College of Business Administration
Department of Management and Organizations
University of Iowa
Iowa City, IA 52242

FINAL REPORT

Modeling Job Performance:
Is There a General Factor?

by

Chockalingam Viswesvaran
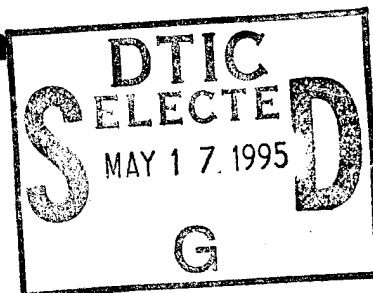
DTIC
SELECTED
MAY 1 7 1995
G

August 1993

Prepared for:

Defense Personnel Security Research Center

Under the

**Office of Naval Research
N00014-92-J-4040**

19950516 050

DTIC QUALITY INSPECTED 5

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS

2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	Unlimited

4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
92-1	

6a NAME OF PERFORMING ORGANIZATION	6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION
University of Iowa		Defense Personnel Security Research and Education Center

6c ADDRESS (City, State, and ZIP Code)	7b ADDRESS (City, State, and ZIP Code)
Iowa City, Iowa 52242	990 Pacific St., Suite 455-E Monterey, CA 93940

8a NAME OF FUNDING/SPONSORING ORGANIZATION Defense Personnel Security Research & Education Center	8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-92-J-4040

8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS			
99 Pacific St., Suite 455-E Monterey, CA 93940	PROGRAM ELEMENT NO	PROJECT NO 36550	TASK NO	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Modeling Job Performance: Is there a general factor?

12. PERSONAL AUTHOR(S)
Viswesvaran, C.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM 9/01/92 TO 6/30/93	June 1993	136

16 SUPPLEMENTARY NOTATION
This report was prepared under an Office of Naval Research Contract

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Personnel selection; Job performance; Construct validity Personnel security; Dimensions of job performance; Intercorrelations

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

Virtually every measurable individual differences variable thought to be relevant to the productivity, efficiency, or profitability of the unit or organization has been used as a measure of job performance in the Industrial-Organizational psychology literature. A question that remains largely unaddressed is the extent to which the

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☑ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Howard Timm, Ph.D.	22b TELEPHONE (Include Area Code) 408-646-2448 22c OFFICE SYMBOL PERSEREC

DD Form 1473, JUN 86 Previous editions are obsolete.

S/N 0102-LF-014-6603

various measures correlate with one another. In this
dissertation, I examined: (a) the true score
intercorrelations between the various dimensions of job
performance; and (b) test whether a hierarchical model
explains the true score correlations between the different
job performance measures.

Psychometric meta-analysis was used to establish the true
score correlation between the conceptually distinct measures
of job performance. Twenty five conceptually distinct
measures were identified by grouping the list of job
performance measures used in the literature into similar
groups. The conceptually distinct measures were further
grouped into five groups. An extensive search of the
literature was conducted to locate all studies that reported
a correlation between the job performance measures. A total
of 2,641 correlations based on 283 independent studies were
found.

A two level hierarchical model in which all job
performance measures were hypothesized to load on a general
factor, and a three level hierarchical model were the job
performance measures were hypothesized to load on one of the
five group factors (which in turn were hypothesized to load
on a second order general factor), were tested.

Results indicated a positive manifold of true score
correlations across the twenty five dimensions of job

performance. Confirmatory factor analyses indicated the presence of a general factor in both the two level and the three level (involving group factors) hierarchical models of job performance. The magnitude of the general factor was larger in the three level hierarchical model. Further, the residuals were smaller in the analysis of the three level hierarchical model (i.e., the three level model better explained the true score correlations between the various dimensions of job performance than the two level model).

Implications of the findings for: (a) theories of job performance, (b) forming composites of job performance measures, (c) differential prediction, and (d) convergent validity of the different sources of performance evaluation, are discussed.

Accesion For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	

Dist	Avail and / or Special
A-1	

Final Technical Report

Modeling Job Performance: Is There a General Factor?

Principal Investigator: Chockalingam Viswesvaran
Thesis Advisor: Professor Frank L. Schmidt

Department of Management & Organizations

College of Business Administration

University of Iowa

Iowa City, 52242

MODELING JOB PERFORMANCE: IS THERE A GENERAL FACTOR

by

Chockalingam Viswesvaran

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Business Administration
in the Graduate College of
The University of Iowa

August 1993

Thesis supervisor: Professor Frank L. Schmidt

Abstract

Virtually every measurable individual differences variable thought to be relevant to the productivity, efficiency, or profitability of the unit or organization has been used as a measure of job performance in the Industrial-Organizational psychology literature. A question that remains largely unaddressed is the extent to which the various measures correlate with one another. In this dissertation, I examined: (a) the true score intercorrelations between the various dimensions of job performance; and (b) test whether a hierarchical model explains the true score correlations between the different job performance measures.

Psychometric meta-analysis was used to establish the true score correlation between the conceptually distinct measures of job performance. Twenty five conceptually distinct measures were identified by grouping the list of job performance measures used in the literature into similar groups. The conceptually distinct measures were further grouped into five groups. An extensive search of the literature was conducted to locate all studies that reported a correlation between the job performance measures. A total of 2,641 correlations based on 283 independent studies were found.

A two level hierarchical model in which all job performance measures were hypothesized to load on a general factor, and a three level hierarchical model were the job performance measures were hypothesized to load on one of the five group factors which in turn were hypothesized to load on a second order general factor , were tested.

Results indicated a positive manifold of true score correlations across the twenty five dimensions of job performance. Confirmatory factor analyses indicated the presence of a general factor in both the two level and the three level involving group factors' hierarchical models of job performance. The magnitude of the general factor was larger in the three level hierarchical model. Further, the residuals were smaller in the analysis of the three level hierarchical model (i.e., the three level model better explained the true score correlations between the various dimensions of job performance than the two level model .

Implications of the findings for: (a) theories of job performance, (b) forming composites of job performance measures, (c) differential prediction, and (d) convergent validity of the different sources of performance evaluation, are discussed.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

A theory is an explanation of a phenomenon in terms of
the processes that cause it (Hunter & Schmidt, 1990a, ch.
1). The development of a theory is an important task in
all sciences (Campbell, 1990a; Schmidt, 1992), and an
explanation of a phenomenon in terms of the processes that
cause it involves identifying functional relations between
constructs. All theories in science mainly concern
statements about constructs rather than specific,
observable variables or measures (Nunnally, 1978, p.85).

An important construct that is widely used in many
theories in industrial/organizational psychology,
organizational behavior, and human resources management
(personnel selection, training, and performance evaluation)
in general, and personnel selection in particular, is the
construct of job performance. Job performance is an
important dependent variable in industrial/organizational
psychology (Schmidt & Hunter, 1992). To date most
researchers focusing on the construct of job performance
have confined themselves to particular situations and
settings with no attempt to generalize the findings. Also
there has been an emphasis on prediction and practical
application rather than explanation and theory building.

The consequence of these two trends has been a proliferation of various measures of job performance in the literature (Smith, 1976). Virtually every measurable individual differences dimension thought to be relevant to the productivity, efficiency, or profitability of the unit or organization has been used as a measure of job performance. Absenteeism, productivity ratings, violence on the job, and teamwork ratings are some examples of the variety of measures used.

A question that remains largely unaddressed is the extent to which the various measures of job performance correlate with one another. The continuing exploration of the wider implications and powers of meta-analytic methods have focused the attention of researchers from situational specificity hypotheses and narrow focus on prediction of job performance to explaining and understanding the psychological processes underlying and determining job performance. Meta-analytic methods have made possible the estimation of population correlation matrices between the various dimensions of job performance. But meta-analysis can go beyond just quantitatively clarifying and summarizing relationships between the various dimensions of job performance. Along with the application of confirmatory factor analysis and latent variable analysis, we can now test hypotheses, theories, and conceptual

explanations for the broad pattern of relationships established by meta-analysis across the various dimensions of job performance.

In this dissertation, I examine: (a) the true score intercorrelations between the various dimensions of job performance; and (b) whether a hierarchical model (to be elaborated later) explains the true score correlations between the different job performance dimensions.

Importance of Establishing the Intercorrelations

Why is it important to estimate the true score correlations between the various job performance dimensions? Establishing the true score correlations between the various dimensions of job performance is important because of the implications for forming composites as well as for the search for common predictors of different dimensions of performance. High correlations imply that the different dimensions of job performance have convergent validity and combining them as one composite measure will be theoretically meaningful (Schmidt & Kaplan, 1971). Further, if the dimensions are highly correlated, then regardless of the particular dimension used as criterion, the same test battery (or set of predictors) will be chosen. That is, differential prediction will be unlikely, thus encouraging the search for common predictors

that can be used as the basis for developing theories of work behavior and human motivation.

Moreover, examining the true score correlations between the various dimensions of job performance is an essential first step in testing and understanding the latent structure of job performance. High intercorrelations between the various dimensions generally support the conclusion that the dimensions have convergent validity (i.e., they measure the same construct) and that the different dimensions will have high factor loadings on one common factor. Examining the true score intercorrelations is an essential first step in testing the latent structure of job performance because mathematically it is possible for two measures to have very low correlation, yet have high factor loadings on a common underlying construct. Nunnally (1978, p. 368) gives an example of two variables each with factor loadings of .50 on the first factor, but with +.50 and -.50, respectively on the second factor; a pattern of loadings that stem from a correlation of zero between the two variables. This underscores the importance of first examining the intercorrelations, before testing for the latent structure of job performance using confirmatory factor analysis. That is, the first and most important step in modeling job performance is to focus on the true score intercorrelations between the various

dimensions of job performance. Using these true score intercorrelations, we can then test competing models of job performance hypothesized to explain the latent structure of job performance.

However, high true score intercorrelations indicating that two variables measure the same construct does not imply that the two variables are conceptually similar. Two variables can correlate highly, measure the same underlying construct, and yet be conceptually distinct. For example, both verbal and quantitative abilities measure General Mental Ability and correlate highly, but are conceptually distinct. Another example is the Raven Progressive Matrices test. The internal consistency measures are in the .90s and factor loadings indicate a highly homogeneous general factor (the factor loadings on perceptual, numerical and spatial factors are negligible) but the items are conceptually distinct. Thus, though high intercorrelations between the measures do not imply that the measures are conceptually similar, researchers who advocate that two highly correlated measures need to be considered separately should provide the justification and rationale for keeping them separate. The key point to note is that there are different levels of analysis. Two measures can be different on one level but interchangeable on another (Hunter & Schmidt, 1990a, p. 481 & pp. 516-524).

The Latent Structure of Job Performance

Why is it important to investigate the latent structure
of job performance? The maturation of a scientific field
is reflected in the relative effort expended on theoretical
analysis and attempts to explain and understand the
underlying processes and mechanisms as compared to the
effort expended on developing practical applications. The
insecurity of a nascent field in the community of sciences
prompts an emphasis on developing practical applications
(Schmidt & Kaplan, 1971). Once the usefulness of a
profession is accepted by the society, researchers indulge
in the luxury of theory building. With increasing emphasis
in I/O psychology on theory building (Campbell, 1990a;
Schmidt, 1992), there is an urgent need to establish the
nomological network of constructs that need to be included
in theories of work behavior and human motivation. Job
performance is perhaps the most important construct and
there is a need to examine the latent structure of job
performance to explicate that construct.

To answer the question whether the different measures
of job performance are manifestations of the same
underlying construct, two lines of evidence are relevant.
First, a positive manifold (a predominance of positive
values in the intercorrelation matrix) of true score

correlations among the different measures would indicate
the presence of a common underlying construct. A positive
manifold of true score correlations between the various
measures indicates the presence of a general factor, the
magnitude of which indicates how well the various measures
"go together" in empirical investigations. Second,
evidence that the various measures have similar
relationships with other well defined constructs would
strengthen the inference that the various measures are
manifestations of a common construct. More specifically,
if other constructs (e.g., General Mental Ability and
Conscientiousness) can be shown to be the common causal
agent of the many different measures of job performance, we
can infer that there is a common underlying construct
across the different measures of job performance (i.e., the
construct of job performance is explicated as referring to
the set of different measures of job performance included
in the analyses). The first line of evidence focuses on
the internal structure of the construct of job performance
whereas the second line of evidence explores the cross-
structure of measures of job performance with measures of
other constructs. In this dissertation, I establish the
true score intercorrelations between the various measures
of job performance, and test whether a hierarchical model
of job performance can explain the intercorrelations.

A distinction is made in this dissertation between an economic construct and a psychological construct. An economic construct of job performance focuses on the economic end products of job behaviors and not on the behaviors themselves. All performance measures can be expressed in economic units (i.e., dollars and cents) and much of recent research in utility analysis (e.g., Judiesch, Schmidt, & Mount, 1992) has focused on identifying the cognitive processes underlying the scaling of performance in different metrics to a common dollar metric.

The ubiquitous, general economic factor in the various measures of job performance is the basis for combining the measures into an overall composite measure regardless of their intercorrelations. In fact, Schmidt and Kaplan (1971) had argued that even when multiple independent measures have been used, for purposes of practical decision making, the assessments of individuals on various dimensions must be combined into an overall composite measure. Failure to develop explicit methods of combination merely relegates the combination of measures to one of ad hoc subjectivity on the part of the decision maker. All decisions in the real world, such as selection, are based on an overall composite measure. A similar idea underlies the arguments of writers (Krug, 1961; Stone &

Kendall, 1956; Tiffin & McCormick, 1958) favoring the composite measure. Terms such as ultimate criterion, job success, standard of reference or yardstick to evaluate overall performance across all measures of job performance, reflect the theme of an underlying economic construct across the different measures.

A psychological or behavioral construct, on the other hand, examines the homogeneity of the various measures of job performance. A psychological construct addresses the question of the extent to which the various measures "go together" or cluster together. Basically, a psychological construct focuses on the common variance represented by the true score intercorrelations between the different measures of the construct under investigation, and examines how this common variance can be defined so as not to cause confusion with the definition of the common variance of a different set of behavioral measures (representing another psychological construct). The focus in this dissertation is not on the common underlying economic construct, but on the question of whether there is a common underlying psychological construct across all measures of job performance. That is, the focus in this dissertation is on examining the psychological construct of job performance.

In investigating whether there is a common underlying psychological construct across the various measures of job

performance a parallel can be drawn to the situation in the domain of abilities, where prior to the work of Spearman, the popular anarchic view postulated that performance on any task is dependent on its own specific ability. This democratization of abilities (Gottfredson, 1988) satisfied many and was expedient to advocate. Individuals scoring low on one ability could always console themselves as being high on other abilities. But based on empirical data indicating the positive manifold of correlations among the abilities, Spearman (1904) established the universal unity of the Intellectual function, showing that the various forms of mental activity constitute a stable interconnected hierarchy according to their different degrees of intellective saturation. Spearman tested a two level hierarchical model involving a general factor (common to all forms of mental activities) at the first or highest level, and specific factors (specific to each form of mental activity) at the second or lower level. The two level theory of Spearman was found to be too simple and subsequent modifications and refinements have converged on a three level hierarchical model involving a general factor, group factors, specific factors, and random measurement error. In a three level hierarchical model, the general factor is at the first or highest level, group factors are at the second level, and specific factors are at the third or lowest level of the hierarchy.

The literature in job performance is fragmented and
mirrors the state of the literature in the predictor domain
that existed before Spearman (1904). One aim in this
dissertation is to examine whether the true score
correlation between the various measures of job performance
can be explained by a two level hierarchical model
involving a general factor common to all measures along
with specific factors for each measure. That is, the
confirmatory model tested in this dissertation views the
various measures of job performance such as quality and
quantity of work performance, absenteeism, turnover,
violence on the job, and teamwork as the manifestations of
a general construct of job performance. This hypothesis
can be stated in factor analytic terms as follows. There
is a general factor common to all the measures of job
performance and an individual's standing on a particular
measure of job performance is the sum of the individual's
standing on this general factor plus that individual's
standing on the specific factor (specific to that measure
of job performance).

In addition to testing a two level hierarchical model,
I also test a three level model. In a three level model--
just as the response of an individual to a questionnaire or
test item is dependent on a general factor, group factor,
specific factor, and a random error component-- the

standing of an individual on any specific performance measure (e.g., absenteeism) is <u>hypothesized</u> to depend on the general factor (i.e., overall job performance), the group factor (e.g., the withdrawal behavior of the employee; absenteeism, tardiness, time theft, turnover all may belong to this group), the specific factor (i.e., absenteeism), and a random error component. That is, the hypothesized latent structure postulates a general factor at the highest or first level, which causes the performance on the group factors at the second level. The group factors, in turn, act as causal antecedents to the various job performance measures. The general factor and the group factors are latent variables. At the third or the lowest level are the various job performance dimensions, which are the observed or manifest variables.

The group factors in the three level hierarchical model were formed based on a review of the job performance literature. However, it should be noted that the group factors identified and included in the three level model tested here reflect my interpretation of the literature. Alternate three level models with a different set of group factors can be hypothesized. Using the true score correlations reported in this dissertation the alternate models can be empirically tested. That is, in this dissertation, I subject the matrix of true score

intercorrelations between the various dimensions of job performance to confirmatory factor analysis to test a two level hierarchical model of job performance; in addition, based on my review of the literature I present and test one possible three level hierarchical model.

An issue to consider in testing the hypothesized hierarchical latent structure of job performance is one that is predominantly discussed in performance evaluation: halo. The various measures of job performance can be broadly grouped as those involving subjective judgments (ratings or rankings) or objective records. The study of subjective judgments (either rankings or ratings) of individuals on various measures has lead researchers to the issue of halo (King, Schmidt, & Hunter, 1980). Halo has been conceptualized in the literature as either: (a) a general impression that produces excessively high correlations between rated dimensions (e.g., a correlation of .90 between personal appearance and performance) when both the ratings are provided by the same individual; or (b) the idiosyncratic overall evaluation of a rater (and this idiosyncrasy is what makes interrater agreement lower than intrarater agreement). To unconfound the effects of halo from a potential general factor across the ratings of various dimensions (when testing for the hypothesized latent structure of job performance), I used the

correlations between (a) job performance measures based on organizational records, and (b) between ratings of various job performance dimensions when the two ratings being correlated are provided by different raters. The appropriate reliability distribution to be used in correcting each type of correlation, along with an elaboration of the issue of halo, is provided in the methods section.

The outline of the dissertation is as follows. Following this introduction, in the second chapter, I review the empirical studies examining the intercorrelations between the various dimensions of job performance. In the third chapter, the methods used in this dissertation are explained. Information regarding the data sources used as well as information on the artifact distributions used are provided. The results of the meta-analyses are presented in the fourth chapter whereas the fifth chapter discusses the implications of the findings for theory and practice.

CHAPTER II

LITERATURE REVIEW

In this chapter, I organize a review of the job performance literature in three sections. In the first section, I summarize the results of studies that examined the correlations between the various dimensions of job performance. In the second section, I summarize the results of previous attempts to explicate the construct of job performance. Finally, in the third section, I outline the limitations of each approach for investigating and drawing inferences about the relationships between the various dimensions of job performance. I conclude this chapter with a discussion of how I addressed these limitations in this dissertation.

At the outset, based on the specification of the domain of interest, it should be obvious that the number of studies that has to be included in this review is numerous. As such this review is necessarily selective and aims to sketch what in my view reflects the important developments in the literature on job performance. In fact, all qualitative reviews are highly idiosyncratic (except maybe when the review focuses on a very narrow and clearly specified domain where it is possible to review all studies). Specifically, in examining the studies that

focused on intercorrelations between dimensions of job performance, no attempt is made here to summarize all the individual studies that report intercorrelations between dimensions of job performance. (To do so would be to conduct a traditional, qualitative, narrative review before conducting a quantitative meta-analytic review of the same database.) Finally, Austin and Villanova (1992) provide a qualitative review of the measures used as criteria in validation studies (including studies conducted in vocational and educational settings in addition to studies conducted in the workplace), whereas this review is limited to studies conducted in workplace settings.

Interrelations Between Job Performance Dimensions

Many researchers have examined the interrelations between the various dimensions of job performance. In this section, I review the studies that either: (a) obtained measures of individuals on more than one dimension of job performance and estimated the correlation between the dimensions; or (b) analyzed the jobs done (by critical incident techniques, task surveys of incumbents, or supervisors, or subordinates, direct observation of incumbents performing the tasks, etc.) and identified the different dimensions of job performance; or (c) used meta analytic techniques to cumulate the results across studies

reporting correlations between a particular pair of job performance dimensions.

Factor Analysis of Correlations

In a typical factor analytic study, individuals are assessed on multiple measures of job performance. Correlations are obtained between the measures of job performance and factor analysis is used to identify the measures that cluster together. Reviewed below are a sample of studies utilizing the framework of factor analysis.

Rush (1953) analyzed 9 rating measures and 3 organizational record based measures of job performance for 100 salespersons and identified four factors of job performance. He named the four factors as: objective achievement, learning aptitude, general reputation, and proficiency of sales techniques. Baier and Dugan (1957) obtained measures for 346 sales agents on 17 variables. Analysis of the 17 X 17 intercorrelation matrix provided support for a general factor of job performance. Various measures such as percentage sales, units sold, tenure, knowledge of products, loaded on this general factor.

Ronan (1963) conducted a factor analysis of 11 job performance measures including accidents and disciplinary actions and found four distinct factors of job performance

using Thurstone's centroid method. One of the four factors reflected the "safe" work habits of the individual (index of injuries, time lost due to accidents are examples of some measures that loaded on this factor). Acceptance of authority and adjustment were two other factors that were identified. Finally, a difficult to interpret fourth factor mirroring overall evaluation also emerged. Prien and Kult (1968) factor analyzed 23 measures of job performance including errors, productivity, knowledge and skills ratings, and found evidence for 7 factors of job performance.

Gunderson and Ryman (1971) examined the job performance in extremely isolated groups (scientists spending the winter in Antarctica) and found three factors of job performance. The three factors they interpreted were task efficiency, emotional stability, and interpersonal relations. An analysis of the measures they used brings home the point that any factor can be identified and interpreted as long as an appropriate set of observable measures are included in the factor analysis. What is required is an a priori (comprehensive) specification of the construct domain, estimation of the true score correlations between conceptually distinct measures, and an examination (using confirmatory factor analysis) of whether

the hypothesized latent structure explains the true score correlations.

Identifying Job Dimensions Based on Inventory of Tasks

This section focuses on studies that examined the relationships between the various dimensions of job performance by describing the domains of organizational behavior thought to be important for organizational effectiveness. Job analysis, critical incident techniques, incumbent interviews, content analysis of performance feedback episodes, observation of incumbents on the jobs, and task surveys are some techniques employed in these studies. Of these, task surveys, critical incident techniques, and observation of incumbents are the widely used techniques. The idea is to generate a list of tasks performed by the incumbents and then attempt to identify (using factor analysis) from these task statements the different dimensions of job performance.

Task Surveys

Task surveys (of incumbents, supervisors, or subordinates) represent a common technique used to identify the various dimensions of job performance. In a typical study, task statements are presented to the sample of respondents (incumbents, or supervisors, or subordinates) who rate the importance and frequency of the tasks. For

example, Brumback and Vincent (1970), based on 4,000 brief, narrative job descriptions extracted 8,000 task statements. A sample of 100 officers from United States Public Health Service (USPHS) representing various positions reviewed the 8,000 task statements and selected 196 duty/work/task statements as descriptive of the various tasks performed by incumbents. Another sample of 5,000 officers (of which 3,719 responded) were then asked in a survey to rate, using a checklist format, the importance, frequency of occurrence, and amount of time spent on each of the 196 duties. A principal components analysis, including a normalized varimax rotation of the product moment correlation matrix of the 196 items, resulted in 26 factors (factor extraction was continued until the latent root of the last extracted factor fell below unity or until the latent root of the last extracted factor accounted for less than one percent of the sum of variance up to that point).

Allan (1981) obtained responses from 1,476 managers (1,550 were surveyed) as to the importance of and time spent on 146 task statements. The 146 task statements were edited by a panel of 52 managers (26 males) and was pretested on 40 managers representing both sexes, various ethnic groups, hierarchical levels and functions. The factor analysis identified 6 factors that were sufficient to describe the construct of job performance.

Observation of Incumbents

Komaki, Zlotnick, and Jensen (1986) used observational techniques to identify the dimensions of job performance. Trained observers were used to obtain information on the time spent by incumbents on various tasks. Questionnaire methods used to identify the dimensions of job performance have employed as respondents either supervisors or incumbents or subordinates. Though each source (supervisors, incumbents, and subordinates) provides a unique perspective, most studies have employed incumbents. Direct observation not only captures what incumbents say and do, but it also avoids self-report biases inherent when the questionnaire method is used with incumbents (cf. Rush, Thomas, & Lord, 1977). Further instead of developing a taxonomy of tasks post-hoc based on raw data using factor analytic techniques or delphi processes, Komaki et al. (1986) developed their factors based on operant conditioning theory. Two field tests using samples of bank managers and theater supervisors identified 7 factors of job performance. One problem with this approach is that not all behaviors of importance to effective job performance are visually observable.

Critical Incident Techniques

Hedberg (1989) content analyzed performance feedback episodes obtained (by employing interview and critical incident methodology) from a sample of financial analysts and project managers. The results paralleled the eight factor latent structure of job performance as hypothesized by Campbell (1990b) and Campbell, McCloy, Oppler, and Sager (1992). Hough (1984) used the method of critical incidents to develop a list of 569 examples of effective and ineffective job behavior for attorneys. When 3 psychologists were asked to sort these 569 critical incidents into tentative categories and describe each category, a taxonomic structure based on 11 dimensions of job performance resulted. Further, in the job analysis phase of the study (Hough, 1984), 326 attorneys rated the time spent and importance of various task activities, factor analysis of which yielded six dimensions of job performance.

Borman and Brush (1993) collected 26 sets of managerial performance dimensions. Several management levels as well as different organizations were included. The 26 sets of managerial performance dimensions resulted in 187 non-redundant task statements which were administered to 25 industrial-organizational psychologists. The I/O

psychologists were required independently to sort the 187
statements into groups based on similarities among the
statements. A 187 X 187 similarity matrix was constructed
(Rosenberg & Sedlak, 1972) based on the proportion of
judges sorting each pair into the same group. Factor
analysis of this similarity matrix identified 18 orthogonal
(and interpretable) dimensions of job performance.

Meta-Analytic Cumulation Over Pairs of Job Performance Measures

In this section, I review the studies that employed
meta-analytic techniques to investigate the convergent
validity between pairs of job performance dimensions. In a
typical study, correlations reported in the literature
between a particular pair of dimensions are obtained and
.mulated across samples. The true score correlation
computed after correcting for the various statistical
artifacts is taken as an index of the convergent validity
between the two dimensions. The literature search
identified five studies that used this approach.

Heneman (1986) cumulated results across 23 studies to
assess the relationship between supervisory ratings and
results-oriented measures of job performance. Based on a
total sample of 3,178, he found the corrected mean
correlation to be .27. Corrections were made for

unreliability in the supervisory ratings (used the value of .60 provided by King et al. [1980]) and for unreliability in the output measures (used a mean--across eight values-- test-retest value of .63).

Harris and Schaubroeck (1988) based on a total sample size of 2,643 (across 23 correlations or samples) found the true score correlation between peer and supervisory ratings of overall job performance to be .62. Corrections were made for measurement error in both peer and supervisory ratings using internal consistency measures that were substantially higher than the value of .60 (the interrater reliability) that is recommended by Schmidt and his colleagues. (Using the value of .60 for reliability, results in a true score correlation of .71.) Harris and Schaubroeck (1988) also found no evidence for the moderating influence of rating method. The true score correlation across 14 samples involving 1,331 data points using the dimensional method of rating was .57. The corresponding correlation across 15 samples involving 1,914 data points using the global method of rating was .65.

McEvoy and Cascio (1987) conducted a meta-analysis across 24 studies involving 7,717 data points and found the true score correlation between supervisory ratings of job performance and voluntary turnover to be -.28. Reliability corrections were applied only to the performance measures.

An upper bound value of .60 for the interrater reliability as reported by King et al. (1980) was used to correct the performance measures (Glenn & McEvoy, 1987, p.749). McEvoy and Cascio's (1987) results indicate a substantial negative relationship (poor performers are more likely to have higher turnover rates) between voluntary turnover and job performance.

Bycio (1992) summarized the relationship between absenteeism and job performance. Across 49 samples involving 15,764 data points he found the true score correlation between supervisory ratings (or rankings) of job performance and organizational records of absenteeism (both time lost and frequency measures) was -.29. Based on 28 samples ($N = 7,704$) that had used the time lost measures of absenteeism the correlation was -.26 with supervisory ratings (or rankings) of job performance, whereas across 21 samples ($N = 8,060$) using frequency measures of absenteeism the correlation with job performance was -.32.

Similarly, the true score correlation between non-rating quality of performance indices and absenteeism measures (both frequency and time lost) was -.24 (based on 23 samples and a total sample size of 5,204). When the 23 samples were subgrouped into those using time lost indices of absenteeism (11 samples with a total sample size of 1,649) and those using frequency based measures of

absenteeism (12 samples with a total sample size of 3,555), the correlation with non-rating quality of performance indices became -.28 and -.22, respectively.

The artifact distribution used by Bycio (1992) to correct the supervisory ratings involved ten interrater reliability values (the mean of the square root of the ten values was .82; comparable to the value of .60 provided by King et al. [1980]). For non-rating quality indices, an artifact distribution involving 4 values were used; the mean of the square roots of the four values was .866. The artifact distribution used to correct time lost measures included 30 values; the mean of the square roots of the 30 estimates was .733. The corresponding value for the frequency measures (36 estimates) was .686.

The results reported by Bycio (1992) indicate that both time lost and frequency based measures of absenteeism are substantially and negatively related to both supervisory ratings (or rankings) of performance as well as non-rating quality of performance indices, thus raising the possibility of a general factor common across these four measures of job performance. In addition to summarizing the relationships between the four measures of job performance which seems to indicate the presence of a general factor, Bycio (1992) also presents some causal mechanisms to explain why a general factor should underlie

these measures. Supervisory annoyance with absentees or disruptions caused to work scheduling could explain the negative relationship between absenteeism and supervisory ratings of job performance. The negative relationship between absenteeism and non-rating quality of performance indices can be explained either in terms of employee dispositions or by invoking the notion of progressive employee withdrawal (Sheridan, 1985). Hogan and Hogan (1989) have argued that poor performance (as measured by overall supervisory ratings and non-rating quality of performance measures) and absenteeism are only reflections of the delinquency construct. Ones, Viswesvaran, and Schmidt (in press) found evidence that integrity tests can predict absenteeism, supervisory ratings of job performance and non-rating measures of job performance (i.e., these different dimensions of job performance have a common predictor or causal antecedent).

Mitra, Jenkins, and Gupta (1992) investigated the relationship between absenteeism and turnover. Based on 33 correlations from 17 studies involving 5,316 data points Mitra et al. (1992) estimated the true score correlation between absenteeism and turnover to be .33. Corrections were made for unreliability in absence measures and for dichotomization of absence measures. Further, turnover measures were corrected for unequal N's (between stayers

and leavers across organizations), but not for
dichotomization in the turnover variable (that is, turnover
was conceptualized as a true dichotomy). The reliability
distributions for absenteeism measures were obtained from
Hackett and Guion's (1985) review of the absenteeism
literature (.46 for absence frequency; .66 for time lost
indices). When the database was subgrouped into those
using frequency based measures of absenteeism (22 samples
with a total sample size of 3,841) and those using time
lost measures of absenteeism (9 samples using 1,159 data
points), the true score correlation with turnover in the
two subgroups were .34 and .32, respectively. The results
of the subgroup analysis support the conclusion from the
Bycio (1992) analyses that frequency and time lost measures
have similar relationships with supervisory ratings of job
performance, non-rating quality of performance measures,
and turnover.

Hypothesized Latent Structures of Job Performance

The earliest conceptualizations of job performance
emphasized the economic value of the behaviors of an
individual to the organization (i.e., the economic
construct of job performance). The value of behaviors, not
the relations between the behaviors themselves were
stressed. Given the restrictions and constraints a nascent
field faces in establishing itself as a useful science

(Schmidt & Kaplan, 1971) such an emphasis on economic value is not surprising. For example, Toops (1944) identifies wages, units of production, absences from work, quality of work, tenure, supervisory and leadership abilities as job performance measures and discusses how these measures affect the value of an individual to the organization. Bechtoldt (1947, p.357) states "a criterion is a means of describing the performance of individuals on a success continuum". Brogden and Taylor (1950) explicitly state, "the criterion should measure the contribution of the individual to the overall efficiency of the organization". Horst (1941, p.20) states that, "The measure of success or failure in an activity is what is technically known as the criterion". Finally, in the same vein, Ghiselli and Brown (1948, p.62) state that, "By criterion is meant any attribute or accomplishment of the worker that can be used as an index of his [sic] serviceability or usefulness to the organization that employees him [sic]". As I/O psychology became more accepted as a profession, the focus shifted from predicting behaviors of economic value to the organization to an understanding of the relations between the various measures of job performance.

In the initial years of examining the relationships between the job performance dimensions, the focus was on job performance measures reflecting the productivity of the

individuals. Over time, the changing demographic
characteristics of the workforce and the increasing
complexities of the job tasks necessitated more job- and
organization- specific training. This lead to an emphasis
on time related measures of job performance such as
absenteeism and turnover. Also the emerging literature on
expectancy theory as well as the increasing difficulties
(due to the complexities of the jobs) in evaluating
behaviors focused the attention of I/O psychologists on
measures of effort expenditure.

Along with an emphasis on time-related measures of job
performance, there was a focus on measures of job
performance reflecting the prosocial behaviors of employees
in an organization. Organizational citizenship behavior is
defined as extra-role, discretionary behavior that helps
other organization members perform their jobs (Bateman &
Organ, 1983; Graham, 1986; Smith, Organ, & Near, 1983).
This builds on the notion of the informal organization
(Barnard, 1938) and spontaneous behavior (Katz & Kahn,
1978). Katz (1964) defines prosocial behavior in terms of
the spontaneous, co-operative, helpful, and altruistic
behaviors that are displayed over and beyond formal role
prescriptions. The literature on prosocial and
organizational citizenship behaviors introduced measures of

job performance such as teamwork, compliance, altruism, and
helping behaviors.

Finally, the last ten years have seen an increase in
the prediction of counterproductive behaviors at work
(Sackett, Burris, & Callahan, 1989; Sackett & Harris,
1984). There has been an increased focus on measures such
as theft of cash or merchandise or property, damaging
merchandise to buy it on discount, time theft, breaking
rules, preventable accidents, substance abuse, misuse of
discount privileges, vandalism, physical or verbal abuse of
peers, supervisors, or customers, and so on.

As could be expected, studies that examine the
intercorrelations between all (organizational citizenship
behaviors, counterproductive behaviors, etc.,) dimensions
of job performance is necessarily a recent phenomenon.
Three attempts at theoretical integration are reviewed
below. The first by Campbell (1990b, 1990c) examines the
relationships between the various dimensions of job
performance in a population of jobs. The second by Borman
and Motowidlo (1992) argues specifically for the inclusion
of contextual performance along with the traditional
measures of job performance in theories of job performance.
Finally, the causal model of job performance presented by
Schmidt and Hunter (1992), examining the relations between

the construct of job performance and their causal determinants, is reviewed.

Campbell (1990b) describes the latent structure of job performance in terms of eight general factors. His eight factors include job specific task proficiency, non-job specific task proficiency, written and oral communication skills, demonstrating effort, maintaining personal discipline, facilitating peer and team performance, supervision, and management or administration.

Job specific task proficiency is defined as reflecting the degree to which the individual can perform the core substantive or technical tasks that are central to a job and distinguish one job from another. Non-job specific proficiency, on the other hand, is used to refer to tasks not specific to a particular job but expected of all members of the organization. Demonstrating effort captures the consistency or perseverance and intensity of the individuals to complete the task, whereas maintenance of personal discipline refers to the eschewment of negative behaviors (such as rule infractions) at work. Management/ administration differs from supervision in that the former includes performance behaviors directed at managing the organization that are distinct from supervisory roles. Written and oral communications reflects that component of job performance which refers to the proficiency of an

incumbent to communicate (written or oral) independent of the correctness of the subject matter. The implications of this eight factor latent structure for personnel selection is further elaborated in Campbell et al. (1992).

According to Campbell (1990b), while the relative relevance of the eight factors varies across jobs (e.g., there could be jobs with minimal supervisory performance), the eight factors or some subset of them can describe the highest order latent variables for every job in the occupational domain. Further, he asserts that three factors--core task proficiency, demonstrated effort, and maintenance of personal discipline--are constituents of job performance in every job. Finally, Campbell (1990b) in addressing the nature of lower order factors points out the lack of any theoretical guideline and states that the description of the lower order factors remains a matter of speculation (at one extreme we can hypothesize as many latent structures of job performance as there are jobs).

Borman and Motowidlo (1992) argue for expanding the criterion domain to include elements of contextual performance in addition to the traditional measures of task (or job specific) performance Drawing from the literature on organizational citizenship behavior (Bateman & Organ, 1983; Smith et al., 1983), prosocial organizational behavior (Brief & Motowidlo, 1986; Graham, 1986; Katz &

Kahn, 1978; Organ, 1988), and a model of soldier
effectiveness developed in Project A (Campbell, 1990c),
Borman and Motowidlo (1992) attempt to define the elements
of contextual performance. According to Borman and
Motowidlo (1992), contextual performance cuts across
different jobs (organization-wide, not job specific) in an
organization and includes performance components such as:
(a) persisting with enthusiasm and extra effort as
necessary to complete own task activities successfully, (b)
volunteering to carry out task activities that are not
formally part of own job, (c) helping and cooperating with
others, (d) following organizational rules and procedures,
and (e) endorsing, supporting and defending organizational
objectives.

Finally, the causal model of job performance developed
by Schmidt and Hunter (1992) views conscientiousness and
general mental ability as common causal antecedents of
effort, citizenship behaviors, job performance, skill
acquisition and performance capability. For example,
general mental ability is hypothesized to have an effect on
job skill acquisition and conscientious employees are
hypothesized to spend more time on their jobs, become more
skilled, and engage more in helping other employees. Thus,
two common causal antecedents (General Mental Ability and

Conscientiousness) are expected to predict a wide variety of job performance measures.

Problems and Limitations

The literature on the interrelationships among the various dimensions of job performance has evolved around four streams of research. First, based on a synthesis of the available knowledge scientists have focused on specifying the domain of the construct of job performance. Essentially, this line of research has focused on identifying and specifying the meaning of words associated with the construct of job performance. The second stream of research has examined the interrelationships between the various dimensions of job performance based on information collected regarding task performed on the jobs. Task surveys (of incumbents, supervisors, or subordinates) of the time spent on and importance of various tasks, content analysis of performance feedback episodes, use of critical incident techniques, and (unobtrusive) observation of job performance are some techniques used to collect information regarding the tasks performed on the jobs. The data obtained is usually analyzed through factor analytic techniques or delphi processes to explicate the relationships between the various dimensions of job performance. The third line of research found in the literature on the relationships between the various

dimensions of job performance focuses on obtaining scores for a sample of individuals on some measures of job performance. An analysis of the intercorrelations between these measures is used to examine the factors that underlie the measures included in the analysis. The final strand of research in the literature relevant to the investigation of the correlations between the various dimensions of job performance are studies that employ meta-analytic techniques to cumulate results across studies reporting correlations between a pair of measures of job performance. These studies address the convergent validity of the two measures of job performance, the correlation between which is being cumulated across studies.

The relevance and value of the four streams of research in investigating the relationships between the various dimensions of job performance can be evaluated in terms of the comprehensiveness of the analysis, the empirical verifiability of the hypothesized relationships, the cumulativeness of the empirical evidence available to date, and the extent to which statistical artifacts are addressed before inferences are drawn. Empirical verification is required to test the hypothesized relationships between the set of measures. The approach used to explicate the relationships between the various dimensions of job performance must also cumulate results across studies

examining the correlation between the same pair of job performance dimensions. The effects of numerous statistical artifacts should be addressed and the results across studies cumulated before firm conclusions can be drawn about the relationships between the various dimensions of job performance.

A comprehensive analysis of the psychological construct of job performance must include all measures of job performance (that have been employed in the literature over the years and for which data is available). A central thesis in this dissertation is that none of the previous attempts to explicate the psychological construct of job performance has been comprehensive. A comprehensive specification of the psychological construct of job performance (encompassing all potential measures of job performance) can be obtained by compiling a list of the different measures of job performance used in the individual studies. Parallel to the lexical hypothesis of Galton (Goldberg, 1992) which states that a complete description of the personality of an individual can be found within the list of words in the dictionary (i.e., the adjectives used in the language), I hypothesize that the psychological construct of job performance can be captured by analyzing the various measures of job performance used in the many individual studies over the years. That is,

each individual study focuses on a subset of all possible manifestations of the underlying construct of job performance, but by cumulating across all individual studies it is possible to cover all manifestations of the underlying construct of job performance.

Though rational analysis can encompass the whole domain of job performance, in actual practice rational analysis has been confined to the idiosyncrasies of the individual researcher synthesizing the available knowledge. The zeitgeist (the spirit of the times) and the ortgeist (the spirit of the place) determine to a great extent what measures are considered in the synthesis. Further, the speculative nature of rational analysis makes it impossible to compare quantitatively the various models of the psychological construct of job performance based on rational analysis. As such cumulation or integration of evidence based on rational analysis has to be largely subjective.

Individual factor analytic studies correlating a subset of the various potential measures of job performance facilitate empirical verification of the hypothesized relationships; but they are not capable of resolving the relationships of interest because of the numerous statistical artifacts affecting the individual studies.

Further they do not provide a comprehensive coverage of the psychological construct of job performance.

Finally, the use of meta-analysis has been very limited in the literature. To date, the convergent validity of only six pairs of measures (supervisory ratings of job performance-production records; absenteeism-supervisory ratings of job performance; absenteeism-turnover; absenteeism-production records; absenteeism-organizational records of quality of performance; peer-supervisory ratings of job performance) have been examined using meta-analysis (Bycio, 1992; Harris & Schaubroeck, 1988; Heneman, 1986; McEvoy & Cascio, 1987; Mitra et al., 1992). Further, establishing the true score correlations between measures is only the first step in testing the psychological construct of job performance. The true score correlations should be used as building blocks with confirmatory factor analysis and latent variable analysis to explicate the psychological construct of job performance. What is needed is a research project that comprehensively (i.e., includes all measures that have been used so far in the literature) specifies the psychological construct of job performance, empirically tests the hypothesized relationships between the various measures, removes the effects of statistical artifacts, cumulates research findings across individual studies, and through the use of confirmatory factor

analysis and latent variable analysis infers conclusions regarding the psychological construct of job performance.

Summary Outline of Dissertation

Despite the problems and limitations, the dimensions of job performance found in the literature revolve around five themes. First, there are job performance measures that focus on productivity. Number of units produced, number of transactions completed are some examples of job performance measures that reflect productivity. Second, there are job performance measures that reflect how conscientious individuals are. The effort expended on the job, the quality of production/work, compliance with rules, not being involved in accidents reflect the same theme of how conscientious and meticulous the individual is. Third, there are job performance measures that assess how well an individual interacts with others. Interpersonal relations with others, leadership, administrative skills are some dimensions of job performance that reflect how well an individual interacts with others. A fourth group of job performance measures captures time related indices such as absenteeism and turnover. Finally, a fifth group of measures focuses on overall evaluation of the individuals (i.e., global measures of performance). I will use these five groups to define the group factors in the second level of a three level hierarchical model. The important point

is not whether the group factors identified are absolute; alternate groups of factors can be postulated and alternate three level hierarchical models can (and should) be tested using the true score correlations reported here. (One could hypothesize different predictors for the different group factors; conscientiousness will be the best predictor of effort, agreeableness will be the best predictor of interpersonal relations, etc.)

In this dissertation, I cumulate the results across individual studies reporting correlations between various measures of job performance. The first step is obtaining studies that measures performance of individual employees on more than one measure of job performance. The second step is identifying the various conceptually distinct measures of job performance that have been used in the literature. The third step is establishing the true score correlations between the identified measures. The fourth and final step is testing whether an hypothesized hierarchical latent structure model fits the pattern of true score correlations between the various measures of job performance. I test both a two- and a three- level hierarchical model. In the next chapter, the methods used are presented.

CHAPTER III

METHOD

Database Description

Inclusion Criteria

To estimate the true score correlations between the
different dimensions of job performance used in the
literature, I searched the literature in industrial/
organizational psychology, organizational behavior, and
human resources management. Studies that report the
correlations between the different dimensions of job
performance were obtained. To be included in the database,
a study had to be from a published source. Conference
presentations and unpublished data were not included.

The exclusion of unpublished studies in this
dissertation is only for feasibility of data management.
It is not possible to identify studies that report
correlations between various measures of job performance
without obtaining a copy of the full study. For example, a
published study comparing grievants with non-grievants
(e.g., Sulkin & Pranis, 1967) in an organization reported
the correlation between absenteeism and tenure. Neither
the title nor the abstract of this published study (Sulkin
& Pranis, 1967) indicates that a correlation that could be

of use for this dissertation is reported; only after scanning through the entire article was such a correlation identified and the study was included in the database. A similar search strategy is impossible with all the dissertations and unpublished studies done in I/O Psychology. Therefore, I did not include dissertations, books, and unpublished reports in this research.

This exclusion of unpublished studies is not an endorsement of misleading and erroneous arguments calling for the exclusion of unpublished studies in all meta-analyses on the grounds that such unpublished studies constitute poor quality data. (The converse argument maintains that published studies have a positive bias that overstate the results. Taken together, these two arguments will lead to scientific nihilism [Hunter & Schmidt, 1990a, p.515].) The hypothesis of methodological inadequacy of unpublished studies (in comparison to published studies) has not been established in any research area. In fact, ample evidence exists to prove the comparability of findings of published and unpublished studies in many research areas (Hunter & Schmidt, 1990a, pp. 507-509).

Hunter and Schmidt (1990a, pp. 509-510) present a hypothetical example that illustrates how differences between published and unpublished studies examining the effectiveness of psychotherapy could have been due to

statistical artifacts. Ones et al. (in press) found that
the correlation between the reported validity of integrity
tests and the dichotomous variable indicating published
versus unpublished studies is negligible. In the
literature on the validity of employment tests, impressive
evidence has been accumulated which indicates that
published and unpublished studies do not differ in the
validities reported (Hunter & Schmidt, 1990a, pp. 507-509).
For example, the data used by Pearlman, Schmidt, and Hunter
(1980) was found to be very similar to the U.S. Department
of Labor (GATB) data base used by Hunter (1983) and other
large sample military data sets. Also the mean validities
in the Pearlman et al. (1980) data base are virtually
identical to Ghiselli's (1966) reported medians. Further,
the percent of nonsignificant studies in the Pearlman et
al. (1980) data base perfectly matches the percent of
nonsignificant published studies reported by Lent,
Auerbach, and Levin (1971). Finally, the percentage of
observed validities that were nonsignificant at the .05
level in the Pearlman et al. (1980) data base (56.1% of the
2,795 observed validities) is consistent with the estimate
obtained by Schmidt, Hunter, and Urry (1976), that the
average criterion-related validation study has statistical
power no greater than .50. If selectivity or bias in
reporting were operating many of the nonsignificant
validities would have been omitted, and the percent

significant should have been higher than 43.9%. On the other hand, if unpublished studies were of poorer quality, not meeting the standards of peer review, then there should have been more non-significant validities among the unpublished studies than 56%. Thus, there is ample evidence arguing for the equivalence of published and unpublished studies. The two data bases are often comparable. The sole reason for confining this dissertation to published studies is feasibility of data management.

In searching the literature I employed both electronic and manual search strategies. Psyclit database was searched for the period 1974-1991, and the keywords used in the search are provided in Appendix A. The same search was carried out in Wils, Infotrac, Oasis, and CLRC databases. However, the most fruitful search strategy was the manual search strategy. I searched the major journals, volume by volume, issue by issue, for articles that report the correlation between any two dimensions of job performance. The references of these articles were also searched for identifying further relevant articles. In short, a snowballing technique, with the articles published in the major journals as the initial kernel, was the most effective strategy in identifying the articles used in this

dissertation. The journals, all issues of which were searched, are listed in Appendix B.

I included in the database the studies that report data on individuals regarding their performance on actual jobs. Studies reporting data on departments or production units were not included. Experimental simulations reporting data on student samples engaged in laboratory work (or ratings of videotaped and manipulated work samples) were excluded. Assessment center ratings as well as interviewer and recruiter ratings were omitted because they reflect scores on predictors rather than job performance.

Self-reports, observer's reports, subordinate ratings, and customer ratings of performance were excluded. It is not enough to observe or record behavior; we must evaluate it (must know what the behavior means). Observing that subjects do a specific task does not reveal the hundreds of subtle differences in how it is done that makes one successful and another not (Schmidt, 1992). Further, the evaluator should know the importance or consequence of a particular behavior to the job. The lack of this expertise in evaluation is the reason for excluding self-reports, observers' reports, subordinate ratings, and customer ratings of performance in this dissertation.

When studies reported only the range (or average) of the intercorrelations between the measures such intercorrelations could not be classified as representing the intercorrelation between a particular pair of job performance dimensions. As such these studies were excluded from the database. Finally, the database also includes studies that report only the interrater reliability and internal consistency estimates (coefficient alphas) for the different dimensions of job performance. The reliability values were compiled to construct artifact distributions (elaborated below) that were used in the psychometric meta-analyses.

The same dimension of job performance can be referred to by different labels. In cumulating results across studies a dimension of job performance could be referred to by different labels in the different studies. The task of identifying the various dimensions of job performance and eliminating redundancies has to be guided by conceptual considerations. At one extreme even the changing of a single letter in the definition of the dimension can be construed as signifying a new and distinct dimension of job performance. Up to a point, it can be proposed that the more fine-grained, narrow, and explicitly defined the dimensions of job performance are, the greater the

conceptual clarity and interpretability of empirical
results.

Any advantage in clarity and interpretability of
empirical results that narrowly defined dimensions have
over more broadly defined dimensions is, however, offset by
considerations of availability of data and robustness of
resulting estimates. When results are cumulated across
studies, intercorrelations between some narrowly defined
dimensions may not be available, necessitating the analysis
at a level where the various dimensions of job performance
are defined more broadly (i.e., necessitating combining of
similar dimensions). Further, if the various dimensions of
job performance are narrowly defined, the number of
correlations for establishing (using the methods of
psychometric meta-analysis) the true score intercorrelation
between two dimensions will be small. The subsequent small
total sample sizes included in the psychometric meta-
analysis will lead to substantial first order sampling
error. Also the small number of correlations used in
psychometric meta-analysis will increase the second order
sampling error. Though the estimated mean true score
intercorrelation may not be greatly affected by the second
order sampling error, the true variability will often be
distorted (either under or overestimated).

Further, the hypothesized high conceptual clarity that accrues from more narrowly defined dimensions should also be balanced against the generality of the dimensions. For example, suppose we define three measures as: (a) supervisory ratings of the frequency with which an employee smiles at his or her peers; (b) smiles at supervisors; and (c) smiles at customers. Now, suppose we include these three measures with other measures of job performance such as productivity, leadership, dependability, etc. The three measures involving smiling will probably have a pattern of correlations that sets them apart from other measures, causing the appearance of a smilability factor. However, the question remains whether such narrowly defined performance measures are important to any job. Even if there is a particular job where frequency of smiling is important, it is highly unlikely that I/O psychologists would consider the inclusion of such narrowly defined measures as worth the labor involved in constructing prediction instruments and theories of job performance tailored for a particular job. Developing theories of job performance for each task (or even job) will hinder the development of a general theoretical understanding of the construct of job performance. As the content generality of the dimensions increase, the value of the dimensions in developing prediction instruments and theories of work behavior increases.

However, the identified dimensions of job performance should be conceptually distinct and completely exhaustive. Conceptual distinctiveness requires that any specific measure used in a study should be unambiguously identified as referring to a particular dimension of job performance. A completely exhaustive scheme of job performance measures ensures that all dimensions of performance in any job are covered.

The various dimensions of job performance analyzed in this dissertation were identified by grouping the list of measures obtained from the studies into conceptually similar groups. The identified dimensions of job performance were cross-validated by having another researcher (Deniz Ones) independently complete the task of grouping the list of measures obtained from the studies into conceptually similar groups. Differences (overall agreement rate was 92.4%) were resolved through mutual discussions until consensus was reached. A list of the various dimensions identified and used in this dissertation is provided in Appendix C.

Some studies reported correlations between different wordings of the same dimensions along with correlations with other dimensions of job performance. A composite measure was formed over the different wordings of the same dimension. The correlation of this composite measure with

other dimensions of job performance was computed. These computed correlations were used in the meta-analyses. Similarly, the reliability of the composite of the different wordings of the same dimension was computed using the Mosier formula and recorded. Unit (equal) weights for the different wordings of the same dimension were used in computing the composite. An alternative type of reliability estimate for the composite is the estimate based on the standardized coefficient alpha, which is the same as the use of the Spearman-Brown formula based on the average intercorrelation between the different wordings of the same dimension. However, the use of either of these would assign the specific factor variance to measurement error. To examine whether there is a psychological construct underlying the various dimensions, it is appropriate to assign the specific factor variance to measurement error. The use of Mosier formula (instead of Spearman-Brown or the standardized coefficient alpha) results in a higher estimate of the reliability (because the specific factor variance is treated as true variance), which lowers the corrected correlation and provides a more rigorous test of the hypothesis that all measures of job performance are manifestations of the same construct. That is, if the correlations corrected with Mosier reliability support the existence of a general factor, the use of the Spearman-Brown reliability (or standardized coefficient

alpha) in which the specific factor variance is removed from the true variance will result in the emergence of a larger general factor.

Data Coded/Extracted from Primary Studies

An identification number was given to each study. The list of studies coded is provided in Appendix D. When more than one sample was reported in a study, a sample within study identification number was given to each sample within that study. Samples were numbered consecutively starting with the number one. If a study reports data on a total sample as well as on subsamples (say blacks and whites, or males and females), I included the data only from the total sample. Subgroup correlations are prone to capitalization on chance and averaging subgroup correlations will introduce a downward bias in the correlation (due to range restriction) as compared to the correlation computed on the total sample (Hunter & Schmidt, 1990a, p.465).

Each sample in a study typically reports more than one intercorrelation (e.g., if the sample in a study examines four dimensions of job performance, there will be six intercorrelations). A separate record was created for each intercorrelation. An identification number was also given to each dimension of job performance (the same number is

also referenced in Appendix C where the dimensions of job performance are listed).

Each record includes a study number, a sample number, and the observed correlation between two dimensions of job performance. Each record also contains, for each of the two dimensions of job performance (the correlation between which is reported in that record), an identification number for the dimension of job performance (the same number as reported in Appendix C), the interrater reliability of that dimension, the internal consistency of the dimension, and the degree of split (if any) in that dimension of job performance. If reliability information is not available, the columns in that record were left blank. If a record involves a study that reports only reliability information for a particular dimension of job performance, the columns for the observed correlation, sample size, degree of split as well as the columns for the second dimension of job performance were left blank. (The sample size and the reliability of dimensions were used to assign appropriate weights to the study results. The degree of split was used to correct the artificially dichotomized correlations.)

In coding the interrater reliabilities, care was taken to adjust the reported values using Spearman-Brown corrections wherever appropriate. For example, whenever a study reported the interrater reliability based on one

rater as .50 but the correlation between the various dimensions of job performance was based on the summed or average ratings of six raters, the reliability was coded as .86.

Intercoder agreement in summarizing or extracting information from the primary studies is a concern in meta-analyses. Haring et al. (1981) present empirical data indicating that intercoder agreement in meta-analyses is a function of the judgmental nature of the items coded. The Haring et al. (1981) review of meta-analyses found that eight of the nine items lowest in coder agreement were judgments (e.g., the quality of the study) as opposed to calculation based variables (e.g., effect sizes, number of subjects). Jackson (1980) and Hattie and Hansford (1982, 1984) also provide data which indicate that problems of intercoder agreement in meta-analyses are negligible for coding computation-based numerical variables. Finally, Whetzel and McDaniel (1988) found no evidence of any coder disagreements in validity generalization data bases. Intercoder agreement in this dissertation was estimated by having another individual (Deniz Ones) code a subset of studies. The percentage agreement between us in coding the selected subset of studies was 97.6%.

Psychometric Meta-Analyses

Data from the sources described in the previous section was cumulated by the methods of psychometric meta-analyses. Depending on the availability of information in the primary studies, we can either correct the observed correlations for the effects of statistical artifacts and cumulate the individually corrected correlations, or use artifact distributions to correct the observed distribution of correlations, or use a combination of individual corrections and artifact distributions.

Because the degree of split for dichotomization is usually given in the research reports, it was possible to correct the correlations individually for the attenuating effects of dichotomization. But to correct for the effects of artifacts such as unreliability, where the information available is sporadic, recourse was made to the use of artifact distributions. That is, a mixed meta-analysis was employed. In the first step, the correlations were corrected individually for the effects of dichotomization. In the second step, the partially corrected distribution obtained from the first step was corrected for unreliability in the measures of the two variables being correlated using artifact distributions (Hunter & Schmidt, 1990a, p.188).

Artificially dichotomized correlations were corrected for discontinuity (Hunter & Schmidt, 1990b). In correcting for dichotomization, care was taken to account for the conceptualization of the measure (Williams, 1990). That is, corrections for dichotomization were undertaken only where it is theoretically defensible. For example, if an absenteeism measure was dichotomized (as high vs. low) then dichotomization corrections were applied. The correlations were corrected individually. When dichotomization corrections were applied to the observed correlations, the sample sizes were adjusted to estimate the correct sampling error variance (Hunter & Schmidt, 1990b). First, the uncorrected correlation and the study sample size were used to estimate the sampling error variance for the observed correlation. This value was then multiplied by the square of the dichotomization factor (the ratio of the corrected to uncorrected correlation), yielding the sampling error variance associated with the dichotomization corrected correlation. This value was then used with the uncorrected correlation in the standard sampling error formula to solve for the adjusted sample size used in the meta-analyses. Entry of this sample size into the meta-analysis calculations results in the correct estimate of the sampling error variance of the corrected correlation in the meta-analysis.

After the correlations were corrected individually for dichotomization, artifact distribution meta-analysis was used to correct for unreliability in the measures. In using artifact distributions for correcting two or more artifacts we have the option to use either the interactive procedure which corrects the observed correlations for the effects of the various statistical artifacts simultaneously, or the noninteractive procedure which corrects the observed correlation for the effects of the statistical artifacts sequentially (one after another). Recent computer simulation studies (e.g., Law, Schmidt, & Hunter, 1992; Schmidt et al., 1993) have shown that among the methods of psychometric meta-analyses the interactive procedure used with certain refinements, such as nonlinear range restriction and mean observed correlation in the sampling error formula, is the most accurate one. As such, the Hunter-Schmidt interactive meta-analytic procedure was used.

In correcting for unreliability in the measures, the use of the correct form of reliability coefficient requires the specification of the nature of the error of measurement in the research domain of interest (Hunter & Schmidt, 1990a, pp. 123-125). The two measures being correlated can: (a) both be organizational records; or (b) both be ratings provided by the same individual; or (c) both be

ratings, but the two ratings are provided by two different
individuals; or (d) one measure is a rating while the other
is an organizational record based measure. When the
correlation is between two organizational records, random
error in responses was corrected by using an artifact
distribution of coefficient alphas and test-retest
reliabilities. To correct the correlation between two
ratings when both the ratings are provided by the same
individual (intra individual correlations), intra rater
reliability estimates were used. Because both ratings are
provided by the same individual, use of interrater
reliability coefficients is inappropriate (interrater
disagreements are not involved in such correlations). Halo
is not corrected for when the true score correlations
between two ratings (when both the ratings are provided by
the same individual--intra individual correlations) are
reported.

When one of the two ratings being correlated is
provided by supervisors whereas the other is provided by
peers, the observed correlations are free of halo.
However, we need to correct the observed correlation for
interrater disagreements. Therefore, when the two ratings
being correlated came from two different individuals, I
used the interrater reliability estimates in the
corrections.

A separate distribution of reliabilities for each dimension of job performance was constructed from the data available in the literature. If no study was available that reported the reliability of some dimension of job performance, then for that dimension of job performance, an appropriate distribution of reliabilities was constructed using the reliability of other similar dimensions of job performance.

The result of a psychometric meta-analysis is an estimate of the distribution of actual correlations i.e., fully disattenuated true score correlations. The mean of the distribution is the true score correlation. The main purpose in using psychometric meta-analyses was to estimate the true score correlations between the various dimensions of job performance. Based on the psychometric meta-analyses, an intercorrelation matrix of true score correlations between the different dimensions of job performance was constructed.

Confirmatory Factor Analysis

The series of psychometric meta-analyses conducted on each pair of job performance dimensions establishes the matrix of intercorrelations between the dimensions of job performance. The intercorrelation matrix was subjected to a confirmatory factor analysis to examine whether the

proposed latent structure of job performance explains the true score correlations among the dimensions of job performance.

Confirmatory factor analysis was conducted on two matrices. The first was a matrix of intercorrelations between the organizational record based dimensions. The second matrix was formed of intercorrelations between ratings of the different dimensions of job performance, where of the two dimensions being correlated, one dimension was rated by supervisors and the other was rated by peers. The reason for restricting the confirmatory factor analysis to correlations between dimensions (where the two dimensions being correlated are provided by two different raters) is halo, which inflates the correlation when both the ratings being correlated are provided by the same rater. The CFA program (in Basic), developed by Hunter as part of a set of computer routines for performing a variety of analysis on correlational data, was used (Hunter, 1992; Hunter & Cohen, 1969).

In both the matrices (one based on organizational records, the other based on ratings) on which confirmatory factor analysis was conducted, I tested a two level hierarchical model involving a general factor and specific factors associated with each of the dimensions. I also tested a three level hierarchical model involving group

factors (where some dimensions were hypothesized to cluster together; that is, correlate more with dimensions in their clusters than with dimensions included in other groups). I examined the residual matrices to draw inferences about the fit of the hypothesized models (Hunter & Gerbing, 1982) to the empirical data.

I formed the group factors based on my review of the literature. Care was taken to ensure that the group factors are theoretically meaningful. The important point is not whether the group factors identified are absolute; alternate hierarchical models can (and should) be tested (as one's theory dictates). The purpose of testing a hierarchical model is only to test whether some structure, could be realized given the plethora of job performance measures.

The parameters estimated in the confirmatory factor analysis of a two level hierarchical model are the factor loadings of the various dimensions of job performance on the general factor as well as estimates of the unique variance in the various dimensions of the job performance. A residual matrix indicating the extent to which the two level model explains the true score correlations between the different dimensions is also obtained.

The parameters estimated in the confirmatory factor analysis of a three level hierarchical model are the factor intercorrelations among the group factors as well as the factor intercorrelations between the group factors and the overall general factor of job performance. Also estimated are the factor loadings of the various dimensions of job performance on the group factors as well as estimates of the unique variance in the various dimensions of the job performance. A residual matrix indicating the extent to which the three level model reproduces the true score correlations between the different dimensions is also obtained.

The extent to which the estimated factor loadings reproduce the intercorrelation matrix between the various dimensions of job performance is taken as a test of the hypothesized latent structure of the psychological construct of job performance (Hunter & Gerbing, 1982). The results of the analyses are presented in the next chapter.

CHAPTER IV

RESULTS

The results are presented in two sections. In the first section, I present the true score correlations between the different dimensions of job performance. I discuss the true score correlations between job performance measures based on organizational records. In the second section, I present the results from confirmatory factor analyses conducted on two intercorrelation matrices (one composed of organizational records, the other of correlations of ratings when the two ratings being correlated are provided by two different individuals) to test the hypothesized hierarchical model of job performance.

True Score Correlations

Of the 25 conceptually distinct dimensions of job performance that were found to have been used in the literature, 5 were based on organizational records, 11 were based on supervisory ratings, and the remaining nine employed peer ratings. The intercorrelations between the job performance dimensions based on organizational records are presented in Table 1.

Insert Table 1 about here

Of importance to note in Table 1 is the relatively high correlation between absenteeism and quality of job performance (ρ = .48). Also of interest is the moderate correlation between accidents and quality of job performance (ρ = .24). These findings might be explainable in terms of common causal antecedents. For example, perhaps conscientious individuals produce more high quality products, are less likely to be absent, and are careful to avoid accidents. Such an inference also meshes with the findings of Ones et al. (in press) who found that integrity tests predict overall job performance as well as absenteeism and accidents.

As could be seen in Table 1, absenteeism and tenure (ρ = .33) are not more highly correlated than absenteeism and quality (i.e., lack of errors; ρ = .48) or absenteeism and accidents (ρ = 1.0). Though a true score correlation of 1.0 should be dismissed as a chance overestimate (probably the erroneous estimate resulted from a small sample of unrepresentative, biased correlations), the results are still surprising given that absenteeism and tenure are hypothesized to fall in a continuum of withdrawal

behaviors. A related point to note in Table 1 is that tenure and accidents have a very low correlation (ρ = .07).

The true score correlation between tenure and quality of performance is .12 and the true score correlation between tenure and productivity is .19. Tenure and accidents correlate only .07. However, given the small number of correlations in each meta-analysis, the conclusions have to be very tentative only. The correlation between productivity (quantity) and quality indices was .37 and this estimate is based on 13 samples involving 2,731 individuals. Absenteeism and productivity correlated .21 lending support to Bycio's (1992) hypothesis that absenteeism disrupts work habits and lowers individual productivity.

In general, based on Table 1, it can be concluded that in forming composites of job performance measures, care should be taken before including measures of tenure. Productivity, absenteeism, quality of performance, and accidents group together, encouraging the search for common causal antecedents such as conscientiousness and general mental ability to explain and predict the various manifestations of job performance.

The intercorrelations between supervisory ratings of 11 job performance measures are presented in Table 2. The

true score correlations between 9 job performance measures involving ratings by peers are presented in Table 3. Each correlation in Tables 2 and 3 represents the correlation between ratings of two different dimensions of job performance when both the ratings are provided by the same individual (intra individual correlations). These correlations are affected by a general impression that the rater has of the ratee (i.e., halo) that affects the ratings in the two dimensions being correlated, thus inflating the correlation between the two performance dimensions.

Insert Tables 2 & 3 about here

The correlations in Table 2 are much higher than in Table 1, indicating the presence of halo in supervisory ratings. These values give a picture of the intercorrelations as seen by individual supervisors, not as agreed on by different supervisors, and not as they necessarily are in the real world.

Further, when the same rater provides rating on two or more dimensions, the intercorrelations reflect the same perspective. Arguments have been advanced that raters at

different levels (peers, supervisors, etc.) of the
organization have different perspectives that will affect
the intercorrelations (Sager, 1990; Sager, Nitti, &
Hazucha, 1993). To address these two concerns, Table 4
provides the intercorrelations between measures of job
performance when the two measures being correlated are
obtained from two different rating sources: supervisors
and peers.

Insert Table 4 about here

The true score correlations between the eleven measures
using supervisory ratings and the five job performance
measures based on organizational records are presented in
Table 5.

Insert Table 5 about here

The correlation between the nine measures of job
performance using peer ratings and the five measures based
on organizational records are presented in Table 6. The
first thing to note is the large number of empty cells.

Further, even in the cells where the true score correlation can be estimated, the sample sizes do not exceed 276, and the number of samples involved range between one and three (with most of the true score correlations based on single samples). Because of the small sample sizes involved, any inferences based on Table 6 has to be very tentative.

Insert Table 6 about here

On reflection, it is not surprising that most of the empty cells found in constructing a matrix of intercorrelations between the various measures of job performance were between peer ratings and organizational records. The literature in I/O psychology has examined: (a) the correlations within supervisory, within peer, and within organizational-record based measures to test how many factors of job performance emerge; (b) between peer and supervisory ratings to test whether the two sources have convergent validity and to assess the extent to which there are differences in the perspectives of the two sources; and (c) between supervisory ratings and organizational records to test whether supervisory ratings are valid (convergent validity) and whether supervisory ratings are more highly correlated with organizational

records for minorities than for the majority group (i.e.,
the hypothesis that supervisors when rating blacks and
women pay more attention to organizational records).
However, there seems to have been no substantive interest
in I/O psychology to examine the intercorrelation between
peer ratings and organizational records. Future research
should address this limitation and examine the
intercorrelation between peer ratings and organizational
records to facilitate the development of comprehensive
models of job performance.

Confirmatory Factor Analyses Results

Two Level Hierarchical Models

To examine whether a general factor emerges from the
true score correlations between the various organization
record based measures of job performance, the 5 by 5 matrix
of intercorrelations (reported in Table 1) was subjected to
a confirmatory factor analysis. All five measures were
hypothesized to load on the same factor. The factor
loadings are provided in Table 7 and the residual matrix is
provided in Table 8. The residual matrix represents the
difference (that is, the discrepancy) between the true
score correlations reported in Table 1 and the correlations
reproduced by the hypothesized latent structure.

Insert Tables 7 & 8 about here

The general factor explained 49.1% of the variance in the five job performance measures. As expected (based on an inspection of the matrix of true score correlations presented in Table 1), measures of tenure had the lowest loading of all the five measures on the general factor. Surprisingly, productivity and quality had lower loadings on the general factor than measures of accidents and absenteeism. This is probably due to the erroneously overestimated true score correlation of 1.00 between absenteeism and accidents. An inspection of the residual matrix indicates that the true score correlations of: (a) accident measures with measures of quality, (b) absenteeism measures with organizational records of accidents, (c) productivity with absenteeism, and (d) accidents and tenure are the four correlations least explained by the hypothesized latent structure (that is, these three correlations have the largest discrepancies between actual and reproduced correlations). These results indicate that the two level hierarchical model does not adequately explain the true score correlations between the job performance measures based on organizational records.

To examine whether a general factor explains the true score correlations between the various job performance dimensions that involve subjective evaluations (ratings or rankings), another confirmatory factor analysis was performed. The true score correlations between ratings of various dimensions of job performance, when both the ratings being correlated are provided by the same individual, are affected by halo. Halo, which reflects the general impression the rater has of the particular ratee, is confounded with any potential general factor across the various dimensions being rated when the rating source is the same individual. Halo cannot be untangled in the analysis of the true score correlations reported in Tables 2 and 3 (within supervisors and within peers). However, the correlations reported in Table 4 can be used to investigate the existence of a general factor. Here of the two ratings being correlated, one is provided by the supervisors and the other is provided by peers. There are eight dimensions of job performance that were rated by both supervisors and peers. Based on these ratings an 8 by 8 matrix can be constructed with supervisors rating one of the dimensions being correlated while the other dimension is rated by peers. The sample size weighted mean of the corresponding cells above and below the diagonal were computed. The values are provided in Table 9.

Insert Table 9 about here

Confirmatory factor analysis of the correlations reported in Table 9 is not affected by halo. The results of the confirmatory factor analysis are reported in Tables 10 and 11. The factor loadings of the eight dimensions of job performance on the general factor are provided in Table 10 and the residual matrix is provided in Table 11.

Insert Tables 10 & 11 about here

The general factor explained 54.9% of the variance in the eight job performance measures. Ratings of overall job performance had one of the highest loading (.79) on the general factor as could be expected. Surprisingly, ratings of administrative skills also had an equally high loading (.79). Ratings of effort and job knowledge had higher loadings (.75 and .76, respectively) than ratings of personality (.69). An examination of the residual matrix shows that the model does not adequately explain the true score correlation between productivity ratings and ratings

of effort as well as the true score correlation between ratings of productivity and ratings of problem solving.

Analysis of the residual matrix as well as the percent variance explained by the general factor in both matrices (Table 1 and Table 9) indicate that a two level hierarchical model of job performance involving one general factor does not explain the true score correlations adequately. Alternate models postulating different groupings of the job performance dimensions based on the true score correlations reported here (and supported by substantive theories) need to be tested.

Three Level Hierarchical Models

Based on a review of the job performance literature, five group factors were identified and all job performance measures were hypothesized to load on one of the five group factors. The five group factors were: Job performance measures focusing on (a) productivity, (b) overall, global evaluation, (c) withdrawal behaviors, (d) interpersonal skills, and (e) conscientious behaviors.

The five job performance measures based on organizational records can be grouped into three of the five group factors. Absenteeism and tenure will load on one group factor (which I term as withdrawal), whereas accidents and quality of production are hypothesized to

load on another factor (which I name, conscientiousness).
Finally, I defined productivity as a separate group by
itself (Prod). Again, to reiterate, the definition of the
three group factors are ad hoc and interested readers can
test their own models of job performance based on the true
score correlations reported in this dissertation.

The factor loadings of the five job performance
measures on the three group factors are presented in Table
12. The factor loadings of the five job performance
measures were as hypothesized with the exception of
absenteeism and accidents. These two measures did not have
the highest loading on the factor they were hypothesized to
measure. This is probably due to the estimated true score
correlation of 1.00 between absenteeism and accidents (an
obviously erroneous overestimate).

Insert Table 12 about here

The intercorrelations between the three group factors
were positive and substantial. The correlation between
prod and withdrawal, between withdrawal and
conscientiousness, between prod and conscientiousness were
.29, .82, .59, respectively. Subjecting this 3 by 3 matrix

to a second order factor analysis (where the three group factors were hypothesized to load on a common general factor) resulted in factor loadings of .75, .84, and .96 for prod, withdrawal, and conscientiousness, respectively on the general factor. This general factor explained 73% of the variance among the three group factors. This represents a 49% (23.9/49.1) improvement in fit over the two level hierarchical model.

A three level hierarchical model can also be tested using the 8 by 8 matrix of true score correlations reported in Table 9. The eight dimensions of job performance are hypothesized to load on four group factors. Ratings of personality, problem solving, and administrative skills can be grouped together, as these three dimensions of job performance reflect interpersonal relations (F1). Ratings of effort, job knowledge, and compliance with rules and regulations can be grouped together (F2) as reflecting the conscientiousness of the individual. Ratings of productivity and ratings of overall job performance are treated separately as the third (F3) and fourth (F4) factors. I expect the group factor involving overall job performance (F4) to have the highest loading on the general factor.

The factor loadings of the eight job performance measures on the four group factors are summarized in Table

13. Even though most of the eight measures had higher factor loadings on the factor they were hypothesized to load than on any other factor, there were two exceptions. First, ratings of job knowledge hypothesized to load on a group factor (F2) along with compliance and effort (all three presumably reflecting how conscientious the individual is) had higher loadings on group factors reflecting productivity (F3) and overall job performance (F4). Second, ratings of effort clusters more with measures of overall job performance than with ratings of compliance and ratings of job knowledge.

Insert Table 13 about here

The intercorrelations among the four group factors were positive and substantial. The correlation of F1 with F2, F3, F4 were .70, .44. and .90, respectively; the correlation of F2 with F3, F4 were .83 and .98, respectively; and, the correlation between F3 and F4 was .73. When this 4 by 4 matrix of intercorrelations was subjected to a second order factor analysis (to test whether the four factors load on a common general factor), the factor loadings of F1, F2, F3, and F4 on the general factor were .85, .98, .84, and 1.01, respectively. The

general factor explained 85.2% of the variance in the four group factors. This represents a 55% improvement in fit over the two level hierarchical model. As anticipated, the group factor reflecting measures of overall job performance (F4) had higher loading on the general factor than any other group factors.

Again, to reiterate, the three level hierarchical models presented and analyzed here reflect my interpretations of the literature. Researchers could (and should) test alternate models based on substantive theories using the true score correlations reported here. Comprehensive tests of three level hierarchical models have to await the development of substantive theories in the criterion domain. The aim in this dissertation was only to estimate the true score correlations between the various dimensions of job performance, test two level hierarchical models for the existence of a general factor across the different dimensions of job performance, and provide an illustration of how a three level hierarchical model could be tested in the future.

CHAPTER V

IMPLICATIONS

Convergence Between Sources of Evaluation

Multiple sources (supervisory ratings, peer ratings, organizational records) have been used in performance evaluation. The convergent validity of the sources has been the focus of many researchers (e.g., Harris & Schaubroeck, 1988; Heneman, 1986; Mount, 1984). The correlations presented in Tables 1 through 4 can be used to test the convergence between different sources of measurement. In fact a more reliable and better estimate of the convergent validity between two sources (e.g., supervisory ratings and organizational records) for overall job performance could be obtained by forming the composite correlation between the measures using one source with the measures using another source. In forming the linear composite, we can use unit weights or weight the measures by their loadings on the general factor.

When we used unit weights and computed (based on the intercorrelations reported in Tables 1 through 4), the composite correlation of a linear combination of organizational records of absenteeism, accidents, quality and quantity of production (I excluded organizational

records of tenure because this measure correlates low with other measures involving organizational records.) with a linear combination of supervisory ratings of 11 job performance measures was .67. Between peer and organizational records, the correlation was .44 whereas the correlation was .62 between supervisory ratings and peer ratings. There seems to be more agreement between supervisors and organizational records than between supervisors and peers. Also the correlation between peer ratings and organizational records is the lowest (of all the three correlations between sources). As such we can infer that organizational records reflect the supervisory view more effectively than does the perspective of peers. If researchers are interested in capturing different perspectives in performance but are constrained by costs to include only two of the three sources, it is more relevant to collect peer ratings and organizational records than to collect supervisory ratings and organizational records. Finally, the correlation of .67 between supervisory ratings and organizational records inspires confidence in the use of supervisory ratings in organizational research.

Implications for Differential Prediction

The findings reported in this dissertation have implications for the investigation of differential prediction. The positive correlations reported in Tables

1-4 encourage the search for common predictors or antecedents of performance. That is, it is likely that the use of different job performance measures will result in the selection of the same set of predictors. Use of job performance measures that are highly correlated will probably result in the selection of the same test battery (and assignment of the same weights), regardless of the criterion used, rendering differential prediction unlikely.

The empirical literature investigating differential prediction using multiple regression strategies has provided conflicting results. In a typical study investigating differential prediction using the multiple regression approach, different regression equations are developed, one with each criterion as the dependent variable, with a set of tests or predictor measures as the independent variables. The purpose is to examine whether the regression weights given to the different tests are different when different criteria are used.

Schmidt (1980) reports that in one large sample study done in the Army it was found that a job sample criterion, supervisory ratings, and a job knowledge measure all resulted in the adoption of the same selection procedures and assignment of essentially identical relative weights. Oppler, Sager, McCloy, and Rosse (1993) found, using a sample of 3,086 soldiers, that prediction composites

developed using job knowledge tests compared favorably in validity to those developed using hands on tests.

Despite these results, researchers have continued to focus on the question of differential prediction. Some research studies claim to have found support for differential prediction. The contradictory findings across studies examining differential prediction may be the result of sampling error and capitalization on chance, which greatly influences the results when multiple regression strategies are used to develop batteries that include correlated predictors (Hunter, Crosson, & Friedman, 1985). The estimated regression weights have large standard errors especially when correlated predictors are used unless sample sizes are large (Helme, Gibson, & Brogden, 1957).

An alternate strategy to examine differential prediction is to examine the intercorrelations between the job performance measures. Differential prediction of different job performance measures requires that the rank ordering of job applicants differ when different job performance measures are used. High intercorrelations between various measures lead to the conclusion that differential prediction is unlikely.

Recently, a hypothesis has been advanced (Campbell et al., 1992) that all job performance measures are determined

by 3 components--declarative knowledge, procedural knowledge, and motivation. This hypothesis holds that differential prediction will be found only if measures of these different performance components are included in the analysis as separate measures. The low true score correlation of .20 (K = 4, N = 386) found between supervisory ratings of job knowledge and peer ratings of effort or of .09 (K = 3, N = 325) between supervisory ratings of effort and peer ratings of job knowledge is consistent with Campbell et al. (1992) hypothesis. However, further research is needed to actually test this hypothesis.

Forming Composites of Different Job Performance Measures

Schmidt (1980) discusses several methods that could be used to weight the job performance measures in forming a composite. The Kellys Bid System, the Dollar Criterion approach, equal weighting, and weighting to maximize the general factor are some approaches. Of these, forming composites by weighting to maximize the general factor (Edgerton & Kolbe, 1936; Nagle, 1953) is analogous to weighting test items by their correlations with total test score. It is appropriate to form composites by weighting the individual measures to maximize the general factor only if the true score correlations between the individual measures indicate support for a large general factor.

The true score correlations presented here can guide researchers in forming composites. For example, in forming a composite of organizational records, care should be taken before including measures of tenure. Further, Campbell (1990c) says that core job performance, demonstrating effort, and maintenance of personal discipline are important job performance measures in all jobs. The true score correlation (reported in Table 4) between supervisory ratings of productivity (or job-specific task performance to use Campbell's terminology) and peer ratings of effort is .50 (K = 2, N = 191); peer ratings of maintaining personal discipline and supervisory ratings of effort correlated .56 (K = 3, N = 378); and, peer ratings of maintaining personal discipline and supervisory ratings of job-specific task performance correlate .53 (K = 1, N = 164). These high correlations suggest that forming composites by weighting to maximize the general factor is acceptable in all jobs.

Finally, the results reported in this dissertation have implications for testing competing models hypothesized for the latent structures of job performance as well as for theory building. The true score correlations reported here form the building blocks that could be used to test alternate latent structures of job performance. The positive manifold of correlations indicate that

differential prediction is unlikely, encouraging the search
for common predictors and antecedents of performance and
general theories of work behavior. Future research should
examine the correlations of the general factor of job
performance with other constructs such as general mental
ability and conscientiousness. More specifically, future
research should examine whether group factors have any
relevance (over and above the general factor) for
evaluating training programs and other human resources
interventions in an organization.

Appendix A

Key Words Used in Searching Electronic Databases

1.Job performance/productivity

2.Theft

3.Absenteeism

4.Turnover

5.Violence on the job

6.Rule breaking

7.Co-worker relations

8.Teamwork

9.Tenure

10.Alcohol use/abuse

11.Drug use/abuse

12.Tardiness

13.Lateness

14.Quality of work

Appendix B

Journals Searched

1. Journal of Applied Psychology

2. Personnel Psychology

3. Academy of Management Journal

4. Human Relations

5. Journal of Business and Psychology

6. Journal of Management

7. Accident Analysis and Prevention

8. International Journal of Intercultural Relations

9. Organizational Behavior and Human Decision Processes

10. Journal of Vocational behavior

11. Journal of Applied Behavioral Analysis

12. Human Resources Management Research

13. Journal of Occupational Psychology

14. Psychological Reports

15. Journal of Organizational Behavior

Appendix C

List of Criteria

1. Organizational records of quality of performance (lack of errors)
2. Organizational records of productivity
3. Supervisory ratings of personality, teamwork etc.
4. Organizational records of absenteeism (lack of)
5. Organizational records of tenure
6. Supervisory ratings of productivity
7. Supervisory ratings of effort
8. Supervisory ratings of overall job performance
9. Peer ratings of administrative skills
10. Supervisory ratings of administrative skills
11. Peer ratings of productivity
12. Supervisory ratings of quality of job performance
13. Supervisory ratings of job knowledge
14. Supervisory ratings of absenteeism
15. Supervisory ratings of problem solving and leadership
16. Organizational records of accidents (lack of)
17. Peer ratings of problem solving and leadership
18. Supervisory ratings of compliance and acceptance of authority
19. Supervisory ratings of communication skills
20. Peer ratings of overall job performance
21. Peer ratings of personality, teamwork
22. Peer ratings of communication skills
23. Peer ratings of effort
24. Peer ratings of job knowledge
25. Peer ratings of compliance and acceptance of authority

Appendix D

<u>List of Studies Included in the Database</u>

Abelson, M. A. (1983). The impact of goal change on prominent perceptions and behaviors of employees. <u>Journal of Management</u>, <u>9</u>, 65-79.

Adler, S., & Golan, J. (1981). Lateness as a withdrawal behavior. <u>Journal of Applied Psychology</u>, <u>66</u>, 544-554.

Albrecht, P. A., Glaser, E. M., & Marks, J. (1964). Validation of a multiple-assessment procedure for managerial personnel. <u>Journal of Applied Psychology</u>, <u>48</u>, 351-360.

Anderson, H. E., Jr., Roush, S. L., & McClary, J. E. (1973). Relationships among ratings, production, efficiency, and the general aptitude test battery scales in an industrial setting. <u>Journal of Applied Psychology</u>, <u>58</u>, 77-82.

Arvey, R. D., Landon, T. E., Nutting, S. M., & Maxwell, S. E. (1992). Development of physical ability tests for police officers: A construct validation approach. <u>Journal of Applied Psychology</u>, <u>77</u>, 996-1009.

Ashford, S. J., & Tsui, A. S. (1991). Self-Regulation for managerial effectiveness: The role of active feedback seeking. <u>Academy of Management Journal</u>, <u>34</u>, 251-280.

Baba, V. V. (1990). Methodological issues in modeling absence: A comparison of least squares and tobit analyses. <u>Journal of Applied Psychology</u>, <u>75</u>, 428-432.

Baehr, M. E., & Williams, G. B. (1968). Prediction of sales success from factorially determined dimensions of personal background data. <u>Journal of Applied Psychology</u>, <u>52</u>, 98-103.

Baier, D. E., & Dugan, R. D. (1957). Factors in sales success. <u>Journal of Applied Psychology</u>, <u>41</u>, 37-40.

Bass, A. R., & Turner, J. N. (1973). Ethnic group differences in relationships among criteria of job performance. Journal of Applied Psychology, 57, 101-109.

Beatty, R. W. (1973). Blacks as supervisors: A study of training, job performance, and employers' expectations. Academy of Management Journal, 16, 196-206.

Beatty, R. W., & Beatty, J. R. (1975). Longitudinal study of absenteeism of hard-core unemployed. Psychological Reports, 36, 395-406.

Beauvais, L. L. (1992). The effects of perceived pressures on managerial and nonmanagerial scientists and engineers. Journal of Business and Psychology, 6, 333-347.

Becker, T. E., & Klimoski, R. J. (1989). A field study of the relationship between the organizational feedback environment and performance. Personnel Psychology, 42, 343-358.

Beehr, T. A., & Gupta, N. (1978). A note on the structure of employee withdrawal. Organizational Behavior and Human Performance, 21, 73-79.

Benson, P. G., & Pond, S. B., III. (1987). An investigation of the process of employee withdrawal. Journal of Business and Psychology, 1, 218-229.

Bernardin, J. H. (1977). The relationship of personality variables to organizational withdrawal. Personnel Psychology, 30, 17-27.

Bernardin, J. H. (1987). Effect of reciprocal leniency on the relation between consideration scores from the leader behavior description questionnaire and perfromance ratings. Psychological Reports, 60, 479-487.

Bhagat, R. S., & Allie, S. M. (1989). Organizational stress, personal life style, and symptoms of life strains: An examination of the moderating role of sense of competence. Journal of Vocational Behavior, 35, 231-253.

Blank, W., Weitzel, J. R., & Green, S. G. (1990). A test of the situational leadership theory. <u>Personnel Psychology</u>, <u>43</u>, 579-597.

Blanz, F., & Ghiselli, E. E. (1972). The mixed standard scale: A new rating system. <u>Personnel Psychology</u>, <u>25</u>, 185-199.

Blau, G. (1981). An empirical investigation of job stress, social support, service length, and job strain. <u>Organizational Behavior and Human Performance</u>, <u>27</u>, 279-302.

Blau, G. (1988). An investigation of the apprenticeship organizational socialization strategy. <u>Journal of Vocational Behavior</u>, <u>32</u>, 176-195.

Blau, G. (1990). Exploring the mediating mechanisms affecting the relationship of recruitment source to employee performance. <u>Journal of Vocational Behavior</u>, <u>37</u>, 303-320.

Blau, G. J. (1985). Relationship of extrinsic, intrinsic, and demographic predictors to various types of withdrawal behaviors. <u>Journal of Applied Psychology</u>, <u>70</u>, 442-450.

Blau, G. J. (1986). Job involvement and organizational commitment as interactive predictors of tardiness and absenteeism. <u>Journal of Management</u>, <u>12</u>, 577-584.

Blau, G. J. (1986). The relationship of management level to effort level, direction of effort, and managerial performance. <u>Journal of Vocational Behavior</u>, <u>29</u>, 226-239.

Bledsoe, J. C. (1981). Factors related to academic and job performance of graduates of practical nursing programs. <u>Psychological Reports</u>, <u>49</u>, 367-371.

Bledsoe, J. C., & Haywood, G. D. (1981). Prediction of job satisfactoriness and job satisfaction of secondary school teachers. <u>Psychological Reports</u>, <u>49</u>, 455-458.

Bolanovich, D. J. (1946). Statistical analysis of an industrial rating chart. <u>Journal of Applied Psychology</u>, <u>30</u>, 23-31.

Borman, W. C. (1974). The rating of individuals in organizations: An alternate approach. <u>Organizational Behavior and Human Performance</u>, <u>12</u>, 105-124.

Borman, W. C., Dorsey, D., & Ackerman, L. (1992). Time-spent responses as time allocation strategies: Relations with sales performance in a stockbrocker sample. <u>Personnel Psychology</u>, <u>45</u>, 763-777.

Borman, W. C., Rosse, R. L., & Abrahams, N. M. (1980). An empirical construct validity approach to studying predictor-job performance links. <u>Journal of Applied Psychology</u>, <u>65</u>, 662-671.

Boyle, A. J. (1980). 'Found Experiments' in accidents research: Report of a study of accident rates and implications for future research. <u>Journal of Occupational Psychology</u>, <u>53</u>, 53-64.

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

Breaugh, J. A. (1981). Relationships between recruiting sources and employee performance, absenteeism, and work attitudes. <u>Academy of Management Journal</u>, <u>24</u>, 142-147.

Brown, C. W., & Ghiselli, E. E. (1947). Factors related to the proficiency of motor coach operators. <u>Journal of Applied Psychology</u>, <u>31</u>, 477-479.

Buckner, D. N. (1959). The predictability of ratings as a function of interrater agreement. <u>Journal of Applied Psychology</u>, <u>43</u>, 60-64.

Buel, W. D., & Bachner, V. M. (1961). The assessment of creativity in a research setting. <u>Journal of Applied Psychology</u>, <u>45</u>, 353-358.

Burke, R. J., & Wilcox, D. S. (1972). Absenteeism and turnover among female telephone operators. <u>Personnel Psychology</u>, <u>25</u>, 639-648.

Bushe, G. R., & Gibbs, B. W. (1990). Predicting organization development consulting competence from the Myers-Briggs type indicator and state of ego development. <u>The Journal of Applied Behavioral Science</u>, <u>26</u>, 337-357.

Butler, M. C., & Ehrlich, S. B. (1991). Positional influences on job satisfaction and job performance: A multivariate, predictive approach. Psychological Reports, 69, 855-865.

Campbell, J. P., Dunnette, M. D., Arvey, R. D., & Hellervik, L. V. (1973). The development and evaluation of behaviorally based rating scales. Journal of Applied Psychology, 57, 15-22.

Campbell, J. T. (1963). Validity information exchange 16-04. Personnel Psychology, 16, 181-183.

Carlson, R. E., Dawis, R. V., & Weiss, D. J. (1969). The effect of satisfaction on the relationship between abilities and satisfactoriness. Occupational Psychology, 43, 39-46.

Cascio, W. F., & Valenzi, E. R. (1977). Behaviorally anchored rating scales: Effects of education and job experience of raters and ratees. Journal of Applied Psychology, 62, 278-282.

Cascio, W. F., & Valenzi, E. R. (1978). Relations among criteria of police performance. Journal of Applied Psychology, 63, 22-28.

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.

Chaney, F. B. (1966). A cross-cultural study of industrial research performance. Journal of Applied Psychology, 50, 206-210.

Cheloha, R. S., & Farr, J. L. (1980). Absenteeism, job involvement, and job satisfaction in an organizational setting. Journal of Applied Psychology, 65, 467-473.

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.

Cleveland, J. N., & Landy, F. J. (1981). The influence of rater and ratee age on two performance judgments. Personnel Psychology, 34, 19-29.

Colarelli, S. M., Dean, R. A., & Konstans, C. (1987). Comparative effects of personal and situational influences on job outcomes of new professionals. Journal of Applied Psychology, 72, 558-566.

Cooper, R. (1966). Leader's task relevance and subordinate behaviour in industrial work groups. Human Relations, 19, 57-84.

Cooper, R., & Payne, R. (1967). Extraversion and some aspects of work behavior. Personnel Psychology, 20, 45-57.

Cortina, J. M., Doherty, M. L., Schmitt, N., Kaufman, G., & Smith, R. G. (1992). The "Big Five" personality factors in the IPI and MMPI: Predictors of police performance. Personnel Psychology, 45, 119-140.

Crouch, A., & Nimran, U. (1989). Office design and the behavior of senior managers. Human Relations, 42, 139-155.

Dailey, R. C. (1979). Group, task, and personality correlates of boundary-spanning activities. Human Relations, 32, 273-285.

Day, D. W., & Silverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. Personnel Psychology, 42, 25-36.

Deadrick, D. L., & Madigan, R. M. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. Personnel Psychology, 43, 717-744.

Denton, J. C. (1963). Validity information exchange. Personnel Psychology, 16, 283-288.

Dicken, C. F., & Black, J. D. (1965). Predictive validity of psychometric evaluations of supervisors. Journal of Applied Psychology, 49, 34-47.

Dickinson, T. L., & Tice, T. E. (1973). A multitrait-multimethod analysis of scales developed by retranslation. Organizational Behavior and Human Performance, 9, 421-438.

Dickinson, T. L., & Tice, T. E. (1977). The dicriminant validity of scales developed by retranslation. Personnel Psychology, 30, 217-228.

Distefano, Jr., M. K. (1988). Work performance, ability, and voluntary turnover among psychiatric aides. Psychological Reports, 63, 875-878.

Distefano, M. K., Jr., Pryer, M. W., & Erffmeyer, R. C. (1983). Application of content validity methods to the development of a job-related performance rating criterion. Personnel Psychology, 36, 621-631.

Dobson, P., & Williams, A. (1989). The validation of the selection of male British army officers. Journal of Occupational Psychology, 62, 313-325.

Dreher, G. F. (1981). Predicting the salary satisfaction of exempt employees. Personnel Psychology, 34, 579-589.

Drory, A. (1982). Individual differences in boredom proneness and task effectiveness at work. Personnel Psychology, 35, 141-151.

Edwards, P. K. (1979). Attachment to work and absence behavior. Human Relations, 32, 1065-1080.

Ekpo-Ufot, A. (1979). Self-Perceived task-relevant abilities, rated job performance, and complaining behavior of junior employees in a government ministry. Journal of Applied Psychology, 64, 429-434.

Erickson, J., Edwards, D., & Gunderson, E. K. E. (1973). Status congruency and mental health. Psychological Reports, 33, 395-401.

Farh, J., Podsakoff, P. M., & Organ, D. W. (1990). Accounting for organizational citizenship behavior: Leader fairness and task scope versus satisfaction. Journal of Management, 16, 705-721.

Farh, J., Werbel, J. D., & Bedeian, A. G. (1988). An empirical investigation of self-appraisal-based performance evaluation. Personnel Psychology, 41, 141-156.

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

Feild, H. S., Bayley, G. A., & Bayley, S. M. (1977). Employment test validation for minority and nonminority production workers. Personnel Psychology, 30, 37-46.

Ferris, G. R., Youngblood, S. A., & Yates, V. L. (1985). Personality, training performance, and withdrawal: A test of the person-group fit hypothesis for organizational newcomers. Journal of Vocational Behavior, 27, 377-388.

Fiedler, F. E. (1970). Leadership experience and leader performance- Another hypothesis shot to hell. Organizational Behavior and Human Decision Processes, 5, 1-14.

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

Flanders, J. K. (1918). Mental tests of a group of employed men showing correlations with estimates furnished by employer. Journal of Applied Psychology, 2, 197-206.

Fox, S., & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. Personnel Psychology, 41, 581-592.

Frost, D. E., Fiedler, F. E., & Anderson, J. W. (1983). The role of personal risk-taking in effective leadership. Human Relations, 36, 185-202.

Gardner, D. G., Dunham, R. B., Cummings, L. L., & Pierce, J. L. (1989). Focus of attention at work: Construct definition and empirical validation. Journal of Occupational Psychology, 62, 61-77.

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

Gavin, J. F., & Ewen, R. B. (1974). Racial differences in job attitudes and performance: Some theoretical considerations and empirical findings. Personnel Psychology, 27, 455-464.

Gaylord, R. H., Russell, E., Johnson, C., & Severin, D. (1951). The relation of ratings to production records: An empirical study. Personnel Psychology, 4, 363-371.

Ghiselli, E. E. (1942). The use of the Strong vocational interest blank and the Pressy senior classification test in the selection of casualty insurance agents. Journal of Applied Psychology, 26, 793-799.

Gordon, M. E., & Fitzgibbons, W. J. (1982). Empirical test of the validity of seniority as a factor in staffing decisions. Journal of Applied Psychology, 67, 311-319.

Gough, H. G., Bradley, P., & McDonald, J. S. (1991). Performance of residents in anesthesiology as related to measures of personality and interests. Psychological Reports, 68, 979-994.

Graen, G., Dansereau, F., Jr., & Minami, T. (1972). An empirical test of the man-in-the-middle hypothesis among executives in a hierarchical organization employing a unit-set analysis. Organizational Behavior and Human Decision Processes, 8, 262-285.

Graen, G., Novak, M. A., & Sommerkamp, P. (1982). The effects of leader-member exchange and job design on productivity and satisfaction: Testing a dual attachment model. Organizational Behavior and Human Performance, 30, 109-131.

Green, S. B., & Stutzman, T. (1986). An evaluation of methods to select respondents to structured job-analysis questionnaires. Personnel Psychology, 39, 543-564.

Griffin, R. W. (1991). Effects of work redesign on employee perceptions, attitudes, and behaviors: A long-term investigation. Academy of Management Journal, 34, 425-435.

Guion, R. M. (1965). Synthetic validity in a small company: A demonstration. Personnel Psychology, 18, 49-63.

Gunderson, E. K. E., & Nelson, P. D. (1966). Criterion measures for extremely isolated groups. Personnel Psychology, 19, 67-80.

Gunderson, E. K. E., & Ryman, D. H. (1971). Convergent and discriminant validities of performance evaluations in extremely isolated groups. Personnel Psychology, 24, 715-724.

Hackman, J. R., & Porter, L. W. (1968). Expectancy theory predictions of work effectiveness. Organizational Behavior and Human Decision Processes, 3, 417-426.

Hammer, T. H., Landau, J. C., & Stern, R. N. (1981). Absenteeism when workers have a voice: The case of employee ownership. Journal of Applied Psychology, 66, 561-573.

Hansen, C. P. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. Journal of Applied Psychology, 74, 81-90.

Hatcher, L., Ross, T. L., & Collins, D. (1989). Prosocial behavior, job complexity, and suggestion contribution under gainsharing plans. The Journal of Applied Behavioral Science, 25, 231-248.

Hater, J. J., & Bass, B. M. (1988). Superiors' evaluations and subordinates' perceptions of transformational and transactional leadership. Journal of Applied Psychology, 73, 695-702.

Hausman, H. J., & Strupp, H. H. (1955). Non-Technical factors in supervisors' ratings of job performance. Personnel Psychology, 8, 201-217.

Hendrix, W. H., & Spencer, B. A. (1989). Development and test of a multivariate model of absenteeism. Psychological Reports, 64, 923-938.

Heneman, R. L. & Cohen, D. J. (1988). Supervisory and employee characteristics as correlates of employee salary increases. Personnel Psychology, 41, 345-360.

Hill, A. B., & Thickett, J. M. B. (1966). Batch size, cycle time and setting time as determinants of productivity in skilled machining work. Occupational Psychology, 40, 83-89.

Hilton, A. C., Bolin, S. F., Parker, J. W., Jr., Taylor, E. K., & Walker, W. B. (1955). The validity of personnel assessments by professional psychologists. Journal of Applied Psychology, 39, 287-293.

Hoffman, C. C., Nathan, B. R., & Holden, L. M. (1991). A comparison of validation criteria: Objective versus subjective performance measures and self- versus supervisor ratings. Personnel Psychology, 44, 601-619.

Hogan, J., Hogan, R., & Busch, C. M. (1984). How to measure service orientation. Journal of Applied Psychology, 69, 167-173.

Hollander, E. P. (1965). Validity of peer nominations in predicting a distant performance criterion. Journal of Applied Psychology, 49, 434-438.

Hollenbeck, J. R., & Williams, C. R. (1986). Turnover functionality versus turnover frequency: A note on work attitudes and organizational effectiveness. Journal of Applied Psychology, 71, 606-611.

Holzbach, R. L. (1978). Rater bias in performance ratings: Superior, self-, and peer ratings. Journal of Applied Psychology, 63, 579-588.

Hough, L. M. (1984). Development and evaluation of the "Accomplishment Record" method of selecting and promoting professionals. Journal of Applied Psychology, 69, 135-146.

Huck, J. R., & Bray, D. W. (1976). Management assessment center evaluations and subsequent job performance of white and black females. Personnel Psychology, 29, 13-30.

Hughes, G. L., & Prien, E. P. (1986). An evaluation of alternate scoring methods for the mixed standard scale. Personnel Psychology, 39, 839-847.

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

Ilgen, D. R., Hollenbeck, J. H. (1977). The role of job satisfaction in absence behavior. Organizational Behavior and Human Performance, 19, 148-161.

Ivancevich, J. M. (1978). The performance to satisfaction relationship: A causal analysis of stimulating and nonstimulating jobs. Organizational Behavior and Human Performance, 22, 350-365.

Ivancevich, J. M. (1980). A longitudinal study of behavioral expectation scales: Attitudes and performance. Journal of Applied Psychology, 65, 139-146.

Ivancevich, J. M. (1983). Contrast effects in performance evaluation and reward practices. Academy of Management Journal, 26, 465-476.

Ivancevich, J. M. (1985). Predicting absenteeism from prior absence and work attitudes. Academy of Management Journal, 28, 219-228.

Ivancevich, J. M., & McMahon, J. T. (1977). Black-white differences in a goal-setting program. Organizational Behavior and Human Performance, 20, 287-300.

Ivancevich, J. M., & McMahon, T. J. (1982). The effects of goal setting, external feedback, and self-generated feedback on outcome variables: A field experiment. Academy of Management Journal, 25, 359-372.

Ivancevich, J. M., & Smith, S. V. (1981). Goal setting interview skills training: Simulated and on-the-job analyses. Journal of Applied Psychology, 66, 697-705.

Ivancevich, J. M., & Smith. S. V. (1982). Job difficulty as interpreted by incumbents: A study of nurses and engineers. Human Relations, 35, 391-412.

Jamal, M. (1981). Shift work related to job attitudes, social participation and withdrawal behavior: A study of nurses and industrial workers. Personnel Psychology, 34, 535-547.

Jamal, M. (1984). Job stress and job performance controversy: An empirical assessment. Organizational Behavior and Human Performance, 33, 1-21.

Jamal, M. (1985). Relationship of job stress to job performance: A study of managers and blue-collar workers. Human Relations, 38, 409-424.

Jamal, M. (1986). Moonlighting: Personal, social, and organizational consequences. Human Relations, 39, 977-990.

James, L. R., & Ellison, R. L. (1973). Creation composites for scientific creativity. Personnel Psychology, 26, 147-161.

Jansen, D. J., Bonk, E. C., & Garvey, F. J. (1973). MMPI characteristics of clergymen in counseling training and their relationship to supervisor's and peers' ratings of counseling effectiveness. Psychological Reports, 33, 695-698.

Jay, R., & Copes, J. (1957). Seniority and criterion measures of job proficiency. Journal of Applied Psychology, 41, 58-60.

Johns, G. (1978). Attitudinal and nonattitudinal predictors of two forms of absence from work. Organizational Behavior and Human Performance, 22, 431-444.

Johnson, J. A., & Hogan, R. (1981). Vocational interests, personality and effective police performance. Personnel Psychology, 34, 49-53.

Jones, J. W., & Terris, W. (1983). Predicting employees' theft in home improvement centers. Psychological Reports, 52, 187-201.

Jordan, J. L. (1989). Effects of race on interrater reliability of peer ratings. Psychological Reports, 64, 1221-1222.

Jurgensen, C. E. (1934). Intercorrelations in merit rating traits. Journal of Applied Psychology, 18, 240-243.

Katz, R. (1978). The influence of job longevity on employee reactions to task characteristics. Human Relations, 31, 703-725.

Keller, R. T. (1984). The role of performance and absenteeism in the prediction of turnover. Academy of Management Journal, 27, 176-183.

King, L. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in multidimensional forced-choice performance evaluation scale. Journal of Applied Psychology, 65, 507-516.

Kingstrom, P. O., & Mainstone, L. E. (1985). An investigation of the rater-ratee acquaintance and rater bias. Academy of Management Journal, 28, 641-653.

Kinicki, A. J., Lockwood, C. A., Hom, P. W., & Griffeth, R. W. (1990). Interviewer predictions of applicant qualifications and interviewer validity: Aggregate and individual analyses. Journal of Applied Psychology, 75, 477-486.

Kirchner, W. K. (1960). Predicting ratings of sales success with objective performance information. Journal of Applied Psychology, 44, 398-403.

Klaas, B. S. (1989). Managerial decision making about employee grievances: The impact of the grievant's work history. Personnel Psychology, 42, 53-68.

Kleiman, L. S., Lounsbury, J. W., & Faley, R. H. (1987). Performance evaluation: A case for the use of consensus ratings. Psychological Reports, 60, 407-413.

Klimoski, R. J., & Hayes, N. J. (1980). Leader behavior and subordinate motivation. Personnel Psychology, 33, 543-555.

Knapp, R. R. (1963). Personality correlates of delinquency rate in a navy sample. Journal of Applied Psychology, 47, 68-71.

Landy, F. J., & Guion, R. M. (1970). Development of scales for the measurement of work motivation. Organizational Behavior and Human Decision Processes, 5, 93-103.

Lane, J., & Herriot, P. (1990). Self-ratings, supervisory ratings, positions and performance. Journal of Occupational Psychology, 63, 77-88.

Latham, G. P., & Wexley, K. N. (1977). Behavioral observation scales for performance appraisal purposes. Personnel Psychology, 30, 255-268.

Latham, G. P., Fay, C. H., & Saari, L. M. (1979). The development of behavioral observation scales for appraising the performance of foremen. Personnel Psychology, 32, 299-311.

Lawshe, C. H., & McGinley, A. D., Jr. (1951). Job performance criteria studies: I. The job performance of proofreaders. Journal of Applied Psychology, 35, 316-320.

Lee, R., Malone, M., & Greco, S. (1981). Multitrait-multimethod-multirater analysis of performance ratings for law enforcement personnel. _Journal of Applied Psychology, 66_, 625-632.

Lifter, M. L., Bass, A. R., & Nussbaum, H. (1971). Effort expenditure and job performance of line and staff personnel. _Organizational Behavior and Human Decision Processes, 6_, 501-515.

Lopez, Jr., F. M. (1966). Current problems in test performance of job applicants: I. _Personnel Psychology, 19_, 10-18.

Love, K. G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. _Journal of Applied Psychology, 66_, 451-457.

Love, K. G., & O'Hara, K. (1987). Predicting job performance of youth trainees under a job training partnership act program (JTPA): Criterion validation of a behavior-based measure of work maturity. _Personnel Psychology, 40_, 323-340.

MacKenzie, S. B., Podsakoff, P. M., & Fetter, R. (1991). Organizational citizenship behavior and objective productivity as determinants of managerial evaluations of salespersons' performance. _Organizational Behavior and Human Decision Processes, 50_, 123-150.

Mathieu, J. E., & Kohler, S. S. (1990). A cross-level examination of group absence influences on individual absence. _Journal of Applied Psychology, 75_, 217-220.

Mathieu, J. E., & Kohler, S. S. (1990). A test of the interactive effects of organizational commitment and job involvement on various types of absence. _Journal of Vocational Behavior, 36_, 33-44.

Matteson, M. T., Ivancevich, J. M., & Smith, S. V. (1984). Relation of type A behavior to performance and satisfaction among sales personnel. _Journal of Vocational Behavior, 25_, 203-214.

Mayfield, E. C. (1970). Management selection: Buddy nominations revisted. _Personnel Psychology, 23_, 377-391.

McCarrey, M. W., & Edwards, S. A. (1973). Organizational climate conditions for effective research scientist role performance. Organizational Behavior and Human Performance, 9, 439-459.

McCauley, C. D., Lombardo, M. M., & Usher, C. J. (1989). Diagnosing management development needs: An instrument based on how managers develop. Journal of Management, 15, 389-403.

McEvoy, G. M., & Beatty, R. W. (1989). Assessment centers and subordinate appraisals of managers: A seven-year examination of predictive validity. Personnel Psychology, 42, 37-52.

McGehee, W., & Owen, E. B. (1940). Authorized and unauthorized rest pauses in clerical work. Journal of Applied Psychology, 24, 605-614.

Meglino, B. M., Ravlin, E. C., & Adkins, C. L. (1989). A work values approach to corporate culture: A field test of the value congruence process and its relationship to individual outcomes. Journal of Applied Psychology, 74, 424-432.

Meredith, G. M. (1990). Dossier evaluation in screening candidates for excellence in teaching awards. Psychological Reports, 67, 879-882.

Meyer, J. P., Paunonen, S. V., Gellatly, I. R., Goffin, R. D., & Jackson, D. N. (1989). Organizational commitment and job performance: It's the nature of the commitment that counts. Journal of Applied Psychology, 74, 152-156.

Miner, J. B. (1970). Executive and personnel interviews as predictors of consulting success. Personnel Psychology, 23, 521-538.

Miner, J. B. (1970). Psychological evaluations as predictors of consulting success. Personnel Psychology, 23, 391-405.

Miner, J. M. (1962). Personality and ability factors in sales performance. Journal of Applied Psychology, 46, 6-13.

Mitchell, T. R., & Albright, D. W. (1972). Expectancy theory predictions of the satisfaction, effort, performance, and retention of naval aviation officers. Organizational Behavior and Human Decision Processes, 8, 1-20.

Morse, J. J., & Wagner, F. R. (1978). Measuring the process of managerial effectiveness. Academy of Management Journal, 21, 23-35.

Motowidlo, S. J. (1982). Relationship between self-rated performance and pay satisfaction among sales representatives. Journal of Applied Psychology, 67, 209-213.

Mount, M. K. (1984). Psychometric properties of subordinate ratings of managerial performance. Personnel Psychology, 37, 687-702.

Nealey, S. M., & Owen, T. W. (1970). A multitrait-multimethod analysis of predictors and criteria of nursing performance. Organizational Behavior and Human Decision Processes, 5, 348-365.

Neider, L. L. (1980). An experimental field investigation utilizing an expectancy theory view of participation. Organizational Behavior and Human Performance, 5, 425-442.

O'Connor, E. J., Peters, L. H., Pooyan, A., Weekly, J., Frank, B., & Erenkrantz, B. (1984). Situational constraint effects on performance, affective reactions, and turnover: A field replication and extension. Journal of Applied Psychology, 69, 663-672.

Oldham, G. R. (1976). The motivational strategies used by supervisors: Relationships to effectiveness indicators. Organizational Behavior and Human Performance, 15, 66-86.

Oldham, G. R., & Miller, H. E. (1979). The effect of other's job complexity on employee reactions to work. Human Relations, 32, 247-260.

Oldham, G. R., Kulik, C. T., Ambrose, M. L., Stepina, L. P., & Brand, J. F. (1986). Relations between job facet comparisons and employee reactions. Organizational Behavior and Human Decision Processes, 38, 28-47.

Organ, D. W., & Konovsky, M. (1989). Cognitive versus affective determinants of organizational citizenship behavior. Journal of Applied Psychology, 74, 157-164.

Orpen, C. (1985). Patterned behavior description interviews versus unstructured interviews: A comparative validity study. Journal of Applied Psychology, 70, 774-776.

Otten, M. W., & Kahn, M. (1975). Effectiveness of crisis center volunteers and the personal orientation inventory. Psychological Reports, 37, 1107-1111.

Parasuraman, S. (1982). Predicting turnover intentions and turnover behavior: A multivariate analysis. Journal of Vocational Behavior, 21, 111-121.

Parsons, C. K., Herold, D. M., & Leatherwood, M. L. (1985). Turnover during initial employment: A longitudinal study of the role of causal attributions. Journal of Applied Psychology, 70, 337-341.

Pierce, J. L., Dunham, R. B., & Blackburn, R. S. (1979). Social systems structure, job design, and growth need strength: A test of a congruency model. Academy of Management Journal, 22, 223-240.

Plag, J. A., & Goffman, J. M. (1967). The armed forces qualification test: Its validity in predicting military effectiveness for naval enlistees. Personnel Psychology, 20, 323-340.

Prien, E. P., & Kult, M. (1968). Analysis of performance criteria and comparison of a priori and empirically-derived keys for a forced-choice scoring. Personnel Psychology, 21, 505-513.

Puffer, S. M. (1987). Prosocial behavior, noncompliant behavior, and work performance among commission salespeople. Journal of Applied Psychology, 72, 615-621.

Pulakos, E. D., & Wexley, K. N. (1983). The relationship among perceptual similarity, sex, and performance ratings in manager-subordinate dyads. Academy of Management Journal, 26, 129-139.

Pulakos, E. D., Borman, W. C., & Hough, L. M. (1988). Test validation for scientific understanding: Two demonstrations of an approach to studying predictor-criterion linkages. Personnel Psychology, 41, 703-716.

Rabinowitz, S., & Stumpf, S. A. (1987). Facets of role conflict, role-specific performance, and organizational level within the academic career. Journal of Vocational Behavior, 30, 72-83.

Reilly, R. R., Zedeck, S., & Tenopyr, M. L. (1979). Validity and fairness of physical ability tests for predicting performance in craft jobs. Journal of Applied Psychology, 64, 262-274.

Rhodes, S. R., & Steers, R. M. (1981). Conventional vs. worker-owned organizations. Human Relations, 34, 1013-1035.

Robinson, D. D. (1970). Predicting police effectiveness from self reports of relative time spent in task performance. Personnel Psychology, 23, 327-345.

Ronan, W. W. (1963). A factor analysis of eleven job performance measures. Personnel Psychology, 16, 255-267.

Rosinger, G., Myers, L. B., Levy, G. W., Loar, M., Mohrman, S. A., & Stock, J. R. (1982). Development of a behaviorally based performance appraisal system. Personnel Psychology, 35, 75-88.

Ross, P. F., & Dunfield, N. M. (1964). Selecting salesmen for an oil company. Personnel Psychology, 17, 75-84.

Rosse, J. G. (1987). Job-Related ability and turnover. Journal of Business and Psychology, 1, 326-336.

Rosse, J. G. (1988). Relations among lateness, absence, and turnover: Is there a progression of withdrawal? Human Relations, 41, 517-531.

Rosse, J. G., & Kraut, A. I. (1983). Reconsidering the vertical dyad linkage model of leadership. Journal of Occupational Psychology, 56, 63-71.

Rosse, J. G., Miller, H. E., & Barnes, L. K. (1991). Combining personality and cognitive ability predictors for hiring service-oriented employees. Journal of Business and Psychology, 5, 431-445.

Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable?. Journal of Applied Psychology, 75, 175-184.

Rush, Jr., C. H. (1953). A factorial study of sales criteria. Personnel Psychology, 6, 9-24.

Russell, C. J. (1990). Selecting top corporate leaders: An example of biographical information. Journal of Management, 16, 73-86.

Russell, C. J., Mattson, J., Devlin, S. E., & Atwater. D. (1990). Predictive validity of biodata items generated from retrospective life experience essays. Journal of Applied Psychology, 75, 569-580.

Russell, G. W. (1974). Machiavellianism, locus of control, aggression, performance, and precautionary behaviour in ice hockey. Human Relations, 27, 825-837.

Saal, F. E., & Landy, F. J. (1977). The mixed standard rating scale: An evaluation. Organizational Behavior and Human Decision Processes, 18, 19-35.

Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. Journal of Applied Psychology, 73, 482-486.

Saunders, D. R., Baehr, M. E., & Froemel, E. C. (1977). Identification of, and correction for, effects of racial bias and job tenure on supervisory ratings of bus operator performance. Psychological Reports, 40, 859-865.

Schippmann, J. S., & Prien, E. P. (1986). Psychometric evaluation of an integrated assessment procedure. Psychological Reports, 59, 111-122.

Schlotzhauer, D. L., & Rosse, J. G. (1985). A five-year study of a positive incentive absence contro program. Personnel Psychology, 38, 575-585.

Schuerger, J. M., Kochevar, K. F., & Reinwald, J. E. (1982). Male and female correction officers: Personality and rated performance. Psychological Reports, 51, 223-228.

Schwab, D. P. (1991). Contextual variables in employee performance-turnover relationships. Academy of Management Journal, 34, 966-975.

Seashore, S. E., Indik, B. P., & Georgopoulos, B. S. (1960). Relationships among criteria of job performance. Journal of Applied Psychology, 44, 195-202.

Seybolt, J. W., & Pavett, C. M. (1979). The prediction of effort and performance among hospital professionals: Moderating effects of feedback on expectancy theory formulations. Journal of Occupational Psychology, 52, 91-105.

Sheridan, J. E. (1985). A catastrophe model of employee withdrawal leading to low job performance, high absenteeism, and job turnover during the first year of employment. Academy of Management Journal, 28, 88-109.

Sheridan, J. E. (1992). Organizational culture and employee retention. Academy of Management Journal, 35, 1036-1056.

Sheridan, J. E., & Vredenburgh, D. J. (1978). Usefulness of leadership behavior and social power variables in predicting job tension, performance, and turnover of nursing employees. Journal of Applied Psychology, 63, 89-95.

Sheridan, J. E., & Vredenburgh, D. J. (1979). Structural model of leadership influence in a hospital organization. Academy of Management Journal, 22, 6-21.

Shore, L. M., & Martin, H. J. (1989). Job satisfaction and organizational commitment in relation to work performance and turnover intentions. Human Relations, 42, 625-638.

Siegel, A. I., Schultz, D. G., Fischl, M. A., & Lanterman, R. S. (1968). Absolute scaling of job performance. Journal of Applied Psychology, 52, 313-318.

Skinner, E. W. (1969). Relationships between leadership behavior patterns and organizational-situational variables. Personnel Psychology, 22, 489-494.

Slocum, J. W., Jr., & Cron, W. L. (1985). Job attitudes and performance during three career stages. Journal of Vocational Behavior, 26, 126-145.

Smith, W. J., Albright, L. E., Glennon, J. R., & Owens, W. A. (1961). The prediction of research competence and creativity from personal history. Journal of Applied Psychology, 45, 59-62.

Sneath, F. A., White, G. C., & Randell, G. A. (1966). Validating a workshop reporting procedure. Occupational Psychology, 40, 15-29.

Soar, R. S. (1956). Personal history as a predictor of success in service station management. Journal of Applied Psychology, 40, 383-385.

South, J. C. (1974). Early career performance of engineers-Its composition and measurement. Personnel Psychology, 27, 225-243.

Spencer, D. G., & Steers, R. M. (1980). The influence of personal factors and perceived work experiences on employee turnover and absenteeism. Academy of Management Journal, 23, 567-572.

Spencer, D. G., & Steers, R. M. (1981). Performance as a moderator of the job satisfaction-turnover relationship. Journal of Applied Psychology, 66, 511-514.

Spitzer, M. E., & McNamara, W. J. (1964). A managerial selection study. Personnel Psychology, 17, 19-40.

Sprecher, T. B. (1959). A study of engineers' criteria for creativity. Journal of Applied Psychology, 43, 141-148.

Staw, B. M., & Oldham, G. R. (1978). Reconsidering our dependent variables: A critique and empirical study. Academy of Management Journal, 21, 539-559.

Steel, R. P., & Mento, A. J. (1986). Impact of situational constraints on subjective and objective criteria of managerial job performance. Organizational Behavior and Human Decision Processes, 37, 254-265.

Steel, R. P., & Ovalle, N. K. (1984). Self-appraisal based upon supervisory feedback. Personnel Psychology, 37, 667-685.

Steel, R. P., Mento, A. J., & Hendrix, W. H. (1987). Constraining forces and the work performance of finance company cashiers. Journal of Management, 13, 473-482.

Steel, R. P., Shane, G. S., & Kennedy, K. A. (1990). Effects of social-system factors on absenteeism, turnover, and job performance. Journal of Business and Psychology, 4, 423-430.

Steers, R. M. (1975). Task-Goal attributes, n-achievement, and supervisory performance. Organizational Behavior and Human Performance, 13, 392-403.

Stout, S. K., Slocum, J. W., Jr., & Cron, W. L. (1987). Career transitions of supervisors and subordinates. Journal of Vocational Behavior, 30, 124-137.

Stumpf, S. A. (1981). Career roles, psychological success, and job attitudes. Journal of Vocational Behavior, 19, 98-112.

Stumpf, S. A., & Dawley, P. K. (1981). Predicting voluntary and involuntary turnover using absenteeism and performance indices. Academy of Management Journal, 24, 148-163.

Stumpf, S. A., & Rabinowitz, S. (1981). Career stage as a moderator of performance relationships with facets of job satisfaction and role perceptions. Journal of Vocational Behavior, 18, 202-218.

Sulkin, H. A., & Pranis, R. W. (1967). Comparison of grievants and non-grievants in a heavy machinery company. Personnel Psychology, 20, 111-119.

Summers, T. P., & Hendrix, W. H. (1991). Development of a turnover model that incorporates a matrix measure of valence-instrumentality-expectancy perceptions. Journal of Business and Psychology, 6, 227-245.

Sutton, C. D., & Woodman, R. W. (1989). Pygmalion goes to work: The effects of supervisor expectations in a retail setting. Journal of Applied Psychology, 74, 943-950.

Svetlik, B., Prien, E., & Barrett, G. (1964). Relationships between job difficulty, employee's attitude toward his job, and supervisory ratings of the employee effectiveness. Journal of Applied Psychology, 48, 320-324.

Swaroff, P. G., Barclay, L. A., & Bass, A. R. (1985). Recruiting sources: Another look. Journal of Applied Psychology, 70, 720-728.

Szilagyi, A. D. (1980). Causal inferences between leader reward behaviour and subordinate performance, absenteeism, and work satisfaction. Journal of Occupational Psychology, 53, 195-204.

Taylor, E. K., Schneider, D. E., & Symons, N. A. (1953). A short forced-choice evaluation form for salesmen. Personnel Psychology, 6, 393-401.

Taylor, R. L., & Wilsted, W. D. (1974). Capturing judgment policies: A field study of performance appraisal. Academy of Management Journal, 17, 440-449.

Taylor, R. L., & Wilsted, W. D. (1976). Capturing judgment policies in performance rating. Industrial Relations, 15, 216-224.

Taylor, S. M., & Schmidt, D. W. (1983). A process-oriented investigation of recruitment source effectiveness. Personnel Psychology, 36, 343-354.

Taylor, S. M., Locke, E. A., Lee, C., & Gist, M. E. (1984). Type A behavior and faculty research productivity: What are the mechanisms? Organizational Behavior and Human Performance, 34, 402-418.

Tenopyr, M. L. (1969). The comparative validity of selected leadership scales relative to success in production management. Personnel Psychology, 22, 77-85.

Thamhain, H. J., & Gemmill, G. R. (1974). Influence styles of project managers: Some performance correlates. Academy of Management Journal, 17, 216-224.

Thompson, D. E., & Thompson, T. A. (1985). Task-Based performance appraisal for blue-collar jobs: Evaluation of race and sex effects. Journal of Applied Psychology, 70, 747-753.

Thomson, H. A. (1970). Comparison of predictor and criterion judgments of managerial performance using the multitrait-multimethod approach. Journal of Applied Psychology, 54, 496-502.

Thornton, G. C. (1968). The relationship between supervisory- and self-appraisals of executive performance. Personnel Psychology, 21, 441-455.

Tsui, A. S., & Ohlott, P. (1988). Multiple assessment of managerial effectiveness: Interrater agreement and consensus in effectiveness models. Personnel Psychology, 41, 779-802.

Turnage, J. J., & Muchinsky, P. M. (1984). A comparison of the predictive validity of assessment center evaluations versus traditional measures in forecasting supervisory job performance: Interpretive implications of criterion distortion for the assessment paradigm. Journal of Applied Psychology, 69, 595-602.

Validity information exchange. No. 7-045. (1954). Personnel Psychology, 7, 279-281.

Validity information exchange. No. 7-089. (1954). Personnel Psychology, 7, 565-566.

Validity information exchange. No. 8-35. (1955). Personnel Psychology, 8, 501-504.

Validity information exchange. No. 11-21. (1958). Personnel Psychology, 11, 264-265.

Validity information exchange. No. 11-27. (1958). Personnel Psychology, 11, 583-584.

Validity information exchange. No. 11-30. (1958). Personnel Psychology, 11, 587-590.

Validity information exchange. (1956). No. 9-32. Personnel Psychology, 9, 375-377.

Validity information exchange. (1958). No. 11-10. Personnel Psychology, 11, 121-122.

Vecchio, R. P. (1987). Situational leadership theory: An examination of a prescriptive theory. Journal of Applied Psychology, 72, 444-451.

Vecchio, R. P., & Gobdel, B. C. (1984). The vertical dyad linkage model of leadership: Problems and prospects. Organizational Behavior and Human Performance, 34, 5-20.

Villanova, P., & Bernardin, J. H. (1990). Work behavior correlates of interviewer job compatibility. Journal of Business and Psychology, 5, 179-195.

Waldman, D. A., Yammarino, F. J., & Avolio, B. J. (1990). A multiple level investigation of personnel ratings. Personnel Psychology, 43, 811-835.

Waters, L. K., & Roach, D. (1979). Job satisfaction, behavioral intention, and absenteeism as predictors of turnover. Personnel Psychology, 32, 393-397.

Wexley, K. N., & Pulakos, E. D. (1982). Sex effects on performance ratings in manager-subordinate dyads: A field study. Journal of Applied Psychology, 67, 433-439.

Wexley, K. N., & Youtz, M. A. (1985). Rater beliefs about others: Their effects on rating errors and rater accuracy. Journal of Occupational Psychology, 58, 265-277.

Wexley, K. N., Alexander, R. A., Greenawalt, J. P., & Couch, M. A. (1980). Attitudinal congruence and similarity as related to interpersonal evaluations in manager-subordinate dyads. Academy of Management Journal, 23, 320-330.

Whitlock, G. H., Barker, J. A., Blackett, F. E., & Stone, C. H. (1963). The relation of accident proneness and the tendency to report injuries. Personnel Psychology, 16, 163-169.

Whitlock, G. H., Clouse, R. J., & Spencer, W. F. (1963). Predicting accident proneness. Personnel Psychology, 16, 35-44.

Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. Journal of Management, 17, 601-617.

Williams, W. E., & Seiler, D. A. (1973). Relationships between measures of effort and job performance. Journal of Applied Psychology, 57, 49-54.

Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. <u>Personnel Psychology</u>, <u>42</u>, 235-261.

Wolpin, J., & Burke, R. J. (1985). Relationships between absenteeism and turnover: A function of measures?. <u>Personnel Psychology</u>, <u>38</u>, 57-74.

Woodmansee, J. J. (1978). Validation of the nurturance scale of the Edwards Personal Preference schedule. <u>Psychological Reports</u>, <u>42</u>, 495-498.

Worbois, G. M. (1975). Validation of externally developed assessment procedures for identification of supervisory potential. <u>Personnel Psychology</u>, <u>28</u>, 77-91.

Yukl, G. A., & Latham, G. P. (1978). Interrelationships among employee participation, individual differences, goal difficulty, goal acceptance, goal instrumentality, and performance. <u>Personnel Psychology</u>, <u>31</u>, 305-323.

Zaccaro, S. J., Craig, B., & Quinn, J. (1991). Prior absenteeism, supervisory style, job satisfaction, and personal characteristics: An investigation of some mediated and moderated linkages to work absenteeism. <u>Organizational Processes and Human Decision Processes</u>, <u>50</u>, 24-44.

Zalesny, M. D., & Kirsch, M. P. (1989). The effect of similarity on performance ratings and interrater agreement. <u>Human Relations</u>, <u>42</u>, 81-96.

Zedeck, S., & Baker, H. T. (1972). Nursing performance as measured by behavioral expectation scales: A multitrait-multirater analysis. <u>Organizational Behavior and Human Decision Processes</u>, <u>7</u>, 457-466.

REFERENCES

Allan, P. (1981). Managers at work: A large scale study of the managerial job in New York city government. Academy of Management Journal, 24, 613-619.

Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917-1992. Journal of Applied Psychology, 77, 836-874.

Baier, D. E., & Dugan, R. D. (1957). Factors in sales success. Journal of Applied Psychology, 41, 37-40.

Barnard, C. (1938). The functions of the executive. Cambridge, MA: Harvard University Press.

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship". Academy of Management Journal, 26, 587-595.

Bechtoldt, H. P. (1947). Problems of establishing criterion measures. In D. B. Stuit (Ed.), Personnel research and test development in the bureau of Navy Personnel (pp. 357-380). Princeton, N.J.: Princeton University Press.

Borman, W. C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. Unpublished manuscript.

Borman, W. C., & Motowidlo, S. J. (1992). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in Organizations (pp. 71-98). San Francisco, CA: Jossey Bass.

Brief, A. P., & Motowidlo, S. J. (1986). Prosocial organizational behaviors. Academy of Management Review, 11, 710-725.

Brogden, H. E., & Taylor, E. K. (1950). The dollar criterion--applying the cost accounting concept to criterion construction. Personnel Psychology, 3, 133-154.

Brumback, G. B., & Vincent, J. W. (1970). Factor analysis of work performed data for a sample of administrative, professional, and scientific positions. Personnel Psychology, 23, 101-107.

Bycio, P. (1992). Job performance and absenteeism: A review and meta-analysis. Human Relations, 45, 193-220.

Campbell, J. P. (1990a). The role of theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology (2nd Edition-Vol. 1) (pp. 39-73). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P. (1990b). Modeling the performance prediction problem in Industrial and Organizational Psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology (2nd Edition-Vol. 1) (pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P. (1990c). An overview of the army selection and classification project. Personnel Psychology, 43, 231-239.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1992). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in Organizations (pp. 35-70). San Francisco, CA: Jossey Bass.

Edgerton, H. A., & Kolbe, L. E. (1936). The method of minimum variation for the combination of criteria. Psychometrika, 1, 183-187.

Ghiselli, E. E. (1966). The validity of occupational aptitude tests. New York: Wiley.

Ghiselli, E. E., & Brown, C. W. (1948). Personnel and industrial Psychology. New York: McGraw Hill.

Goldberg, L. R. (1992). What the hell took so long? Donald Fiske and the big-five factor structure. To appear in P. E. Shrout & S. T. Fiske (Eds.), Advances in personality research, methods, and theory: A festschrift honoring Donald W. Fiske. New York, NY: Erlbaum.

Gottfredson, L. S. (1988). Reconsidering fairness: A matter of social and ethical priorities. _Journal of Vocational Behavior, 33_, 293-319.

Graham, J. W. (1986). _Organizational citizenship informed by political theory_. Paper presented at the Academy of Management meetings, Chicago, IL.

Gunderson, E. K. E., & Ryman, D. H. (1971). Convergent and discriminant validities of performance evaluations in extremely isolated groups. _Personnel Psychology, 24_, 715-724.

Hackett, R. D., & Guion, R. M. (1985). A re-evaluation of the absenteeism-job satisfaction relationship. _Organizational Behavior and Human Decision Processes, 35_, 340-381.

Haring, M. J., Okun, M. A., Stock, W. A., Miller, W., Kinney, C., & Ceurvorst, W. R. (1981, April). _Reliability issues in meta-analyses_. Paper presented at the annual meeting of the American Education Research Association, Los Angeles.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. _Personnel Psychology, 41_, 43-62.

Hattie, J. A., & Hansford, B. C. (1982). Self-Measures and achievement: Comparing a traditional review of literature with a meta-analysis. _Australian Journal of Education, 26_, 71-75.

Hattie, J. A., & Hansford, B. C. (1984). Meta-Analysis: A reflection on problems. _Australian Journal of Psychology, 36_, 239-254.

Hedberg, P. R. (1989). _Individual perceptions of performance feedback in a work setting_. Unpublished Doctoral Dissertation, Minneapolis, MN: University of Minnesota

Helme, W. E., Gibson, N. A., & Brogden, H. E. (1957). _An empirical test of shrinkage problems in personnel classification research_. Personnel Board, Technical Research Note 84.

Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. Personnel Psychology, 39, 811-826.

Hogan, J., & Hogan, R. (1989). How to measure employee reliability. Journal of Applied Psychology, 74, 273-279.

Horst, P. (1941). The prediction of personnel adjustment. Social Sciences Research Council Bulletin, 48, 20-36.

Hough, L. M. (1984). Development and evaluation of the "Accomplishment Record" method of selecting and promoting professionals. Journal of Applied Psychology, 69, 135-146.

Hunter, J. E. (1983). Test validation for 12,000 jobs: An application of job classification and validity generalization analysis to the general aptitude test battery (GATB) (Test Research Report No. 45). Washington, DC: U.S. Employment Service, U. S. Department of Labor.

Hunter, J. E. (1992). CFA: A subroutine in Basic [computer program]. East Lansing, MI: Michigan State University.

Hunter, J. E., & Cohen, S. H. (1969). PACKAGE: A system of computer routines for the analysis of correlational data. Educational and Psychological Measurement, 29, 697-700.

Hunter, J. E., Crosson, J. J., & Friedman, D. H. (1985). The validity of ASVAB for civilian and military job performance. Technical Report.

Hunter, J. E., & Gerbing, D. W. (1982). Unidimensional measurement, second order factor analysis, and causal models. In B. M. Staw & L. L. Cummings (Eds.), Research in organizational behavior (Vol. 4) (pp. 267-320). Greenwich, CT: JAI Press.

Hunter, J. E., & Schmidt, F. L. (1990a). Methods of meta-analysis: Correcting for error and bias in research findings. Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (1990b). Dichotomization of continuous variables: The implications for meta-analysis. Journal of Applied Psychology, 75, 334-349.

Jackson, G. B. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.

Judiesch, M. K., Schmidt, F. L., & Mount, M. K. (1992). Estimates of the dollar value of employee output in utility analyses: An empirical test of two theories. Journal of Applied Psychology, 77, 234-250.

Katz, D. (1964). The motivational basis of organizational behavior. Behavioral Science, 9, 131-146.

Katz, D., & Kahn, R. L. (1978). The social psychology of organizations. New York: Wiley.

King, L. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in multidimensional forced-choice performance evaluation scale. Journal of Applied Psychology, 65, 507-516.

Komaki, J. L., Zlotnick, S., & Jensen, M. (1986). Development of an operant-based taxonomy and observational index of supervisory behavior. Journal of Applied Psychology, 71, 260-269.

Krug, R. E. (1961). Personnel selection. In B. V. H. Gilmer (Ed.), Industrial Psychology (pp. 91-129). New York: McGraw Hill.

Law, K. S., Schmidt, F. L., & Hunter, J. E. (1992). Nonlinearity of range restriction corrections in meta-analysis: A test of an improved procedure. Manuscript submitted for publication.

Lent, R. H., Auerbach, H. A., & Levin, L. S. (1971). Predictors, criteria, and significant results. Personnel Psychology, 24, 519-533.

McEvoy, G. M., & Cascio, W. F. (1987). Do good or poor performers leave? A meta-analysis of the relationship between performance and turnover. Academy of Management Journal, 30, 744-762.

Mitra, A., Jenkins, Jr., G. D., & Gupta, N. (1992). A meta-analytic review of the relationship between absence and turnover. Journal of Applied Psychology, 77, 879-889.

Mount, M. K. (1984). Psychometric properties of subordinate ratings of managerial performance. Personnel Psychology, 37, 687-702.

Nagle, B. F. (1953). Criterion development. Personnel Psychology, 6, 271-289.

Nunnally, J. C. (1978). Psychometric Theory. New York: McGraw Hill.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (in press). Meta-Analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance [Monograph]. Journal of Applied Psychology.

Oppler, S. H., Sager, C. E., McCloy, R. A., & Rosse, R. L. (1993, May). The role of performance determinants in the development of prediction equations. In F. L. Schmidt (Chair), Job Performance: Theories of determinants and factor structure. Symposium conducted at the eighth annual meeting of the Society of Industrial and Organizational Psychologists, San Francisco, CA.

Organ, D. W. (1988). Organizational citizenship behavior: The good soldier syndrome. Lexington, MA: Lexington Books.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.

Prien, E. P., & Kult, M. (1968). Analysis of performance criteria and comparison of a priori and empirically-derived keys for a forced-choice scoring. Personnel Psychology, 21, 505-513.

Ronan, W. W. (1963). A factor analysis of eleven job performance measures. Personnel Psychology, 16, 255-267.

Rosenberg, S., & Sedlak, A. (1972). Structural representations of perceived personality trait relationships. In A. K. Rommey, R. J. Shepard, & S. B. Nerlave (Eds.), Multidimensional scaling. New York: Seminar Press.

Rush, C. H., Jr. (1953). A factorial study of sales criteria. Personnel Psychology, 6, 9-24.

Rush, M. C., Thomas, J. C., & Lord, R. G. (1977). Implicit leadership theory: A potential threat to the internal validity of leader behavior questionnaires. Organizational Behavior and Human Decision Process, 20, 93-110.

Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. Personnel Psychology, 42, 491-529.

Sackett, P. R., & Harris, M. M. (1984). Honesty testing for personnel selection: A review and critique. Personnel Psychology, 37, 221-246.

Sager, C. E. (1990). A component model of halo: Peer and supervisory ratings of job performance. Unpublished Doctoral Dissertation, University of Minnesota, Minneapolis.

Sager, C. E., Nitti, T. A., & Hazucha, J. F. (1993, May). The factor structure of job performance ratings: The effect of correcting for idiosyncratic influences on the relationships among dimensions. In F. L. Schmidt (Chair), Job Performance: Theories of determinants and factor structure. Symposium conducted at the eighth annual meeting of the Society of Industrial and Organizational Psychologists, San Francisco, CA.

Schmidt, F. L. (1980). The measurement of job performance. Unpublished manuscript.

Schmidt, F. L. (1992). Psychology at the cutting edge. In N. Schmitt & W. C. Borman (Eds.), Personnel Selection in Organizations (pp. 497-515). San Francisco, CA: Jossey-Bass.

Schmidt, F. L., & Hunter, J. E. (1992). Development of a causal model of processes determining job performance. Current Directions in Psychological Science, 1, 89-92.

Schmidt, F. L., Hunter, J. E., & Urry, J. E. (1976). Statistical power in criterion-related validation studies. Journal of Applied Psychology, 61, 473-485.

Schmidt, F. L., & Kaplan, L. B. (1971). Composite versus multiple criteria: A review and resolution of the controversy. Personnel Psychology, 24, 419-434.

Schmidt, F. L., Law, K., Hunter, J. E., Rothstein, H. R., Pearlman, K., & McDaniel, M. (1993). Refinements in validity generalization methods: Implications for the situational specificity hypothesis. Journal of Applied Psychology, 78, 3-12.

Sheridan, J. E. (1985). A catastrophe model of employee withdrawal leading to low job performance, high absenteeism, and job turnover during the first year of employment. Academy of Management Journal, 28, 88-109.

Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. Journal of Applied Psychology, 68, 653-663.

Smith, P. C. (1976). Behavior, results, and organizational effectiveness: The problem of criteria. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology (pp. 745-775). Chicago: Rand McNally.

Spearman, C. (1904). "General intelligence", objectively determined and measured. American Journal of Psychology, 15, 201-293.

Stone, C. H., & Kendall, W. E. (1956). Effective personnel selection procedures. Englewood Cliffs, N. J.: Prentice-Hall.

Sulkin, H. A., & Pranis, R. W. (1967). Comparison of grievants with non-grievants in a heavy machinery company. Personnel Psychology, 20, 111-119.

Thompson, D. E., & Thompson, T. A. (1985). Task-Based performance appraisal for blue-collar jobs: Evaluation of race and sex effects. Journal of Applied Psychology, 70, 747-753.

Tiffin, J., & McCormick, E. J. (1958). Industrial Psychology. Englewood Cliffs, N. J.: Prentice-Hall.

Toops, H. A. (1944). The criterion. Educational and psychological Measurement, 4, 271-297.

Whetzel, D. L., & McDaniel, M. A. (1988). Reliability of validity generalization data bases. _Psychological Reports, 63_, 131-134.

Williams, C. R. (1990). Deciding when, how, and if to correct turnover correlations. _Journal of Applied Psychology, 75_, 732-737.

Table 1. True Score Intercorrelations Between Job Performance Measures Based on Organizational Records

	ERRORS	PRODUCT	ABSENT.	TENURE	ACCIDENT
Lack of errors (Quality) (**ERRORS**)	$.78^a$				
Productivity (**PRODUCT**)	$\rho = .37$ $K = 13$ $N = 2,731$	$.88^a$			
Absenteeism (Lack of) (**ABSENT.**)	$\rho = .48$ $K = 12$ $N = 1,290$	$\rho = .21$ $K = 7$ $N = 1,825$	$.76^a$		
Tenure (**TENURE**)	$\rho = .12$ $K = 7$ $N = 1,216$	$\rho = .19$ $K = 7$ $N = 1,522$	$\rho = .33$ $K = 26$ $N = 5,888$	$.76^a$	
Accidents (Lack of) (**ACCIDENT**)	$\rho = .24$ $K = 2$ $N = 1,061$	$\rho = .39$ $K = 1$ $N = 975$	$\rho = 1.00$ $K = 3$ $N = 409$	$\rho = .07$ $K = 4$ $N = 806$	$.63^a$

Note. ρ = True score correlation; K= number of correlations included in the psychometric meta-analysis; N= Total sample size involved in that psychometric meta-analyses.
[a]The mean of the internal consistency estimates used in the corrections.

Table 2. True Score Intercorrelations Between Job Performance Measures (**Supervisory Ratings Only**)

SUPERVISOR RATINGS	PERS.	PROD.	EFFORT	OVERALL	ADM.SK.	QUALITY	J.K.	ABS.	PROB& LEAD	COMPLI.	COMMUN.
Personality (**PERS.**)	.77a										
Productivity (**PROD.**)	ρ = .77 K = 64 N= 11,936	.71a									
Effort (**EFFORT**)	ρ = .82 K = 94 N = 21,739	ρ = .90 K = 12 N = 4,360	.73a								
Overall Job Performance (**OVERALL**)	ρ = .71 K = 74 N= 12,388	ρ = .92 K = 37 N = 6,006	ρ = .84 K = 45 N = 8,425	.82a							
Administrat. Skills (**ADM.SK.**)	ρ = .72 K = 60 N 14,393	ρ = .89 K = 14 N = 1,282	ρ = .85 K = 24 N = 4,420	ρ = .64 K = 14 N = 1,479	.82a						
Quality (**QUALITY**)	ρ = .63 K = 37 N = 5,161	ρ = .81 K = 25 N = 4,459	ρ = .77 K = 15 N = 2,774	ρ = .90 K = 17 N = 2,792	ρ = .92 K = 8 N = 613	.82a					
Job Knowledge (**J. K.**)	ρ = .57 K = 76 N = 20,390	ρ = .75 K = 17 N = 2,108	ρ = .86 K = 36 N = 7,232	ρ = .75 K = 28 N = 7,687	ρ = .79 K = 20 N = 3,842	ρ = .74 K = 9 N = 1,528	.72a				
Absenteeism (**ABS.**)	ρ = .51 K = 6 N = 687	ρ = .29 K = 1 N = 86	ρ = .47 K = 2 N = 229	ρ = .37 K = 3 N = 379	ρ = .69 K = 1 N = 143	ρ = .78 K = 1 N = 113	ρ = .45 K = 1 N = 183	.72a			
Problem Solving & Leadership (**PROB&LEAD**)	ρ = .82 K = 130 N= 39,556	ρ = .89 K = 22 N = 4,495	ρ = .92 K = 44 N= 12,020	ρ = .79 K = 43 N = 8,141	ρ = .99 K = 41 N= 12,918	ρ = .78 K = 16 N = 2,008	ρ = .82 K = 38 N = 8,429	ρ = .72 K = 2 N = 229	.82a		
Compliance & Acceptance of Authority (**COMPLI.**)	ρ = .86 K = 51 N= 12,488	ρ = .90 K = 16 N = 3,440	ρ = .91 K = 17 N = 4,144	ρ = .55 K = 26 N = 7,574	ρ = .79 K = 5 N = 912	ρ = .65 K = 5 N = 770	ρ = .68 K = 13 N = 4,026	ρ = .33 K = 3 N = 381	ρ = .91 K = 15 N = 4,453	.79a	
Communication Skills (**COMMUN.**)	ρ = .78 K = 54 N= 16,446	ρ = .87 K = 6 N = 1,904	ρ = .94 K = 24 N = 6,835	ρ = .57 K = 10 N = 1,790	ρ = .87 K = 13 N = 3,370	ρ = .75 K = 4 N = 654	ρ = .81 K = 21 N = 4,741	1	ρ = .90 K = 35 N =10,861	ρ = .94 K = 6 N = 2,006	.83a

Note. ρ = True score correlation; K= number of correlations included in that psychometric meta-analysis; N = Total sample size involved in that psychometric meta-analysis. aThe mean of the intrarater reliability estimates.

Table 3. True Score Intercorrelations Between Job Performance Measures: **Peer Ratings Only**

	ADM. SKIL.	PROD.	PROB. SOLVE	OVERALL	PERSON.	COMMUN	EFFORT	J. K.	COMPLI.
Administ. Skills **(ADM. SKIL.)**	.78[a]								
Productivity **(PROD.)**	$\rho=.83$ K = 4 N =395	.55[a]							
Problem Solving & Leadership **(PROB. SOLVE)**	$\rho=.52$ K = 7 N=1649	$\rho=.74$ K = 4 N = 202	.80[a]						
Overall Job Performance **(OVERALL)**	$\rho=.83$ K = 4 N = 491	$\rho=.95$ K = 6 N = 554	$\rho=.67$ K = 9 N = 1345	.78[a]					
Personality **(PERSON.)**	$\rho=.60$ K = 8 N = 1434	$\rho=.76$ K = 5 N = 321	$\rho=.62$ K = 23 N = 4512	$\rho=.81$ K = 13 N = 2606	.78[a]				
Communication Skills **(COMMUN.)**	$\rho=.75$ K = 2 N = 750	$\rho=.72$ K = 1 N = 72	$\rho=.53$ K = 6 N = 1644	$\rho=.09$ K = 1 N = 72	$\rho=.62$ K = 4 N = 1166	.77[a]			
Effort **(EFFORT)**	$\rho=.76$ K = 5 N = 657	$\rho=1.00$ K = 2 N = 191	$\rho=.91$ K = 6 N = 1029	$\rho=.87$ K = 5 N = 1341	$\rho=.83$ K = 19 N = 2970	$\rho=.75$ K = 1 N = 344	.78[a]		
Job Knowledge **(J. K.)**	$\rho=.74$ K = 3 N = 494	$\rho=1.00$ K = 2 N = 48	$\rho=.89$ K = 8 N = 1981	$\rho=1.00$ K = 5 N = 1374	$\rho=.83$ K = 12 N = 2573	$\rho=.52$ K = 1 N = 406	$\rho=.95$ K = 6 N=1338	.78[a]	
Compliance & Acceptance of authority **(COMPLI.)**	$\rho=.90$ K = 1 N = 164	$\rho=1.00$ K = 1 N = 164	$\rho=.63$ K = 1 N = 192	$\rho=1.00$ K = 2 N = 328	$\rho=.99$ K = 3 N = 378	I	$\rho=.89$ K = 5 N =711	I	.61[a]

Note. ρ = True score correlation; K= number of correlations included in that psychometric meta-analysis; N = Total sample size involved in that psychometric meta-analysis.
[a]The mean of the intrarater reliability estimates.

Table 4. True Score Correlations Between Performance Measures, One Rating From Peers the Other From Supervisors

PEER RATINGS

SUPERVISOR RATINGS	ADM.SK. (.54)a	PROD. (.48)a	PROB. & LEAD (.75)a	OVERALL (.47)a	PERS. (.45)a	COMMUN. (.53)a	EFFORT (.60)a	J.K. (.28)a	COMPLI. (.53)a
Personality (PERS.) (.58)a	ρ = .84, K = 7, N = 937	ρ = .34, K = 6, N = 311	ρ = .45, K = 7, N = 269	ρ = .63, K = 10, N = 1172	ρ = .62, K = 31, N = 2731	I	ρ = .48, K = 14, N = 1519	ρ = .29, K = 5, N = 413	ρ = .54, K = 5, N = 762
Productivity (PROD.) (.59)a	ρ = .54, K = 4, N = 395	ρ = .86, K = 7, N = 488	ρ = .09, K = 5, N = 151	ρ = .52, K = 5, N = 492	ρ = .36, K = 7, N = 411	I	ρ = .50, K = 2, N = 191	ρ = .83, K = 1, N = 27	ρ = .53, K = 1, N = 164
Effort (EFFORT) (.55)a	ρ = .41, K = 2, N = 191	ρ = .37, K = 3, N = 218	ρ = .03, K = 1, N = 27	ρ = .62, K = 4, N = 677	ρ = .42, K = 14, N = 1470	I	ρ = .69, K = 11, N = 1216	ρ = .09, K = 3, N = 325	ρ = .56, K = 3, N = 378
Overall Job Performance (OVERALL) (.56)a	ρ = .83, K = 4, N = 613	ρ = .53, K = 6, N = 431	ρ = .48, K = 11, N = 1436	ρ = .61, K = 29, N = 4566	ρ = .48, K = 14, N = 1695	ρ = .59, K = 1, N = 72	ρ = .62, K = 5, N = 927	ρ = .73, K = 3, N = 489	ρ = .51, K = 2, N = 356
Administrative Skills (ADM.SK.)(.60)a	ρ = .65, K = 5, N = 431	ρ = .51, K = 4, N = 395	ρ = .03, K = 1, N = 27	ρ = .44, K = 3, N = 430	ρ = .46, K = 4, N = 318	I	ρ = .52, K = 2, N = 191	ρ = .28, K = 1, N = 27	ρ = .43, K = 1, N = 164
Quality (QUAL.)(.63)a	ρ = .21, K = 1, N = 164	ρ = .30, K = 1, N = 164	ρ = .76, K = 1, N = 49	ρ = .31, K = 2, N = 328	ρ = .01, K = 1, N = 164	I	ρ = .26, K = 1, N = 164	I	ρ = .26, K = 1, N = 164
Job Knowledge (J. K.) (.58)a	ρ = .44, K = 1, N = 27	ρ = .84, K = 1, N = 27	ρ = .47, K = 5, N = 679	ρ = .66, K = 3, N = 483	ρ = .26, K = 11, N = 1314	I	ρ = .20, K = 4, N = 386	ρ = .65, K = 7, N = 844	I
Absenteeism (ABS.)(.62)a	I	I	I	I	I	I	I	I	I
Problem Solving & Leadership (PROB. & LEAD) (.59)a	ρ = .88, K = 4, N = 435	ρ = .20, K = 4, N = 120	ρ = .64, K = 11, N = 535	ρ = .73, K = 7, N = 599	ρ = .60, K = 15, N = 1189	I	ρ = .70, K = 5, N = 533	ρ = .80, K = 3, N = 149	ρ = .37, K = 2, N = 384
Compliance & Acceptance of Authority (COMPLI.)(.60)a	ρ = .46, K = 1, N = 164	ρ = .51, K = 1, N = 164	ρ = .31, K = 1, N = 435	ρ = .44, K = 4, N = 603	ρ = .51, K = 4, N = 478	I	ρ = .50, K = 3, N = 378	I	I
Communication Skills (COMMUN.) (.59)a	I	I	I	I	I	I	I	I	I

Note. ρ = True score correlation; K = number of correlations included in that psychometric meta-analysis; N = Total sample size involved in that psychometric meta-analysis.
[a] The mean of the interrater reliabilities.

Table 5. True Score Correlations Between Supervisory Ratings and Organizational Records

Sup ratings	Organizational Records				
	Lack of Errors (Quality)	Productivity	Absenteeism (Lack of)	Tenure	Accidents (Lack of)
Personality	$\rho = .34$ K = 12 N = 1290	$\rho = .26$ K = 16 N = 2201	$\rho = .33$ K = 15 N = 2753	$\rho = .11$ K = 13 N = 1832	$\rho = .44$ K = 13 N = 2704
Productivity	$\rho = .06$ K = 7 N = 931	$\rho = .44$ K = 10 N = 1387	$\rho = .27$ K = 12 N = 1349	$\rho = .17$ K = 8 N = 975	$\rho = .76$ K = 4 N = 652
Effort	$\rho = .39$ K = 8 N = 1366	$\rho = .66$ K = 6 N = 508	$\rho = .54$ K = 10 N = 1615	$\rho = .09$ K = 8 N = 1383	$\rho = 1.00$ K = 4 N = 526
Overall Job Performance	$\rho = .32$ K = 20 N = 4365	$\rho = .35$ K = 43 N = 8467	$\rho = .10$ K = 37 N = 13302	$\rho = .25$ K = 47 N = 23294	$\rho = .25$ K = 9 N = 2843
Administrative Skills	$\rho = .34$ K = 1 N = 100	$\rho = .39$ K = 6 N = 776	$\rho = .41$ K = 2 N = 164	$\rho = .23$ K = 3 N = 204	I
Quality	$\rho = .17$ K = 6 N = 763	$\rho = .14$ K = 7 N = 674	$\rho = .30$ K = 8 N = 1153	$\rho = .15$ K = 3 N = 681	$\rho = .74$ K = 3 N = 566
Job Knowledge	$\rho = .06$ K = 1 N = 100	$\rho = .33$ K = 5 N = 491	$\rho = .37$ K = 5 N = 567	$\rho = .41$ K = 4 N = 235	$\rho = .63$ K = 5 N = 923
Absenteeism	$\rho = .36$ K = 3 N = 428	$\rho = .21$ K = 1 N = 133	I	$\rho = .09$ K = 3 N = 607	$\rho = .11$ K = 1 N = 86
Problem Solving & Leadership	$\rho = .19$ K = 3 N = 554	$\rho = .39$ K = 11 N = 1674	$\rho = .63$ K = 2 N = 291	$\rho = .21$ K = 5 N = 746	$\rho = .74$ K = 5 N = 860
Compliance & Acceptance of Authority	$\rho = .18$ K = 3 N = 381	$\rho = .33$ K = 5 N = 711	$\rho = .12$ K = 3 N = 1091	$\rho = .41$ K = 6 N = 1059	$\rho = .38$ K = 6 N = 1471
Communication Skills	I	$\rho = .25$ K = 1 N = 40	$\rho = 1.00$ K = 1 N = 142	$\rho = -0.02$ K = 1 N = 100	$\rho = .57$ K = 3 N = 566

Note. ρ = True score correlation; K= number of correlations included in that psychometric meta-analysis; N = Total sample size involved in that psychometric meta-analysis. Interrater reliabilities were used to correct supervisory ratings; internal consistency measures to correct organizational records.

Table 6. True Score Correlations Between Peer Ratings and Organizational Records

Peer ratings	Organizational Records				
	Lack of Errors (Quality)	Productivity	Absenteeism (Lack of)	Tenure	Accidents (Lack of)
Personality	$\rho = .19$ K = 2 N = 172	$\rho = .18$ K = 2 N = 172	$\rho = .09$ K = 1 N = 149	$\rho = .14$ K = 1 N = 149	I
Productivity	$\rho = -.13$ K = 1 N = 72	$\rho = .47$ K = 2 N = 174	I	I	I
Effort	I	I	$\rho = .39$ K = 1 N = 149	$\rho = .34$ K = 1 N = 149	I
Overall Job Performance	$\rho = .23$ K = 1 N = 72	$\rho = .66$ K = 3 N = 276	I	$\rho = .84$ K = 1 N = 149	I
Administrat Skills	I	$\rho = .07$ K = 1 N = 102	I	I	I
Job Knowledge	I	I	I	I	I
Problem Solving & Leadership	$\rho = .17$ K = 2 N = 144	$\rho = .37$ K = 3 N = 243	$\rho = .67$ K = 1 N = 149	$\rho = .44$ K = 1 N = 149	I
Compliance & Acceptance of Authoriity	I	I	I	I	I
Commun. Skills	$\rho = .14$ K = 1 N = 72	$\rho = .15$ K = 1 N = 72	I	I	I

Note. ρ = True score correlation; K= number of correlations included in that psychometric meta-analysis; N = Total sample size involved in that psychometric meta-analysis. Interrater reliabilities were used to correct Peer ratings; internal consistency measures to correct organizational records.

Table 7. Factor Loadings: Cofirmatory Factor Analysis of
Organizational Record Based Measures

Performance Dimension	Loading on General Factor
Lack of Errors	.64
Productivity	.63
Absenteeism (Lack of)	.88
Tenure	.50
Accidents (Lack of)	.79

Table 8. Residual Matrix: Confirmatory Factor Analysis on
Organizational Record Based Measures

	ERRORS	PRODUCT	ABSENT.	TENURE
Lack of errors (Quality)(**ERRORS**)				
Productivity (**PRODUCT**)	.03			
Absenteeism(Lack of)(**ABSENT.**)	.08	.34		
Tenure (**TENURE**)	.20	.13	.11	
Accidents (Lack of) (**ACCIDENT**)	.26	.11	-.30	.33

Table 9. Correlations of Ratings subjected to Confirmatory Factor Analysis

SUPERVISOR RATINGS	PEER RATINGS							
	PERS. (.45)[a]	PROD. (.48)[a]	EFFORT (.60)[a]	OVERALL (.47)[a]	J.K. (.28)[a]	PROB. & LEAD (.75)[a]	COMPLI. (.53)[a]	ADM.SK. (.54)[a]
Personality (PERS.) (.58)[a]	.62							
Productivity (PROD.) (.59)[a]	.35	.86						
Effort (EFFORT) (.55)[a]	.45	.43	.69					
Overall Job Performance (OVERALL) (.56)[a]	.54	.53	.62	.61				
Job Knowledge (J. K.) (.58)[a]	.27	.84	.57	.69	.65			
Problem Solving & Leadership (PROB. & LEAD) (.59)[a]	.57	.14	.67	.55	.53	.64		
Compliance & Acceptance of Authority (COMPLI.)(.60)[a]	.53	.52	.53	.47	.60	.34	.79	
Administrative Skills (ADM.SK.)(.60)[a]	.74	.53	.47	.67	.36	.83	.45	.65

[a]The mean of the interrater reliabilities.

Table 10. Factor Loadings: Confirmatory Factor Analysis of
Ratings Based Measures

Performance Dimension	Loading on General Factor
Personality	.69
Productivity	.71
Effort	.75
Compliance	.71
Problem Solving Skills	.72
Job Knowledge	.76
overall Job Performance	.79
Administrative Skills	.79

Table 11. Residual Matrix From Confirmatory Factor Analysis of Ratings

	PERS.	PROD.	EFFORT	OVERALL	J.K.	PROB. & LEAD	COMPLI.
Personality (PERS.)							
Productivity (PROD.)	.14						
Effort (EFFORT)	.07	.10					
Overall Job Performance (OVERALL)	.01	.03	-.03				
Job Knowledge (J. K.)	.25	-.30	.00	-.09			
Problem Solving & Leadership (PROB. & LEAD)	-.07	.37	-.13	.02	.02		
Compliance & Acceptance of Authority (COMPLI.)	-.04	-.02	.00	.09	-.06	.17	
Administrative Skills (ADM. SK.)	-.19	.03	.12	-.05	.24	-.26	.11

Table 12. Factor Loadings of the 5 Organizational Record Based Job Performance Measures on 3 Group Factors

Variables	G1-PROD	G2-WITHDRAWAL	G3-CONSCIENTIOUSNESS
Productivity (**PRODUCT**)	.94	.27	.55
Absenteeism (Lack of) (**ABSENT.**)	.22	.74	1.07
Tenure (**TENURE**)	.20	.74	.14
Lack of errors (Quality) (**ERRORS**)	.40	.41	.73
Accidents (Lack of) (**ACCIDENT**)	.42	.73	.66

Table 13. Factor Loadings: Confirmatory Factor Analysis of Eight
Job Performance Ratings Loading on Four Group Factors.

SUPERVISOR RATINGS	G1-INTERPERSONAL	G2-CONSCIENTIOUSNESS	G3-PROD	G4-OVERALL
Personality	.77	.54	.38	.69
Leadership	.82	.66	.15	.70
Administrative Skills	.89	.55	.57	.86
Compliance & Acceptance of Authority	.53	.83	.56	.60
Job Knowledge	.46	.78	.91	.88
Effort	.64	.77	.46	.79
Productivity	.41	.77	.93	.68
Overall Job Performance	.70	.77	.57	.78