	109.88	0.17	-0.27	0.78
Mean Sev.		29.92	89.02		39.82	88.34	0.93	-0.90	0.37
Long Leg. Absence	370			531					
Frequency		0.02	0.12		0.02	0.12	1.00	0.00	1.00
Severity		6.33	71.14		12.16	97.28	0.00	-0.55	0.58
Mean Sev.		6.33	71.14		12.16	97.28	0.00	-0.55	0.58
Sht. Non-Leg. Ab.	370			531					
Frequency		1.07	1.56		1.28	1.46	0.41	-1.11	0.26
Severity		31.59	85.68		53.02	122.59	0.00	-1.63	0.10
Mean Sev.		18.65	66.03		34.44	106.97	0.00	-1.43	0.15

B-3: T-Tests of Long-Term Change During the Time that the Program Was Fully in Force

B-3a: Discussion of the T-Tests for Year-to-Year Change in the 100-Day Frequency and Severity Measures During the Program Period (See Table 6 and Table 7 below)

Between the summer of 1980 and the summer of 1981, both the frequency indices and severity indices show significant shifts in time worked (down from 376 hours to 264 hours), in time scheduled off (up from 250 hours to 394 hours), and in missing values (up from 38 hours to 246 hours).

This increase in missing values was startling. On the average, an increase of only one occurrence in the frequency measure co-occurred with this anomaly, which indicates a severe one-episode time-use occurred for each employee. Comparison between day 1347 and day 1390 suggests that this event must have occurred sometime between July 1 and August 12, 1981, since the event was not included in the 100-day base for day 1390. The pattern here suggested a strike or massive shutdown during the summer of 1981. The frequency indices also indicate significant year-to-year increases for other periods in overall frequency of legitimate absences and a corresponding overall increase in the frequency of work intervals, and these two findings are both consistent with the this work interruption hypothesis. An anomaly so pervasive and unusual could neither be an accident nor be ignored, so the DOT Personnel Director was telephoned for

Table 6: T-Test Results for Hypotheses about Long-Term Year-to-Year Change in the 100-Day Measures, Frequency Indices Only During the Program Period

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	2.95	2.22	1531	4.02	2.52	0.12	-3.83	0.00
	982	3.37	2.04	1347	2.62	1.87	0.31	3.29	0.00
	1025	3.45	2.38	1390	4.06	2.69	0.13	-2.06	0.04
	1210	3.01	2.17	1575	4.02	2.50	0.09	3.69	0.00
Routine Work	1166	2.89	2.28	1531	3.93	2.43	0.43	-3.75	0.00
	982	3.34	2.02	1347	2.60	1.82	0.22	3.27	0.00
	1025	3.42	2.28	1390	3.99	2.59	0.12	-1.96	0.05
	1210	2.97	2.25	1575	4.00	2.52	0.17	-3.68	0.00
All Sch. Off	1166	13.56	6.67	1531	13.38	5.15	0.00	0.26	0.79
	982	12.91	4.30	1347	9.30	3.82	0.16	7.61	0.00
	1025	13.64	4.74	1390	12.28	5.04	0.46	2.37	0.01
	1210	12.75	6.65	1575	13.54	5.25	0.01	1.12	0.26
Missing Values	1166	0.15	0.43	1531	0.28	1.76	0.00	-0.82	0.41
	982	0.17	0.41	1347	1.10	0.94	0.00	-10.85	0.00
	1025	0.13	0.37	1390	0.21	1.48	0.00	-0.68	0.49
	1210	0.15	0.37	1575	0.30	1.74	0.00	-1.06	0.28
Routine Off	1166	13.42	6.87	1531	13.31	5.29	0.00	0.15	0.88
	982	12.77	4.41	1347	9.39	4.26	0.69	6.67	0.00
	1025	13.54	4.87	1390	12.20	5.19	0.44	2.27	0.02
	1210	12.58	6.87	1575	13.47	5.39	0.00	-1.24	0.21
All Time Absent	1166	2.47	2.04	1531	3.27	2.31	0.14	-3.12	0.00
	982	3.18	1.97	1347	2.54	1.81	0.32	2.89	0.00
	1025	3.40	2.32	1390	3.59	2.59	0.20	-2.66	0.50
	1210	2.48	1.97	1575	3.45	2.38	0.02	-3.80	0.00
Sht. Leg. Ab.	1166	1.81	1.69	1531	2.51	2.09	0.01	-3.12	0.00
	982	2.36	1.69	1347	2.26	1.78	0.56	0.50	0.61
	1025	2.70	2.00	1390	3.30	2.63	0.00	-2.16	0.03
	1210	1.73	1.75	1575	2.67	2.13	0.02	-4.10	0.00
Long Leg. Ab.	1166	0.22	0.54	1531	0.33	0.78	0.00	-1.36	0.17
	982	0.60	0.62	1347	0.33	0.59	0.57	3.78	0.00
	1025	0.59	0.64	1390	0.34	0.65	0.74	3.30	0.00
	1210	0.21	0.48	1575	0.37	0.76	0.00	-2.18	0.03
Sht Non-Leg. Ab.	1166	0.68	1.23	1531	0.73	1.17	0.58	-0.39	0.69
	982	0.45	0.96	1347	0.20	0.45	0.00	2.89	0.00
	1025	0.30	0.93	1390	0.17	0.43	0.00	1.52	0.12
	1210	0.75	1.19	1575	0.74	1.17	0.82	0.03	0.97

Table 7: T-Test Results for Hypotheses about Long-Term Year-to-Year Change in the 100-Day Measures, Severity Indices Only, During the Program Period

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	412.55	204.59	1531	436.37	184.62	0.22	-1.05	0.29
	982	376.40	151.03	1347	264.89	126.99	0.04	6.84	0.00
	1025	415.21	172.11	1390	403.56	182.65	0.47	0.56	0.57
	1210	418.51	214.19	1575	435.45	191.03	0.17	-0.71	0.47
Routine Work	1166	411.81	204.68	1531	434.87	184.84	0.22	-1.01	0.31
	982	375.34	150.62	1347	264.43	126.71	0.04	6.83	0.00
	1025	414.52	171.82	1390	403.33	182.38	0.47	0.54	0.59
	1210	417.77	213.92	1575	434.43	190.77	0.17	0.70	0.48
All Sch. Off	1166	282.32	197.86	1531	257.06	161.11	0.01	1.20	0.23
	982	250.52	135.30	1347	394.01	105.51	0.00	-10.14	0.00
	1025	261.13	146.87	1390	222.48	146.30	0.96	2.26	0.02
	1210	270.39	203.50	1575	255.13	168.09	0.02	0.70	0.48
Missing Values	1166	92.03	256.02	1531	56.06	193.95	0.00	1.36	0.17
	982	38.78	157.57	1347	246.07	85.02	0.00	-14.11	0.00
	1025	59.60	186.22	1390	30.92	137.55	0.00	1.50	0.13
	1210	95.74	256.74	1575	63.06	206.91	0.01	1.20	0.23
Routine Off	1166	186.75	89.00	1531	193.50	72.37	0.01	-0.71	0.47
	982	207.50	79.67	1347	137.74	65.84	0.02	8.18	0.00
	1025	200.09	71.03	1390	178.97	79.69	0.17	2.38	0.01
	1210	168.17	81.55	1575	186.91	72.52	0.16	-2.08	0.03
All Time Absent	1166	110.01	190.81	1531	110.21	164.91	0.08	-0.01	0.99
	982	168.99	173.15	1347	140.04	170.53	0.86	1.44	0.15
	1025	126.99	172.50	1390	178.52	218.35	0.00	-2.23	0.0
	1210	115.05	195.41	1575	112.53	165.16	0.05	0.60	0.54
Sht. Leg. Ab.	1166	44.68	95.66	1531	50.62	78.71	0.02	-0.58	0.56
	982	65.06	76.12	1347	60.10	82.12	0.36	0.53	0.59
	1025	68.29	80.62	1390	77.89	102.19	0.00	-0.89	0.37
	1210	46.20	103.29	1575	42.50	58.71	0.00	0.38	0.70
Long Leg. Ab.	1166	55.63	170.79	1531	49.73	150.92	0.14	0.31	0.75
	982	97.39	172.49	1347	75.74	166.34	0.66	1.09	0.27
	1025	53.74	168.58	1390	97.41	216.84	0.00	-1.91	0.05
	1210	58.01	170.66	1575	68.35	161.32	0.50	-0.02	0.98
Sht Non-Leg. Ab.	1166	9.70	25.08	1531	9.85	17.38	0.00	-0.06	0.95
	982	6.54	19.36	1347	4.19	13.35	0.00	1.21	0.22
	1025	4.97	18.08	1390	3.22	11.13	0.00	1.00	0.31
	1210	10.83	23.32	1575	11.69	27.13	0.03	-0.29	0.77

comment on unusual events during the period. The explanation fits very well. Apparently, the mayor of Detroit gave those employees an extra 80 hours of legitimate paid vacation in a concession bargaining decision at that time. This was not appropriately coded as short-term legitimate absence and therefore became part of "missing values." In addition, a 70 hour reduction in routine time off and a 30 hour reduction in all time absent contributed their effect to account for the rest of these missing hours.

Except for the anomaly during the summer of 1981, the only significant change in overall absence for these tests was in long-term legitimate absence which increased from 53 to 97 hours per employee immediately after the summer anomaly occurred.

Hypothesis 6 was supported for the frequency of both absence and time worked. Both increased. On the other hand, this did not significantly affect severity of either index which suggests that the employees were developing a long-term habit of using their casual leave days in several occurrences instead of just a few. Overall, it appears that except for some increase in long-term absence, the 100-day measures do not show evidence of major shifts in absence within the time span of the fully implemented absence control program.

These findings are highly vulnerable to short-term intra-annual effects, however, and despite the comparison of corresponding year-to-year dates, they should be considered with caution.

B-3b: Discussion of T-Test Results for Hypotheses about Holiday Effects in the 100-Day Measures During the Program Period

Across the period when the program itself was fully implemented, both the frequency and severity indices were T-tested to detect significant holiday effects. But, the results were essentially the same for both so that a detailed presentation of only the severity shifts is included in Table 8 below. In brief, two significant shifts have occurred across major holiday periods within the span of the program. These occurred for the time-use measures during the summers of 1981 and of 1980. One large shift (an increase in time worked, decrease in time scheduled off, and in missing values) occurred just after the interval for which the anomaly which affected Hypothesis 6 showed up (Day 1347 to Day 1390). Apparently things went back to normal after the sudden increase of paid time off had been used by employees. A smaller effect, a decrease in long term legitimate absence was also associated with a summer (July 4th) holiday period in 1980 and was also noticeable and significant.

Table 8: T-Test Results for Holiday Effects During the Program Period on the 100-Day Data, Using Severity Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	412.55	204.59	1210	418.51	214.19	0.58	-0.25	0.80
	1531	436.37	184.62	1575	435.41	191.03	0.69	0.04	0.96
	982	376.40	151.03	1025	415.21	172.11	0.11	-2.08	0.03
	1347	264.89	126.99	1390	403.56	182.65	0.00	-7.40	0.00
Routine Work	1166	411.81	204.68	1210	417.77	213.92	0.59	-0.25	0.80
	1531	434.87	184.84	1575	434.43	190.77	0.71	0.02	0.98
	982	375.41	150.52	1025	414.52	171.82	0.11	-2.11	0.03
	1347	264.43	126.71	1390	403.33	182.38	0.00	7.43	0.00
All Sch. Off	1166	282.32	197.86	1210	270.39	203.50	0.73	0.52	0.60
	1531	257.06	161.11	1575	255.13	168.09	0.62	0.10	0.92
	982	250.52	135.30	1025	261.23	146.87	0.32	-0.66	0.51
	1347	394.01	105.51	1390	222.48	146.30	0.00	11.29	0.00
Missing Values	1166	92.03	256.02	1210	95.74	256.74	0.97	-0.13	0.90
	1531	56.06	193.95	1575	63.06	206.91	0.45	-0.29	0.76
	982	38.78	157.57	1025	59.60	186.22	0.04	-1.05	0.29
	1347	246.07	85.02	1390	30.92	137.55	0.00	15.80	0.00
Routine Off	1166	186.75	89.00	1210	168.17	81.55	0.29	1.89	0.05
	1531	193.50	72.47	1575	186.91	72.52	0.99	0.76	0.44
	982	207.50	79.67	1025	200.09	71.03	0.16	0.85	0.39
	1347	137.74	65.84	1390	178.97	79.69	0.02	-4.74	0.00
All Time Absent	1166	110.01	190.81	1210	115.05	195.41	0.77	-0.23	0.82
	1531	110.21	164.91	1575	112.53	165.16	0.99	-0.12	0.90
	982	168.99	173.15	1025	126.99	175.50	0.96	2.11	0.03
	1347	140.04	170.53	1390	178.52	218.35	0.00	-1.65	0.10
Sht. Leg. Ab.	1166	44.68	95.66	1210	46.20	103.29	0.35	-0.13	0.89
	1531	50.62	78.71	1575	42.50	58.71	0.00	0.98	0.32
	982	65.06	76.12	1025	68.29	80.62	0.48	-0.36	0.72
	1347	60.10	82.12	1390	77.89	102.19	0.01	-1.61	0.10
Long Leg. Ab.	1166	55.63	170.79	1210	58.01	170.66	0.99	-0.12	0.90
	1531	49.73	150.92	1575	58.35	161.32	0.43	-0.46	0.64
	982	97.38	172.49	1025	53.74	188.58	0.78	2.22	0.02
	1347	75.74	166.34	1390	97.41	216.84	0.00	-0.94	0.34
Sht Non-Leg. Ab.	1166	9.70	25.08	1210	10.83	23.32	0.37	-0.41	0.68
	1531	9.86	17.38	1575	11.69	27.13	0.00	-0.67	0.50
	982	6.54	19.36	1025	4.97	18.08	0.40	0.73	0.46
	1347	4.19	13.35	1390	3.22	11.13	0.03	0.66	0.50

No holiday effects were apparent across the Christmas periods however, and while this one finding does offer weak support for Hypothesis 7 as it relates to the summer holidays, this holiday effect may have been at least partly the effect of annual renewal or vacation. There was, for example, no significant effect across the winter (Christmas) holiday season when renewal would not be expected to have much effect and holidays might. Once again, some rather anomalous shifts and reductions in time-worked as well as time scheduled off occurred during the summer of 1981.

B-3c: Discussion of Results for the T-Tests of Combined Seasonal, Holiday and Annual Renewal Effects During the Program Period on the 100-Day Measures

Across the span of the program period itself, the 100-day measures were also examined for combined cyclical effects. Several T-test comparisons for the 100-day measures offer general support for Hypothesis 8. Combined seasonal, holiday and renewal effects on absence do appear to take place across surrounding periods not contaminated by the anomaly in the early summer of 1981, producing significant decreases in levels of absence and increases in time worked from summer to winter. Since the same summer of 1981 anomaly (Date 1 = 1347) which was evident in earlier T-tests was also presumably contributing to combination effects in that year, the 1981 results (using Day 1347) really don't reflect cyclical change (see Table 9 below).

Comment on the Measures Used to Examine Intra-Annual T-Test Results

In general, the frequency and mean severity indices co-vary with the severity indices and tend to provide significant results across the same intervals, although in this there were some exceptions. For convenience, in the present analysis the results on the frequency indices have not been included. Detailed examination of the relationship between severity and frequency of absence is definitely an interesting problem. Indeed, the parallelism in the 100-day indices between severity and frequency is in the reverse direction

Table 9: T-Test Results for Hypotheses of Combined Holiday, Seasonal, and Renewal Effects During the Program Period on the 100-Day Measures, Severity Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	982	376.40	151.03	1166	412.55	204.59	0.00	-.175	0.08
	1347	264.89	126.99	1531	436.37	184.62	0.00	-9.09	0.00
	1025	415.21	172.11	1210	418.51	214.19	0.01	-0.15	0.88
	1390	403.56	182.65	1575	435.45	191.03	0.60	-1.43	0.15
Routine Work	982	375.34	150.52	1166	411.81	204.68	0.00	-1.76	0.07
	1347	264.43	126.71	1531	434.87	184.84	0.00	-9.03	0.00
	1025	414.52	171.82	1210	417.77	213.92	0.01	-0.15	0.16
	1390	403.33	182.38	1575	434.43	190.77	0.60	-1.40	0.16
All Sch. Off	982	250.52	135.30	1166	282.32	197.86	0.00	-1.63	0.10
	1347	394.01	105.57	1531	257.06	161.11	0.00	-8.44	0.00
	1025	261.23	146.87	1210	270.39	203.50	0.00	-0.45	0.65
	1390	222.48	146.30	1575	255.13	168.09	0.10	-1.74	0.08
Missing Values	982	38.78	157.57	1166	92.03	256.02	0.00	-2.18	0.03
	1347	246.07	85.02	1531	56.01	193.95	0.00	10.65	0.00
	1025	59.60	186.22	1210	95.74	256.74	0.00	-1.40	0.162
	1390	30.92	137.55	1575	63.06	206.91	0.00	-1.54	0.12
Routine Off	982	207.50	79.67	1166	186.75	89.00	0.18	2.13	0.03
	1347	137.74	65.84	1531	193.50	72.47	0.26	-6.76	0.00
	1025	200.09	71.03	1210	168.17	81.55	0.09	3.63	0.00
	1390	178.97	79.69	1575	186.91	72.52	0.27	-0.88	0.38
All Time Absent	982	168.99	173.15	1166	110.01	190.81	0.24	2.81	0.00
	1347	140.04	170.53	1531	110.21	164.91	0.69	1.49	0.13
	1025	126.99	172.50	1210	115.05	195.41	0.13	0.56	0.57
	1390	178.52	218.35	1575	112.53	165.16	0.00	2.86	0.00
Sht. Leg. Ab.	982	65.06	76.12	1166	44.68	95.66	0.00	2.05	0.04
	1347	60.10	82.12	1531	50.62	78.71	0.62	0.98	0.32
	1025	68.29	80.62	1210	46.20	103.29	0.00	2.07	0.03
	1390	77.89	102.19	1575	42.50	58.71	0.00	3.57	0.00
Long Leg. Ab.	982	97.38	172.49	1166	55.63	170.79	0.90	2.11	0.03
	1347	75.74	166.34	1531	49.73	150.93	0.25	1.38	0.17
	1025	53.74	168.58	1210	58.01	170.66	0.88	-0.22	0.82
	1390	97.41	216.84	1575	58.35	161.32	0.00	1.72	0.08
Sht Non-Leg. Ab.	982	6.54	19.36	1166	9.70	25.08	0.00	-1.23	0.22
	1347	4.19	13.35	1531	9.85	17.38	0.00	-3.07	0.00
	1025	4.97	18.08	1210	10.83	23.32	0.00	-2.44	0.01
	1390	3.22	11.13	1575	11.69	27.13	0.00	-3.43	0.00

to an apparent co-variation suggested by the 30-day measures, where frequency went down as severity went up. It is possible that this latter effect is made evident by the use of a shorter time period as the base for these measures, and that both patterns can and do operate at the same time so that they covary positively over long periods and negatively over short periods. Further insight might be forthcoming from the use of time-series correlations between measures across relatively short time periods such as one week, but no support for the inverse relationship appears when the 100-day measures are correlated (see Table 13).

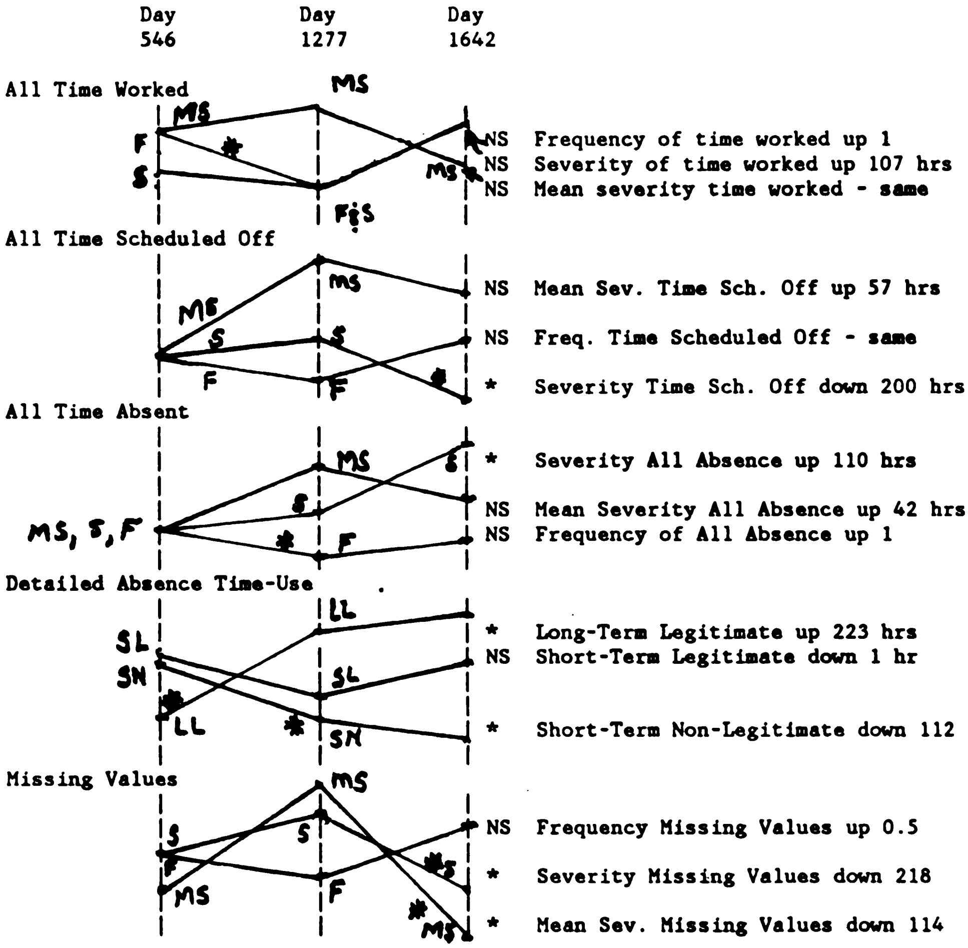
B-4: Summary and Discussion of the T-Tests on Overall Program-Induced Change in the 100-Day Measures

The program induced changes between the baseline year and later periods were not entirely as predicted for the 100-day data, but were unquestionably strong and apparently significant. Time Scheduled Off, Time Worked, and All Time Absent decreased significantly from the end of the baseline year to the end of the first program year. Then all three significantly increased again back toward their original levels by the end of the second program year. In the mean time, missing values were of course affected by the anomaly in the late spring or early summer of 1981 and increased noticeably between the end of the baseline year and the end of the first program year, then decreased dramatically back to much less than their original level by the end of the second program year. Overall, between the last 100 days of the baseline year and the last 100 days of the second program year, time worked apparently showed no significant change, nor did time scheduled off or overall time absent. However, these overall year-to-year tests of program induced changes in the 100-day measures were likely to be affected both by the anomaly during the summer of 1981 and by other short-term fluctuations for the 100-day measures, and in any case are only comparisons of the fourth quarter values in each respective year. The 365-day measures below were used instead to assess overall program impact on annual time-detailed tabular presentation of evidence of long-term inter-annual change in the 100-day measures has been omitted here.

B-5: Discussion of the T-Tests on Overall Long-Term Program Induced Changes Indicated by the 365-Day Measures, A Direct Evaluation of the Program

Figure 7 below offers a rough graphical summary of overall changes in the major annual time-use indices. Full details have been presented in Table 10. Overall, time worked per employee per year went up by 107 hours to 1490 hours per year. Time scheduled off was down by 200 hours to 915 hours. And an overall increase in time absent of 110 hours per employee per year took place, all of which was in long-term legitimate absence since short-term non-legitimate absence dropped 112 hours to 27 hours per employee per year from 139 hours, and short-term legitimate absence also dropped 40 hours (non significant). Most of the observed change in absence behavior had occurred by the end of the first complete year of the program.

Also worthy of note is a substantial decrease in the level of missing values present in the data, over the second year of the program especially. It appears likely that many of the 80 hours given away in concession bargaining during the first year were recorded as missing values and that a certain amount of time must have been needed after program start-up to streamline the administration of data collection by supervisors, data-entry clericals, etc.



NOTE: The relative scales of sketches above are merely roughly "eyeballed" pictures of trends in each line, clustered for convenience.

Figure 7:
Approximate Graphic Profiles of Long-Term Impact on the Overall 365-Day Time-Use Measures (Relative Scales)

Table 10: T-Test Results of Program Induced Change in the 365-Day Measures

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	546			1277					
Frequency		11.17	5.44		9.34	5.29	0.74	2.81	0.00
Severity		1383.33	517.67		1322.63	591.38	0.12	0.91	0.36
Mean Sev.		151.10	108.38		155.07	106.95	0.87	-0.30	0.76
Routine Work	546			1277					
Frequency		10.90	5.33		9.20	5.16	0.69	2.69	0.00
Severity		1379.21	517.23		1320.09	590.91	0.12	0.89	0.37
Mean Sev.		154.12	109.31		156.97	107.91	0.87	-0.22	0.82
All Scheduled Off	546			1277					
Frequency		47.79	18.37		44.21	18.61	0.88	1.61	0.10
Severity		1130.93	501.61		1160.86	588.36	0.07	-0.46	0.64
Mean Sev.		40.09	70.25		164.61	573.57	0.00	-2.64	0.00
Missing Values	546			1277					
Frequency		1.29	2.27		1.09	0.36	0.00	1.02	0.31
Severity		419.68	703.82		506.07	753.95	0.43	-0.98	0.32
Mean Sev.		227.95	541.41		471.54	710.64	0.00	-3.23	0.00
Routine Off	546			1277					
Frequency		47.28	18.85		43.70	18.79	0.97	1.57	0.11
Severity		701.85	245.18		641.01	269.04	0.28	1.97	0.50
Mean Sev.		15.09	2.47		14.09	5.02	0.00	2.14	0.03
All Time Absent	546			1277					
Frequency		10.76	5.50		8.92	5.11	0.39	2.87	0.00
Severity		396.29	459.44		431.03	560.36	0.02	-0.57	0.57
Mean Sev.		52.92	141.63		103.63	338.48	0.00	-1.67	0.09
Sht. Leg. Absence	546			1277					
Frequency		7.44	3.90		6.95	3.99	0.80	1.02	0.31
Severity		222.06	277.14		180.44	256.47	0.36	1.29	0.19
Mean Sev.		39.41	105.95		32.11	79.58	0.00	0.64	0.52
Long Leg. Absence	546			1277					
Frequency		0.04	0.23		0.95	1.05	0.00	-10.35	0.00
Severity		34.60	274.80		221.22	509.37	0.00	-3.88	0.00
Mean Sev.		23.05	159.09		127.80	333.81	0.00	-3.42	0.00
Sht. Non-Leg. Ab.	546			1277					
Frequency		3.94	3.80		1.75	2.55	0.00	5.49	0.00
Severity		139.63	275.64		29.36	61.49	0.00	4.40	0.00
Mean Sev.		35.86	90.91		8.33	13.94	0.00	3.37	0.00

Table 10 Continued

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	546			1642					
Frequency		11.17	5.44		12.21	6.80	0.01	-1.39	0.16
Severity		1383.33	517.67		1490.70	591.42	0.13	-1.58	0.11
Mean Sev.		151.10	108.38		147.33	127.49	0.06	0.27	0.78
Routine Work	546			1642					
Frequency		10.90	5.33		11.94	6.56	0.02	-1.43	0.15
Severity		1379.21	517.23		1487.47	591.21	0.13	-1.60	0.11
Mean Sev.		154.12	109.31		148.20	125.65	0.11	0.41	0.68
All Scheduled Off	546			1642					
Frequency		47.79	18.37		48.18	16.45	0.20	-0.18	0.85
Severity		1130.93	501.61		915.97	510.73	0.84	3.47	0.00
Mean Sev.		40.09	70.25		92.32	387.79	0.00	-1.57	0.11
Missing Values	546			1642					
Frequency		1.29	2.27		1.70	5.14	0.00	-0.84	0.40
Severity		419.68	703.82		201.50	597.66	0.06	2.71	0.00
Mean Sev.		227.95	541.41		113.30	396.57	0.00	1.95	0.05
Routine Off	546			1642					
Frequency		47.28	18.85		47.81	16.62	0.15	-0.24	0.81
Severity		701.85	245.18		683.84	234.47	0.61	0.61	0.54
Mean Sev.		15.09	2.47		13.75	3.87	0.00	3.41	0.00
All Time Absent	546			1642					
Frequency		10.76	5.50		11.74	6.58	0.04	-1.33	0.18
Severity		396.29	459.44		506.01	564.39	0.02	-1.75	0.08
Mean Sev.		52.92	141.63		94.15	350.52	0.00	-1.28	0.20
Sht. Leg. Absence	546			1642					
Frequency		7.44	3.90		10.01	5.82	0.00	-4.29	0.00
Severity		222.06	277.14		221.06	224.37	0.02	0.03	0.97
Mean Sev.		39.41	105.95		34.91	123.56	0.08	0.32	0.74
Long Leg. Absence	546			1642					
Frequency		0.04	0.23		1.03	1.61	0.00	7.22	0.00
Severity		34.60	274.80		257.50	534.78	0.00	-4.35	0.00
Mean Sev.		23.05	159.09		107.46	286.28	0.00	-3.02	0.00
Sht. Non-Leg. Ab.	546			1642					
Frequency		3.94	3.80		1.75	2.25	0.00	5.62	0.00
Severity		139.63	275.64		27.44	54.52	0.00	4.49	0.00
Mean Sev.		35.86	90.91		7.80	11.44	0.00	3.44	0.00

Table 10 Continued

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1277			1642					
Frequency		9.34	5.28		12.21	6.80	0.00	-4.00	0.00
Severity		1322.63	591.38		1490.70	591.42	1.00	-2.43	0.21
Mean Sev.		155.07	106.95		147.33	127.49	0.03	0.56	0.57
Routine Work	1277			1642					
Frequency		9.20	5.16		11.94	6.56	0.00	-3.96	0.00
Severity		1320.09	510.91		1487.47	591.21	0.99	-2.42	0.01
Mean Sev.		156.97	107.91		148.20	125.65	0.07	0.64	0.52
All Sch. Off	1277			1642					
Frequency		44.27	18.61		48.18	16.45	0.14	-1.93	0.05
Severity		1160.86	588.36		915.97	510.73	0.09	3.80	0.00
Mean Sev.		164.61	573.57		92.32	387.79	0.00	1.27	0.20
Missing Values	1277			1642					
Frequency		1.09	0.36		1.70	5.14	0.00	-1.40	0.16
Severity		506.07	753.95		201.50	597.66	0.00	3.84	0.00
Mean Sev.		471.54	710.64		113.30	396.57	0.00	5.36	0.00
Routine Off	1277			1642					
Frequency		43.70	18.79		47.81	16.62	0.14	-1.98	0.04
Severity		641.01	269.04		683.84	234.47	0.10	-1.45	0.14
Mean Sev.		14.09	5.02		13.75	3.87	0.00	0.66	0.50
All Time Absent	1277			1642					
Frequency		8.92	5.11		11.74	6.58	0.00	-4.08	0.00
Severity		431.03	560.36		506.01	564.39	0.93	-1.14	0.25
Mean Sev.		103.63	338.48		94.15	350.52	0.67	0.23	0.81
Sht. Leg. Ab.	1277			1642					
Frequency		6.95	3.99		10.01	5.82	0.00	-5.21	0.00
Severity		180.44	256.47		221.06	224.37	0.11	-1.44	0.15
Mean Sev.		32.11	79.58		34.91	123.56	0.00	-0.23	0.81
Long Leg. Ab.	1277			1642					
Frequency		0.95	1.05		1.03	1.61	0.00	-0.47	0.64
Severity		221.22	509.37		257.50	534.78	0.58	-0.59	0.55
Mean Sev.		127.80	333.81		107.46	286.28	0.07	0.56	0.57
Sht Non-Leg. Ab.	1277			1642					
Frequency		1.75	2.55		1.75	2.25	0.14	0.01	0.99
Severity		29.36	61.48		27.44	54.52	0.15	0.28	0.77
Mean Sev.		8.33	13.94		7.80	11.44	0.02	0.36	0.72

Overall, it appears from mere comparison of the means that the program was clearly successful although the T-test probability statistics are of questionable value. Once again, a caveat is critical here. These T-test values are based on distinctly non-normal data distributions and come from a 7% sample of the workforce, a sampling procedure necessitated by the sheer volume of computer time and memory space requirements needed to prepare the derived data for detailed analysis.

Hypothesis 9 is supported. See the over-time plots in Figure 7 for a rough graphical picture of the program impact over time.

Aside from the anomaly in the early summer of 1981, several issues become apparent from analysis of the T-test data.

1. There was an apparent overall decrease in short-term non-legitimate absence over the long term as intended.
2. This was absorbed by an increase in long-term legitimate absence instead of producing a larger increase in the number of hours worked, but was accompanied by an increase in hours worked which was large enough for practical if not statistical significance (an estimated 7% increase).

3. There was a small but significant effect induced by annual renewal of the number of casual leave days granted to every employee, such that in early July a surge of short-term legitimate absence occurred. No significant evidence of this effect occurred during the Christmas holiday period, which suggests that the holiday effect only happened in association with the annual renewal, basically a cyclical version of the legitimacy effect.
4. Seasonal shifts in time worked and time scheduled off do occur surrounding holiday periods and the annual renewal period in July, but seasonal cycles seem to have little effect on absence itself, a finding that was counter to prior research.
5. As in prior research, the T-tested data distributions are non-normal to the point where probability estimates are of questionable merit, and although findings in the present work are generally consistent with much of the prior work, this leaves the problem of how to improve upon the cross-sectional comparisons unresolved.

C. The Split-Half Time-Series Reliability Correlations

The procedure for calculating split-half reliability correlations has been presented and analyzed in detail in Appendix C. The results have been described below. Table 11 presents reliabilities for those measures that were central to the substantive questions to be investigated, and identifies the conditions which brought about a few anomalously low reliabilities. A list of such conditions, any of which could cause low reliabilities, follows the results.

The Reliability of Measures for the Substantive Hypotheses

The results below have been reported for three sets of time-series correlations between split-halves of the time-series data for specific measures central to analysis of the substantive hypotheses. The three time-series data sets were: (1) using the 100-day data, an interval in the baseline year from day 281 to day 546; (2) also using the 100-day data, an interval during the post-installation period, from day 982 to day 1680; and (3) using the 365-day data, an interval starting 365 days after data collection began for the post-installation period and continuing until day 1680. For each of the three sets, reliabilities for 27 different time-use indices have been presented. In general, these reliability correlations were unusually good, for data related to absenteeism. See Table 11 and the following discussion for details.

Table 11: Time Series Correlations for Split-Half Reliability

Variable or Index	Pre-Prog. r:100-Day	Post-Inst r:100-Day	Post-Inst r:365-Day
All Time Worked			
Frequency	0.75	0.88	0.99
Severity	0.98	0.92	0.63
Mean Sev.	0.88	0.50	0.50
Routine Work			
Frequency	0.76	0.89	0.98
Severity	0.97	0.91	0.61
Mean Sev.	0.91	0.57	0.64
All Scheduled Off			
Frequency	0.93	0.99	0.98
Severity	0.99	0.96	0.87
Mean Sev.	0.97	0.82	0.92
Missing Values			
Frequency	0.89	0.88	0.46
Severity	0.99	0.96	0.87
Mean Sev.	0.98	0.96	0.84
Routine Off			
Frequency	0.95	0.99	0.97
Severity	0.97	0.98	0.99
Mean Sev.	0.86	0.92	Variation too small
All Time Absent			
Frequency	0.82	0.89	0.98
Severity	0.74	0.91	0.74
Mean Sev.	0.58	0.73	0.64
Sht. Leg. Absence			
Frequency	0.91	0.92	0.97
Severity	Sample too small	0.57	Variation too small
Mean Sev.		0.21	
Long Leg. Absence			
Frequency	No Values in 1/2 Prior to Installation	0.84	0.88
Severity		0.94	0.87
Mean Sev.		0.89	0.69
Sht. Non-Leg. Ab.			
Frequency	Variation too small	0.76	Sample became too small
Severity	0.96	0.36	
Mean Sev.	0.68	0.56	

As Table 11 shows, reliabilities for most of the time-use measures were high compared with conventional absence measures, typically ranging between 0.5 and 0.99 with most values between 0.8 and 0.98. Correlations for indices of All Time Worked and for Routine Work (which was the largest single component of "All Time Worked") were both very strong and both very similar, almost identical when compared across the indices of Frequency, Severity and Mean Severity. The same strong results were also found for All Time Scheduled Off and for All Time Absent. Not surprisingly, Mean Severity uniformly had somewhat lower reliabilities than Frequency or Severity. The procedure for calculation of values for mean severity computes a ratio each day for Severity divided by Frequency, thereby allowing any measurement error for either Severity or Frequency to affect the daily value of Mean Severity.

"Missing values" were an unusual but apparently reliable measure. As Tables 2 through 10 above have shown, the frequency of missing periods across the measurement intervals for DOT employees was quite low (less than one per employee) while the severity of missing value occurrences tended to be high, especially in the baseline period and early in the program when the incidence of undocumented periods of employee time-use behavior was much greater. These conditions could very easily have brought a sporadic distribution of the time-use across the two split halves and led to low reliabilitis for the missing values. But, apparently this time-use was common to many employees and was adequately represented in both split halves. The reliability estimates of the missing values shown in Table 11 are quite strong and consistent, although they do show slightly smaller reliabilities for frequency than severity.

Anomalous Values in the Split-Half Reliabilities

Among the 78 split half reliabilities on measures for the substantive hypotheses were fifteen unusually low or anomalous values (see Table 11). Conditions under which these could be expected to occur have been described in detail in Appendix C. A brief outline of those conditions follows.

Conditions Conducive to Anomalously Low Reliability in the Time-Use Measures

1. The sample size can be too small, including too few employees.
2. Very few employees may participate in the respective time use.
3. The time-use itself may rarely occur, even for those who did choose it.
4. The time-use may be so carefully regulated or may otherwise have occurred with such consistency that the amount of observed variation underlying the split-half reliability estimate was too tiny.

5. The amount of time-use in one split half or the other may be constant over time causing a standard deviation of zero and making calculation of a correlation impossible.
6. The subdivision of overall categories of measurement into more detailed ones (e.g., three detailed absence categories are components of All Time Absent) may introduce too many distinctions for the amount of behavior in the sample and so cause the subsample of behavior assigned to one or more of the detailed time uses to become too small for effective study.
7. The length of the time-series may be too short to effectively establish reliability for the series.
8. The length of the observation period for collecting each day's values on each measure may be too short to smooth out the sporadic daily fluctuations.

All eight of the conditions discussed above were potential threats to the reliability of the time-use measures as the detailed discussion in Appendix C has shown. In examining the reliabilities reported in Table 11, some of these conditions became evident from cross-referencing related statistics that correspond to each anomaly. The reasons behind each anomaly were quite clear and these have been noted in Table 11. In general they may be reduced to one or both of two problems: (1) small sample size, or (2) very little variation in the underlying behaviors being measured.

For the measures with a large enough sample of variation in the underlying behavior, the reliabilities were quite high. The anomalous reliabilities all occurred in the detailed categories of behavior where sample sizes of behavior were bound to be much smaller, and generally occurred at times when the behavior was unusually scarce.

Both Hypothesis 10 and Hypothesis 10(a) were generally supported for the broad overall measures. The detailed small sample measures do not all lend themselves well to assessment of the reliability hypotheses because of sample size. But, where the reliability estimates for both frequency and severity were useable, they were generally quite strong and neither was clearly superior.

D-1 & D-2: Discussion of Time-Series Correlations Testing the Substantive Hypothesis, Using Severity Indices

Time-series correlations were calculated within selected time intervals for both the 100-day and 365-day datasets to examine various substantive hypotheses as outlined in Chapter IV and in Appendix C.

Results for Hypotheses 11, 12, 13, and 14 were examined in that order for each of the five selected combinations of: 100-day or 365-day measure, and time interval of analysis. The results reported to show association between time intervals, correlations and associated significance levels, datasets, and related hypotheses are summarized in Table 12.

Table 12: Time-Series Correlations Testing Substantive Hypotheses for Severity Indices Only

Variable 1	Variable 2	Interval	r	r:Prob	Dataset	Related Hypothesis
		281- 546	* 0.45	0.0001	100-day	
Short-Term Legitimate Absence	Long-Term Legitimate Absence	982-1277	0.36	0.0001	100-day	Hypothesis
		1278-1680	0.56	0.0001	100-day	11
		982-1680	0.52	0.0001	100-day	Progression
		1247-1680	* 0.53	0.0001	365-day	
		281- 546	*-0.17	0.0058	100-day	
Short-Term Legitimate Absence	Short-Term Non-Leg. Absence	982-1277	-0.89	0.0001	100-day	Hypothesis
		1278-1680	-0.93	0.0001	100-day	12
		982-1680	-0.92	0.0001	100-day	Legitimacy
		1247-1680	*-0.67	0.0001	365-day	
		281- 546	* 0.69	0.0058	100-day	
Short-Term Non-Leg. Absence	Long-Term Legitimate Absence	982-1277	-0.44	0.0001	100-day	Hypothesis
		1278-1680	-0.55	0.0001	100-day	13
		982-1680	-0.53	0.0001	100-day	Progression
		1247-1680	*-0.56	0.0001	365-day	Legitimacy
		281- 546	0.97	0.0058	100-day	
Short-Term Non-Leg. Absence	Time Worked	982-1277	-0.02	0.0001	100-day	Hypothesis
		1278-1680	0.25	0.0001	100-day	14
		982-1680	0.06	0.0001	100-day	Program
		1247-1680	*-0.10	0.0001	365-day	Impact
		281- 546	-0.98	0.0058	100-day	
Missing Values	Time Worked	982-1277	-0.91	0.0001	100-day	As
		1278-1680	-0.75	0.0001	100-day	Above
		982-1680	-0.96	0.0001	100-day	
		1247-1680	-0.97	0.0001	365-day	
		281- 546	*-0.71	0.0058	100-day	
Missing Values	Long-Term Legitimate Absence	982-1277	-0.22	0.0001	100-day	As
		1278-1680	-0.45	0.0001	100-day	Above
		982-1680	-0.18	0.0001	100-day	
		1247-1680	-0.40	0.0001	365-day	
		281- 546	* 0.08	0.0058	100-day	
Missing Values	Short-Term Non-Leg. Absence	982-1277	-0.38	0.0001	100-day	Hypothesis
		1278-1680	-0.63	0.0001	100-day	14
		982-1680	-0.32	0.0001	100-day	Program
		1247-1680	*-0.47	0.0001	365-day	Impact

Note: * indicates a finding in which one or both of the measures correlated showed low reliability for the split-half samples. Since the overall sample size is twice that in the split-halves, such findings may or or may not be based upon a reliable series of values (see Appendix II

There is evidence of a very strong legitimacy effect after the program installation occurred, but not before. In fact, before installation, short-term non-legitimate absence correlated +0.97 with time worked, while after installation it correlated -0.92 with short-term legitimate absence, and was uncorrelated with time worked. Hypothesis 12 is definitely supported (see Table 12). Ancedotal evidence from the DOT Personnel Director indicated that before installation a great part of absence was informally excused by the former Personnel Director as a distribution of these personal favors. The more an employee was there at work, the greater the chance of personal favors, which in any case were unpunished absences at that time. After the program, the legitimate portion of this was given formal recognition (and so was the non-legitimate absence), so that after the program became routine the record includes fewer missing values.

The group-level progression hypothesis (Hypothesis 11) is clearly not supported, however. Short-term and long-term forms of legitimate absence are positively correlated, both on the 100-day and 365-day measures, although the reliability of the 365-day measure is questionable due to sample size on the short-term legitimate absence measures.

A combined progression and legitimacy effect (Hypothesis 13) is supported, but very probably the apparent effect is largely induced by legitimacy alone. As noted above, there is no support for any independent progression effect at the group level, and in fact, the evidence even indicates the reverse -- that an increase in more severe forms co-occurs with more of the less severe forms of absence.

Hypothesis 14 is not supported, presumably because the overall impact of the program was allowed to dissipate in long-term legitimate absence which had been deliberately placed outside the purview of the program.

A Comment on Frequency-Severity Relationships

For the sake of curiosity, the relationship between frequency and severity of the three major time-use categories was examined briefly. Table 13 below summarizes the results. Time scheduled off shows a strong negative correlation between frequency and severity. But, for the other two time uses, the relationship is strong and positive. As the extremely low variability of the indices for Time Scheduled Off indicated, the DOT regulated Scheduled Time Off very thoroughly. The negative frequency-severity correlation presumable just reflects careful scheduling of a fixed total amount of scheduled time off per employee. This inverse relationship might also obtain for shorter-span-of-observation measures (i.e., a one week or one day base for the moving values). But, to reliably test that would require a larger sample than the present one since the smoothing effect of the longer moving averages would be lost.

Table 13: Time-Series Correlations To Inspect Relations Between Frequency Indices and Severity Indices

Variable 1	Variable 2	Interval	r	r:Prob	Dataset
Time-Worked Severity	Time-Worked Frequency	982-1680	0.70	0.0001	100-day
		1246-1680	0.60	0.0001	365-day
Time Sch Off Severity	Time Sch off Frequency	982-1680	-0.78	0.0001	100-day
		1246-1680	-0.74	0.0001	365-day
Time Absent Severity	Time Absent Frequency	982-1680	0.51	0.0001	100-day
		1246-1680	0.73	0.0001	365-day

Chapter VI
Conclusions and Implications

General Conclusions:

The findings presented in Chapter V lead to seven general conclusions.

1. The dissertation set fundamental methodological goals of reconceptualizing the definition of absenteeism to overcome problems of theoretical inconsistency and low reliability encountered in prior research and of establishing an appropriate procedure to test reliability for those measures. Absence has been redefined as one among many forms of time-use, and time-series reliability tests have demonstrated a high and generally consistent reliability for these measures where sample size and variability were sufficient for an informative test.

On the whole, the time-series reliability correlations show a stability in these measures that is about twice as high as findings in prior research, especially for severity (see Breaugh, 1981; Hammer and Landau, 1981). Part of this is probably due to the care taken in design, installation and administration of the absence data collection itself after the program started. Part is due to the use of a time-series approach instead of the cross-sectional procedures used in prior research. In particular, the time-series correlations, although very simple, are a necessary change in procedure for such work. They provide a reliable base for hypothesis testing, while the cross-sectional T-tests did not, despite their utility for comparing the means and providing a handy test of program impact.

2. The findings also shed light on the question of whether measures of absence frequency or severity (duration) are more reliable for absence research. The controversial use of frequency measures in replacement of severity measures is not necessary. Both indices are highly reliable, but while absence frequency can tell us what starts to happen and how often, absence (or time-use) severity can tell us with far greater precision exactly what does happen, at least with respect to overall impact on work and the flow of time expenditure from one time-use to another.

3. Hill and Trist and others have hypothesized that absence as a form of withdrawal may, over time, occur with increasing severity. Although the analysis implemented after derivation of the time-use measures is not suitable for examination of this effect at an individual level, there was no evidence of such a progression effect at a group level.

4. There was some minimal evidence of seasonal or holiday shifts in absence occurring together with strong renewal effects, and in the baseline year, evidence of other shifts occurring in seasonal patterns but not affecting

absence. In general, the T-test data support seasonality, but not for absence itself, only for scheduled work and time off.

5. The legitimacy hypothesis, that employees will substitute legitimate absence for non-legitimate absence, is very strongly supported, both as an intra-annual and a long-term trend. Employees emit absence behavior in patterns that are allowed by the rules and the absence control policy. This appears to have dominated the impact of the program far beyond the level of a mere side effect. It was the main effect and in this respect, somewhat surprising. An hypothesis has been introduced in the present work that annually renewing the bank of legitimate days of discretionary absence an employee may "take" would result in an annual renewal of substitution effects. This was clearly supported by the T-test results. There is some evidence for seasonal fluctuation, a very clear renewal effect, and a minor holiday effect associated with the renewal effect. The renewal effects strongly illustrate the impact of collectively bargained time-use allowance upon employee behavior, a sort of special condition legitimacy effect.

6. The dissertation was also intended to identify the overall impact of the absence control program, if any occurred. There was a significant decrease in short-term non-legitimate absence, and some increase in the overall

amount of time worked (about 7%, a return on the investment that was practically but not statistically significant), plus an overall increase in long-term legitimate absence (significant). Prior to the program the archival data showed long-term absence to be minimal so the program deliberately included a decision not to focus on long-term absence until later. As a result, due to the legitimacy effect, much of the program impact was apparently absorbed in the long-term absence shift. The 7% gain in time-worked was a very clear practical success in that the success criterion set by DOT before the program started was a 1% gain. Further comment on the program impact follows below.

7. On a methodological level, the evidence brought to bear on procedural hypotheses clearly supported the use of group level time-series analysis, and in particular supported the time series correlation technique employed here. The evidence also supported the contention that daily cross-sections of individual absence data are distributed highly non-normally. Despite this finding, T-tests comparing daily cross-sections of individual data over time still showed results that are consistent with the results of group-level time series analyses. The comparison of daily cross sections has obvious practical merit since it allows estimation of costs and benefits. But for inferential tests to analyze the significance of trends in time-use behavior, conventional cross-sectional procedures are at best tenuous and

where possible these should probably be set aside in favor of the time-series procedure.

A Detailed Evaluation of Long-Term Program Impact

On the whole, the absence control program had several long-term impacts on employee time-use.

1. The program was, by contract, supposed to decrease short-term non-legitimate absence significantly. A significant decrease of 112 hours per employee per year did occur. Allowing for turnover and assuming an average size for the DOT workforce of 2000 employees, this would correspond to 224,000 hours per year. At an average of $20 per hour for wages and benefits, this constitutes a savings to the public of about $4.5 million per year.

2. The program was not intended to change the amount of time scheduled off, but during the program a significant 200 hours per employee per year decrease in scheduled time off did take place. This would correspond to about 400,000 employee hours per year, or $8 million per year at $20 per hour, presumably an unexpected side effect of the more precise knowledge the program information system provided for scheduling. This discovery would not have surfaced without the detailed statistics about time-use shifts that were possible due to reconceptualizing both absence and attendance and other employee activities as alternative forms of time-use.

3. A massive increase of 223 hours per employee per year in long-term legitimate absence occurred, mostly during the first year following installation.

4. There was also: (a) a strong negative correlation between short-term non-legitimate absence and long-term legitimate absence; (b) a moderate positive correlation between short-term legitimate absence and long-term legitimate absence (which refutes the progression hypothesis at a group level); and (c) no correlation worthy of note between time worked and short-term non-legitimate absence after the program began.

5. In order for the program to pay for itself, it was expected to move 1% of employee time-use from other categories into time worked. Number of hours worked should have increased; and they apparently did increase by 107 hours per employee per year, an amount which closely corresponds to the saving of hours generated by reducing short-term non-legitimate absence, and a 700% return on the investment within two years.

6. The number of hours hidden in missing values was reduced by 218 hours per employee per year. These went into both work and absence.

Overall, these six events in concert indicate that the program did four types of things to time-use behavior.

a. It moved about 110 hours per employee per year from short-term non-legitimate absence into legitimate time-use behaviors, primarily long-term legitimate absence.

b. It moved a large amount of time formerly hidden in the records as missing values into time worked, and some time into long-term legitimate absence.

c. It increased the amount of time worked overall.

d. The program also induced a very strong effect attributable to annual renewal. Immediately after the renewal date itself on July 1st, an annual cycle occurs in which time worked and short-term non-legitimate absence decreases while short-term legitimate absence increases. This impact of the program offers further support for the legitimacy hypothesis.

On the face of these results, the project worked well. But, between the concession bargaining giveaway of 80 hours per employee per year and the lack of any effective restraint on long-term legitimate absence, much of this impact was dissipated in substitution toward forms of absence that remained legitimate. DOT employees are apparently even more resistant to giving up their absence time than the program planners anticipated.

The positive results obtained are largely attributable to the thorough groundwork prepared by the consultants and the enthusiastic and widespread support from the DOT top management and general supervisors. That support began in the planning phase and carried throughout group workshops and supervisory training in how to operate the program, computerization, support of the attendance incentive and day-to-day program administration, all in full cooperation with the consultants.

Implications of the Present Research for Future Work

There are a number of implications which follow from limitations and insights discovered through the present work. The time interval between the baseline year and the start of the program presents an unfortunate unknown period, and although the 30-day sample was small, there were significant changes in absence evident in the 30-day cross-sectional comparisons, as well as other changes. For a fully detailed analysis, this period of time should also be sampled. The data required are apparently not as clean as data recorded later, but are all still available in the DOT payroll records and this task could very conveniently be completed. Furthermore, data between August 8, 1982, and the present are now available and these will have captured the long-term impact of the program in fuller detail. Those data should also be incorporated into the database before further analysis is carried out.

The absence literature has recognized variations in absence across demographic and occupational subgroups. This work still needs to be done. Very possibly the correlations reported here would be even higher if other trends were not confounded by demographic variations. Also very possible is an accompanying increase in normality if non-absence proneness varies by demographics of employees and this would strengthen the T-tests. Given additional computer resources, the rest of the workforce, excluded from the present work, could be included so that the problems introduced by small sample sizes would be overcome. This would also permit a comparison of the sample results with results for the entire workforce, an intriguing methodological curiosity.

Furthermore, the measurement technique used in this work should be introduced and tested in other settings, both for methodological and substantive insights. There may be variation due to cultural differences or industrial sectors and these have policy and measurement implications concerning the validity of both the absence control program itself and the derived measures.

Shorter time spans of observation such as one day or one week measures or 30 day time-series{measures should be examined using a time-series correlation procedure. There is need to check for an inverse relationship between frequency and severity of absence time-uses that could become evident across the narrow time interval.

Using the same methods applied here, a more detailed analysis of substitution across additional pairs of time-use variables could be done, supplemented by cross-log panel correlations to assess causality. Only major trends were examined here.

Causal relationships and induced changes attributable to the program could be inferred to a greater extent if a formal time-series analysis were conducted to establish functional forms of change and interdependence. These procedures use exploratory research tools to isolate underlying patterns that are of theoretical interest and are consistent in mathematical form. From that point, it is a short step to synthesize the separated patterns described by each isolated component of absence into a general model.

A more detailed examination of the distributional forms of time-use behaviors should be completed. The present work substantiates prior findings in that absence behavior is distinctly non-normal in its distribution. The implications of this for applying transformation procedures to make absence data more tractable for use in inferential statistics have yet to be clarified. It is not clear, for example, just what the mean of a transformed distribution of absence data actually means. Does it have to be transformed back to the original units before its practical significance may be interpreted? An empirical check on the result of doing that, or of comparing the results with those for the untransformed data from the entire workforcce would be instructive.

The use of techniques presented here combined with time-series research on attitude changes should definitely be carried out in respect to absenteeism. The hypothetical relationship between attitude change and behavioral change has yet to be established, both in general and specifically in respect to absence behavior.

The time-series correlation procedure has promise, particularly for measurement and analysis of events in the personnel field. That possibility should be explored. In particular, the associations between time-use, monetary use and other resource use, and performance on the job has sweeping implications, both for modeling the performance system management tool and for fundamental research on human work behavior.

A more technical insight from the present research contributes in a direct and pragmatic way. The type of work presented here creates from merely large amounts of raw data enormous amounts of derived data and these data require computerization to make them manageable. Anyone planning to use the technique should be prepared to resolve the associated problems of database construction and information management. Despite these problems, using a detailed time-series procedure for program evaluation of absence control is clearly so much superior to the cross-sectional approach in scope, precision and reliability that the effort is warranted, especially since many of the technical problems involved have already had to be resolved for completion of the present work.

Implications for Personnel Practice

1. The phenomena of employee attendance, absence or other time uses constitute a complex and interdependent set of events which covary over time in measurable and predictable ways. This interdependence has been reduced to a tractable form.

Four requirements were met to bring this about. (A) Program evaluation to measure change in those events was based upon a set of time-use definitions that are custom tailored and fundamentally anchored in the absence control policy. (B) The administrative precision and consistency of data collection was sufficient to provide a thorough record of the employee time uses. This was one of the functions served by pre-program supervisory training. (C) The entire information set and procedures to create summary statistics from it were computerized so that the plethora of interdependent facts became a coherent, interpretable system. (D)

There was extensive administrative and supervisory support for the program throughout the work system. Even the unions cooperated.

It is probable that in the absence of any one of these four elements, both the absence control program itself and the information system which supported it would have failed.

2. The dominant influence on employee absence and attendance is the implementation of control practices. Absence or attendance is apparently a highly calculated decision. If allowed to be absent with no associated cost then employees will be absent. The legitimacy effect is very clear and very strong.

For absence control these two issues mean that every loophole in the absence control policy and its implementation in practice must be plugged. These behaviors covary as a system. To suppress one aspect of absence will induce an increase in some other aspect unless the system specifically prevents the substitution.

APPENDIX A

THE DETAILED DESCRIPTION OF THE WORKFORCE DEMOGRAPHICS

Appendix A

The Detailed Description of the Workforce Demographics

Unpacking: The initial unprocessed raw data tapes received from the research site contained, by construction, some coding idiosyncracies. To save space in computer memory, the data for each employee had been packed into 6-month records within which the position of any character indicated what it represented, or what date it referred to. This created, in effect, about four hundred different variables, before any new derived variables could even be created. Yet most of those variables were really the same variable taking on a value that had occurred on a different day.

Before any further analysis could be done, all of the raw absence data were split off from the demographics and both were rewritten onto new tapes. In the process, a unique (and confidential) case number was introduced to replace each employee social security number. This case number provides a convenient link between the demographics for each employee and the absence data for that employee. The raw absence data were changed in the procedure from 6-month records to daily strings containing the employee case number, the relative date of the time-use (absence) behavior (starting with day 1 on January 1, 1978), the time-use (absence) code, and the duration of the time-use in hours. The demographic data were (selectively) written out to a separate file each time they appeared on the raw data tapes (about 5 to 8 times per employee). These were later edited to create one clean, nonredundant file of all 2336 employee demographic records.

The statistics below summarize the demographic characteristics of the workforce. They were generated by submitting the file of demographics to analysis using the Biomedical Dataprocessing Package (1982), subroutine BMDP2D. The first such run generated an overall description of the entire workforce and those statistics are presented below in detail. That set of statistics provided numerical values for the composition of the entire workforce by demographic subgroup. A fortran subroutine was then written to randomly select a sample group representing all employees in the DOT workforce. Demographic statistics for the sample were then generated using BMDP2D. On the following page, in Figure 8, is a list of the demographic variables analyzed.

Variable	Description of the Data Item
Case Number	A 4-digit integer case label used to link the absence data tapes and the employee demographics.
Bargaining Unit Code	A 4-digit alphanumeric code which may allow later data analysis comparing results for different collective bargaining constituencies.
Year	1 2-digit integer for the last two digits in the year the data represent.
First Month	A 2-digit integer for the first month the data represent ... 1 through 12.
Year Hired	A 2-digit integer value.
Month Hired	A 2-digit integer value.
Birth Year	A 2-digit integer value.
Birth Month	A 2-digit integer value.
Birth Day	A 2-digit integer value.
Job Class	A 6-digit integer code representing the payroll job category of each employee.
Cost Center	A 4-digit integer code identifying the organizational profit-center associated with each employee. May be used in later work for comparison across occupational subgroups.
Marital Status	M for married, S for single. A 1-digit alphabetic code.
Dependents	A 2-digit integer variable.
Race	A 1-digit integer code.
Sex	M or F, a 1-digit alphabetic code.

Figure 8: The Demographic Data Variables

Continued on next page

Figure 8 Continued

Location Group	A 2-digit integer code which identifies the work site at which each employee is stationed. Could be used in validation across locations if the coding documentation for it was available, but at present, it was not.
Job Class Group	A 2-digit brief occupational Category used for Affirmative Action. The coding documentation for these values is also unavailable.
Quit Month Quit Day	A 2-digit integer code for the month and the day of termination. In most cases this field is blank.

Detailed Statistics for the DOT Workforce Demographics

Case Number: These values were generated to provide a unique link between the demographic data and the time-use data for each respective employee, a link which intrinsically protects the confidentiality of data for each employee. One-by-one they range from 1001 up to 2336.

Bargaining Unit: The bargaining unit code for each employee represents the subdivision of each given union which holds representation rights for the employee. There are 1342 different subdivisions of this kind represented in the DOT demographics, plus 60 employees for whom no code was recorded (missing values). These results have been clustered to provide the following summary statistics. The largest employee group, Transportation Equipment Operators (bus drivers) were represented by the Analgamated Transit Union Local 26. A substantial segment of the DOT workforce was represented by various AFSCME Locals. Two of these, AFSCME Local 312, representing 533 mechanics and general transportation employees over the research period, and AFSCME Local 214, representing 251 Clerical Transportation employees, together encompassed 34.4% of the DOT employees. A further 21.8% of the workforce (479 employees) was represented by locals of the Detroit Building Trades Council. 11.3% of the employees were members of a variety of unions, including Teamsters Local 214, all non-supervisory. 3.0% of the workforce (68 employees) were members of supervisors' and technical/supervisory employees' unions. The remaining 72 employees are non-unionized.

Year and Month of Data Collection: The variables were a vital component of the raw absence data, but in the demographics they are merely coincidental and therefore uninformative.

Year and Month and Day Hired (Seniority Data): The median and modal values for each of these is reported in the table below. For the month and day the mean is also included.

Table 14: Year and Month and Day Hired

Variable	Year	N	Month	N	Day	N
Mean	not reported		June	139	15th	74
Median	1974	196	July	210	15th	74
Mode	1979	319	August	311	6th	114
N	2240		2241		2241	
Range	1917-82		1-12		1-31	
Missing	96		95		95	

From these statistics, several patterns in hiring at DOT are indicated. 1979 was the biggest hiring year for DOT (and 1974 was second). This is consistent with a sizable purchase of new busses at about that time. The strongest hiring month appears to be August; September was second, with 214 hires. Day of hire during the month is fairly variable, but quite evenly distributed in general with a slight modal peak on the 6th and a fluctuation on approximately weekly intervals that suggests a weekly pattern, perhaps attributable to hiring practices. The median seniority was about 8 years (starting in June 1974) at the end of the program year in June, 1982.

<u>Year</u>, <u>Month</u> <u>and</u> <u>Day</u> <u>of</u> <u>Birth</u> (<u>Age</u>): Birth year for DOT employees ranged from 1914 to 1962, so that at the end of the study in 1982, the oldest employees were about 68 years of age and the youngest employees about 20 years of age. The median age was about 38 years and the mode (n of 107) was 32 years. These figures again reflect a strong upsurge

in recruitment levels in recent years, in which the new recruits tended to be somewhat younger. Day and Month of birth are both quite evenly distributed across the year for the DOT workforce. Summary statistics are presented in the table below. These are included here for comparison with the sample characteristics.

Table 15: Birth Dates

Variable	Birth Year	Birth Month	Birth Day
Mean	not reported	June	15th
Median	1945	July	16th
Mode	1950	August	23rd
Range	1914-1962	Jan. - Dec.	1 - 31
N	2335	2336	2335
Missing	1	0	1

Job Class (Occupational Payroll Category): Two types of occupational data were recorded for the two types of DOT workforce. The first type described here was recorded as a six-digit, detailed payroll code assigned to each specific job. The second type is described later. There were 115 different detailed occupational categories for the DOT employees. These range from the Director and Deputy Director of the department to clerical employees and office workers, transportation terminal workers, transportation equipment operators (TEO's ... 1335 of these), mechanics, and other craft workers. For convenience of interpretation, these detailed occupational categories have been summarized into the seven clusters below.

Table 16: Job Class

Job Class Cluster	No. of Employees	No. of Jobs	% of Total
Managerial - Executive, Supervisory	105	23	4.5
Administrative & Technical/Professional	28	18	1.2
Craft Workers	401	23	17.2
Transportation Equipment Operators	1335	1	57.1
Guards	42	3	1.8
Clerical Employees	182	23	7.8
Terminal Service Workers	234	16	10.0
Miscellaneous Others	9	8	0.4
TOTALS	2336	107	100.0

<u>Cost Center</u> (<u>Geographic/Functional Distribution of Employees</u>): 40 distinct cost centers were used to identify the specific work location and type of work of DOT employees. These are listed with summary statistics showing: the number of employees in each geographic location from all job categories and the number of salaried and non-salaried employees from all locations and job categories. The codes used to represent the "miscellaneous" category seem to suggest that the jobs involved are all salaried, but since the coding manuals available from DOT were not specific in this

regard, no use can be made of those particular coding categories yet. The other 4-digit cost-center codes, nonetheless, offer an excellent sampling-frame for later subsampling from the database.

# Salaried Employees per Cost Center	Cost Center	Cost Center Code	N	# of Employees by Location
10	Gilbert Garage Salaried	1166	10	
	Gilbert Terminal Workers	1230	41	Gilbert
	Gilbert TEO's	5420	420	539
	Gilbert Garage Workers	6390	89	
9	Shoemaker Garage Salaried	1167	9	
	Shoemaker Terminal Workers	1200	28	Shoemaker
	Shoemaker TEO's	5460	461	572
	Shoemaker Garage Workers	6380	100	
13	Coolidge Garage Salaried	1168	13	
	Coolidge Terminal Workers	1220	36	Coolidge
	Coolidge TEO's	5480	452	593
	Coolidge Garage Workers	6370	119	
16	Administration Salaried	1100	16	
55	Personnel Salaried	1110	55	
15	Auditing Salaried	1120	15	
3	Stationery Stores Salaried	1121	3	
30	Cashiers Salaried	1140	30	
41	Technical Services Salaried	1164	41	
23	Rolling Stock Office Salar.	1165	23	
8	Shops Salaried	1161	8	
23	Plant Maintenance Salaried	1160	23	
	Plant Maintenance Trades	6361	38	
	Plant Mainten. Electricians	6362	7	
	Plant Maintenance Drivers	6363	22	Main
	Heating Plant Maintenance	6360	21	Office
	Technical Services, Maint.	6345	5	477
	Unit Repair, Maintenance	6300	35	
	Heavy Repair, Maintenance	6330	56	
	Body Shop, Maintenance	6340	20	
	Security	6350	39	Loc.
	TEO's on L.T. Disability	5499	2	Indet-
	Miscellaneous (1118, 6367, 6377, 6387, 6397, 7370)		41	ermin. 80
6	Charter Services Salaried	1170	6	
20	Transp. Planning & Schedul.	1180	20	
32	Transport. Office Salaried	1190	32	

TOTALS:
304 All Salaried Employees
1991 Non-Salaried Employees
41 Miscellaneous Employees
2336 All Employees in the DOT

Figure 9:
Distribution of All Employees in the DOT Workforce by Geographic & Functional Cost Centers

Marital Status: Marital status of DOT employees was only recorded as either married or single for payroll purposes. Out of 2336 cases read, there were 1940 single employees, 392 married employees, and 4 cases with missing values. In short, 83% of the workforce is single.

Dependents: The number of dependents for members of the DOT workforce ranged from 0 to 10 and 12, with a mean of 0.75 and both median and mode of 0. 1592 employees had no dependents (so presumably about 352 employees had dependents but were not married). For those 744 employees who did have dependents, the mean number of dependents was 2.34, the mode was 1 dependent, and the median was 2. There were 5 missing values.

Race: Racial characteristics of the DOT workforce are defined as follows. "Black" means persons have origins in any of the black racial groups of Africa and does not include those of Hispanic origin. "White" means all persons have origins in any of the original peoples of Europe, North Africa, or the Middle East and does not include those of Hispanic origin. "Asian or Pacific Islanders" means all persons having origins in any of the peoples of the Far East, Southeast Asia, the Indian Subcontinent, or the Pacific Islands. This category includes, for example, Chinese, Japanese, Korean, the Phillipine Islands and Samoa. "American Indian or Alaskan native" includes all persons having origins in any of the original peoples of North America and maintaining a cultural identification through tribal affiliation or community recognition. The "Hispanic" category means all persons of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish culture or origin, regardless of race. The category "other" was included to catch any misfits or unknowns. The distribution of DOT employees by category was as follows.

Table 17: Race

Category	#	% of DOT	Category	#	% of DOT
Black	1856	79.5%	Hispanic	3	0.1%
White	442	18.9%	Am. Indian	1	0.0%
Asiatic	7	0.3%	Other	3	0.1%

There were 23 missing values.

<u>Sex</u>: Out of 2336 employees on the DOT Demographic File, 1981, or 84.8% were male. 354 of the employees were female, or 15.2%. There was 1 missing value.

<u>Location Group and Job Class Group</u>: Two of the most potentially useful categories of demographic data were flawlessly recorded throughout the entire measurement period. Unfortunately, the coding manuals are not available for these groups at this time. But, on the other hand, the constructs of theoretical interest described by those data are also described by combinations of the cost center codes above,

and those 4-digit codes can be used to provide the necessary sampling frame. Each value is a 2-digit code ranging from 0 to 14 and 99.

<u>Year and Month of Quit</u>: 400 employees quit between July 1, 1978 and June 30, 1982; but this figure should be taken more as an indicator for June 1, 1980 through August 10, 1982 since no values appear in the demographic record from 1978 or 1979. This amounts to about 17.1% of all employees on the payroll over the period of study, and calculated across 26 months, it approximates an annual rate of 7.9%. 1981 was the largest year for quits (180), but the number of terminations in 1982 would probably have exceeded those for 1981 (182) by December since the 1982 quits recorded only represent 7 months of 12 in 1982. On that basis, the rate of terminations was rising.

A look at the monthly quits distribution shows a definite pattern. The monthly distribution of the 400 quits on record by percentage is noted on the following page. Winter and Spring (January to July) are about 3.4 times as likely a quit period as Summer and Fall (August to December). On the basis of progressive levels of alienation as a theory explaining absence and related turnover, January would therefore be a likely month in which to watch for increases in long-term absence prior to termination.

Table 18: Year and Month of Quit

Month	Count	% of Total	Histogram of % by Month		
Jan	157	39.3%	XXXXXXXXXXXXXXXXXXXX		
Feb	25	6.3	XXX		Winter
Mar	54	13.5	XXXXXX	312	&
Apr	33	8.3	XXX	or	Spring
May	27	6.8	XXX	78%	
June	16	4.0	XX		
July	17	4.3	XX		
Aug	11	2.8	X		
Sep	33	8.3	XXXX	88	Summer
Oct	11	2.8	X	or	&
Nov	4	1.0	X	22%	Fall
Dec	12	3.0	XX		

Brief Demographics for the Subsample Group

Since the computer time required to create the derived data for one case alone is about 30 seconds, it was quite impossible to carry out an analysis of all 2336 cases in the DOT database. Instead, the statistics that were generated by analyzing demographic data for the entire DOT workforce were used to calculate a group-specific "sampling fraction." This in turn was incorporated into a fortran program, which generated a random selection of cases representing the entire employee population. The target sample size for this group was 160 cases. Due to the way in which the random selection program operated, the actual sample generated includes 164 cases. This sample was quite adequate in size to accommodate both attrition and replacement over time due to turnover. It was also randomly selected and just small enough to accommodate the computer processing time constraints. Statistics for the most salient demographic characteristics of each subpopulation sample were generated using BMDP2D again, and are summarized below. Note that values for events in 1982 represent only the first seven months, while employees terminating before the study began are not represented at all. Demographic figures which were merely incidental have been omitted from the subgroup statistics.

Demographics for the Sample from All DOT Employees

There were 2336 employees in the entire DOT database. The computer resources available could handle about 160 cases at a time, enough to cover sample size decreases due to turnover, etc. Selecting about 160 cases required a sampling fraction of 160/2336 or 0.0684932. This limit was used in the random case selection program and generated a set of 164 cases randomly selected from the data for all DOT employees. The deviation from 160 cases to 164 is presumably a result of rounding error.

The median year of hire was 1974 and by 1974 56.1% of the employees in this sample were already on the payroll with 9% hired in 1974. By 1979 (the modal year) 95.5% of the sample had been hired and 16.1% were hired in 1979 alone. The seniority year of these employees ranged from

1942 to 1982. The modal month of hire (13.5% of employees) was August, with a mean in June and median in July. Hiring also seemed to have peaked from these employees in the first and last few days of each month.

The median year of birth for employees in this sample was 1944, with 51.8% of the employees born in or before 1944. The mode was 1950 and the range was from 1916 to 1960. In general, the age distribution of this sample is moderately skewed in favor of younger employees. Day and month of birth are evenly distributed.

According to the Cost-Center and Job-Class codes, 55.5% of the sample for all DOT employees were Transportation Equipment Operators (TEO's), i.e., 91 bus drivers.

83.5% or 137 employees in the sample were single. 15.9% or 26 employees were married. 73% or 119 of the employees had no dependents. For those 27% who had dependents, the modal number of dependents was one, the median was two, and the range was from one to ten.

Most employees in the overall sample were black (78% or 126 employees). The remainder (22% or 36 employees) were white. 87.2% or 143 employees were male. The remaining 12.8% or 21 employees were female.

During the three years spanned by the study, 28 out of the 164 employees in the sample quit, or about 17%, an annual turnover of about 2.5% in 1980 and about 7.3% in both 1981 and 1982. Most of these terminations occurred in January (42.9%, July (14.3%), and September (14.3%), and there was a definite tendency to quit in the first half of the year, not in the late summer or fall.

APPENDIX B

DERIVATION OF THE RAW TIME-USE DATA AND CREATION OF THE DERIVED DATA STRUCTURE

Appendix B

Derivation of the Raw Time-Use Data and Creation of the Derived Data Structure

The raw absence data tape as described in Figure 10 below contains daily values for the case number, relative date, absence type, and "absence" severity in hours. The first two values are assigned to the second two and identify to which employee and on what day those last two data values relate. The meanings of these absence data values are dependent on several things.

1. Of primary importance is the reason for absence (or worked time) which was originally coded for the specific employee on a given day.

2. Second, the duration of the occurrence of the time-use over time may persist until is forces a recoding of the entire occurrence or at least of part of it. Indeed, the duration of an absence episode may initially be understood to be short and yet later may turn out to last for weeks, or even months. Reassignment of such time-use occurrences to corrected categories from their onset is the first of two main functions served by the Fortran program which generates the derived database.

3. Third in importance but critical to the integrity of the derived data are a number of systematic coding errors

which must be removed from the data before they can be interpreted and used.

Variable	Description of the Data Item
Case Number	A unique 4-digit integer case-label used also on the demographics tape.
Relative Date	A 4-digit integer for the relative date on which respective absences occur. January 1, 1978 was Day 1. January 2, 1978 was Day 2, etc. The data extend from Day 182 to Day 1680.
Absence Type	A 2-digit integer code as listed in Figure 1. Each code represents how the time was spent on a given day by the employee.
Absence Severity	A 1-digit integer for the number of hours expended in the type of absence behavior coded under absence type above.

Figure 10: The Raw Absence Data Tape Variables

The logical structure of relations between absence reasons has been used to create a program that converts each primitive "absence reason" and duration coded in the raw data into derived measures for one or more of the 13 time-use indices. Figure 11 is a flowchart of the recategorization procedures. The recategorization rules used have been explained following the flowchart. Read the flowchart from right to left, following the arrows. Note that where the condition associated with a given arrow in the flowchart was satisfied, the data values in the raw data categories (on the right) will have been processed as contributing elements and added into the associated higher-level "resultant" category of time use (on the left). "Resultant" categories are each marked with an asterisk in the respective box of the flowchart.

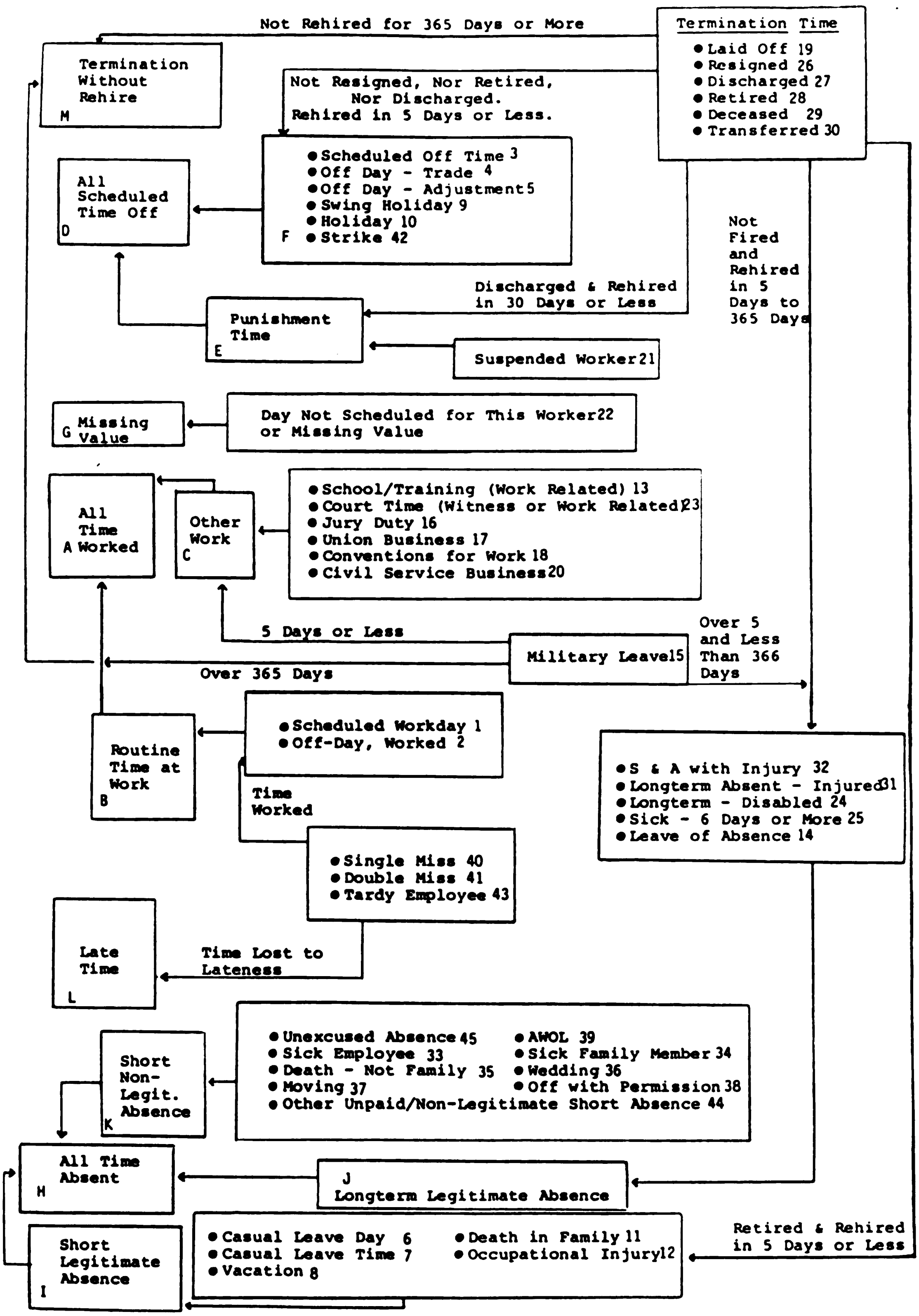


Figure 11:
Logical Relations Among Different Uses of Employee Time

Type 0: Time Worked on the Job

Computer Code #	Code Definition
00	Created for generating the derived database as a code for Day Present and Worked.
04	Off Day - Worked

Type 1: Scheduled/Contractually Provided Absenteeism

Computer Code #	Code Definition
01	Regularly Scheduled Day Off
02	Off Day - Trade
03	Off Day - Adjustment
11	Casual Leave Day
12	Casual Leave Time
13	Vacation
14	Swing Holiday
15	Holiday
23	Death in Family
25	Occupational Injury
30	School/Training (DOT Related)
32	Leave of Absence
34	Military Leave
35	Jury Duty
36	Union Business
37	Conventions (DOT approved)
38	Laid Off
39	Civil Service Business
42	Suspended
62	Not Scheduled
33	Court Time (Witness or DOT related)

Figure 12: Summary of the Raw Absence Data as Recorded in the Department of Transport Absence Management Database or as Modified and then Used to Generate the Derived Data

Continued on next page

Figure 12 Continued

Type 2: Long-Term/Non-Recurring Absenteeism

Computer Code #	Code Definition
19	Long-Term Disability
20	Sick Employee (6 days or more)

(These codes may be used to handle catastrophic illnesses approved by the Personnel Division without the 6 day restriction.)

51	Resigned
52	Discharged
53	Retired
54	Deceased
55	Transferred
18	Long-Term with Injury
17	S & A with Injury

Type 3: Excessive/Recurring Absenteeism

Computer Code #	Code Definition
21	Sick Employee
22	Sick Family Member
24	Death - Not Family
26	Wedding
27	Moving
31	Off with Permission
41	Absent without Permission
43	Single Miss (TEO's only)
44	Double Miss (TEO's only)
46	Strike
47	Tardy
61	Other
40	Unexcused Absence

Notes:

1. According to the Attendance Control Policy, employee absenteeism has been grouped into three categories, with time worked on the job there are four:
   Type 0 - Time Worked on the Job
   Type 1 - Scheduled/Contractually Provided Absenteeism
   Type 2 - Long-Term/Non-Recurring Absenteeism
   Type 3 - Excessive/Recurring Absenteeism
2. These types are reworked as shown in Figure __ to create the various levels and categories of derived data.

Derivation of the Time-Use Data Values

The programs which generate the derived time-use data calculate them in two separate datasets from the raw absence types and severities and from the respective relative dates. This is done in seven stages for each dataset.

1. First, all the data for a single individual employee is read into an array and the size of that array is determined in the process.

2. Second, the program goes through the array of raw values and discovers and recodes each of the several types of systematically miscoded data values wherever they occur in the array.

3. Third, the data are scanned to detect any scheduled days off intervening within longer span time-use occurrences. The "second" portion of the respective occurrence long-span time-use is then flagged to allow calculations of frequency to be corrected (return from days off is not counted as a second occurrence). This "flag" is a value that tells the computer how many days back to go looking for the "first" portion whenever it encounters a flagged "second" portion. Instead of incrementing the counter by one for the second portion of the occurrence, the program just works around the days off for treatment of that specific time-use.

4. Long term absences for reasons of termination are checked to see if the employee was afterwards rehired. Rehires within 365 days or less are allocated to the long-term absence time-use category. Military leave and long-term sickness are similarly treated. For fired and rehired employees, the first 30 days are recoded as punishment time, the remainder as long-term legitimate absence unless they are off for over one year. Employees coded as deceased and later as rehired are recoded for that interval as missing values.

5. Once all the raw absence codes have been checked and recoded or cleaned as necessary, the program sorts and aggregates occurrences from the original 45 different absence-types into 13 time-use categories. Seven of these categories are intermediate time-use categories, and when clustered together in groups of one, two, or three, these in turn contribute their effect to create the six final or overall, outside-level time-use categories. The pattern of this aggregation and reasons for each decision are both indicated in Figure 11, and the 13 derived time-use indices are listed in Figure 13.

6. Two different computer runs are then executed. For each time-use index in each day of the study period for each individual employee, two sets of three derived values (100-day frequency, 100-day severity, mean 100-day severity,

and annual frequency, annual severity, annual mean severity) are then generated as the derived values in two arrays of three types of measures across the 13 time-use indices for the respective employee. Estimates for the two sets of three types of measures, on all 13 indices, were generated on a daily basis for every day of the program. Numeric codes for the 13 time use indices are listed below in Figure 13, and these correspond to the numbers on boxes in Figure 11.

Code for Time Use	Time Use Category/Index
10	All time worked on or off the job
11	Time worked on the job
12	Time worked off the job
20	All scheduled time off
22	Time off due to punishment
23	Routine normal time scheduled off or strike
21	Missing value or unscheduled day
30	Time absent for all reasons
31	Short-term legitimate absence
32	Long-term legitimate absence
33	Short-term non-legitimate absence (except late)
34	Late time
40	Unended termination time (no rehire)

Note: All codes are 2-digit integer values.

Figure 13: Derived Data Types

7. For each of the two sets of three measures, all 13 time-use indices for every day are written onto tape in a

comprehensive array which constitutes the track record of time-use for that employee throughout the employee's participation in the study. These two arrays are created daily, a piece at a time, as the new values are derived. When the calculations are completed, each entire array of daily values for the employee is written out onto tape in daily records 219 characters long. The two derived arrays are described in Figure 14 and Figure 15 below. The selective use of these individual derived data values to generate statistics for each subgroup in later methodological and substantive analyses is described in Chapter IV. Selection of subsample groups of cases from the overall 2336 cases in the DOT database is done automatically during derivation of the time-use categories. An early step in the process is a fortran subroutine that rejects all cases whose case numbers don't match those in the respective subgroup demographics.

Variable	Description
Case Number	A 4-digit code linking the individuals' derived statistics for time usage with the demographics tape. Used for sampling purposes.
Relative Date	A 4-digit number indicating number of days from the starting date of the program. Range is from 182 to about 1686.
Absence Type	A 2-digit code for the raw absence type used to sort the data in creating the derived statistics.
Absence Severity	A 1-digit code for the number of hours consumed in the time usage indicated by the absence type above on a given day.
Day Type A	A 2-digit code (see Figure __) for the highest level of derived category into which the sorted absence types above become categorized.
Day Type B	A 2-digit code (see Figure __) for the alternate data type on a given day. For example: if Day Type A is "late" then Day Type B would be "worked time".
Day Type 1	A 2-digit code (see Figure __) for am intermediate data type also derived by categorically sorting the absence types above according to the logical structure noted in Figure __.
Day Type 2	A 2-digit code which alternates for Day Type 1 ... like Day type B.

Figure 14: The Individual "Derived" 100-Day Data Types

Continued on next page

Figure 14 Continued

Variable	Description
Day Severity A Day Severity B Day Severity 1 Day Severity 2	1-digit codes for the number of hours spent in the respective time usage on a given day.
100-Day Frequency 10 100-Day Frequency 11 100-Day Frequency 12 100-Day Frequency 20 100-Day Frequency 21 100-Day Frequency 22 100-Day Frequency 23 100-Day Frequency 30 100-Day Frequency 31 100-Day Frequency 32 100-Day Frequency 33 100-Day Frequency 34 100-Day Frequency 40	4-digit values for the frequency of occurrence derived for each respective form (see Figure __ and Figure __) of time usage in the past 100 days, as calculated for the specific day indicated by relative date.
100-Day Severity 10 100-Day Severity 11 100-Day Severity 12 100-Day Severity 20 100-Day Severity 21 100-Day Severity 22 100-Day Severity 23 100-Day Severity 30 100-Day Severity 31 100-Day Severity 32 100-Day Severity 33 100-Day Severity 34 100-Day Severity 40	4-digit integer values for the cumulative severity of each respective form (see Figure __ and Figure __) of time usage in the past 100 days, as calculated for the specific day indicated by relative date.
100-Day Mean Severity 10 100-Day Mean Severity 11 100-Day Mean Severity 12 100-Day Mean Severity 20 100-Day Mean Severity 21 100-Day Mean Severity 22 100-Day Mean Severity 23 100-Day Mean Severity 30 100-Day Mean Severity 31 100-Day Mean Severity 32 100-Day Mean Severity 33 100-Day Severity 34 100-Day Severity 40	6-digit (plus a 7th for decimal point) real values calculated for each respective type of time usage each day by dividing 100-Day Severity by annual frequency. There are 2 digits to the right of the decimal place.

Variable	Description
Case Number	A 4-digit code linking the individuals' derived statistics for time usage with the demographics tape. Used for sampling purposes.
Relative Date	A 4-digit number indicating number of days from the starting date of the program. Range is from 182 to about 1686.
Absence Type	A 2-digit code for the raw absence type used to sort the data in creating the derived statistics.
Absence Severity	A 1-digit code for the number of hours consumed in the time usage indicated by the absence type above on a given day.
Day Type A	A 2-digit code (see Figure __) for the highest level of derived category into which the sorted absence types above become categorized.
Day Type B	A 2-digit code (see Figure __) for the alternate data type on a given day. For example: if Day Type A is "late" then Day Type B would be "worked time".
Day Type 1	A 2-digit code (see Figure __) for am intermediate data type also derived by categorically sorting the absence types above according to the logical structure noted in Figure __.
Day Type 2	A 2-digit code which alternates for Day Type 1 ... like Day type B.

Figure 15: The Individual "Derived" Annual Data Types

Continued on next page

Figure 15 Continued

Variable	Description
Day Severity A Day Severity B Day Severity 1 Day Severity 2	1-digit codes for the number of hours spent in the respective time usage on a given day.
Annual Frequency 10 Annual Frequency 11 Annual Frequency 12 Annual Frequency 20 Annual Frequency 21 Annual Frequency 22 Annual Frequency 23 Annual Frequency 30 Annual Frequency 31 Annual Frequency 32 Annual Frequency 33 Annual Frequency 34 Annual Frequency 40	4-digit values for the frequency of occurrence derived for each respective form (see Figure __ and Figure __) of time usage in the past 365 days, as calculated for the specific day indicated by relative date.
Annual Severity 10 Annual Severity 11 Annual Severity 12 Annual Severity 20 Annual Severity 21 Annual Severity 22 Annual Severity 23 Annual Severity 30 Annual Severity 31 Annual Severity 32 Annual Severity 33 Annual Severity 34 Annual Severity 40	4-digit integer values for the cumulative severity of each respective form (see Figure __ and Figure __) of time usage in the past 365 days, as calculated for the specific day indicated by relative date.
Ann. Mean Severity 10 Ann. Mean Severity 11 Ann. Mean Severity 12 Ann. Mean Severity 20 Ann. Mean Severity 21 Ann. Mean Severity 22 Ann. Mean Severity 23 Ann. Mean Severity 30 Ann. Mean Severity 31 Ann. Mean Severity 32 Ann. Mean Severity 33 Ann. Mean Severity 34 Ann. Mean Severity 40	6-digit (plus a 7th for decimal point) real values calculated for each respective type of time usage each day by dividing Annual Severity by annual frequency. There are 2 digits to the right of the decimal place. These digits are significant due to the way raw values for absence frequency and severity were recorded.

APPENDIX C

METHODOLOGICAL NOTES

Appendix C

Methodological Notes

(A): Procedures for Assessing Normality and Distributional Form

Various statistics offer insight into the relative normality of data distributions. In the use of inferential statistics, a tacit assumption is made that there is a consistency between the shape of empirical data distribution and the shape of some theoretical distribution ... such as the normal curve. The merit of T-tests and like procedures depends upon the correctness of that assumption. In the present work, T-tests on both the 100-day and 365-day data are used to examine substitution and long-term effects by comparing daily cross-sections. The assumption of an underlying normal curve is accordingly worthy of verification, especially since the absenteeism literature has reported a history of asymetric data distributions in cross sectional studies of employee absence behavior.

Various EDF (empirical data function) statistics have been developed to assess departure from normality and to assess the direction of its impact. For a first approximation, conventional estimates for the mean, median and mode, when compared, may be quite different, and this difference would indicate departure from normality. The SAS software which was used here to generate EDF statistics also calculates customary measures for skewness and kurtosis. However, SAS also offers two other statistics that provide a complimentary set of overall indices of normality: the Shapiro-Wilk W statistic and the Kolomogorov D statistic. Each of these two statistics is a number created by making a ratio of a theoretical variance over the observed variance for the data, assuming a theoretical distribution that is uniform and normal with the mean and standard deviation observed in the empirical sample.

Values for the W and D statistics range between 0 and 1 with smaller values indicating less normality. Unfortunately, interpreting the values of D is not yet a very convenient procedure. For sample sizes as large as 126 cases (the smallest case T-tested here), a D-statistic of less than about 0.8 would indicate a significant departure from normality. The degree to which this departure affects the meaning of the T-test statistic will tend to vary with the type of departure from normality (kurtosis, skewness, bimodality, etc.). Tables for comparison of these D values with the results of monte-carlo studies (M.A. Stephens, 1974) can be used to estimate the probability that the population distribution underlying the observed distribution departs from normality with a .05 significance level. Low probabilities mean a low probability of drawing a sample with a better W or D-statistic from the respective distribution. The W statistic is a theoretically better estimate than D, and can be used for samples of 50 observation or less, while the D statistic is still calculable and interpretable for larger samples where W becomes inconvenient.

In use, the normality indices become helpful in two ways. (1) Low values constitute a warning not to depend heavily on inferential statistics (like T-tests) from the relevant dataset. (2) Increasingly higher values may be generated by trial and error, transforming the data toward symmetry in successive approximations. Given enough computer time, this may be continued until a transformation is discovered that most effectively normalizes the distribution. However, in the present research situation there are limitations on time and problems associated with interpreting the transformed data relative to its external reference frame. The present research will not go beyond assessing normality of the untransformed data to check the degree to which assumptions of normality underlying the T-tests may be correct. The data for each date involved in direct cross-sectional comparisons were accordingly subjected to SAS procedure UNIVARIATE and W or D statistics for each date, plus the associated probability of getting a higher test statistic (significance level), were generated as appropriate. Selected results for these are listed with other EDF statistics in Table 19, Table 20 and Table 21.

While it might be feasible to create a set of meaningful comparison data for a Monte-Carlo type study of normality for a distribution of the means used in the correlations, the computer resources and time this requires are not available at present. It should be noted, however, that since the standard error of the means is by construction = 1/n (the standard deviation of the raw data), the means used in time-series correlations should be substantially less affected than the individual data by extreme values and, as group means they will rarely have zero values so that the truncation problem is bypassed by construction.

To provide a frame of reference with more traditional measures of the empirical data function and insight into the nature of departures from normality, the respective means, medians, ranges, skewness, and kurtosis are also reported in Tables 19 through 21.

(B): A Methodological Note on the Time-Series Correlation Procedures

Two sets of correlations were calculated using the SAS "CORR" procedure. The first set of these were done to establish the split-half reliability of the 39 derived time-series measures (time-use indices) for three sets of data. The two sets of daily means for corresponding indices for both pre-program and post-installation data on the 100-day data, and for data from day 1247 onward (May 21, 1981) for the 365-day data, were correlated between the split-halves over time generating three sets of 39 time-series correlations.

The second set of time-series correlations was done to examine the substantive relationships between changes in the various types of employee time-use behaviors. These internal relations between time-series of means for the two overall datasets (164 cases for each measure involved) were also compared. Frequencies, severities and mean severities for each of the two datasets were analyzed separately, creating essentially six sets of time-series correlations comparing the 13 different time-use variables. For each set of 13 time-use indices, a 13-by-13 matrix of time-series correlations was created. By time-series correlation between selected pairs of the 13 different indices across selected intervals of time, both intra-annual and inter-annual evidence of progression effects, substitution effects, and general shifts between work and absence were examined.

Introduction to the Time-Series Correlation Procedure

Split-Half Reliabilities: Traditional procedures to assess the reliability of absence measures have been complicated by contamination from the unreliability of absence behavior itself. Prior work has characteristically compared a cross-section of data from a group of employees at one point in time with another cross-section of data from the same group of employees at some other time. The fundamental procedure used in the present work does not compare cross-sections at all. Instead, a mean for each index is calculated for each day across all employees in each split-half of the overall sample. This is done for every day of the program and the exact same procedure is applied to the other split-half. This generates two series of daily means for each time-use index for every day of the program (one for each split half), and these two subsets of means ought to be very nearly congruent if the two subsamples are each approximately representative of the overall sample. If these two sets of means are then correlated, day 1 with day 1, day 2 paired with day 2, etc., the result containes two basic sources of incongruence: error of measurement, and any variation attributable to unintended sampling effects of group composition. As sample size increases, such sampling effects should become tiny by comparison with measurement error as the law of large numbers enhances representativity of the means.

This incongruence or unreliability should then appear as a reduction for the value of the correlation. Perfectly reliable measures should generate a split-half reliability of 1. Lower values indicate less reliability. Any value of 0.7 to 0.8 or larger is particularly worthy of note. Traditional absence research has very seldom produced reliability correlations greater than 0.35 for severity data and 0.65 for frequency data. It is possible that by using the time-series reliability approach such findings could be raised for the traditional studies even with traditional measures. On the other hand, this would require a choice to sacrifice the insights possible from the detailed time-use data. While the conventional measures would probably be reliable if appropriately tested, the question which is of bottom line interest is whether they would be informative. Without the level of detail in the data toward which the careful observation and recategorization procedures can move the measurement process, such detail seems doubtful at the stage of analysis.

It was necessary to prevent tautological inflation of the reliabilities due to variation while the moving values were still aggregating from a base period of 1 day up to the full 100-day or 365-day starting point. For this reason, the calculation of time-series split-half reliability correlations excludes data from "early" periods in which moving values for these measures had not yet aggregated across a period equal to the full number of days for which the measures were defined. Since the correlation is defined as covariance divided by the product of the standard deviations, any similarity between the overall absolute size of the means across the two samples would not affect the size of the correlations, while day-to-day changes can and will.

Substantive Correlational Tests: The same basic procedure that was used for split-half reliabilities can be applied to compare changes in two different types of time-use over time. In this way, for example, short-term non-legitimate absence can be correlated with long-term legitimate absence or with routine time worked to look for long-term program impacts. All that is needful is to calculate daily means for each index of the two measures each day of the program and then correlate the sets of means over time.

Underlying the calculation of time-series reliabilities was a simple plan to create daily summary statistics and compare those instead of comparing the unsummarized raw data directly. The resulting correlations are not nearly as sensitive to two fundamental problems that have plagued conventional cross-sectional reliabilities.

Conventional cross-sectional correlations compare two sets of values from the same individual's behavior at two points in time. Differences in those values between time 1 and time 2 can occur either (1) because of a fundamental unreliability in the measurement procedure, or (2) because the observed behavior of those individuals itself tends to change over time. Absence levels for individuals do change over time, however, which means that prior absence research has in fact found low reliabilities because the two series of values compared weren't from comparable samples of observed behavior, even though they were in all probability collected using the same measurement procedure. It might even be as close to the truth to call those findings estimates of the reliability of individual absence behavior than estimates of reliability for the measures used.

To overcome this limitation in the conventional paradigm, the time-series correlation procedure had to begin with two series of values that were: (a) representative of the absence behavior of all the workers over time, and (b) series from comparable samples of behavior over time. The second problem was resolved first -- by splitting the overall random sample of DOT employees into two subsamples of randomly selected individuals. The time-series of values for those two subsample groups should then have been comparable samples of behavior over time. Each series, however, was still made up of values from many employees each day and to compared these values as correlated series, a single value to represent the values for all employees each day had to be found for every day in each of the two series. The group mean value was selected. Then these daily group means for each series were sorted into ascending order by date to create two time-series of comparable group means. Those two time-series were then correlated by date. The procedure was followed for each of the three indices on each measure for each of the three datasets tested.

Two theoretically comparable series were thereby established for every index in each of the datasets. Calculation of reliability for the time-use measures over time should have been straightforward. Thee empirical results, however, showed consistently high reliabilities for the majority of the measures and erratically low reliabilities for the others (see Table 31).

Such anomalies could have resulted from the research procedure if any one or more of the following conditions occurred.

1. The workforce or the sample drawn from it could be too small to provide an adequate sample of time-use behavior for study or to properly represent the occurrence of behavior for the workforce over time. In effect, this would mean that one or both series of means came from too small a group of employees and would be overly affected by the behavior of any individual employee, a behavior past reserch has shown to be generally unreliable. The result would be low reliability estimates for all time-uses.
2. Most of the employees may have spent very little or no time in the respective time-use behavior so that for them the time-use value was at or very close to zero most of the time. A mean would then still be calculable for the entire sample of employees, but only the idiosyncratic time-use behavior of relatively few employees

would have contributed to the time-series of means underlying the estimates of variance and covariance for each split half. Given a small enough number of contributing employees, the pattern of means in the time-series would then start to become as idiosyncratically variable as the time use behavior of the individual employee, while this variation in the total amount of the time use would then be divided by the number of employees in the total sample in calculating the mean level of time use for the entire group.

If these conditions occurred, it could cause anomalous reliabilities in two ways. First, the idiosyncratic pattern of means in one split half could then be quite different than that in the other split half. To the extent that this occurred, the split half reliability would have been lower. Second, if the total amount of time-use in a given category were to be small enough, then after calculation of the group mean for each day, the level of measurement error or even of rounding error in calculation could be large enough to mask this real variation in the time use. This, too, would dilute the actual observed pattern and cause low estimates of reliability. In effect, the cause of such low reliabilities boils down to another result of too small a smaple of the actual time-use behavior.

3. The amount of variation in either one or both series of means might be so low that the standard deviation of one or both series was computed as zero. Such a zero variation for some time-use could occur either because the extremely regular occurrence of that time-use behavior so that the value of the mean never varied, or because there was none of the behavior occurring in one or both of the split halves. In either situation, the procedure to calculate a correlation would require division by zero, an operation that has no defined meaning. The SAS software used in the present research will simply report a zero value for the correlation in such a case.
4. In a less extreme but potentially more confusing case, the sample of variation observed in the mean level of a time-use for the group could be quite tiny but non-zero. This would occur if very little of the behavior occurred at all, or if it was regulated to an extreme degree so that the level of the mean scarcely varied. Contrary to the result in condition 3 above, this would result in a calculation of the correlation as a ratio of two very small non-zero values, the covariance divided by the product of standard deviations. Those very small values would be distorted severely by surrounding measurement error and their covariance could very well even be negative, tiny, and quite definitely meaningless but still negative. Negative values for the covariance would

result in a negative estimate for the reliability correlation.

5. There may be a need to subdivide overall categories of time-use in order to discover patterns in the more detailed but less inclusive time-uses that are contributory to the overall category. This also subdivides the sample of time-use behavior in the overall category into smaller samples of behavior in each of the more detailed categories. This, in effect, can bring into play one or more of conditions 2, 3 or 4 above and induce anomalous reliability estimates.
6. The length of the time-series across time could be too short so that the number of means in each series was too small to capture the day-to-day variation in group-level time-use and too small to produce a reliable correlation.
7. The length of the observation period used to accumulate the moving values can also cause the reliability correlations to be larger or smaller. On any given day the time-use of an individual employee may vary from one category to another. If the period of observation were limited to one day only, this would result in zero values for every time-use category the employee did not use on that given day. Quite apart from giving no indication of time-use rates across longer periods of special interest (such as annual absence), using these one day

values would also cause the truncation effect in the data distribution for each day (Hammer and Landau, 1981) to become very severe for all of the time-use measures. In fact, the more detailed the number of time-use categories become, the more severely the truncation effect would distort the values of the daily group means for every time-use. Instead of permitting the effect of daily differences to be smoothly integrated with daily fluctuations from other employees by averaging the moving values over time, the direct use of such daily fluctuations to calculate the means would transfer the truncation effect into the series of means and make the means more easily affected by sporadic individual employee behavior. In effect, using a period of time longer than one day initiates a moving average procedure that is later completed by calculating the daily group means. The series of means is then less affected by idiosyncratic local fluctuations from individual employees, although sensitivity to daily fluctuations at a group level is preserved.

The choice of a length for the time period of observation was made as 100 days and 365 days for the present research. A seven day period and 30-day period were also considered because the business use of weekly and monthly scheduling periods would make them convenient for program evaluation. The longer time spans chosen here are equivalent to the length of the quarterly planning period and the business year.

8. Another time related condition can also affect the variance in the series of means and result in anomalous reliability values. As the length of the time period of observation approaches 365 days, recurring cyclical effects that are locked into an annual cycle will tend to make the annual moving values settle around a total annual value which from then on would only show variation from year to year. In effect, this would mean that the 365 day measures, which by their construction damp out intra-annual cycles, may not retain enough of a sample of inter-annual variation to produce a meaningful reliability estimate. Because annual absence level and long-term inter-annual changes are of particular interest in the present work, the 365-day interval was selected anyway on a gamble that the other conditions listed above would not combine to make the sample size too small and the measures unreliable.

Finally, it should be noted that the split-half procedure itself requires splitting the overall sample in half prior to any other steps. Measurement error for the original sample can be shown to approximate that of the split half multiplied by the square root of two. In effect, the split-half reliability is a value that relates to a representative half of the overall sample. This means that while a high split-half reliability indicates reliability for the measure in the overall sample (where sample size would be twice as large), a low value merely indicates that no such reliability has been demonstrated with the smaller sample size.

APPENDIX D

A REVIEW AND DISCUSSION OF THE DETAILED RESULTS

Appendix D

A Review and Discussion of the Detailed Results

The results presented here and in Chapter V are both organized to parallel the discussion of methodology presented in Chapter IV and Appendix II. Sections A, B, C, D, and E in the methodology described the procedures which have produced corresponding Sections A, B, C, D, and E in the results. See Chapter IV, Figure 6 for an outline of these procedures.

To minimize the number of tables necessary (and surrounding explanatory material), results for all hypotheses which rely upon a specific analytical procedure are presented together, either in a single table or in a cluster of consecutive tables or time-series graphs. Where results are extremely repetitive, only a selective summary was reported here. Procedures for deriving each one of the three forms of each of the 13 time-use indices have been presented in Appendix III. A brief description and guidelines for interpreting the various results and summary statistics presented in the tables are described briefly preceding each table or set of tables. Following each respective block of tables, results are discussed and interpreted relative to the specific related hypotheses presented in Chapter IV. The extent of support for each hypothesis is assessed and discussed briefly in the light of other co-occurrent findings.

Results of the (A) normality tests and statistics describing distributional form are presented first; followed by (B) the T-tests; (C) the time-series split-half reliability correlations; (D) the time-series correlational tests of substitution effects; and (E) the graphical time-series analysis.

Note: For all results presented here, no findings on indexes for termination time have been reported because in the randomly selected sample of 164 cases, which is the basis of the present report, no occurrences of termination were reported in the raw data.

A: Normality and EDF Statistics for the T-Test Data

The selected EDF statistics described below are tabulated here to illustrate the distributional forms typical of the T-test data. Since there was an overwhelming repetitiveness in these EDF statistics, only sample statistics for the three overall indices of All-Time-Worked, All-Time-Absent, and All-Time-Scheduled-Off have been reported here, and only for the dates involved in T-tests of year-to-year overall change. Those particular dates were selected because they provided a basis for comparing the 100-day and 365-day data distributions. EDF statistics for the 30-day, 100-day and 365-day datasets respectively have been presented in Table 19, Table 20 and Table 21.

Date:

Refers to the relative date to which the daily cross-section of T-test data pertain. Day 182 means July 1, 1978, day 183 means July 2, 1978, etc.

N:

Refers to the number of cases for which data existed on the respective day. Variation in N is indicative of the extent of turnover across the time duration of this study.

Mean and Med:

Refer to the group mean value of the selected time-use index on a particular day, or to the group median value, respectively. The size and direction of differences between mean and median are crude but useful indicators of the direction and magnitude for departures from normality. Where the mean is greater than the median, the distribution will generally be positively skewed.

Std. Dev.:

Refers to the standard deviation for data in a particular daily cross-section.

Skew:

Refers to the skewness of the data distribution, ideally, in a normal distribution, skew = 0. Any negative skewness in time-use data will tend to be rather small because such a result is limited by minimum values at zero. On the other hand, positive skewness in time-use data can get quite large, depending on the number and size of unusually large outliers above the mean.

Kurtosis:

The "peakedness" of the distribution relative to a normal distribution, ideally = 0. Negative values mean a flat distribution, positive values mean a high modal concentration.

D:

Refers to the Kolmogorov "D" statistic described in Chapter IV. The D statistic varies between 0 and 1, with lower values indicating less normal distribution. In a normal distribution, D = 1.0.

Prob > D:

Means "the probability of this distribution producing a higher D statistic." The D and Prob > D statistics occur as pairs.

Table 19: Normality and EDF Statistics for Selected 30-Day T-Test Data Distributions

Index	Date	N	Mean	Med.	Std. Dev.	Skew	Kur- tosis	Prob. D	> D
				ALL TIME WORKED					
30-Day	211	129	1.64	2.0	1.14	0.61	0.68	0.19	<0.01
Frequency	912	151	1.19	1.0	1.05	0.90	0.98	0.23	<0.01
30-Day	211	129	103.30	120.0	58.12	-0.69	-0.84	0.15	<0.01
Severity	912	151	125.58	152.0	61.59	-1.10	-0.02	0.19	<0.01
30-Day Mean	211	129	67.25	56.0	53.31	0.65	-0.60	0.16	<0.01
Sev./Occur.	912	151	73.77	74.5	65.06	0.38	-1.14	0.17	<0.01
				ALL TIME ABSENT					
30-Day	211	129	1.23	1.0	1.07	0.95	1.28	0.23	<0.01
Frequency	912	151	1.08	1.0	0.96	1.09	1.76	0.28	<0.01
30-Day	211	129	31.73	16.0	45.96	2.63	8.14	0.24	<0.01
Severity	912	151	43.78	20.0	61.31	1.90	2.95	0.24	<0.01
30-Day Mean	211	129	22.50	8.0	41.03	3.85	16.69	0.29	<0.01
Sev./Occur.	912	151	36.59	8.0	60.20	2.25	4.27	0.29	<0.01
				ALL TIME SCHEDULED OFF					
30-Day	211	129	4.32	5.0	2.11	-0.30	-0.33	0.26	<0.01
Frequency	912	151	4.49	5.0	1.50	-1.21	3.14	0.32	<0.01
30-Day	211	129	103.61	80.0	60.25	1.42	0.95	0.30	<0.01
Severity	912	151	69.48	64.0	41.83	2.89	10.40	0.31	<0.01
30-Day Mean	211	129	52.02	16.0	78.11	1.82	1.59	0.43	<0.01
Sev./Occur.	912	151	23.24	14.4	47.85	4.31	16.98	0.48	<0.01

Table 20: Normality and EDF Statistics for Selected 100-Day T-Test Data Distributions

Index	Date	N	Mean	Med.	Std. Dev.	Skew	Kur-tosis	D	Pro > D
				ALL TIME WORKED					
100-Day Frequency	546	126	3.44	3.0	2.09	1.04	1.64	0.17	<0.0
	1277	151	1.74	2.0	1.51	1.07	2.55	0.17	<0.0
	1642	141	3.77	3.0	2.44	0.51	0.08	0.14	<0.0
100-Day Severity	546	126	419.30	477.0	156.33	-1.59	1.53	0.25	<0.0
	1277	151	281.83	336.0	151.65	-0.96	-0.44	0.20	<0.0
	1642	141	412.06	504.0	197.96	-1.13	-0.05	0.23	<0.0
100-Day Mean Sev./Occur.	546	126	144.95	124.0	108.28	1.94	5.45	0.16	<0.0
	1272	151	134.43	114.3	127.11	1.03	0.49	0.15	<0.0
	1642	141	130.59	103.2	122.12	2.12	5.53	0.17	<0.0
				ALL TIME ABSENT					
100-Day Frequency	546	126	2.92	3.0	1.93	1.33	2.45	0.18	<0.0
	1277	151	1.86	2.0	1.53	0.97	2.10	0.20	<0.0
	1642	141	3.33	3.0	2.29	0.68	0.09	0.16	<0.0
100-Day Severity	546	126	131.35	64.0	173.33	2.38	5.34	0.29	<0.0
	1277	151	88.77	40.0	143.65	2.51	5.69	0.28	<0.0
	1642	141	136.25	72.0	169.46	2.29	5.26	0.25	<0.0
100-Day Mean Sev./Occur.	546	126	75.75	21.3	158.64	3.36	11.26	0.36	<0.0
	1277	151	54.36	16.0	123.21	3.69	13.17	0.34	<0.0
	1642	141	59.24	22.7	135.67	4.43	20.07	0.33	<0.0
				ALL TIME SCHEDULED OFF					
100-Day Frequency	546	126	14.31	15.0	4.27	-0.50	5.26	0.31	<0.0
	1277	151	9.72	11.0	4.99	-0.22	0.21	0.20	<0.0
	1642	141	13.06	14.0	5.22	-0.78	2.55	0.33	<0.0
100-Day Severity	546	126	254.37	236.0	96.04	2.77	13.95	0.24	<0.0
	1277	151	433.97	400.0	154.71	1.54	1.71	0.28	<0.0
	1642	141	255.19	220.0	176.60	2.46	5.32	0.39	<0.0
100-Day Mean Sev./Occur.	546	126	26.50	16.0	73.30	9.96	105.58	0.43	<0.0
	1277	151	141.16	37.1	256.78	2.18	2.91	0.42	<0.0
	1642	141	74.45	15.4	205.26	3.29	9.11	0.50	<0.0

Table 21: Normality and EDF Statistics for Selected 365-Day T-Test Data Distributions

Index	Date	N	Mean	Med.	Std. Dev.	Skew	Kurtosis	D	Prob. > D
				ALL TIME WORKED					
Annual Frequency	546	126	11.17	11.0	5.44	0.53	0.35	0.12	<0.01
	1277	151	9.34	9.0	5.29	0.23	-0.23	0.08	<0.04
	1642	141	12.21	12.0	6.80	0.39	0.27	0.06	<0.15
Annual Severity	546	126	1383.33	1594.0	517.67	-0.95	-0.17	0.16	<0.01
	1277	151	1322.63	1583.0	591.38	-1.34	0.31	0.27	<0.01
	1642	141	1490.70	1727.0	591.42	-1.42	1.10	0.20	<0.01
Annual Sev./Occur.	546	126	151.20	128.0	108.38	2.68	9.79	0.19	<0.01
	1272	151	155.07	135.4	106.95	1.59	4.46	0.11	<0.01
	1642	141	147.33	122.1	127.49	3.30	17.74	0.13	<0.01
				ALL TIME ABSENT					
Annual Frequency	546	126	10.76	10.5	5.50	0.49	0.24	0.09	<0.02
	1277	151	8.92	9.0	5.11	0.37	0.16	0.07	<0.04
	1642	141	3.33	3.0	2.29	0.68	0.09	0.16	<0.01
Annual Severity	546	126	396.29	264.0	459.44	3.08	11.04	0.26	<0.01
	1277	151	431.03	240.0	560.3	2.78	7.54	0.25	<0.01
	1642	141	506.01	289.0	564.40	2.70	7.89	0.24	<0.01
Annual Sev./Occur.	546	126	52.91	24.3	141.63	8.28	78.74	0.36	<0.01
	1277	151	103.63	27.4	388.48	6.14	41.40	0.42	<0.01
	1642	141	59.24	22.7	135.67	4.43	20.07	0.33	<0.01
				ALL TIME SCHEDULED OFF					
Annual Frequency	546	126	47.79	53.0	18.37	-0.01	1.21	0.22	<0.01
	1277	151	44.21	48.0	18.61	-0.98	0.95	0.27	<0.01
	1642	141	48.18	52.0	16.45	-1.14	3.57	0.24	<0.01
Annual Severity	546	126	1130.93	918.0	501.61	1.44	1.73	0.27	<0.01
	1277	151	1160.86	996.0	588.36	1.96	3.23	0.32	<0.01
	1642	141	915.97	835.0	510.73	2.52	7.10	0.33	<0.01
Annual Sev./Occur.	546	126	40.09	16.6	70.25	4.90	28.14	0.33	<0.01
	1277	151	164.61	21.0	573.57	4.54	19.40	0.43	<0.01
	1642	141	92.32	15.8	387.79	6.23	41.31	0.46	<0.01

Summary of Results for the Normality and EDF Statistics

Between July 30, 1978, and June 30, 1980, there was a 25% increase in the size of the means for all time worked, a 30% increase in the mean for all time absent, and a corresponding 35% decrease in all time scheduled off. The D statistics are uniformly less than 0.5 and most are less than 0.4, indicating that the samples were non-normal, while the associated probabilities of the underlying distribution producing a higher D statistic were without exception less than 0.01 (D was highly significant). Skew and kurtosis were generally positive and a clear relationship exists between the sizes of the mean and standard deviation. This set of distributions is definitely non-normal and positively skewed. The relatively large standard deviations in the values of 30-day severity per occurrence (compared to the respective means) are a consequence of compounding two sources of variation in measurement. The severity per occurrence (mean severity) was calculated as a function of 30-day severity divided by 30-day frequency.

For the 100-day data, three sets of data were examined: June 30, 1979, June 30, 1981, and June 30, 1982. There was a 30% drop in the mean amount of time worked per hundred days from 1979 to 1981, followed by a return to the 1978 level by 1982. This was accompanied by a 30% reduction in all time absent from 1979 to 1981, which also returned to the 1979 level by 1982. These two fluctuations were possibly due to a 40% increase in the mean of time scheduled off between June 30, 1979, and June 30, 1981, which also settled back down to the 1979 level by June 30, 1982. The differences between 100-day and 30-day values suggest that in the last month of the pre-program period a lot of scheduled time off may have been formally granted. The D statistics for the 100-day data are uniformly less than 0.50 with associated probabilities of a higher D value that are all less than 0.01. Skew and kurtosis are similar to the 30-day data. The mean and standard deviation are strongly and positively correlated.

For the 365-day data, there is a fluctuation in the mean that follows the changes for corresponding dates in the short-term 100-day data, but is much less severe. There is also a generalized long-term 8% increase in the number of hours worked per year and a 20% increase in number of hours absent, with a corresponding 20% decrease in the number of hours scheduled off. Skewness for frequency of time scheduled off was small and negative, as was skewness for severity of all time worked. The D statistics are uniformly less than 0.50 with accompanying probabilities for getting a higher D value of less than 0.01.

In general, these distributions are non-normal, positively skewed, and indicate some sizeable changes from year-to-year. Those differences have been T-tested despite the non-normality of their associated distributions and the results are reported below.

Hypothesis 1 is definitely not supported. The T-test results should be considered, but estimates of confidence levels associated with those results are not at all accurate. These results are entirely consistent with prior findings in the absenteeism literature.

B: Results for the T-Tests of Cross-Sectional Comparisons

T-tests were implemented on differences between the means of 17 different, specific daily time datasets to examine five sets of hypotheses (see page 57 and Figure 6). An explanation of the reasons for choosing these particular dates has been presented in Chapter IV. Results of the normality and EDF analyses for these data distributions indicate very serious non-normality, just as prior research would imply. Nonetheless, the T-test results are presented below. In most cases the variances of the two distributions were different enough (F-tested) that the T-test for distributions with unequal variance was used. This decision was made separately for each T-statistic, however, since the F-test values and probabilities were all available.

For the T-tests, the table or tables of statistics associated with each hypothesis have been presented, one hypothesis at a time, followed immediately by a brief interpretation of the results and a discussion of support for the hypothesis. Values presented in the T-test tables include:

- **Var/Index:** the variable and time-use index tested
- **Date 1:** the first of the two dates for which the T-test was done
- **Mean 1:** the mean value of the index variable on Date 1
- **Std.Dev. 1:** the standard deviation for the data distribution at Date 1
- **Date 2:** the second of the two relative dates for which the T-test was done
- **Mean 2:** the mean value of the index variable on Date 2
- **Std.Dev. 2:** the standard deviation for the data distribution at Date 2
- **F-Prob:** the probability of equal variance in the two data distributions
- **T:** the T-test statistic, in most cases calculated for unequal variance
- **T-Prob:** the probability that difference between the means is not significant

Table 22: T-Test Results of the Hypothesis that a Pre-Program Shift in Time-Use Behaviors Occurred - 30-Day Data

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	211			912					
Frequency		1.64	1.14		1.19	1.05	0.37	3.47	0.00
Severity		103.30	58.12		125.58	61.59	0.50	-3.11	0.00
Mean Sev.		67.25	53.31		73.77	65.06	0.02	-0.92	0.35
Routine Work	211			912					
Frequency		1.65	1.16		1.15	1.04	0.19	3.80	0.00
Severity		103.09	58.07		125.11	61.52	0.50	-3.01	0.00
Mean Sev.		67.60	54.01		72.91	65.16	0.03	-0.75	0.45
Other Work	211			912					
Frequency		0.03	0.17		0.03	0.27	0.00	-0.08	0.93
Severity		0.22	1.26		0.48	4.64	0.00	-0.66	0.51
Mean Sev.		0.22	1.25		0.23	1.77	0.00	-0.07	0.94
All Scheduled Off	211			912					
Frequency		4.31	2.11		4.49	1.49	0.00	-0.77	0.43
Severity		103.60	60.25		69.48	41.83	0.00	5.14	0.00
Mean Sev.		52.02	78.11		23.24	47.85	0.00	3.64	0.00
Missing Values	211			912					
Frequency		0.26	0.50		0.09	0.29	0.00	3.24	0.00
Severity		43.53	86.20		10.65	46.34	0.00	3.88	0.00
Mean Sev.		40.50	82.19		10.65	46.34	0.00	3.66	0.00
Punishment Time	211			912					
Frequency		0.05	0.34		0.06	0.24	0.00	-0.15	0.88
Severity		0.81	5.18		2.60	20.31	0.00	-1.04	0.29
Mean Sev.		0.44	2.61		2.60	20.31	0.00	-1.28	0.19
Routine Off	211			912					
Frequency		4.10	2.45		4.39	1.61	0.00	-1.15	0.25
Severity		59.26	34.29		56.23	21.71	0.00	0.87	0.38
Mean Sev.		12.02	6.26		11.84	4.46	0.00	0.26	0.79
All Time Absent	211			912					
Frequency		1.23	1.07		1.08	0.96	0.18	1.25	0.21
Severity		31.73	45.96		43.78	61.31	0.00	-1.88	0.06
Mean Sev.		22.50	41.03		36.59	60.21	0.00	-2.32	0.02
Sht. Leg. Absence	211			912					
Frequency		0.89	0.97		0.66	0.86	0.17	2.07	0.03
Severity		22.36	35.17		20.30	33.51	0.57	0.50	0.61
Mean Sev.		16.94	31.33		17.86	32.23	0.74	-0.24	0.80
Long Leg. Absence	211			912					
Frequency		0.02	0.12		0.11	0.32	0.00	-3.46	0.00
Severity		3.72	29.77		18.75	58.13	0.00	-2.78	0.00
Mean Sev.		3.72	29.77		18.75	58.13	0.00	-2.78	0.00
Sht. Non-Leg. Ab.	211			912					
Frequency		0.38	0.66		0.36	0.70	0.58	0.27	0.78
Severity		5.64	13.63		4.72	10.91	0.01	0.62	0.53
Mean Sev.		4.20	9.14		3.47	7.86	0.08	0.71	0.47
Late Time	211			912					
Frequency		0.14	0.54		0.15	0.48	0.17	0.10	0.92
Severity		0.53	2.42		0.43	1.91	0.01	0.37	0.71
Mean Sev.		0.33	1.48		0.32	1.30	0.13	0.04	0.97

B-1: T-Tests on Overall Pre-Program Shifts

Contrary to the prediction in Hypothesis 2, some apparently highly significant shifts in time-use behavior were recorded between the start of the baseline year and the month prior to program commencement. It should be emphasized once again that these data, although randomly selected, constitute less than a 7% sample, and were in each case collected across only 30 days of observation. D-statistics indicate a serious non-normality in the respective distributions. Accordingly, the statistics reported here are extremely tenuous. It should also be noted that for time-use reported in Other Work, Missing Values, Punishment Time, and Late Time, the index values are based upon scarce behavior from a much smaller sample of employees, so that results for those time-use variables are even more questionable.

The severity and frequency indices for all time worked both changed by about 25%, but in reverse direction, indicating that more time was worked with fewer interuptions, mean "severity" of time worked was up from about 67 to 73 hours per occurrance. The same was true for routine time worked which is the largest single contributing element in all time worked. Significance levels of 0.002 were reported. At the same time, scheduled time-off dropped to 69 hours from 104 hours per worker, with a frequency that did not change much, and a drop in mean severity from 52 hours per occurrence to 23 hours. Absence levels increased from 31 hours per employee in 30 days to 43 hours per employee. The improvement in quality of data recorded is evident in the number of hours of missing values, which dropped from 43 hours to 10 hours per employee.

Hypothesis 2 is definitely supported. There was a significant increase in both time worked (.002) and time absent (.06), with a corresponding significant decrease in time scheduled off (.0001). It is also notable that **frequency** indices and **severity** indices do not always vary, either in the same directions or by corresponding amounts. Indeed, they do have some apparent tendency to do exactly the reverse for these 30-day measures, a possibility that will be investigated and discussed later, based on the time-series correlations and the time-series plots. While this tendency would imply a need to stop using frequency and severity as alternative but parallel measures for the same "absence," the relationship is quite congruent with theories of substitution and progression effects.

There is a noteworthy internal consistency among the changes of severity values especially, one of which demonstrates the integral construction of the <u>set</u> of measures. There is a general "fund" of eight hours available in each employee's day, seven days a week. Since the combined measures <u>must</u> account for how all those hours get used, an interval total of hours per employee, about 240 hours per 30 days, provides a common base or account that all measures of time-use basically draw against. Because the moving values include rounding errors and are calculated against a variable number of employee "cases" due to turnover, the total of (all-time-worked) plus (all-time-scheduled-off) plus (all-time-absent) will tend to vary to some extent above and below this target level. But, a rough equivalence should exist and indeed it does, (total = 239 hours instead of 240).

Table 23: T-Test Results of Seasonal Shifts in the Pre-Program 100-Day Measures: October 20, 1978 to April 14, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	294			469					
Frequency		3.65	2.44		3.23	2.08	0.07	1.48	0.13
Severity		342.92	192.61		409.95	165.98	0.09	-3.00	0.00
Mean Sev.		103.83	102.19		156.40	136.01	0.00	-3.51	0.00
Routine Work	294			469					
Frequency		3.70	2.54		3.14	1.94	0.00	1.98	0.04
Severity		340.27	191.79		409.51	166.05	0.10	-3.10	0.00
Mean Sev.		101.53	96.80		157.66	135.75	0.00	-3.82	0.00
Other Work	294			469					
Frequency		0.19	1.25		0.05	0.25	0.00	1.30	0.19
Severity		2.65	16.43		0.43	2.66	0.00	1.51	0.13
Mean Sev.		0.99	7.24		0.37	2.37	0.00	0.92	0.35
All Scheduled Off	294			469					
Frequency		12.58	6.49		13.65	5.29	0.02	-1.45	0.14
Severity		355.48	203.25		287.87	139.84	0.00	3.11	0.00
Mean Sev.		98.66	206.04		48.59	138.99	0.00	2.29	0.02
Missing Values	294			469					
Frequency		0.67	1.23		0.40	0.76	0.00	0.21	0.03
Severity		175.69	287.29		90.23	190.11	0.00	2.81	0.00
Mean Sev.		115.12	225.31		71.03	170.35	0.00	1.77	0.07
Punishment Time	294			469					
Frequency		0.11	0.44		0.13	0.49	0.20	-0.40	0.68
Severity		2.05	9.02		2.67	9.40	0.64	-0.54	0.58
Mean Sev.		1.30	5.15		1.99	7.42	0.00	-0.87	0.38
Routine Off	294			469					
Frequency		12.12	6.99		13.45	5.46	0.01	-1.71	0.08
Severity		177.74	94.66		194.97	66.93	0.00	-1.69	0.09
Mean Sev.		13.92	4.86		14.15	3.70	0.00	-0.42	0.67
All Time Absent	294			469					
Frequency		3.32	2.28		2.87	2.16	0.52	1.63	0.10
Severity		106.61	126.05		107.23	154.33	0.02	-0.04	0.97
Mean Sev.		41.64	83.67		56.70	140.03	0.00	-1.05	0.29
Sht. Leg. Absence	294			469					
Frequency		2.51	1.94		1.77	1.64	0.06	3.33	0.00
Severity		67.45	83.39		52.60	97.88	0.07	1.31	0.19
Mean Sev.		30.99	70.53		33.28	91.39	0.00	0.22	0.82
Long Leg. Absence	294			469					
Frequency		0.03	0.21		0.02	0.12	0.00	0.71	0.47
Severity		9.61	79.78		8.56	75.57	0.54	0.11	0.91
Mean Sev.		6.51	51.63		8.56	75.57	0.00	-0.25	0.79
Sht. Non-Leg. Ab.	294			469					
Frequency		1.09	1.55		1.22	1.36	0.15	-0.68	0.49
Severity		29.55	64.91		46.08	109.13	0.00	-1.48	0.14
Mean Sev.		19.48	56.43		28.21	93.21	0.08	-0.91	0.36
Late Time	294			469					
Frequency		0.40	1.17		0.32	0.99	0.06	0.63	0.52
Severity		1.28	4.49		0.71	2.85	0.00	1.21	0.22
Mean Sev.		0.73	2.41		0.37	1.43	0.00	1.43	0.15

B-2a: T-Test Results on Pre-Program Seasonal Shifts

During the baseline year there was no evidence of any significant seasonal change in absence behavior whatsoever. But, highly significant changes occurred in both the amount of time worked (up from 342 hours per employee to 409 hours, T-Prob = 0.003), and in the amount of time scheduled off (down from 355 hours per employee to 287 hours, T-Prob = 0.002). These differences presumably reflect summer vacation time. It is interesting to note that short-term non-legitimate absence became a much more variable time-use behavior in the winter months than it was during the summer. Furthermore, there was a very large decrease (from 174 to 90 hours) in the amount of time accounted for by missing values across the period.

Hypothesis 3 is supported for time worked and for time scheduled off (changes at least partly controlled by the employer), but not supported at all for absence behavior which occurs at the discretion of the employee. This result is clearly different from conventional findings in absence research.

Once again, note that the EDF statistics show little evidence of normality in the distributions compared here, and that these T-test statistics should be considered with caution, especially since missing values account for about half as much employee time as does routine time worked.

Table 24: T-Test Results of Holiday Shifts in the 100-Day Pre-Program Time-Use Measures: January 5, 1979 to April 14, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	370			469					
Frequency		3.45	2.30		3.23	2.08	0.24	0.79	0.42
Severity		354.70	201.76		409.95	165.98	0.03	-2.40	0.01
Mean Sev.		113.94	107.86		156.40	136.01	0.01	-2.78	0.00
Routine Work	370			469					
Frequency		3.38	2.19		3.14	1.94	0.17	0.93	0.35
Severity		353.93	201.73		409.51	166.05	0.03	-2.42	0.01
Mean Sev.		115.53	111.97		157.66	135.75	0.03	-2.72	0.00
Other Work	370			469					
Frequency		0.09	0.53		0.05	0.25	0.00	0.75	0.45
Severity		0.78	4.42		4.34	2.65	0.00	0.75	0.45
Mean Sev.		0.30	1.80		0.37	2.37	0.00	-0.28	0.77
All Scheduled Off	370			469					
Frequency		12.63	6.68		13.65	5.29	0.01	-1.35	0.17
Severity		355.79	203.72		287.87	139.84	0.00	3.12	0.00
Mean Sev.		109.67	228.24		48.59	138.99	0.00	2.60	0.01
Missing Values	370			469					
Frequency		0.71	1.35		0.40	0.76	0.00	2.32	0.02
Severity		174.57	292.18		90.23	190.11	0.00	2.75	0.00
Mean Sev.		112.70	234.37		71.03	170.35	0.00	1.63	0.10
Punishment Time	370			469					
Frequency		0.09	0.34		0.13	0.49	0.00	-0.74	0.46
Severity		1.55	5.92		2.67	9.40	0.64	-1.14	0.25
Mean Sev.		1.33	4.98		1.99	7.42	0.00	-0.84	0.40
Routine Off	370			469					
Frequency		12.26	7.15		13.45	5.46	0.00	-1.50	0.13
Severity		177.67	100.00		194.97	66.93	0.00	-1.44	0.15
Mean Sev.		12.99	5.80		14.15	3.70	0.00	-1.91	0.05
All Time Absent	370			469					
Frequency		3.12	2.20		2.87	2.16	0.80	0.94	0.34
Severity		95.71	136.53		107.23	154.33	0.17	-0.64	0.52
Mean Sev.		48.81	127.11		56.70	140.03	0.27	-0.47	0.63
Sht. Leg. Absence	370			469					
Frequency		2.30	1.74		1.77	1.64	0.48	2.54	0.01
Severity		57.79	97.35		52.60	97.88	0.95	0.43	0.66
Mean Sev.		29.92	89.02		33.28	91.39	0.77	0.30	0.76
Long Leg. Absence	370			469					
Frequency		0.02	0.12		0.02	0.12	1.00	0.00	1.00
Severity		6.33	71.14		8.56	75.57	0.49	-0.24	0.80
Mean Sev.		6.33	71.14		8.56	75.57	0.49	-0.24	0.80
Sht. Non-Leg. Ab.	370			469					
Frequency		1.07	1.56		1.22	1.36	0.12	-0.81	0.42
Severity		31.59	85.68		46.08	109.13	0.01	-1.19	0.23
Mean Sev.		18.65	66.03		28.21	93.21	0.00	-0.95	0.34
Late Time	370			469					
Frequency		0.28	0.90		0.32	0.99	0.28	-0.33	0.74
Severity		0.82	4.22		0.71	2.85	0.00	0.24	0.80
Mean Sev.		0.53	3.67		0.37	1.43	0.00	0.47	0.64

B-2b: Results of the T-Tests on Pre-Program Holiday Shifts

The Christmas Holidays had no effect at all on long-term legitimate absence, nor any significant effect on any of the absence measures except the frequency of short-term legitimate absence, which decreased from the holiday to non-holiday period from 2.30 to 1.77, T-Prob = 0.01. It appears very likely that the non-significant decrease of five hours in the severity index for short-term legitimate absence, which also occurred at this time, was the result of many employees taking a legitimate casual-leave day at Christmas, but that except for this very minor fluctuation, the holiday had little effect on absence per se.

However, significant shifts did occur in time worked and in time scheduled off. Time worked went up from 354 hours per employee to 409 hours, T-Prob = 0.01, and time scheduled off went down by a corresponding amount from 355 hours to 287 hours per employee after the holiday period (T-Prob = 0.002). Apparently both the employees and the employer cooperated to arrange legitimate time off during the holiday period.

Hypothesis 4 is supported for shifts in time-use that were at least partly subject to control by the employer, but not for absence in general.

Table 25: T-Test Results for the Combined Seasonal and Holiday Effects on the 100-Day Pre-Program Measures: January 5, 1979 to June 15, 1979

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	370			531					
Frequency		3.45	2.30		3.40	2.20	0.60	0.17	0.86
Severity		354.70	201.76		423.01	159.02	0.01	-3.02	0.00
Mean Sev.		113.94	107.86		150.06	123.72	0.12	-2.50	0.01
Routine Work	370			531					
Frequency		3.38	2.19		3.24	2.03	0.40	0.53	0.59
Severity		353.93	201.73		422.51	158.70	0.01	-3.03	0.00
Mean Sev.		115.53	111.97		157.04	124.26	0.24	-2.82	0.00
Other Work	370			531					
Frequency		0.09	0.53		0.04	0.23	0.00	0.91	0.36
Severity		0.78	4.42		0.50	2.96	0.00	0.60	0.55
Mean Sev.		0.30	1.80		0.43	2.70	0.00	-0.48	0.63
All Scheduled Off	370			531					
Frequency		12.64	6.69		13.65	4.51	0.00	-1.43	0.15
Severity		355.79	203.72		255.44	104.66	0.00	4.98	0.00
Mean Sev.		109.67	228.24		28.59	76.95	0.00	3.82	0.00
Missing Values	370			531					
Frequency		0.71	1.35		0.23	0.54	0.00	3.75	0.00
Severity		174.57	292.18		48.25	131.87	0.00	4.47	0.00
Mean Sev.		112.70	234.37		39.94	112.14	0.00	3.18	0.00
Punishment Time	370			531					
Frequency		0.09	0.34		0.13	0.56	0.00	-0.67	0.50
Severity		1.55	5.92		2.36	10.79	0.00	-0.74	0.45
Mean Sev.		1.33	4.98		1.39	5.47	0.29	-0.09	0.93
Routine Off	370			531					
Frequency		12.26	7.15		13.57	4.61	0.00	-1.74	0.08
Severity		177.67	100.00		204.84	53.32	0.00	-2.52	0.01
Mean Sev.		12.99	5.80		15.15	3.03	0.00	-3.76	0.00
All Time Absent	370			531					
Frequency		3.12	2.20		2.85	2.07	0.48	1.02	0.30
Severity		95.71	136.53		126.46	174.34	0.01	-1.58	0.11
Mean Sev.		48.81	127.11		71.03	154.34	0.03	-1.26	0.20
Sht. Leg. Absence	370			531					
Frequency		2.30	1.74		1.63	1.69	0.73	3.15	0.01
Severity		57.79	97.35		61.29	109.88	0.17	-0.27	0.78
Mean Sev.		29.92	89.02		39.82	88.34	0.93	-0.90	0.37
Long Leg. Absence	370			531					
Frequency		0.02	0.12		0.02	0.12	1.00	0.00	1.00
Severity		6.33	71.14		12.16	97.28	0.00	-0.55	0.58
Mean Sev.		6.33	71.14		12.16	97.28	0.00	-0.55	0.58
Sht. Non-Leg. Ab.	370			531					
Frequency		1.07	1.56		1.28	1.46	0.41	-1.11	0.26
Severity		31.59	85.68		53.02	122.59	0.00	-1.63	0.10
Mean Sev.		18.65	66.03		34.44	106.97	0.00	-1.43	0.15
Late Time	370			531					
Frequency		0.28	0.90		0.22	0.86	0.59	0.50	0.62
Severity		0.82	4.22		0.56	2.86	0.00	0.59	0.55
Mean Sev.		0.53	3.67		0.30	1.43	0.00	0.65	0.51

B-2c: Results of the T-Tests on Combined Seasonal and Holiday Effects in the Pre-Program 100-Day Measures

Once again there was an increase in time worked, from 354 hours per employee to 423 hours, T-Prob = 0.003. There was also a significant decrease in time scheduled off, from 355 to 255 hours per employee, T-Prob = 0.0001; and a very large decrease in time accounted for by missing values, 174 hours to 48 hours, T-Prob = 0.0001. Again, the change in time absent was non-significant, except for the frequency of short-term legitimate absence.

Hypothesis 5 is therefore supported to the extent that it is evident in time use behavior that is at least partly under the control of the employer. However, no significant evidence exists that a significant holiday/seasonal effect on discretionary non-legitimate absence occurred.

Table 26: T-Test Results for Hypotheses about Long-Term Year-to-Year Change in the 100-Day Measures, Frequency Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	2.95	2.22	1531	4.02	2.52	0.12	-3.83	0.00
	982	3.37	2.04	1347	2.62	1.87	0.31	3.29	0.00
	1025	3.45	2.38	1390	4.06	2.69	0.13	-2.06	0.04
	1210	3.01	2.17	1575	4.02	2.50	0.09	3.69	0.00
Routine Work	1166	2.89	2.28	1531	3.93	2.43	0.43	-3.75	0.00
	982	3.34	2.02	1347	2.60	1.82	0.22	3.27	0.00
	1025	3.42	2.28	1390	3.99	2.59	0.12	-1.96	0.05
	1210	2.97	2.25	1575	4.00	2.52	0.17	-3.68	0.00
Other Work	1166	0.05	0.24	1531	0.06	0.39	0.00	-0.27	0.78
	982	0.07	0.32	1347	0.04	0.24	0.00	0.72	0.46
	1025	0.05	0.24	1390	0.03	0.24	0.88	0.64	0.52
	1210	0.07	0.30	1575	0.07	0.41	0.00	-0.11	0.91
All Sch. Off	1166	13.56	6.67	1531	13.38	5.15	0.00	0.26	0.79
	982	12.91	4.30	1347	9.30	3.82	0.16	7.61	0.00
	1025	13.64	4.74	1390	12.28	5.04	0.46	2.37	0.01
	1210	12.75	6.65	1575	13.54	5.25	0.01	1.12	0.26
Missing Values	1166	0.15	0.43	1531	0.28	1.76	0.00	-0.82	0.41
	982	0.17	0.41	1347	1.10	0.94	0.00	-10.85	0.00
	1025	0.13	0.37	1390	0.21	1.48	0.00	-0.68	0.49
	1210	0.15	0.37	1575	0.30	1.74	0.00	-1.06	0.28
Punishment Time	1166	0.11	0.37	1531	0.13	0.50	0.00	-0.42	0.67
	982	0.15	0.46	1347	0.11	0.35	0.00	0.96	0.33
	1025	0.09	0.33	1390	0.12	0.42	0.00	-0.62	0.53
	1210	0.12	0.40	1575	0.11	0.36	0.23	0.13	0.89
Routine Off	1166	13.42	6.87	1531	13.31	5.29	0.00	0.15	0.88
	982	12.77	4.41	1347	9.39	4.26	0.69	6.67	0.00
	1025	13.54	4.87	1390	12.20	5.19	0.44	2.27	0.02
	1210	12.58	6.87	1575	13.47	5.39	0.00	-1.24	0.21
All Time Absent	1166	2.47	2.04	1531	3.27	2.31	0.14	-3.12	0.00
	982	3.18	1.97	1347	2.54	1.81	0.32	2.89	0.00
	1025	3.40	2.32	1390	3.59	2.59	0.20	-2.66	0.50
	1210	2.48	1.97	1575	3.45	2.38	0.02	-3.80	0.00
Sht. Leg. Ab.	1166	1.81	1.69	1531	2.51	2.09	0.01	-3.12	0.00
	982	2.36	1.69	1347	2.26	1.78	0.56	0.50	0.61
	1025	2.70	2.00	1390	3.30	2.63	0.00	-2.16	0.03
	1210	1.73	1.75	1575	2.67	2.13	0.02	-4.10	0.00
Long Leg. Ab.	1166	0.22	0.54	1531	0.33	0.78	0.00	-1.36	0.17
	982	0.60	0.62	1347	0.33	0.59	0.57	3.78	0.00
	1025	0.59	0.64	1390	0.34	0.65	0.74	3.30	0.00
	1210	0.21	0.48	1575	0.37	0.76	0.00	-2.18	0.03
Sht Non-Leg. Ab.	1166	0.68	1.23	1531	0.73	1.17	0.58	-0.39	0.69
	982	0.45	0.96	1347	0.20	0.45	0.00	2.89	0.00
	1025	0.30	0.93	1390	0.17	0.43	0.00	1.52	0.12
	1210	0.75	1.19	1575	0.74	1.17	0.82	0.03	0.97
Late Time	1166	0.37	1.43	1531	0.58	1.16	0.01	-1.39	0.16
	982	0.35	1.05	1347	0.43	0.94	0.21	-0.70	0.48
	1025	0.33	0.96	1390	0.50	1.07	0.19	-1.45	0.14
	1210	0.33	0.88	1575	0.52	0.98	0.20	-1.71	0.08

Table 27: T-Test Results for Hypotheses about Long-Term Year-to-Year Change in the 100-Day Measures, Severity Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	412.55	204.59	1531	436.37	184.62	0.22	-1.05	0.29
	982	376.40	151.03	1347	264.89	126.99	0.04	6.84	0.00
	1025	415.21	172.11	1390	403.56	182.65	0.47	0.56	0.57
	1210	418.51	214.19	1575	435.45	191.03	0.17	-0.71	0.47
Routine Work	1166	411.81	204.68	1531	434.87	184.84	0.22	-1.01	0.31
	982	375.34	150.62	1347	264.43	126.71	0.04	6.83	0.00
	1025	414.52	171.82	1390	403.33	182.38	0.47	0.54	0.59
	1210	417.77	213.92	1575	434.43	190.77	0.17	0.70	0.48
Other Work	1166	0.74	4.56	1531	1.50	12.31	0.00	-0.69	0.48
	982	1.06	6.38	1347	0.45	2.49	0.00	1.08	0.28
	1025	0.69	4.13	1390	10.23	1.90	0.00	1.24	0.21
	1210	0.74	3.38	1575	1.02	7.15	0.00	-0.42	0.67
All Sch. Off	1166	282.32	197.86	1531	257.06	161.11	0.01	1.20	0.23
	982	250.52	135.30	1347	394.01	105.51	0.00	-10.14	0.00
	1025	261.13	146.87	1390	222.48	146.30	0.96	2.26	0.02
	1210	270.39	203.50	1575	255.13	168.09	0.02	0.70	0.48
Missing Values	1166	92.03	256.02	1531	56.06	193.95	0.00	1.36	0.17
	982	38.78	157.57	1347	246.07	85.02	0.00	-14.11	0.00
	1025	59.60	186.22	1390	30.92	137.55	0.00	1.50	0.13
	1210	95.74	256.74	1575	63.06	206.91	0.01	1.20	0.23
Punishment Time	1166	3.54	17.63	1531	7.50	38.41	0.00	-1.12	0.26
	982	4.24	22.28	1347	10.20	56.25	0.00	-1.18	0.24
	1025	1.54	7.71	1390	12.58	74.47	0.00	-1.75	0.08
	1210	6.49	30.73	1575	5.16	25.32	0.02	0.40	0.68
Routine Off	1166	186.75	89.00	1531	193.50	72.37	0.01	-0.71	0.47
	982	207.50	79.67	1347	137.74	65.84	0.02	8.18	0.00
	1025	200.09	71.03	1390	178.97	79.69	0.17	2.38	0.01
	1210	168.17	81.55	1575	186.91	72.52	0.16	-2.08	0.03
All Time Absent	1166	110.01	190.81	1531	110.21	164.91	0.08	-0.01	0.99
	982	168.99	173.15	1347	140.04	170.53	0.86	1.44	0.15
	1025	126.99	172.50	1390	178.52	218.35	0.00	-2.23	0.02
	1210	115.05	195.41	1575	112.53	165.16	0.05	0.60	0.54
Sht. Leg. Ab.	1166	44.68	95.66	1531	50.62	78.71	0.02	-0.58	0.56
	982	65.06	76.12	1347	60.10	82.12	0.36	0.53	0.59
	1025	68.29	80.62	1390	77.89	102.19	0.00	-0.89	0.37
	1210	46.20	103.29	1575	42.50	58.71	0.00	0.38	0.70
Long Leg. Ab.	1166	55.63	170.79	1531	49.73	150.92	0.14	0.31	0.75
	982	97.38	172.49	1347	75.74	166.34	0.66	1.09	0.27
	1025	53.74	168.58	1390	97.41	216.84	0.00	-1.91	0.05
	1210	58.01	170.66	1575	68.35	161.32	0.50	-0.02	0.98
Sht Non-Leg. Ab.	1166	9.70	25.08	1531	9.85	17.38	0.00	-0.06	0.95
	982	6.54	19.36	1347	4.19	13.35	0.00	1.21	0.22
	1025	4.97	18.08	1390	3.22	11.13	0.00	1.00	0.31
	1210	10.83	23.32	1575	11.69	27.13	0.03	-0.29	0.77
Late Time	1166	0.79	2.94	1531	1.19	3.28	0.18	-1.09	0.27
	982	1.32	4.74	1347	0.72	2.22	0.00	1.39	0.16
	1025	1.25	3.69	1390	1.18	3.77	0.79	0.15	0.87
	1210	0.60	2.25	1575	1.18	3.18	0.00	-1.79	0.07

B-3a: Results of the T-Tests for Year-to-Year Change in the 100-Day Measures

Both the frequency indices and severity indices show significant shifts in time worked (down from 376 hours to 264 hours), in time scheduled off (up from 250 hours to 394 hours), and in missing values (up from 38 hours to 246 hours between summer of 1980 and summer of 1981). The pattern here suggests a strike or plant shutdown during the summer of 1981. Comparison between day 1347 and day 1390 suggests that this event must have occurred between July 1 and August 12, 1981, since the event was not included in the 100-day base for day 1390. The frequency indices however, also indicate significant year-to-year increases for other periods in overall frequency of legitimate absences and a corresponding overall increase in the frequency of work intervals, and these two findings are both internally consistent with the progression hypothesis, although no corresponding increase in time used in absence was observed.

Furthermore, except for the anomaly during the summer of 1981, the only significant change in overall absence for these tests was in long-term legitimate absence which increased from 53 to 97 hours per employee just after the summer anomaly occurred and which suggests that some workers may have taken a long-term leave during those summer months.

Hypothesis 6 is only supported for the frequency of both absence and time worked, workers increased both. On the other hand, this did not significantly affect severity of either index which suggests that the employees were developing a long-term habit of using their casual leave days in several occurrences instead of just a few. Some weak support for an hypothesis on long-term absence increase is evident.

These findings are highly vulnerable to short-term intra-annual effects, however, and despite the comparison to corresponding year-to-year dates, they should be considered with caution.

Table 28: T-Test Results for Holiday Effects on the 100-Day Data, Severity Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1166	412.55	204.59	1210	418.51	214.19	0.58	-0.25	0.80
	1531	436.37	184.62	1575	435.41	191.03	0.69	0.04	0.96
	982	376.40	151.03	1025	415.21	172.11	0.11	-2.08	0.03
	1347	264.89	126.99	1390	403.56	182.65	0.00	-7.40	0.00
Routine Work	1166	411.81	204.68	1210	417.77	213.92	0.59	-0.25	0.80
	1531	434.87	184.84	1575	434.43	190.77	0.71	0.02	0.98
	982	375.41	150.52	1025	414.52	171.82	0.11	-2.11	0.03
	1347	264.43	126.71	1390	403.33	182.38	0.00	7.43	0.00
Other Work	1166	0.74	4.56	1210	0.74	3.38	0.00	0.00	1.00
	1531	1.50	12.31	1575	1.02	7.15	0.00	0.40	0.68
	982	1.06	6.38	1025	0.69	4.13	0.00	0.60	0.54
	1347	0.45	2.49	1390	0.23	1.90	0.00	0.86	0.39
All Sch. Off	1166	282.32	197.86	1210	270.39	203.50	0.73	0.52	0.60
	1531	257.06	161.11	1575	255.13	168.09	0.62	0.10	0.92
	982	250.52	135.30	1025	261.23	146.87	0.32	-0.66	0.51
	1347	394.01	105.51	1390	222.48	146.30	0.00	11.29	0.00
Missing Values	1166	92.03	256.02	1210	95.74	256.74	0.97	-0.13	0.90
	1531	56.06	193.95	1575	63.06	206.91	0.45	-0.29	0.76
	982	38.78	157.57	1025	59.60	186.22	0.04	-1.05	0.29
	1347	246.07	85.02	1390	30.92	137.55	0.00	15.80	0.00
Punishment Time	1166	3.54	17.63	1210	6.49	30.73	0.00	-1.02	0.30
	1531	7.50	38.41	1575	5.16	25.32	0.00	0.60	0.54
	982	4.24	22.28	1025	1.54	7.71	0.00	1.41	0.16
	1347	10.20	56.25	1390	12.58	74.47	0.00	-0.30	0.76
Routine Off	1166	186.75	89.00	1210	168.17	81.55	0.29	1.89	0.05
	1531	193.50	72.47	1575	186.91	72.52	0.99	0.76	0.44
	982	207.50	79.67	1025	200.09	71.03	0.16	0.85	0.39
	1347	137.74	65.84	1390	178.97	79.69	0.02	-4.74	0.00
All Time Absent	1166	110.01	190.81	1210	115.05	195.41	0.77	-0.23	0.82
	1531	110.21	164.91	1575	112.53	165.16	0.99	-0.12	0.90
	982	168.99	173.15	1025	126.99	175.50	0.96	2.11	0.03
	1347	140.04	170.53	1390	178.52	218.35	0.00	-1.65	0.10
Sht. Leg. Ab.	1166	44.68	95.66	1210	46.20	103.29	0.35	-0.13	0.89
	1531	50.62	78.71	1575	42.50	58.71	0.00	0.98	0.32
	982	65.06	76.12	1025	68.29	80.62	0.48	-0.36	0.72
	1347	60.10	82.12	1390	77.89	102.19	0.01	-1.61	0.10
Long Leg. Ab.	1166	55.63	170.79	1210	58.01	170.66	0.99	-0.12	0.90
	1531	49.73	150.92	1575	58.35	161.32	0.43	-0.46	0.64
	982	97.38	172.49	1025	53.74	188.58	0.78	2.22	0.02
	1347	75.74	166.34	1390	97.41	216.84	0.00	-0.94	0.34
Sht Non-Leg. Ab.	1166	9.70	25.08	1210	10.83	23.32	0.37	-0.41	0.68
	1531	9.86	17.38	1575	11.69	27.13	0.00	-0.67	0.50
	982	6.54	19.36	1025	4.97	18.08	0.40	0.73	0.46
	1347	4.19	13.35	1390	3.22	11.13	0.03	0.66	0.50
Late Time	1166	0.79	2.94	1210	0.60	2.25	0.00	0.66	0.51
	1531	1.19	3.28	1575	1.17	3.18	0.70	0.04	0.97
	982	1.32	4.74	1025	1.25	3.69	0.00	0.14	0.89
	1347	0.72	2.22	1390	1.18	3.77	0.00	-1.25	0.21

B-3b: T-Test Results for Hypotheses about Holiday Effects in the 100-Day Measures

Both the frequency and severity indices were T-tested to detect significant holiday effects, but the results were essentially the same for both so that a detailed presentation of only the severity shifts is included in Table 28. In brief, two significant shifts have occurred across major holiday periods. These occurred for the time-use measures during the summers of 1980 and of 1981. A large shift occurredd across the same interval for which an anomaly was discovered which affected Hypothesis 6 and showed up as a year-to-year overall change. A smaller shift also associated with a summer (July 4th) holiday period in 1980 is also apparent and significant. No holiday effects were apparent across the Christmas periods.

Hypothesis 7 is supported for summer holidays, but this holiday effect may be at least in part the effect of annual renewal or vacation. There is no significant effect across the winter (Christmas) holiday season when renewal would not be expected to have much effect and holidays might. Once again, some rather anomalous shifts and reductions in time-worked as well as time scheduled off occurred during the summer of 1981.

Table 29: T-Test Results for Hypotheses of Combined Holiday, Seasonal, and Renewal Effects on the 100-Day Measures: Severity Indices Only

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	982	376.40	151.03	1166	412.55	204.59	0.00	-.175	0.08
	1347	264.89	126.99	1531	436.37	184.62	0.00	-9.09	0.00
	1025	415.21	172.11	1210	418.51	214.19	0.01	-0.15	0.88
	1390	403.56	182.65	1575	435.45	191.03	0.60	-1.43	0.15
Routine Work	982	375.34	150.52	1166	411.81	204.68	0.00	-1.76	0.07
	1347	264.43	126.71	1531	434.87	184.84	0.00	-9.03	0.00
	1025	414.52	171.82	1210	417.77	213.92	0.01	-0.15	0.16
	1390	403.33	182.38	1575	434.43	190.77	0.60	-1.40	0.16
Other Work	982	1.06	6.38	1166	0.74	4.56	0.00	0.50	0.61
	1347	0.45	2.49	1531	1.50	12.31	0.00	-0.99	0.32
	1025	0.69	4.13	1210	0.74	3.38	0.01	-0.12	0.90
	1390	0.23	1.90	1575	1.02	7.15	0.00	-1.28	0.20
All Sch. Off	982	250.52	135.30	1166	282.32	197.86	0.00	-1.63	0.10
	1347	394.01	105.57	1531	257.06	161.11	0.00	-8.44	0.00
	1025	261.23	146.87	1210	270.39	203.50	0.00	-0.45	0.65
	1390	222.48	146.30	1575	255.13	168.09	0.10	-1.74	0.08
Missing Values	982	38.78	157.57	1166	92.03	256.02	0.00	-2.18	0.03
	1347	246.07	85.02	1531	56.01	193.95	0.00	10.65	0.00
	1025	59.60	186.22	1210	95.74	256.74	0.00	-1.40	0.16
	1390	30.92	137.55	1575	63.06	206.91	0.00	-1.54	0.12
Punishment Time	982	4.24	22.28	1166	3.54	17.63	0.00	0.30	0.76
	1347	10.20	56.25	1531	7.50	38.41	0.00	0.47	0.63
	1025	1.54	7.71	1210	6.49	30.73	0.00	-1.92	0.05
	1390	12.58	74.47	1575	5.16	25.32	0.00	1.12	0.26
Routine Off	982	207.50	79.67	1166	186.75	89.00	0.18	2.13	0.03
	1347	137.74	65.84	1531	193.50	72.47	0.26	-6.76	0.00
	1025	200.09	71.03	1210	168.17	81.55	0.09	3.63	0.00
	1390	178.97	79.69	1575	186.91	72.52	0.27	-0.88	0.38
All Time Absent	982	168.99	173.15	1166	110.01	190.81	0.24	2.81	0.00
	1347	140.04	170.53	1531	110.21	164.91	0.69	1.49	0.13
	1025	126.99	172.50	1210	115.05	195.41	0.13	0.56	0.57
	1390	178.52	218.35	1575	112.53	165.16	0.00	2.86	0.00
Sht. Leg. Ab.	982	65.06	76.12	1166	44.68	95.66	0.00	2.05	0.04
	1347	60.10	82.12	1531	50.62	78.71	0.62	0.98	0.32
	1025	68.29	80.62	1210	46.20	103.29	0.00	2.07	0.03
	1390	77.89	102.19	1575	42.50	58.71	0.00	3.57	0.00
Long Leg. Ab.	982	97.38	172.49	1166	55.63	170.79	0.90	2.11	0.03
	1347	75.74	166.34	1531	49.73	150.93	0.25	1.38	0.17
	1025	53.74	168.58	1210	58.01	170.66	0.88	-0.22	0.82
	1390	97.41	216.84	1575	58.35	161.32	0.00	1.72	0.08
Sht Non-Leg. Ab.	982	6.54	19.36	1166	9.70	25.08	0.00	-1.23	0.22
	1347	4.19	13.35	1531	9.85	17.38	0.00	-3.07	0.00
	1025	4.97	18.08	1210	10.83	23.32	0.00	-2.44	0.01
	1390	3.22	11.13	1575	11.69	27.13	0.00	-3.43	0.00
Late Time	982	1.32	4.74	1166	0.79	2.93	0.00	1.15	0.25
	1347	0.72	2.22	1531	1.19	3.28	0.00	-1.31	0.19
	1025	1.25	3.69	1210	0.60	2.25	0.00	1.86	0.06
	1390	1.18	3.77	1575	1.18	3.18	0.04	0.02	0.98

B-3c: Summary of Results for the T-Tests of Combined Seasonal and Annual Renewal Effects on the 100-Day Measures

Several T-test comparisons for the 100-day measures offer general support for Hypothesis 8. Combined seasonal and renewal effects on absence do appear to take place, producing significant decreases in levels of absence and increases in time worked from summer to winter. However, the same summer of 1981 anomally (Date 1 = 1347) which was evident in earlier T-tests would also presumably be contributing to this pattern, so to some extent these results too are suspect. Short-term absence does account for part of the change, both legitimate and non-legitimate short-term absence. Long-term absence was also affected, but not to the same extent.

Comment on the Measures Used to Examine Intra-Annual T-Test Results

In general, the frequency indices and mean severity co-vary with the severity indices and tend to provide significant results across the same intervals, although in this there were some exceptions. For convenience, in the present analysis the results on the frequency indices have not been included. Detailed examination of the relationship between severity and frequency of absence is definitely an interesting problem. Indeed, the parallelim in the 100-day indices between severity and frequency is in the reverse direction to an apparent co-variation suggested by the 30-day measures, where frequency went down as severity went up. It is likely that this latter effect is present by construction of the measures and that both patterns can and do operate at the same time.

The internal consistency and detail built into the integral set time-use measures should allow such an analysis to be carried out. The respective mathematical forms should be distinguishable and any relationships between them may be as well. Formal procedures for detailed time-series analysis need to be applied. But, that problem is beyond the scope of this present inquiry due to limitations on time and computer resources. Some insight may be forthcoming from the time-series correlations and graphical analysis.

B-4: Results of the T-Tests on Overall Program-Induced Change in the 100-Day Measures

The so-called program induced changes predicted in the 100-day data unquestionably appear significant. Time scheduled off, time worked, and time absent all decreased significantly from the end of the baseline year to the end of the first program year. Then all three again significantly increased back to their original level by the end of the second program year. Meantime, missing values increased dramatically between the end of the baseline year and the end of the first program year, then decreased back toward their original level by the end of the second program year.

Overall, between the end of the baseline year and the end of the second program year, time worked apparently showed no significant change, nor did time scheduled off or time absent. These overall year-to-year tests of program induced changes were likely to be affected by short-term fluctuations for the 100-day measures, however. Rather than focusing on them, the 365-day measures were used instead to assess overall program impact.

Table 30: T-Test Results of Program Induced Change in the 365-Day Measures

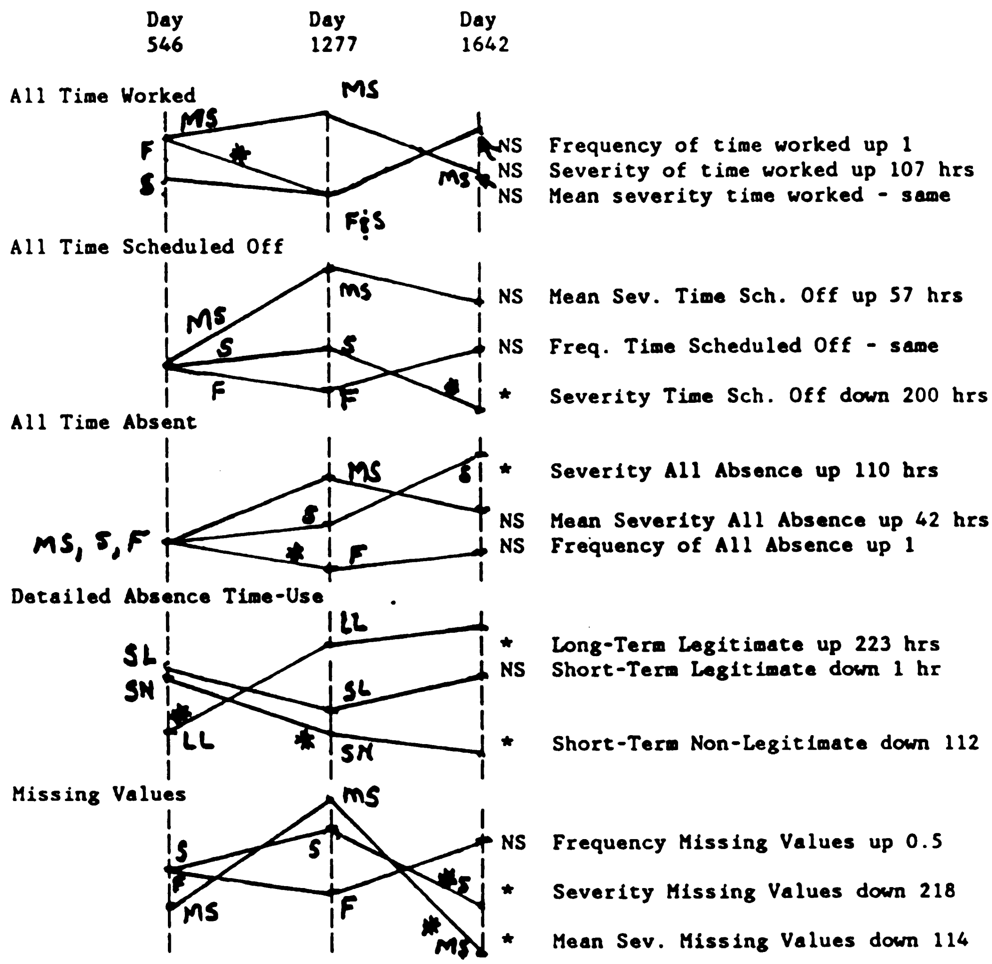
Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	546			1277					
Frequency		11.17	5.44		9.34	5.29	0.74	2.81	0.00
Severity		1383.33	517.67		1322.63	591.38	0.12	0.91	0.36
Mean Sev.		151.10	108.38		155.07	106.95	0.87	-0.30	0.76
Routine Work	546			1277					
Frequency		10.90	5.33		9.20	5.16	0.69	2.69	0.00
Severity		1379.21	517.23		1320.09	590.91	0.12	0.89	0.37
Mean Sev.		154.12	109.31		156.97	107.91	0.87	-0.22	0.82
Other Work	546			1277					
Frequency		0.34	1.67		0.17	0.53	0.00	1.13	0.25
Severity		4.11	19.09		2.54	8.33	0.00	0.86	0.39
Mean Sev.		1.63	7.87		2.04	7.02	0.18	-0.46	0.64
All Scheduled Off	546			1277					
Frequency		47.79	18.37		44.21	18.61	0.88	1.61	0.10
Severity		1130.93	501.61		1160.86	588.36	0.07	-0.46	0.64
Mean Sev.		40.09	70.25		164.61	573.57	0.00	-2.64	0.00
Missing Values	546			1277					
Frequency		1.29	2.27		1.09	0.36	0.00	1.02	0.31
Severity		419.68	703.82		506.07	753.95	0.43	-0.98	0.32
Mean Sev.		227.95	541.41		471.54	710.64	0.00	-3.23	0.00
Punishment Time	546			1277					
Frequency		0.45	1.07		0.36	0.82	0.00	0.81	0.41
Severity		9.40	27.64		13.79	48.23	0.00	-0.95	0.34
Mean Sev.		4.17	8.85		6.88	17.61	0.00	-1.65	0.09
Routine Off	546			1277					
Frequency		47.28	18.85		43.70	18.79	0.97	1.57	0.11
Severity		701.85	245.18		641.01	269.04	0.28	1.97	0.50
Mean Sev.		15.09	2.47		14.09	5.02	0.00	2.14	0.03
All Time Absent	546			1277					
Frequency		10.76	5.50		8.92	5.11	0.39	2.87	0.00
Severity		396.29	459.44		431.03	560.36	0.02	-0.57	0.57
Mean Sev.		52.92	141.63		103.63	338.48	0.00	-1.67	0.09
Sht. Leg. Absence	546			1277					
Frequency		7.44	3.90		6.95	3.99	0.80	1.02	0.31
Severity		222.06	277.14		180.44	256.47	0.36	1.29	0.19
Mean Sev.		39.41	105.95		32.11	79.58	0.00	0.64	0.52
Long Leg. Absence	546			1277					
Frequency		0.04	0.23		0.95	1.05	0.00	-10.35	0.00
Severity		34.60	274.80		221.22	509.37	0.00	-3.88	0.00
Mean Sev.		23.05	159.09		127.80	333.81	0.00	-3.42	0.00
Sht. Non-Leg. Ab.	546			1277					
Frequency		3.94	3.80		1.75	2.55	0.00	5.49	0.00
Severity		139.63	275.64		29.36	61.49	0.00	4.40	0.00
Mean Sev.		35.86	90.91		8.33	13.94	0.00	3.37	0.00
Late Time	546			1277					
Frequency		1.17	3.04		1.05	2.81	0.37	0.32	0.74
Severity		3.44	9.80		2.95	7.79	0.01	0.46	0.64
Mean Sev.		1.17	2.76		1.04	2.19	0.01	0.44	0.66

Table 30 Continued

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	546			1642					
Frequency		11.17	5.44		12.21	6.80	0.01	-1.39	0.16
Severity		1383.33	517.67		1490.70	591.42	0.13	-1.58	0.11
Mean Sev.		151.10	108.38		147.33	127.49	0.06	0.27	0.78
Routine Work	546			1642					
Frequency		10.90	5.33		11.94	6.56	0.02	-1.43	0.15
Severity		1379.21	517.23		1487.47	591.21	0.13	-1.60	0.11
Mean Sev.		154.12	109.31		148.20	125.65	0.11	0.41	0.68
Other Work	546			1642					
Frequency		0.34	1.67		0.18	0.90	0.00	0.98	0.32
Severity		4.11	19.09		3.23	19.73	0.71	0.37	0.71
Mean Sev.		1.63	7.87		1.24	4.97	0.00	0.48	0.63
All Scheduled Off	546			1642					
Frequency		47.79	18.37		48.18	16.45	0.20	-0.18	0.85
Severity		1130.93	501.61		915.97	510.73	0.84	3.47	0.00
Mean Sev.		40.09	70.25		92.32	387.79	0.00	-1.57	0.11
Missing Values	546			1642					
Frequency		1.29	2.27		1.70	5.14	0.00	-0.84	0.40
Severity		419.68	703.82		201.50	597.66	0.06	2.71	0.00
Mean Sev.		227.95	541.41		113.30	396.57	0.00	1.95	0.05
Punishment Time	546			1642					
Frequency		0.45	1.07		0.34	0.86	0.01	0.93	0.35
Severity		9.40	27.64		30.63	142.68	0.00	-1.73	0.08
Mean Sev.		4.17	8.85		21.57	133.91	0.00	-1.54	0.12
Routine Off	546			1642					
Frequency		47.28	18.85		47.81	16.62	0.15	-0.24	0.81
Severity		701.85	245.18		683.84	234.47	0.61	0.61	0.54
Mean Sev.		15.09	2.47		13.75	3.87	0.00	3.41	0.00
All Time Absent	546			1642					
Frequency		10.76	5.50		11.74	6.58	0.04	-1.33	0.18
Severity		396.29	459.44		506.01	564.39	0.02	-1.75	0.08
Mean Sev.		52.92	141.63		94.15	350.52	0.00	-1.28	0.20
Sht. Leg. Absence	546			1642					
Frequency		7.44	3.90		10.01	5.82	0.00	-4.29	0.00
Severity		222.06	277.14		221.06	224.37	0.02	0.03	0.97
Mean Sev.		39.41	105.95		34.91	123.56	0.08	0.32	0.74
Long Leg. Absence	546			1642					
Frequency		0.04	0.23		1.03	1.61	0.00	7.22	0.00
Severity		34.60	274.80		257.50	534.78	0.00	-4.35	0.00
Mean Sev.		23.05	159.09		107.46	286.28	0.00	-3.02	0.00
Sht. Non-Leg. Ab.	546			1642					
Frequency		3.94	3.80		1.75	2.25	0.00	5.62	0.00
Severity		139.63	275.64		27.44	54.52	0.00	4.49	0.00
Mean Sev.		35.86	90.91		7.80	11.44	0.00	3.44	0.00
Late Time	546			1642					
Frequency		1.17	3.04		1.79	3.03	0.99	-1.67	0.09
Severity		3.44	9.80		4.22	10.50	0.43	-0.62	0.53
Mean Sev.		1.17	2.76		0.99	2.09	0.00	0.59	0.55

Table 30 Continued

Variable Index	Relative Date 1	Mean 1	Std. Dev.1	Relative Date 2	Mean 2	Std. Dev.2	F-Prob	T	T-Pr
All Time Worked	1277			1642					
Frequency		9.34	5.28		12.21	6.80	0.00	-4.00	0.00
Severity		1322.63	591.38		1490.70	591.42	1.00	-2.43	0.21
Mean Sev.		155.07	106.95		147.33	127.49	0.03	0.56	0.57
Routine Work	1277			1642					
Frequency		9.20	5.16		11.94	6.56	0.00	-3.96	0.00
Severity		1320.09	510.91		1487.47	591.21	0.99	-2.42	0.01
Mean Sev.		156.97	107.91		148.20	125.65	0.07	0.64	0.52
Other Work	1277			1642					
Frequency		0.17	0.54		0.18	0.90	0.00	-0.13	0.89
Severity		2.54	8.33		3.23	19.73	0.00	-0.39	0.70
Mean Sev.		2.04	7.02		1.24	4.97	0.00	1.14	0.25
All Sch. Off	1277			1642					
Frequency		44.27	18.61		48.18	16.45	0.14	-1.93	0.05
Severity		1160.86	588.36		915.97	510.73	0.09	3.80	0.00
Mean Sev.		164.61	573.57		92.32	387.79	0.00	1.27	0.20
Missing Values	1277			1642					
Frequency		1.09	0.36		1.70	5.14	0.00	-1.40	0.16
Severity		506.07	753.95		201.50	597.66	0.00	3.84	0.00
Mean Sev.		471.54	710.64		113.30	396.57	0.00	5.36	0.00
Punishment Time	1277			1642					
Frequency		0.36	0.82		0.34	0.86	0.55	0.17	0.86
Severity		13.79	48.23		30.63	142.68	0.00	-1.33	0.18
Mean Sev.		6.87	17.61		21.57	133.91	0.00	-1.29	0.19
Routine Off	1277			1642					
Frequency		43.70	18.79		47.81	16.62	0.14	-1.98	0.04
Severity		641.01	269.04		683.84	234.47	0.10	-1.45	0.14
Mean Sev.		14.09	5.02		13.75	3.87	0.00	0.66	0.50
All Time Absent	1277			1642					
Frequency		8.92	5.11		11.74	6.58	0.00	-4.08	0.00
Severity		431.03	560.36		506.01	564.39	0.93	-1.14	0.25
Mean Sev.		103.63	338.48		94.15	350.52	0.67	0.23	0.81
Sht. Leg. Ab.	1277			1642					
Frequency		6.95	3.99		10.01	5.82	0.00	-5.21	0.00
Severity		180.44	256.47		221.06	224.37	0.11	-1.44	0.15
Mean Sev.		32.11	79.58		34.91	123.56	0.00	-0.23	0.81
Long Leg. Ab.	1277			1642					
Frequency		0.95	1.05		1.03	1.61	0.00	-0.47	0.64
Severity		221.22	509.37		257.50	534.78	0.58	-0.59	0.55
Mean Sev.		127.80	333.81		107.46	286.28	0.07	0.56	0.57
Sht Non-Leg. Ab.	1277			1642					
Frequency		1.75	2.55		1.75	2.25	0.14	0.01	0.99
Severity		29.36	61.48		27.44	54.52	0.15	0.28	0.77
Mean Sev.		8.33	13.94		7.80	11.44	0.02	0.36	0.72
Late Time	1277			1642					
Frequency		1.05	2.81		1.79	3.03	0.36	-2.14	0.03
Severity		2.95	7.79		4.22	10.50	0.00	-1.17	0.24
Mean Sev.		1.04	2.19		0.99	2.09	0.57	0.18	0.85



NOTE: The relative scales of sketches above are merely roughly "eyeballed" pictures of trends in each line, clustered for convenience.

Figure 16:
Approximate Graphic Profiles of Long-Term Program Impact on the Overall 365-Day Time-Use Measures: Relative Scales

B-5: Results of the T-Tests on Overall Program Induced Changes Indicated by the 365-Day Measures

Figure 16 offers a rough graphical summary of overall changes in the major annual time-use indices. Full details have been presented in Table 30. Overall, time worked per employee per year is up by 107 hours to 1490 hours per year. Time scheduled off was down by 200 hours to 915 hours. And an overall increase in time absent of 110 hours per employee per year took place, all of which was in long-term legitimate absence since short-term non-legitimate absence dropped 112 hours to 27 hours per employee per year from 139 hours. Most of the observed change in absence behavior occurred before the end of the first complete year of the program.

Also worthy of note is a substantial decrease in the level of missing values present in the data, over the second year of the program especially. It appears likely that a certain amount of time must have been needed after program start-up to streamline the administration of data collection by supervisors, data-entry clericals, etc.

Overall, it appears that the program was clearly successful if the T-test statistics are to be believed. Once again, a caveat is critical here. These T-test values are based on distinctly abnormal data distributions.

Hypothesis 9 is supported. See the time-series plots (Figure 16 above) for a more detailed picture of the program impact over time.

C. The Split-Half Time-Series Reliability Correlations

Results for three sets of time-series correlations between split-halves of the time-use datasets are summarized in Table 31. A reliability correlation and the associated probability of that correlation occurring by mere coincidence (significance level) has been listedd for each of the 36 time-use indices (time due to termination of employees was dropped because sample size was equal to 0). The three datasets for which reliability correlations were generated are: (1) during the baseline year from day 281 to day 546, using the 100-day data; (2) during the post-installation period from day 982 to day 1680, using the 100-day data; and (3) during the post-installation period from day 1247 to day 1680, using the 365-day data.

The time-series split-half reliability correlations have been discussed above in Chapter IV. An interpretation and discussion of the results follows.

Table 31: Time Series Correlations for Split-Half Reliability

Variable Index	Pre-Prog r 100-Day	Post-Inst r 100-Day	Post-Inst r 365-Day
All Time Worked			
Frequency	0.75	0.88	0.99
Severity	0.98	0.92	0.63
Mean Sev.	0.88	0.50	0.50
Routine Work			
Frequency	0.76	0.89	0.98
Severity	0.97	0.91	0.61
Mean Sev.	0.91	0.57	0.64
Other Work			
Frequency	0.53	Sample too small	Sample too small
Severity	0.91	Sample too small	Sample too small
Mean Sev.	Sample too small	Sample too small	0.70
All Scheduled Off			
Frequency	0.93	0.99	0.98
Severity	0.99	0.96	0.87
Mean Sev.	0.97	0.82	0.92
Missing Values			
Frequency	0.89	0.88	0.46
Severity	0.99	0.96	0.87
Mean Sev.	0.98	0.96	0.84
Punishment Time			
Frequency	Sample too small	Sample too small	Sample too small
Severity	Sample too small	Sample too small	Sample too small
Mean Sev.	Sample too small	Sample too small	Sample too small
Routine Off			
Frequency	0.95	0.99	0.97
Severity	0.97	0.98	0.99
Mean Sev.	0.86	0.92	Var. too small
All Time Absent			
Frequency	0.82	0.89	0.98
Severity	0.74	0.91	0.74
Mean Sev.	0.58	0.73	0.64
Sht. Leg. Absence			
Frequency	0.91	0.92	0.97
Severity	Sample too small	0.57	Var. too small
Mean Sev.	Sample too small	0.21	Var. too small
Long Leg. Absence			
Frequency	Sample too small	0.84	0.88
Severity	No values in 1/2	0.94	0.87
Mean Sev.	during baseline	0.89	0.69
Sht. Non-Leg. Ab.			
Frequency	0.08	0.76	Sample too small
Severity	0.96	0.36	Sample too small
Mean Sev.	0.68	0.56	Sample too small
Late Time			
Frequency	0.39	Sample too small	0.51
Severity	Sample too small	Sample too small	Sample too small
Mean Sev.	Sample too small	0.57	0.57

Summary of Results of the Time-Series Split-Half Reliability Correlations

Except for two peculiarities, the reliabilities are extremely high for all measures (see Table 31).

The construction of the time-series split-half reliability correlations is such that reliability is based upon the actual sample size of time-use behaviors underlying the correlated series of means as noted in Appendix III. Where the time-use behavior involved becomes an idiosyncratic and sporadic event associated with relatively few employees, then the size of the respective mean will start to vary unpredictably instead of systematically, and the two "halves" of the overall sample may be anomalously (but meaninglessly) correlated as has occurred for small sample behaviors in a few of the present measures.

Five types of the 12 reported time-use measures were apparently affected by this type of problem. These were: other work, late time, disciplinary time-off (punishment time), short-term absence (both legitimate and non-legitimate after program installation), and long-term legitimate absence before program installation. Where the amount of available behavior (and sample size) became larger (even for these five exceptions), the anomalous correlations disappeared.

The second anomaly in the split-half reliability correlations was an unusual value for reliability of the routine time-off mean severity index in the 365-day measures. Examination of the covariance and prodduct of standard deviations which contributed to that result showed that both had extremely tiny values (less than .0001 of the size of the mean). The same period for the 365-day measures were high and consistent. Routine scheduled time off is controlled by the employer and is otherwise probably the least likely time-use to show a low reliability. For consistency with frequency and severity, it should be about +0.94.

Setting these anomalies aside, the results of the split-half reliability correlations are overwhelmingly exactly what they should be by construction, very high and all positive. Hypothesis 10 is unquestionably supported, and this is nothing remarkable but merely the result of measurement design.

Comparison of the reliabilities for the 100-day measures and 365-day measures separately to examine the frequency-severity controversy shows an intriguing result. When the "small sample" behaviors were excluded in the 100-day measures, severity was marginally more reliable than frequency in 9 out of 12 cases. When only the "small sample" behaviors were examined, severity was more reliable than frequency in 7 out of 12 cases. For the 365-day measures neither measure was more reliable. Frequency was a more reliable measure than severity in four out of six cases for the "large sample" behaviors, and more reliable than severity for three out of six of the "small sample" behaviors.

Hypothesis 10(a) is very well supported although it does require an adequate sample size, andboth the frequency and severity of the time-use behaviors can be highly reliable. Overall, neither was preferable on the basis of reliability.

Results of the Time-Series Correlations on Substitution Effects on the Severity Indices

Time-series correlations were calculated within selected time intervals for both the 100-day and 365-day datasets to examine various substantive hypotheses as outlined in Chapter IV.

Results for Hypotheses 11, 12, 13, and 14 were examined in that order for each of the five selected combinations of: 100-day or 365-day measure, and time interval of analysis. Intervals, correlations and associated significance levels, datasets, and related hypotheses are summarized in Table 32.

Table 32: Time-Series Correlations Testing Substantive Hypotheses for Severity Indices Only

Variable 1	Variable 2	Interval	r	r:Prob	Dataset	Related Hypothesis
		281- 546	* 0.45	0.0001	100-day	
Short-Term	Long-Term	982-1277	0.36	0.0001	100-day	Hypothesis
Legitimate	Legitimate	1278-1680	0.56	0.0001	100-day	11
Absence	Absence	982-1680	0.52	0.0001	100-day	Progression
		1247-1680	* 0.53	0.0001	365-day	
		281- 546	*-0.17	0.0058	100-day	
Short-Term	Short-Term	982-1277	-0.89	0.0001	100-day	Hypothesis
Legitimate	Non-Leg.	1278-1680	-0.93	0.0001	100-day	12
Absence	Absence	982-1680	-0.92	0.0001	100-day	Legitimacy
		1247-1680	*-0.67	0.0001	365-day	
		281- 546	* 0.69	0.0058	100-day	
Short-Term	Long-Term	982-1277	-0.44	0.0001	100-day	Hypothesis
Non-Leg.	Legitimate	1278-1680	-0.55	0.0001	100-day	13
Absence	Absence	982-1680	-0.53	0.0001	100-day	Progression
		1247-1680	*-0.56	0.0001	365-day	Legitimacy
		281- 546	0.97	0.0058	100-day	
Short-Term	Time	982-1277	-0.02	0.0001	100-day	Hypothesis
Non-Leg.	Worked	1278-1680	0.25	0.0001	100-day	14
Absence		982-1680	0.06	0.0001	100-day	Program
		1247-1680	*-0.10	0.0001	365-day	Impact
		281- 546	-0.98	0.0058	100-day	
Missing	Time	982-1277	-0.91	0.0001	100-day	As
Values	Worked	1278-1680	-0.75	0.0001	100-day	Above
		982-1680	-0.96	0.0001	100-day	
		1247-1680	-0.97	0.0001	365-day	
		281- 546	*-0.71	0.0058	100-day	
Missing	Long-Term	982-1277	-0.22	0.0001	100-day	As
Values	Legitimate	1278-1680	-0.45	0.0001	100-day	Above
	Absence	982-1680	-0.18	0.0001	100-day	
		1247-1680	-0.40	0.0001	365-day	
		281- 546	* 0.08	0.0058	100-day	
Missing	Short-Term	982-1277	-0.38	0.0001	100-day	Hypothesis
Values	Non-Leg.	1278-1680	-0.63	0.0001	100-day	14
	Absence	982-1680	-0.32	0.0001	100-day	Program
		1247-1680	*-0.47	0.0001	365-day	Impact

Note: * indicates a finding in which one or both of the measures correlated showed low reliability for the split-half samples. Since the overall sample size is twice that in the split-halves, such findings may or or may not be based upon a reliable series of values (see Appendix II

Summary of Time-Series Correlations Testing the Substantive Hypothesis, Using Severity Indices

There is evidence of a very strong legitimacy effect after the program installation occurred, but not before. In fact, before installation, short-term non-legitimate absence correlated +0.97 iwth time worked, while after installation it correlated -0.92 with short-term legitimate absence, and was uncorrelated with time worked. Hypothesis 12 is definitely supported.

The progression hypothesis (Hypothesis 11) is clearly not supported, however. Short- and long-term forms of legitimate absence are positively correlated, both on the 100-day and 365-day measures.

A combined progression and legitimacy effect (Hypothesis 13) is supported, but very probably the apparent effect is largely induced by legitimacy alone.

Hypothesis 14 is not supported, presumably because the overall impact of the program was allowed to dissipate in long-term legitimate absence.

A Comment on Frequency-Severity Relationships

For the sake of curiosity, the relationship between frequency and severity of the three major time-use categories was examined briefly. Table 33 below summarizes the results. Time scheduled off shows a strong negative correlation between frequency and severity. But, for the other two time uses, the relationship is equally strong and positive.

Table 33: Time-Series Correlations To Inspect Relations Between Frequency Indices and Severity Indices

Variable 1	Variable 2	Interval	r	r:Prob	Dataset
Time-Worked Severity	Time-Worked Frequency	982-1680 1246-1680	0.70 0.60	0.0001 0.0001	100-day 365-day
Time Sch Off Severity	Time Sch off Frequency	982-1680 1246-1680	-0.78 -0.74	0.0001 0.0001	100-day 365-day
Time Absent Severity	Time Absent Frequency	982-1680 1246-1680	0.51 0.73	0.0001 0.0001	100-day 365-day

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, Steven G. "Compensation, Safety, and Absenteeism: Evidence From The Paper Industry," Industrial and Labor Relations Review, Vol. 34, No. 2 (January 1981).

Baum, John F. "Effectiveness of an Attendance Control Policy in Reducing Chronic Absenteeism," Personnel Psychology, 1978, Vol. 31.

Beatty, Richard W. & Beatty, James R. "Longitudinal Study of Absenteeism of Hard-Core Unemployed," Psychological Reports, 1975, Vol. 36, 395-406.

Bloomfield, Peter. Fourier Analysis of Time-Series: An Introduction, John Wiley & Sons, 1976.

Breagh, James A. "Predicting Absenteeism from Prior Absenteeism and Work Attitudes," Journal of Applied Psychology, 1981, Vol. 66, No. 5, 555-560.

Behrend, Hilde & Pocock, Stuart, "Absence and the Individual: A Six-Year Study in One Organization," Industrial and Labor Review, Vol. 114, No. 3, Nov-Dec, 1976, pp. 311-327.

Campbell, Edward J. "Sick Leave Abuse and What To Do About It: A Look at Government Employees," Personnel, Vol. 57, Nov-Dec 1970.

Carlson, John G. & Hill, Kenneth D. "The Effect of Gaming - Attendance and Attitude," Personnel Psychology, 1982, Vol. 35.

Chadwick-Jones, J.K., Nicholson, Nigel, and Brown, Colin. Social Psychology of Absenteeism, New York: Praeger Publishers, 1982.

Cheloha, Randall S. & Farr, James L. "Absenteeism, Job Involvement, and Job Satisfaction in an Organizational Setting," Journal of Applied Psychology, 1980, Vol. 65, No. 4. pp. 467-473.

Elkouri, Frank and Elkouri, Edna Asper. How Arbitration Works, Third Edition, Washington, D.C.: The Bureau of National Affairs, Inc., 1960.

Garrison, Kathleen R. & Muchinsky, Paul M. "Evaluating the Concept of Absentee-Proneness with Two Measures of Absence," Personnel Psychology, 1977, Vol. 30.

Glass, Gene V., Willson, Victor L. & Gottman, John M. Design and Analysis of Time-Series Experiments, Colorado Associated University Press, Boulder Colorado, 1975.

Glasser, Jay H. "A Stochastic Model for Industrial Illness Absenteeism," American Journal of Public Health, Vol. 60, No. 10, 1970.

Hamblin, Robert L. "Social Attitudes: Magnitude Measurement and Theory," Chapter 3 in Measurement in the Social Sciences, H. M. Blalock, Jr., ed., 1974.

Hammer, Tove Helland & Landay, Jaqueline. "Methodological Issues in the Use of Absence Data," Journal of Applied Psychology, 1981, Vol. 66, No. 5. pp. 574-581.

Hedges, Janice Niepert. "Special Labor Force Report: Absence from Work - Measuring the Hours Lost," Monthly Labor Review, Oct. 1977, Vol. 100, No. 10.

Ilgen, Daniel R. "Attendance Behavior: A Reevaluation of Latham and Pursell's Conclusions," Journal of Applied Psychology, 1977, Vol. 62, No. 2, pp. 230-233.

Ilgen, Daniel R. and Hollenback, John H. "The Role of Job Satisfaction in Absence Behavior," Organizational Behavior and Human Performance, Vol. 19, pp. 148-161 (1977).

Ivancevich, John M. & McMahon, J. Timothy. "Black-White Differences in a Goal-Setting Program," Organizational Behavior and Human Performance, Vol. 20, pp. 287-300, (1977).

Johns, Gary. "A Multivariate Study of Absence from Work," Academy of Management Proceedings, 1978, pp. 69-73.

Kim, Jay S. & Campagna, Anthony F. "Effects of Flextime on Employee Attendance and Performange: A Field Experiment," Academy of Management Journal, 1981, Vol. 24, No. 4, 729-741.

Kuzmito, Frank E. "How Much is Absenteeism Costing Your Organization," The Personnel Administrator, June 1979.

Latham, Gary P. & Pusell, Elliott D. "Measuring Absenteeism From the Opposite Side of the Coin," Journal of Applied Psychology, 1975, Vol. 60, No. 3, 369-471.

Latham, Gary P. & Pursell, Elliott D. "Measuring Attendance: A Reply to Ilgen," Journal of Applied Psychology, 1977, Vol. 62, No. 2, 234-236.

Liu, Lon-Mu. "Box-Jenkins Time -- Series Analysis," BMDP Statistical Software, 1978 Edition, Ch. 20.2.

Luthans, Fred & Mario, Terry L. "Evaluating Personnel Programs Through the Reversal Tchnnique," Personnel Journal, October 1979.

Macy, Barry A. & Mirvis, Phillip H. "A Methodology for Assessment of Quality of Work Life and Organizational Effectiveness in Behavioral - Economic Terms," Administrative Science Quarterly, June 1976, Vol. 21.

Markham, Steven E.; Dansereau, Fred Jr. & Alutto, Joseph A. "Group Size and Absenteeism Rates: A Longitudinal Analysis," Academy of Management Journal, 1982, Vol. 25, No. 4, 921-927.

Markham, Steven E.; Dansereau, Fred & Alutto, Joseph A. "On the Use of Shift as an Independent Variable in Absenteeism Research," Journal of Occupational Psychology, 1982, Vol. 55, 225-231.

Markham, Steven E.; Dansereau, Fred & Alutto, Joseph A. "Female vs. Male Absence Rates: A Temporal Analysis," Personnel Psychology, 1982, Vol. 35.

Mikalachki, A. & Chapple, D. C. "Absenteeism and Overtime: Double Jeopardy," Relations Industrielles, 1977, Vol. 32.

Miner, Mary Green. "Job Absence and Turnover: A New Source of Data," Monthly Labor Review, 1977.

Mirvis, Philip H. & Lawler, Edward E., III. "Measuring the Financial Impact of Employee Attitudes," Journal of Applied Psychology, 1977, Vol. 62, No. 1, 1 - 8.

Mobley, William H. "Some Unanswered Questions in Turnover and Withdrawal Research," Academy of Management Review, 1982, Vol. 7, No. 1, 111 - 116.

Montgomery, Douglas C. & Johnson, Lynwood A. Forecasting and Time-Series Analysis, McGraw-Hill, Inc., 1976.

Mowday, Richard T. & Spencer, Daniel G. "The Influence of Task and Personality Characteristics on Employee Turnover and Absenteeism Incidents," Academy of Management Journal, 1981, Vol. 24, No. 3, 634 - 642.

Nicholson Nibel; Brown, Colin A. & Chadwick-Jones, J.K. "Absence from Work and Personal Characteristics," Journal of Applied Psychology, 1977, Vol. 62, No. 3, 319 - 327.

Nicholson, Nigel; Brown, Colin A. & Chadwick-Jones J.K. "Absence from Work and Job Satisfaction," Journal of Applied Psychology, 1976, Vol. 61, No. 6, 727-737.

Nicholson, Nigel & Goodge, Peter M. "The Influence of Social, Organizational and Biographical Factors on Female Absence," Journal of Management Studies, October 1976, Vol. 13.

Nicholson, Nigel; Jackson, Paul & Howes Gillian. "Shiftwork and Absence: An Analysis of Temporal Trends," Journal of Applied Psychology, 1978, Vol. 51, 127 - 137.

Nicholson, Nigel; Wall, Toby & Lischeron, Joe. "The Predictability of Absence and Propensity to Leave from Employees' Job Satisfaction and Attitudes Toward Influence in Decision-Making," Human Relations, Vol. 30, No. 6, 1977, 499-514.

Parasuraman, Saroj. "Predicting Turnover Intentions and Turnover Behavior: A Multivariate Analysis," Journal of Vocational Behavior, Vol. 21, 111 - 121, 1982.

Popp, Paul O. & Behlohlav, James A. "Absenteeism in a Low Status Work Environment," Academy of Management Journal, 1982, Vol. 25, No. 3, 677 - 683.

Rousseau, Denise M. "Relationship of Work to Nonwork," Journal of Applied Psychology, 1978, Vol. 63, No 4, 513 - 517.

SAS Institute Inc., SAS User's Guide: Basics, 1982 Edition, Cary, NC: SAS Institute Inc., 1982.

Scott, Dow & Markham, Steve. "Absenteeism Control Methods: A Survey of Practices and Results," Personnel Administrator, June 1982.

Shapiro, S.S. and Wilk, M.B. "An Analysis of Variance Test for Normality (Complete Samples)," Biometrika, 52, 1965, pp. 591-611.

Smulders, Peter G. W. "Comments on Employee Absence/Attendance as a Dependent Variable in Organizational Research," Journal of Applied Psychology, 1980, Vol. 65, No. 3, 368 - 371.

Steers, Richard M. "Employee Turover and Absenteeism: A Future Research Agenda," Invited address, Michigan State University, April 24, 1981.

Steers, Richard M. & Rhodes, Susan R. "Major Influences on Employee Attendance: A Process Model," Journal of Applied Psychology, 1978, Vol. 63, No. 4, 391 - 407.

Steers, Richard M. & Rhodes, Susan R. "A New Look at Absenteeism," Personnel, November - December 1980.

Stephens, M.A. "EDF Statistics for Goodness of Fit and Some Comparisons," Journal of the American Statistical Association, 69, pp. 730-737.

Stevens, S.S. "Mathematics, Measurement and Psychophysics," Handbook of Experimental Psychology, 1951.

Taylor, Daniel E. "Absences from Work Among Full-Time Employees," Monthly Labor Review, Vol. 104, 1981.

Terborg, James R.; Thomas W. Lee; Frank J. Smith; Gregory A. Davis & Mark S. Burbin. "Extension of the Schmidt and Hunter Validity Generalization Procedure to the Prediction of Absenteeism Behavior from Knowledge of Job Satisfaction and Organizational Commitment," Journal of Applied Psychology, 1982, Vol. 67, No. 4, 440 - 449.

Thrall, Tony & Engelman, Laszlo. "Univariate and Bivariate Spectral Analysis," BMDP Statistical Software, 1981 Edition, Ch. 20.1.

Tukey, John W. Exploratory Data Analysis, Addison - Wesley, 1977.