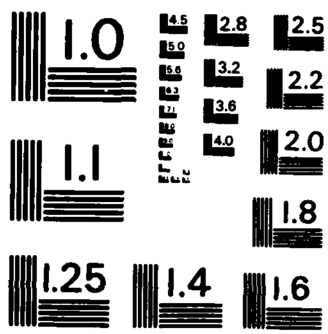
AD-A161 744 RELIABILITY: A COMPARISON OF ABSENCE MEASURES(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYSTEMS AND LOGISTICS K A KENNEDY SEP 85 1/1

UNCLASSIFIED AFIT/GSM/LSB/85S-19 F/G 5/9 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A161 744

RELIABILITY: A COMPARISON OF

ABSENCE MEASURES

THESIS

Kenneth A. Kennedy
Captain, USAF

AFIT/GSM/LSB/85S-19

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GSM/LSB/85

RELIABILITY: A COMPARISON OF

ABSENCE MEASURES

THESIS

Kenneth A. Kennedy
Captain, USAF

AFIT/GSM/LSB/85S-19

Approved for public release; distribution unlimited

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GSM/LSB/85S-19

RELIABILITY: A COMPARISON OF ABSENCE MEASURES

THESIS

Presented to the Faculty of the School of Systems and Logistics of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Kenneth A. Kennedy, B.S.

Captain, USAF

September 1985

Acknowledgments

This study would not have been possible without the help of others. First and most importantly, I would like to thank my advisor, Dr. Robert P. Steel, for providing the topic, data, guidance, and encouragement to finish. Further, I would like to thank my wife, Denise, for all her patience and my children, Kevin, Dawna, and David for the time and attention they sacrificed.

Kenneth A. Kennedy

Table of Contents

List of Figures

List of Tables

Abstract

Absenteeism is an important and costly problem for all organizations. This study investigates the reliability of different absence measures for a sample of Air Force civil service employees to see if they are consistent in the way they use leave. The measures evaluated are sick leave, annual leave, and total leave absence in the form of both time lost and frequency. Internal consistency reliability estimates for each measure are calculated using the Spearman-Brown prophesy formula.

The results show the frequency indices to be very reliable, but that a carefully defined and measured index like sick leave hours lost can provide comparable reliability estimates. The results are also compared to past absence research and evaluated on the basis of measurement differences. The implications of these results are then discussed in terms of absence behavior and organization policy. Finally, an alternative cost cutting recommendation for policy makers and suggestions for future research are presented.

RELIABILITY: A COMPARISON OF ABSENCE MEASURES

I. Literature Review

Introduction

Absenteeism may be defined as undesired work absence. This withdrawal behavior is an important organizational problem because of probable reduced labor productivity and increased training costs. "Absenteeism is almost certainly determined by a variety of social, organizational, economic, and individual factors" (Fitzgibbons & Moch, 1980, p. 349). Some of these factors include worker sex, age, tenure, job level, job satisfaction, group cohesiveness, marital status, number of children, self esteem, and health locus of control (Keller, 1983).

Carefull collection of worker absence data can help to predict temporary labor and training requirements, avoid work schedule conflicts, provide early management signals about organizational problems, and identify habitual absentee workers for appropriate corrective action.

These uses depend heavily on the validity and reliability of the absence measure. Both of these properties relate to the problem of measurement error, but while validity is concerned with actually measuring what is desired, reliability is concerned with the repeatability of measurement (Guion, 1965). A measure may be reliable, but still not valid. Thus, "reliability is a necessary, but not sufficient condition for validity" (Nunnally, 1978, p. 192). A measure cannot predict another variable any better than itself (Guion, 1965). Reliability "is important primarily because it imposes a limitation on validity of measurement and the accuracy of prediction" (Guion, 1965, p. 49).

Operationally, a measure is reliable "to the extent that it supplies consistent results" (Emory, 1980, p. 132) and "one can safely generalize from the results obtained by a measurement method to people in one situation at one point in time to the application of the same, or supposedly comparable, measure of the same trait to the same people in a similiar situation at another point in time" (Nunnally, 1970, p. 108). In other words, reliable measures should be repeatable. "Repeatability of measurement is a fundamental necessity in all areas of science" (Nunnally, 1970, p. 108) and is dependent on freedom from measurement error.

Reliability Theory

Reliability theory is based on the theory of measurement error. The basic assumption of psychological measurement is "that any measure contains an element of error and an element of truth" (Guion, 1965, p. 28). True scores have been defined as containing no errors of measurement, as real but unobserved values obscured by errors of measurement, as an average of scores after infinite measurements, and as determined by the relationship between true and error scores (Nunnally, 1978; Cascio, 1978). "The difference between observed score and true score is the error of measurement" (Cronbach, 1984, p. 159). This error of measurement can be further divided into constant and random components. Constant errors effect each observation the same and are systematic in nature while random errors effect each observation differently and are unpredictable. Thus an observed score (X) can be represented mathematically as a composite of the true score and systematic constant error (s), plus any unsystematic random error (e) (Guion, 1965).

$$X = s + e \qquad (1)$$

This formula can summarize many observations if expanded to breakout score variance. As constant measurement errors produce no variance in the distribution of scores, the composite variance reduces to true variance (Cascio, 1978). Expanded then, the formula becomes

$$S_x^{\,2} = S_t^{\,2} + S_e^{\,2} \qquad (2)$$

where $S_x^{\,2}$ is the total observed variance, $S_t^{\,2}$ is the true variance, and $S_e^{\,2}$ is the unsystematic random error variance (Cascio, 1978). This formula assumes that errors can be positive or negative, will have an expected mean of zero, and will not correlate with true scores or errors from a parallel form of the same test (Guilford, 1954).

Reliability theory is concerned exclusively with random errors and is the degree "to which a set of measurements is free from random-error variance" (Guion, 1965, p. 30). Thus reliability (r_{xx}) can be expressed as the ratio of true variance (S_t^2) to total observed variance (S_x^2) (Cascio, 1978).

$$r_{xx} = S_t^2 / S_x^2 \tag{3}$$

Alternatively, reliability (r_{xx}) can be expressed as one minus the ratio of error variance (S_e^2) to total observed variance (S_x^2) (Guilford, 1954).

$$r_{xx} = 1 - S_e^2 / S_x^2 \tag{4}$$

These formulas show that reliability is interpreted as a coefficient of determination (Guilford, 1954). The conceptual difference is that reliability concerns the amount of error between true scores and observed scores while the coefficient of determination concerns the amount of error between prediced scores and observed scores. However, neither true scores nor error scores can be directly measured (Guion, 1965). Thus, reliability must be estimated from observed scores. This requires repeated measurements "in order to determine the relative proportions of true and error variance in a distribution of scores"

(Cascio, 1978, p. 72). But in practice, infinite retests are infeasible.

To get around this obstacle, reliability may be estimated by the correlation between parallel tests. Parallel tests do not have to be identical, but they must be comparable or equivalent such that they contain the same number of equally difficult items with nonsignificant differences between means, variances, and intercorrelations (Cascio, 1978). The resultant reliability coefficient is traditionally symbolized as $r_{xx'}$ because it "equals the correlation of one measure x with a similiar measure x'" (Cronbach, 1984, p. 160). The square root of the reliabilty coefficient is called the index of reliability and represents the theoretical correlation between true scores and observed scores (Guion, 1965).

With reliability theory introduced, it is now appropriate to discuss how reliability is estimated.

Reliability Estimation

There are three basic forms of reliability estimates. They all involve deriving two sets of scores from the same sample, but they differ in how those scores are obtained and how reliability is interpreted (Guilford, 1954).

The most conservative method is the alternate-forms procedure. It involves administering two parallel or equivalent forms of a measure in close succession. The reliability estimate is called the coefficient of equivalence and identifies the consistency of scores between equivalent or alternate forms. This method can increase measurement error and lower reliability estimates because alternate forms of a measure will contain more items than a single form and there are potential opportunities for changes between administering conditions. It is best used for measures of well-defined traits like math ability or mechanical aptitude (Cascio, 1978). In fact, "Were it not for the expense and the difficulties of making up and administering alternative forms, this method of reliability estimation would almost always be preferable" (Nunnally, 1970, p. 124).

A simpler method is the test-retest procedure. It also requires two test administrations, but only of one form. The reliability coefficient it yields is called the coefficient of stability and identifies the consistency of a measure over time. Differences between administrative conditions may still increase measurement error, but the smaller number of sample questions and the potential memory of first test administration answers due to practice may reduce measurement error and over estimate reliability (Cascio, 1978). This procedure is justified when the development of alternate-forms is not feasible, when test characteristics minimize the effect of memory, and when repeatability is more important than content sampling (Nunnally, 1970). This procedure is not appropriate for most psychological measurements, but may be used for sensory discrimination and psychomotor tests (Cascio, 1978).

The most popular method of reliability estimation is the split-half procedure. This involves subdividing the scores from one administration of a test into two rationally equivalent halves. It assumes functional unity or that "all parts of the total measure are so interrelated that they must all be interpreted as measuring the same thing" (Guion, 1965, p. 42). Random subdivisions are preferred, but odd and even subdivisions are traditionally used. This procedure results in spuriously high reliability estimates because temporary individual characteristics will influence both sets of scores used for correlation (Guion, 1965). The estimate must also be corrected to estimate the reliability of a test of full length because test length effects r_{xx}. The Spearman-Brown prophecy formula used for this purpose is

$$r_{ff} = 2r_{hh} / 1 + r_{hh} \tag{5}$$

where r_{ff} is the prophesied coefficient for a full-length test and r_{hh} is the computed correlation between two half-test measures (Guion, 1965). Other methods of estimating internal consistency, such as Kuder-Richardson formula 20, involve analysis of item variance and are more rigorous, but "produce essentially the same result when used appropriately" (Cronbach, 1984, p. 169). The split-half method is appropriate when using alternate-forms or retesting are not practical, the measure is untimed, and success on one item does not improve success on another (Cronbach, 1984).

Although the most appropriate method depends upon the characteristics of the sample and the aims of the research, reliability estimates may also vary with different measures.

Reliability of Absence Measures

This is an important consideration of any absence study because it sets an upper bound for predictive analysis. Unfortunately, the reliability of absence measures and even the type of index of absence measurement is seldom reported. In fact, Munchinsky (1977) reports that only six of over 70 early studies on absenteeism reported absence measure reliability. As shown in Table 1, these estimates are extremely varied and inconsistent. However, the reliability of frequency indices "appears to be the highest and most consistent across studies" (Munchinsky, 1977, p. 318).

Since these early studies, six more absence studies have specifically reported test-retest relationships which may be construed as absence measure reliability data. As shown in Table 2, frequency indices again provide higher reliability estimates than time lost measures, but the results are still inconsistent across studies.

Some inconsistency in study findings has been attributed to differences between absence definitions, samples of workers, and level of aggregation (Muchinsky, 1977; Spencer & Steers, 1980). However, other causes of

TABLE 1

Early Absence Measure Reliability Estimates

Investigator	Index	Reliability	Method
Turner (1960)	Frequency	.74 (plant 1)	Spearman-Brown
	Frequency	.60 (plant 2)	
Huse & Taylor (1962)	Frequency	.61	Test-Retest
	Attitudinal	.52	
	Severity	.23	
	Medical	.19	
Ronan (1963)	Time Lost	.70	Estimated From Factor Analysis
Chadwick-Jones et al. (1971)	Frequency	.43	Test-Retest
	Attitudinal	.38	
	Other Reasons	.27	
	Worst Day	.20	
	Time Lost	.19	
	Lateness	.16	
	Blue Monday	.00	
Farr et al. (1971)	Times Absent	.39	Spearman-Brown
	Days Absent	.35	
Latham & Pursell (1975)	Frequency	.51	Test-Retest

(Source: Munchinsky, 1977, p. 318)

TABLE 2

Recent Absence Measure Reliability Estimates

Investigator	Index	Reliability	Method
Fitzgibbons & Moch (1980)	Sickness Freq	.66	Test-Retest
	Excused Freq	.50	
	Unexcused Freq	.29	
Hammer & Landau (1981)	Voluntary Freq	.14-.58	Test-Retest
	Vol Days Lost	.12-.53	
	Vol Hours Lost	.08-.51	
	Involuntary Freq	.13-.46	
	Invol Days Lost	.09-.26	
	Invol Hours Lost	.09-.24	
Breaugh (1981)	Frequency	.49-.81	Test-Retest
	Days Absent	.29-.87	
	Supervisor Rating	.50-.69	
Keller (1983)	Unexcused Frequency	.31	Test-Retest
Clegg (1983)	Unauthorized Frequency	.54	Test-Retest
Ivancevich (1985)	Unexcused Freq	.72-.79	Test-Retest
	Co-Worker Rating	.52-.64	
	Unexcused Days	.27-.32	

these inconsistencies result from methodological issues in absence measurement.

Issues In Absence Measurement

Five methodological issues related to absence research are measurement and interpretation, sample normality, criterion contamination, cross sectional versus longitudinal data collection, and bivariate versus multivariate data analysis.

The first and "single, most vexing problem associated with absenteeism as a meaningful concept involves the metric or measure of absenteeism" (Muchinsky, 1977, p. 317). Absenteeism has been defined and measured many different ways. These include the Time-Lost Index representing absence hours or days lost per period, the Frequency Index which reports the number of absence occurrences per period, the Blue Monday index calculated as the difference between the number of workers absent on Friday and Monday, the Worst Day Index calculated as the difference between the most and least workers absent in a week, the Attitude Index calculated as the number of one day absences, and the Lateness Index reported as the number of lateness incidents (Chadwick-Jones, Brown, Nicholson, & Sheppard, 1971).

The two most common indices are time lost and frequency while all other measures have been viewed as just variations or combinations of these two (Smulders, 1980). However, even these indices have measurement and interpretation problems. The frequency index does not distinguish between a one-day, one-week, or one-month absence. This lowers the significance of a long absence and thus provides a better representation of voluntary behavior. It may help to predict occurrence of absence, but not duration. On the other hand, the time lost index does not distinguish between voluntary and involuntary absence as each incident has a different significance depending on actual time lost (Breaugh, 1981; Garrison & Muchinsky, 1977). Although it does not provide an inherent absence interpretation, it may help to identify the duration of absence. Thus, both indices are required to determine absence frequency and duration, but either can be used for predictive correlation. As the causes of voluntary absence are more likely to be consistent over time than the causes of other absence, the frequency index is expected to be a more stable or reliable measure (Breaugh, 1981). This is confirmed by most research, but the reasons may have more to do with the properties of sample absence data than the behavioral causes.

Hammer and Landau (1981, p. 580) found frequency indices to have "fewer psychometric deficiencies" than time lost indices. These deficiencies stem from the sample distribution characteristics of absence measures. Absence data is usually positively skewed and truncated because most people are seldom absent. This means that sample statistics may be significantly different from population statistics and that the normality assumption used in correlation and regression models is violated. This can constrain or depress correlation coefficients, make regression weight significance tests meaningless, and produce regression lines that predict negative absence. Thus, alternative regression models may be appropriate, even if only to check the results for differences from ordinary least squares regression models (Hammer & Landau, 1981).

A third problem is criterion contamination. This occurs because personal absence may be due to a myriad of reasons, but must be reclassified into company categories (Latham & Pursell, 1975). The two major categories are voluntary and involuntary absence. Other categories of absence include excused or unexcused, authorized or unauthorized, certified or noncertified, and legitimate or illegitimate absence. Voluntary absence has been related to worker motivation while involuntary absence has been related to the worker's ability to attend (Steers & Rhodes, 1978). The company must try to correctly classify the real reason for absence based on worker control or necessity, while "the ultimate definition of inability to attend rests with the worker" (Hammer & Landau, 1981, p. 575). However, even an apparent voluntary absence may really be involuntary due to mounting family pressures or work stress (Fitzgibbons & Moch, 1980; Hammer & Landau, 1981). Thus, short term absence is especially susceptible to classification errors and should be interpreted carefully.

The fourth issue concerns research design. Many studies have investigated cross sectional relationships between absence and other variables, but few studies have used longitudinal designs to investigate the stability of findings over time (Fitzgibbons & Moch, 1980). This is important because the effect of any short term factor is minimized and the reliability of a measure has been shown to increase over longer time intervals (Hammer & Landau, 1981).

The last issue concerns data analysis. Although absenteeism is a function of many variables, most cross sectional studies analyze only one predictor at a time. The alternative, multivariate analysis, could also be used to determine the relative influence of two or more predictors (Munchinsky, 1977).

Clearly, these internal issues must be resolved through future research, but there are also external measurement problems that prevent interpretation and generalization.

Effects of Organizational Policy on Absence

Organizational policies toward absence may be another source of absence study inconsistency. These policies are the result of financial, technological, strategic, and moral concerns (Johns & Nicholson, 1982) while individual absence decisions involve deterrent and motivating consequences (Morgan & Herman, 1976; Dalton & Perry, 1981). Policy deterrents can include reprimands and counseling, loss of pay, loss of benefits, loss of promotional opportunities, and even employment termination. Motivating consequences may include income level, proof of illness requirement, sick leave accumulation rate, and unused sick leave disposal policies. If the motives for absence are stronger than the perceived deterrents, then absence may occur more often. However, Morgan and Herman (1976) found that organizational policies do not act as deterrents to absenteeism. This can happen if company policy outlines one position on absence, but management action conveys another position. On the other hand, Dalton and Perry (1981) found that policy motivators do make a difference in that organizations with higher wage rates, faster accumulation of sick benefits, and uncompensated but earned sick leave had higher absence rates.

Further, differences in organization policy on types of absence may signal employees to adopt favored excuses. For example, organizations that issue strong policies against absence, but offer employees liberal sick days may encourage absence (Johns & Nicholson, 1982). This would be expected if medical certificates were not required, and was confirmed in two different organizations by Larson and Fukami (1985) who found that paid sick days sanctioned excused absence.

Thus, different organizational policies, the extent to which they are enforced, the potential for compensation, and the way employees perceive the consequences may deter or encourage absence.

Research Objectives

The objective of this thesis was to see if government workers are consistent in the way they use leave by:

1. Tabulating sample absence data for federal civil service employees from management timekeeper records.

2. Calculating Spearman-Brown reliability coefficients from odd and even pay periods.

3. Comparing and analyzing the reliability estimates for time lost and frequency indices.

II. Method

Sample

The subjects is this study were 162 full-time federal civil service employees from an Air National Guard Unit involved in aircraft maintenance. Complete demographics for the total sample were not available, but this data was available for a subset of 103 survey respondents. Of this group, 95 percent were male. Also, the typical respondent had an an average age between 26 and 30 years, and an average tenure between 18 and 24 months.

Civilian personnel absence in the Department of the Air Force is classified as authorized or unauthorized in accordance with Air Force Regulation 40-631 (USAF, 1973). Authorized absence may either be excused or charged to annual leave, sick leave, or leave without pay (LWOP), while unauthorized absence is charged as absence without leave (AWOL). This study only investigated annual leave and sick leave absence as the potential for LWOP or AWOL is small in view of civilian personnel leave administration and absence policies.

Policy

Leave administration for Department of the Air Force civilian personnel is defined in Air Force Regulation 40-631 (USAF, 1971). Generally, full-time Air Force civil service employees can earn up to 208 hours or 26 days of annual leave per year at different hourly rates per pay period depending on seniority and basic workweek hours. Based on a normal 40-hour workweek, employees with less than 3 years of service can earn 104 hours or 13 days per year, employees with more than 3 years but less than 15 years of service can earn 160 hours or 20 days per year, and employees with over 15 years of service can earn the full 26 days per year. This leave becomes available to the employee at the start of the leave year, but maximum annual leave carry over to the next leave year is 30 days. This creates an incentive to plan short term absence under the disguise of annual leave. However, unused accumulated annual leave may be included in a lump sum leave payment upon federal service separation.

All Air Force civilian personnel are authorized 13 days of sick leave per year regardless of seniority. This leave also becomes available at the beginning of the leave year and up to 30 days of advance sick leave may be granted in cases of serious illness or disability. However, there is <u>no absolute accumulation limit</u> on the amount of sick leave employees may accrue, and they may "bank" accrued sick leave for other purposes. This means that accumulated sick leave can be used to retire early without reducing service time annuity payment calculations or can even be used to increase the service time used in calculating retirement and survivor annuity payments. Thus, this policy effectively takes the "use or loose" incentive out of sick leave decision making by creating a strong incentive to avoid sick leave use as the opportunity costs are higher than for annual leave use.

Both annual and sick leave are charged in increments of one hour. Thus, a two year employee with 26 days of accumulated annual leave and 13 days of advance annual leave could take 39 days or 312 hours of annual leave that year. This same employee with 26 days of accumulated sick leave and 30 days of advance sick leave could also take 56 days or 448 hours of sick leave for a total of 95 days or 760 hours in the same year. This means that after only two years, an employee could be absent 19 full work weeks or almost 10 pay periods.

Annual leave use must be approved in advance, but sick leave can be taken without advance approval and does not require a medical certificate for less than three days duration except when sick leave abuse is suspected. Annual leave is authorized for vacations, personal or emergency purposes while sick leave is authorized when incapacitated for performance of duties, to attend medical, dental, and optical examinations or treatment, and to care for immediate family members afflicted with a contagious disease. If an employee does not have enough sick leave to cover the charge, any available annual leave or compensatory time may be substituted. Further, absence for brief periods and tardiness of less than one hour may be excused, charged to annual leave, or written off against earned compensatory time.

With these liberal policies, the lack of leave without pay or absence without leave charges is not surprising.

This means that all reported absence was considered authorized and employees suffered no loss of pay.

<u>Measures</u>

Based on these absence policies, it might appear that the distinctions between annual leave and sick leave classifications could be viewed as component absence variables that combined would be similar to other absence measures found in the research literature. However, a closer look reveals that while other research studies used direct or derived measures based on multiple cause management judgements, Air Force civil service absence policy provides measures based only on medical concerns, advance approval for other reasons, and leave availability. Further, few studies have identified whether vacation time or sick days were included in their measures, and only Garrison and Munchinsky (1977) specifically investigated paid versus unpaid absence. As Hammer and Landau (1981) point out, even measures based on formal contract agreement or organization policy are not perfect. Cross contamination between sick leave and annual leave can occur if unearned sick leave is charged as annual leave or short term personal absence is charged as sick leave, but the opportunity for compounded errors through management judgement is minimized.

Further, regardless of absence classifications or actual causes, all absence is a problem for management (Morgan & Herman, 1976) and must be absorbed or covered by the organization. Thus, to provide a measure of this impact and to create a bridge for comparison against other researh, annual leave and sick leave measures were combined to form a total leave measure. The three basic leave measures of annual leave, sick leave, and total leave were then expressed as both time lost and frequency indices for each employee.

Absence records for each employee were divided into 26 biweekly pay periods. All measures were aggregated into these 26 units of data for every case across the three basic leave variables. Time lost indices were calculated by summing hours of leave across pay periods. Frequency indices were created by counting the loss of any hours during a pay period as one absence incident. Thus, an employee who was absent one hour and another employee who was absent 80 hours would each have a frequency index of one for that pay period. This procedure underestimates the frequency index for an employee who was absent more than once during a pay period, but the transformation seems reasonable because reliability is dependent on long term repeatability and consistency, not short term amplitude.

By collapsing data into biweekly pay periods rather than monthly or yearly periods, total measure variance should be reduced. Hence, the reliability estimates are probably underestimates of the true reliabilities of these absence variables.

Procedure

Absence data were collected from unit time keeper records in the form of hours lost for 26 biweekly pay periods from January 1, 1983 to December 31, 1983. This data was collected in conjunction with a larger investigation.

Analyses

Spearman-Brown internal consistency reliability estimates were computed for annual leave hours lost, sick leave hours lost, total leave hours lost, annual leave frequency, sick leave frequency, and total leave frequency indices as follows.

The data from the 26 pay periods for each employee were divided into two rationally equivalent halves by summing across pay periods partitioned into odd and even pay period sequences. The Pearson product-moment correlations calculated between odd and even scores were then corrected to yield an internal consistency reliability estimate using the Spearman-Brown prophesy formula

$$r_{ff} = 2r_{hh} / 1 + r_{hh} \tag{5}$$

where r_{ff} is the prophesied coefficient for a full-length test and r_{hh} is the computed correlation between two half-test measures (Guion, 1965).

III. Results and Discussion

Total Leave

Tables 3 and 4 summarize the results for total leave hours lost and frequency indices. The raw correlation between odd and even pay periods for total leave hours lost was .54, and the Spearman-Brown correction was .70. The corresponding values for total leave frequency were .77 and .87 respectively. Thus, total leave frequency appeared more reliable than the hours lost index.

Annual Leave

Tables 5 and 6 summarize the results for annual leave hours lost and frequency indices. The even and odd pay period correlation for annual leave hours lost was .33 and the Spearman-Brown prophesy formula correction was .49. An uncorrected Pearson r of .72 was calculated between odd and even pay periods for total leave frequency. The corrected Spearman-Brown r was .84. Here again, the frequency index appeared more reliable, but by a larger difference.

Sick Leave

Tables 7 and 8 summarize the results for sick leave hours lost and frequency indices. The Pearson r for sick leave hours lost was .70 and the Spearman-Brown correction calculated to be .82. The uncorrected sick leave frequency r was .68 and the corrected reliability coefficient was .81. Surprisingly, the time lost metric was marginally greater than the frequency index in this computation.

TABLE 3

Total Leave Hours Lost Index

Statistic	Value
N	162
Odd Mean	89.06
Even Mean	90.56
Odd Std Dev	58.83
Even Std Dev	57.42
Odd Range	0-239
Even Range	0-298
Correlation	.54
Reliability	.70

TABLE 4

Total Leave Frequency Index

Statistic	Value
N	162
Odd Mean	6.12
Even Mean	6.42
Odd Std Dev	3.27
Even Std Dev	3.32
Odd Range	0-13
Even Range	0-13
Correlation	.77
Reliability	.87

TABLE 5

Annual Leave Hours Lost Index

Statistic	Value
N	162
Odd Mean	66.44
Even Mean	66.16
Odd Std Dev	48.93
Even Std Dev	46.54
Odd Range	0-203
Even Range	0-237
Correlation	.33
Reliability	.49

TABLE 6

Annual Leave Frequency Index

Statistic	Value
N	162
Odd Mean	4.75
Even Mean	4.99
Odd Std Dev	2.87
Even Std Dev	2.97
Odd Range	0-13
Even Range	0-12
Correlation	.72
Reliability	.84

TABLE 7

Sick Leave Hours Lost Index

Statistic	Value
N	162
Odd Mean	22.62
Even Mean	24.40
Odd Std Dev	24.46
Even Std Dev	28.99
Odd Range	0-169
Even Range	0-235
Correlation	.70
Reliability	.82

TABLE 8

Sick Leave Frequency Index

Statistic	Value
N	162
Odd Mean	2.40
Even Mean	2.45
Odd Std Dev	2.08
Even Std Dev	2.14
Odd Range	0-08
Even Range	0-10
Correlation	.68
Reliability	.81

Overall, the reliability of time lost indices ranged from .49 to .82 while the reliability of frequency indices ranged from .81 to .87. These results show the frequency indices of absence for Air Force civil service workers to be highly reliable and more consistent than time lost indices. This result agrees with the findings of past research (Munchinsky, 1977; Hammer & Landau, 1981; Breaugh, 1981), but with some interesting differences.

The major surprise was that contrary to the findings of past research (Chadwick-Jones et al., 1971) the sick leave time lost index was found to be quite reliable. Further, unlike all other studies reporting reliability estimates for multiple absence measures, the reliability estimates for sick leave hours lost and frequency were essentially equal. Although the reliability estimate for sick leave hours lost at .82 was actually larger than the .81 estimate for sick leave frequency, the difference is negligible.

Furthermore, the reliability estimates for both frequency and time lost were higher than psychometric tradition in absence research has assumed (Munchinsky, 1977). In fact, this data suggest that sick leave absences are no less reliable than other major criterion variables such as performance ratings.

Another interesting result was that the .87 reliability estimate for total leave frequency was the largest. In comparison, Morgan and Herman (1976) calculated a .70 reliability estimate for total absence. However, these measures may be unsuitable for use as predictor criteria because "it is not legitimate to treat all types of absences as a homogeneous set of behaviors" (Smulders, 1980, p. 369).

Further, although annual leave hours lost had the lowest reliability estimate at .49, the frequency index reliability at .84 was still high. This estimate is even higher than for sick leave frequency and seems unusual considering the broad reasons for annual leave use.

Partial explanation of these results may be due to criterion contamination and the narrow focus of civil service absence classification policies discussed in Chapter II, but other causes may be the index sample distribution characteristics, and the method of reliability estimation.

Figures 1 through 6 show hours lost and frequency distributions for total leave, annual leave, and sick leave variables. The frequency index figures show the distribution of pay periods absent in one unit increments, and the hours lost index figures show the distribution of hours lost as zero or in ten unit increments. The figures also include mean, standard deviation, range, skewness/ symmetry, and kurtosis/peakedness values for each index.

A visual check of each figure shows that all have many zero values and a few large extreme values while the sick leave indices appear most skewed. A further evaluation of each distribution for statistical skewness and kurtosis revealed only the sick leave hours lost data had significant

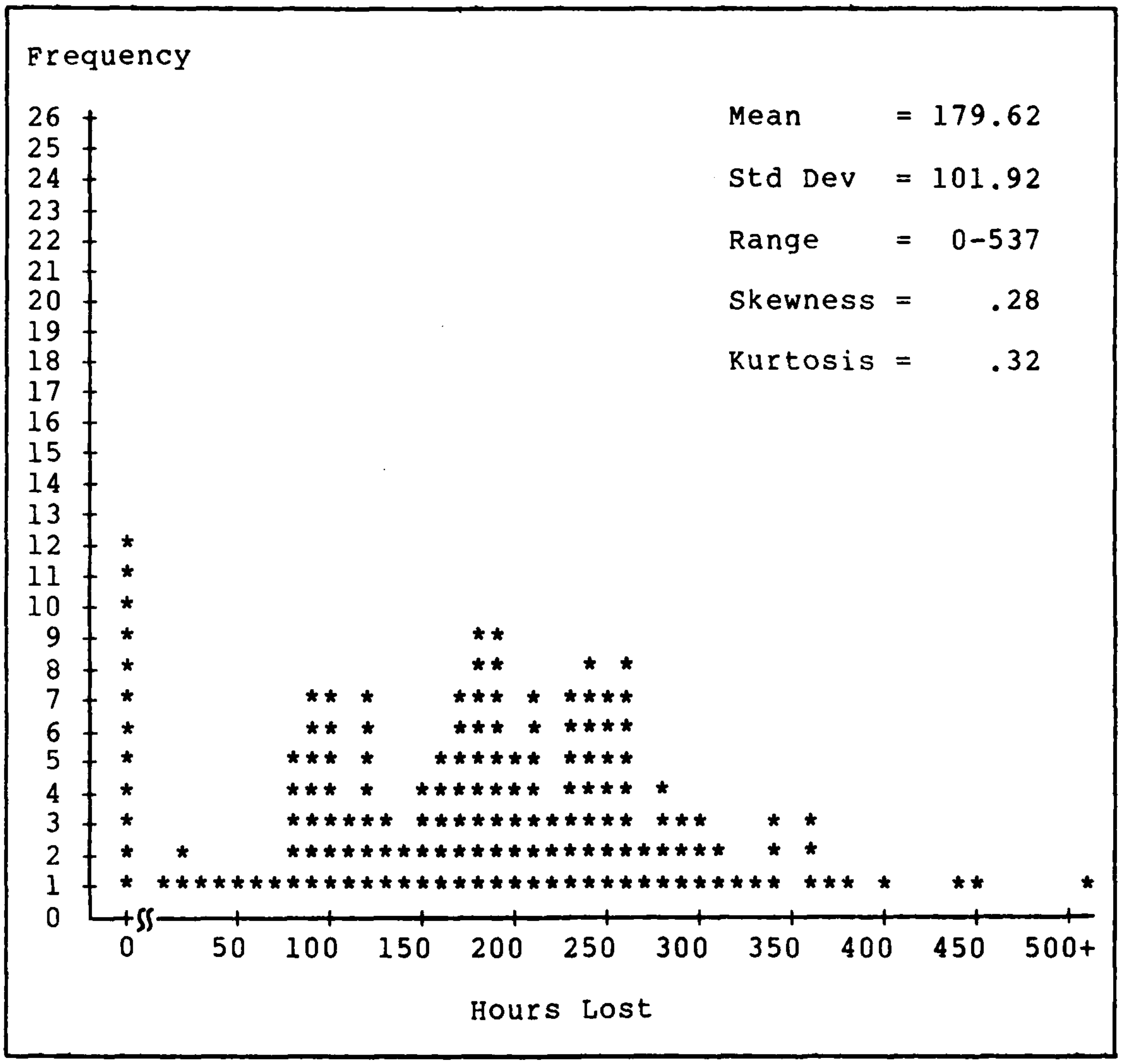


FIGURE 1. Total Leave Hours Lost Distribution

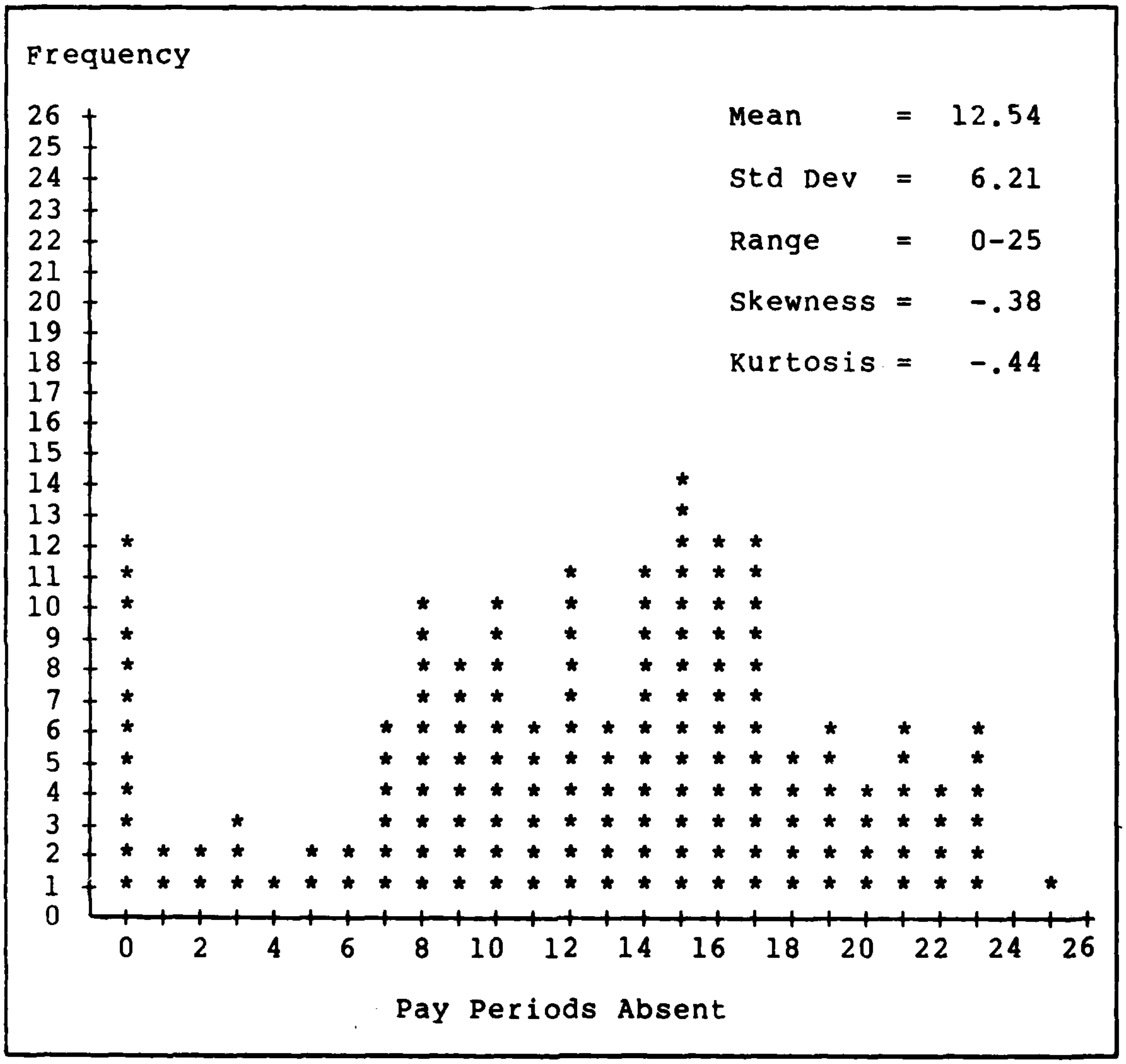


FIGURE 2. Total Leave Frequency Distribution

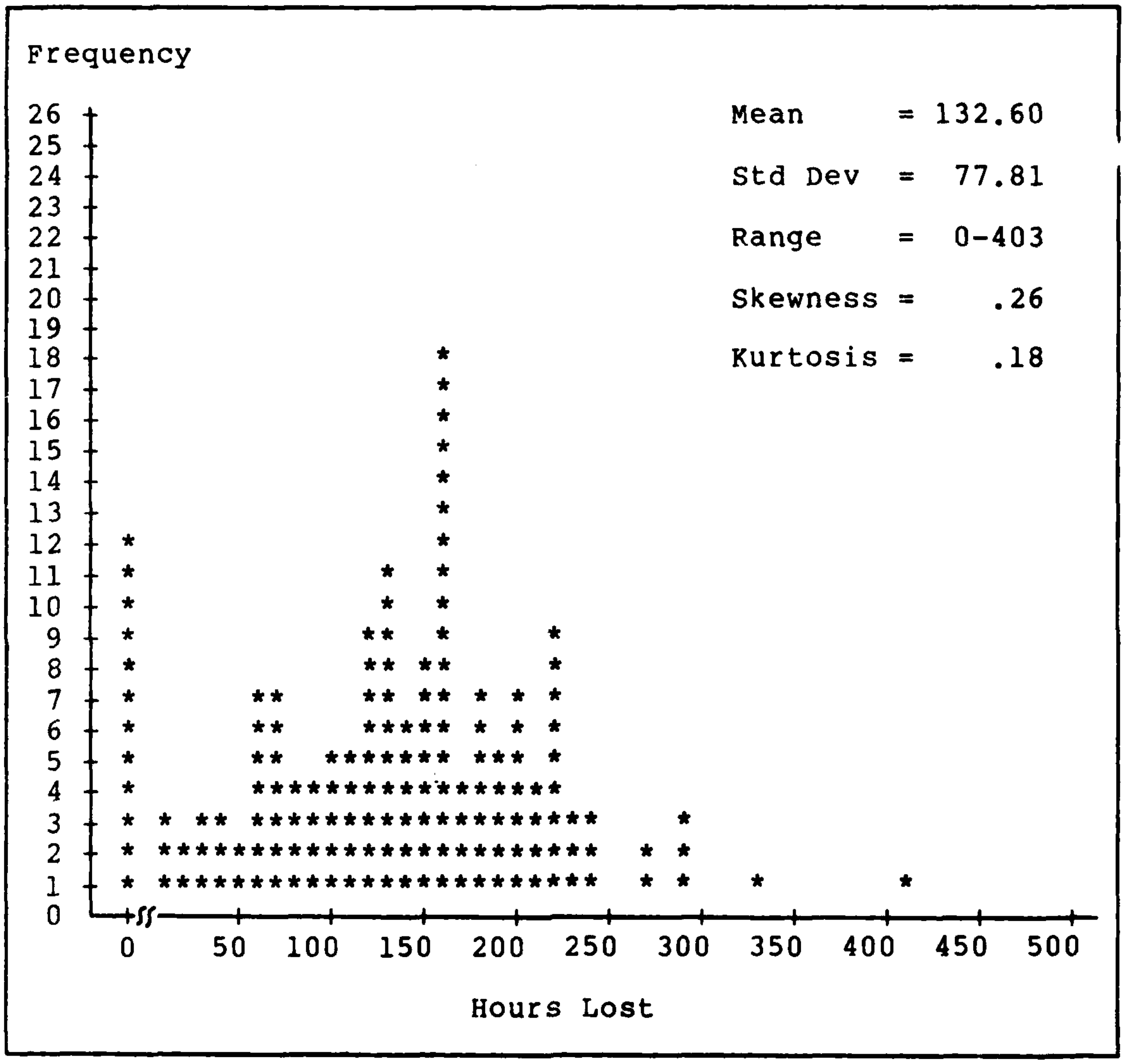


FIGURE 3. Annual Leave Hours Lost Distribution

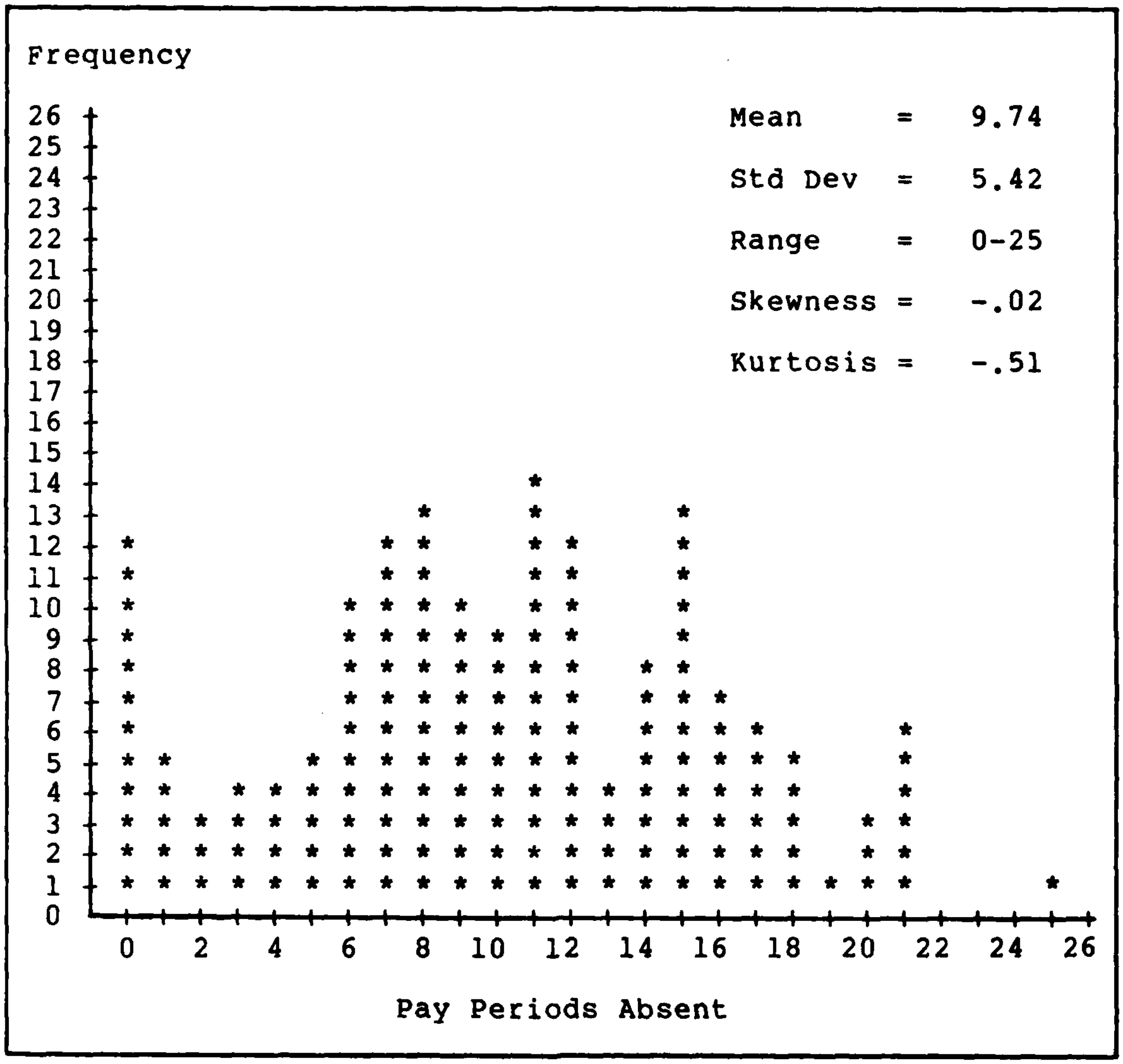


FIGURE 4. Annual Leave Frequency Distribution

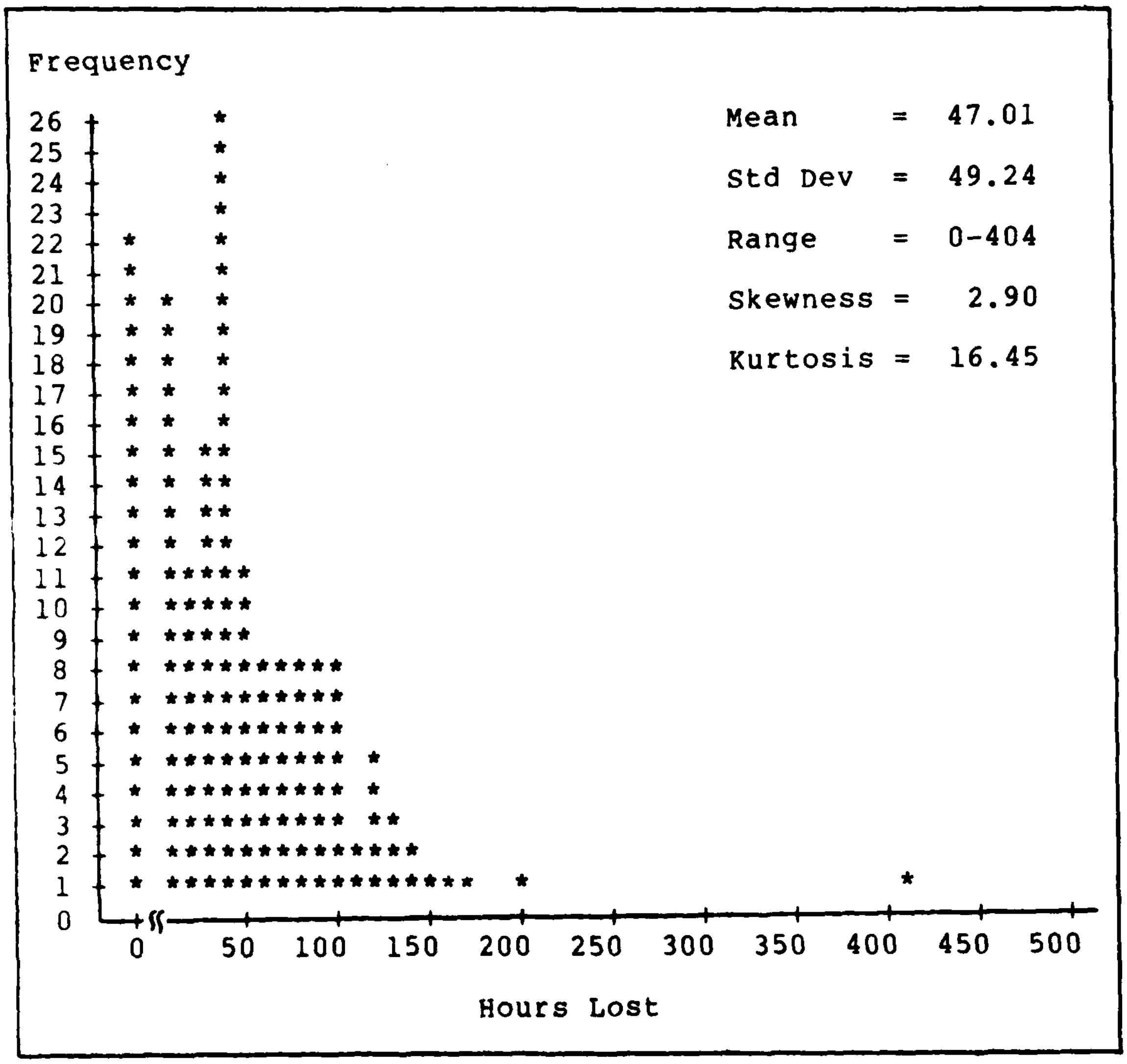


FIGURE 5. Sick Leave Hours Lost Distribution

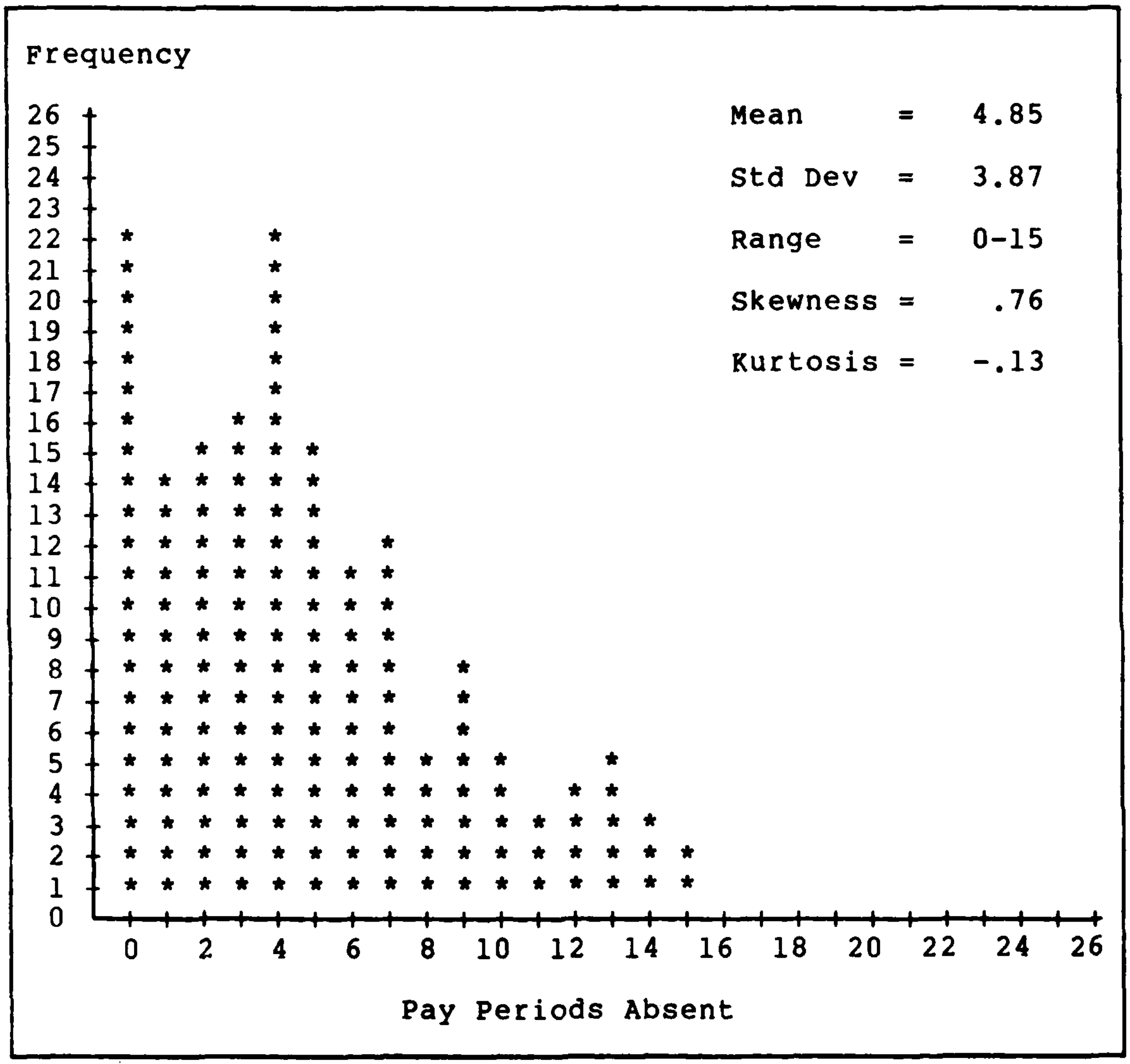


FIGURE 6. Sick Leave Frequency Distribution

deviations from normality. As this can attenuate correlation coefficients rather than inflate them, it does not explain why the sick leave hours lost and frequency reliability estimates were almost equal, but it does imply that the high sick hours lost reliability estimate may be inaccurate because the normality assumption used in correlation was violated. As such, the sick leave hours lost criteria may benefit from some type of statistical transformation prior to conventional predictive correlation. Annual leave hours lost and total leave hours lost data did not contain significant departures from normality. In contrast, Hammer and Landau (1981) found voluntary hours lost, involuntary hours lost, and involuntary days lost indices to all contain considerable skewness and kurtosis.

Another possible explanation for the high reliability estimates obtained in the present study is the method of reliability estimation employed. With the exception of reliability estimates by Latham and Purcell (1975), most past reliability studies have utilized the test-retest method of estimation over two sets of monthly, quarterly, or yearly aggregated absence measures. However, this method of estimation may not be technically appropriate because social, organizational, economic, and individual factors may change significantly between periods (Steel, 1985)

The present study calculated an internal consistency reliability estimate for each absence index by dividing the

26 pay periods of data into an odd-even split, correlating the halves, and correcting for length. This procedure created two rationally equivalent halves that equally distributed the effect of any extraneous factors over the two halves. Although Latham and Purcell (1975) also computed absence reliability this way, their data spaned only 12 weeks and their reliability estimates were based on different 6 week data subsets.

Thus, the high reliability estimates for total leave, annual leave, and sick leave frequency may be due to fewer non-normal sample distribution problems, shorter data aggregation sub-periods, full use of all available data, equal distribution of extraneous factors, and the clear distinction between absence classifications.

IV. Conclusion

Implications

These results lead to several important implications. First, the reliability of absence frequency indices are not inherently higher than time lost measures. In fact, the high reliability estimates for sick leave hours lost and frequency imply that if absence is carefully defined and contamination sources are minimized, both measures can provide comparable reliability estimates.

Further, these results confirm that measures like annual leave and total leave frequency can be highly reliable, yet still unsuitable as criterion variables because of loose definition and poor interpretation.

Finally, the high reliability of sick leave hours lost implies that civil service workers consistently use sick leave as intended or abuse it within the limits of management tolerance. The low sick leave mean of 47 hours lost and the one extreme incident of 404 hours lost implies that Air Force policy incentives to minimize sick leave abuse and reward attendance are generally effective while still providing income protection for those with serious medical problems. In contrast, Morgan and Herman (1976) found that organizational policies did not act as deterrents to absenteeism. However, Dalton and Perry (1981, p. 428) found organizations that "do not remunerate earned, but unused sick leave have higher absence rates". Likewise, Larson and Fukami (1985) found that organizational policies did deter chronic unexcused absenteeism, but that sick leave sanctioned excused absence if attendance was uncompensated. Thus, any Air Force policy changes that attempt to limit sick leave accumulation or create a "use or loose" policy may significantly change the honesty of employee sick leave behavior and increase sick leave absence. The implication for Air Force policy makers is that if serious budget cuts force changes, a more effective alternative might be to reduce the accumulation rate from 13 to 10 days per year for current and/or new employees, but not to limit maximum sick leave accumulation or end of service compensation.

Recommendation

The three greatest limitations of this research were the lack of employee demographic data to match absence records, the small sample size compared to total Air Force civil service force size, and the short data period of one year. In order to broaden future generalizations, future studies should cover multiple years of absence data and larger samples--preferably samples representative of a cross-section of federal civil service personnel.

Bibliography

Breaugh, J.A., (1981). Predicting absenteeism from prior absenteeism and work attitudes. Journal of Applied Psychology, 66, 555-560.

Cascio, W.F., (1978). Applied psychology in personnel management. Reston VA: Reston Publishing Company.

Chadwick-Jones, J.K., Brown, C.A., Nicholson, N., & Sheppard, C. (1971). Absence measures: Their reliability and stability in an industrial setting. Personnel Psychology, 24, 463-470.

Clegg, C.W., (1983). Psychology of employee lateness, absence, and turnover: A methodological critique and an empirical study. Journal of Applied Psychology, 68, 88-101.

Cronbach, L.J. (1984). Essentials of psychological testing. New York: Harper & Row, Publishers.

Dalton, D.R. & Perry, J.L. (1981). Absenteeism and the collective bargaining agreement: An empirical test. Academy of Management Journal, 24, 425-431.

Emory, W.C. (1980). Business research methods. Homewood IL: Richard D. Irwin, Inc.

Farr, J.L., O'Leary, B.S., & Bartlett, C.J. (1971). Ethnic group membership as a moderator of the prediction of job performace. Personnel Psychology, 24, 609-636.

Fitzgibbons, D., & Moch, M. (1980). Employee absenteeism: A multivariate analysis with replication. Organizational Behavior and Human Performance, 26, 349-372.

Garrison, K.R. & Muchinsky, P.M. (1977). Evaluating the concept of absentee-proneness with two measures of absence. Personnel Pyschology, 30, 389-393.

Guilford, J.P. (1954). Psychometric methods. New York: McGraw-Hill Book Company.

Guion, R.M. (1965). Personnel testing. New York: McGraw-Hill Book Company.

Hammer, T.H., & Landau, J. (1981). Methodological issues in the use of absence data. Journal of Applied Psychology, 66, 574-581.

Huse, E.F., & Taylor, E.K. (1962). Reliability of absence measures. Journal of Applied Psychology, 46, 159-160.

Ivancevich, J.M. (1985). Predicting absenteesim from prior absence and work attitudes. Academy of Management Journal, 28, 221-228.

Johns, G., & Nicholson, N. (1982). The meanings of absence: New strategies for theory and research. In B. M. Staw & L.L. Cummings (Eds.), Research in Organizational Behavior, 4, 127-172. Greenwich CT: JAI Press.

Keller, R.T. (1983). Predicting absenteeism from prior absenteeism, attitudinal factors, and nonattitudinal factors. Journal of Applied Psychology, 68, 536-540.

Larson, E.W., & Fukami, C.V. (1985). Employee absenteeism: The role of ease of movement. Academy of Management Journal, 2, 464-471.

Latham, G.P., & Pursell, E.D. (1975). Measuring Absenteeism from the opposite side of the coin. Journal of Applied Psychology, 60, 369-371.

Morgan, L.G., & Herman, J.B. (1976). Perceived consequences of absenteeism. Journal of Applied Psychology, 61, 738-742.

Muchinsky, P.M. (1977). Employee absenteeism: A review of the literature. Journal of Vocational Behavior, 10, 316-340.

Nunnally, J.C. (1978). Psychometric theory. New York: McGraw-Hill Book Company.

Nunnally, J.C. (1970). Introduction to psychological measurement. New York: McGraw-Hill Book Company.

Ronan, W.W. (1963). A factor analysis of eleven job performance measures. Personnel Psychology, 16, 255-268.

Smulders, P.G.W. (1980). Comments on employee absence/attendance as a dependent variable in organizational research. Journal of Applied Psychology, 65, 368-371.

Spencer, D.G. & Steers, R.M. (1980). The influence of personal factors and perceived work experiences on employee turnover and absenteeism. Academy of Management Review, 3, 567-572.

Steel, R.P. (1985). Associate Professor of Organizational Behavior, Air Force Institute of Technology. _Psychometric theory and the problem of relating prior and subsequent absences_. Manuscript submitted for publication.

Steers, R.M. & Rhodes, S.R. (1978). Major influences on employee attendance: A process model. _Journal of Applied Psychology_, _63_, 391-407.

Turner, W.W. (1960). Dimensions of Foreman Performance: A factor analysis of criterion measures. _Journal of Applied Psychology_, _44_, 216-223.

United States Air Force (1971). _Civilian personnel: Leave administration_. AFR 40-630. Washington: HQ USAF.

United States Air Force (1973). _Civilian personnel: Policies relating to specific types of absence_. AFR 40-631. Washington: HQ USAF.

Waters, L.K., & Roach, D. (1979). Job satisfaction, behavioral intention, and absenteeism as predictors of turnover. _Personnel Psychology_, _32_, 393-397.

VITA

Captain Kenneth A. Kennedy was born on 25 August 1954 in Wichita, Kansas. He graduated from Macanaquah High School at Bunker Hill, Indiana, in 1972 and enlisted in the USAF as a missile systems analyst specialist in October 1973. He attented Northern Michigan University part time while assigned to K.I. Sawyer AFB, Michigan and graduated in December 1980 under program bootstrap. After graduation, he received a USAF commission through OTS in April 1981. He served as a propulsion system test manager in the Deputy for Propulsion on the F110 Alternate Fighter Engine and in the Deputy for F-16 as the F110 populsion system integration program manager at Aeronautical Systems Division, Wright Patterson AFB, Ohio, until entering the School of Systems and Logistics, Air Force Institute of Technology, in May 1984.

Permanent address: 4041 Alconbury Cr.
Pensacola, FL 32514

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A161744

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GSM/LSB/85S-19	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: School of Systems and Logistics; 6b. OFFICE SYMBOL (If applicable): AFIT/LSB	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): See Box 19	

12. PERSONAL AUTHOR(S): Kenneth A. Kennedy, B.S., Captain, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
MS Thesis	FROM ____ TO ____	1985 September	54

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	10	
14	04	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Psychology, Behavioral Science, Correlation Techniques, Reliability, Consistency

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: RELIABILITY: A COMPARISON OF ABSENCE MEASURES

Thesis Chairman: Robert P. Steel
Associate Professor of Organizational Behavior

Approved for public release: IAW AFR 190-1/.
LYNN E. WOLAVER 11 Sept 85
Dean for Research and Professional Development
Air Force Institute of Technology (ATC)
Wright-Patterson AFB OH 45433

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Robert P. Steel	22b. TELEPHONE NUMBER (Include Area Code): 513-255-4549; 22c. OFFICE SYMBOL: AFIT/LSB

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Absenteeism is an important and costly problem for all organizations. This ~~study~~ thesis investigates the reliability of different absence measures for a sample of Air Force civil service employees to see if they are consistent in the way they use leave. The measures evaluated are sick leave, annual leave, and total leave absence in the form of both time lost and frequency. Internal consistency reliability estimates for each measure are calculated using the Spearman-Brown prophesy formula.

The results show the frequency indices to be very reliable, but that a carefully defined and measured index like sick leave hours lost can provide comparable reliability estimates. The results are also compared to past absence research and evaluated on the basis of measurement differences. The implications of these results are then discussed in terms of absence behavior and organization policy. Finally, an alternative cost cutting recommendation for policy makers and suggestions for future research are presented.

U.S Government Printing Office: 1985 — 646-047/20350

END

FILMED

1-86

DTIC