To the Graduate Council:

I am submitting herewith a dissertation written by John R. Burnett-Doering entitled "A Meta-Analysis of the Validity of Personality/Test Scores in the Prediction of Employee Turnover." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

John W. Lounsbury, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Vice Provost
and Dean of The Graduate School

A META-ANALYSIS OF THE VALIDITY OF PERSONALITY/TEST SCORES IN THE PREDICTION OF EMPLOYEE TURNOVER

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

John R. Burnett-Doering

December 1987

ACKNOWLEDGMENTS

In the course of dissertation research, a number of individuals have provided assistance. I would like to thank the members of my committee: Drs. William H. Calhoun, Dean J. Champion, Michael G. Johnson, and J. Albert Wiberley, for their careful and helpful critiques. I am particularly grateful to my chairperson Dr. John W. Lounsbury, whose help and guidance made me a better psychologist, and whose friendship is greatly valued. A special thanks goes to Dr. Dean Champion for all the good advice and good times. I would also like to thank all the people at Aptitude Testing for Industry who provided assistance during the research project.

Also, thanks of a special kind go to my wife, Tara, for her encouragement, support, and for being the super-mom that she is, and to Chenin and Ashley for being the good little girls that they were during the project. Finally, I would like to thank my parents for the love and support that they have always given me.

ABSTRACT

The literature would suggest that the relationship between personality and turnover is weak and contradictory--leading numerous researchers to downplay the importance of personality constructs/variables. The Hunter, Schmidt, and Jackson (1982) Bayesian validity generalization procedure was employed to determine whether this state of affairs might be explained by statistical and/or methodological artifacts. Rather than relying on published data as in most meta-analytic research, data in this study were collected on 24 personality test scores (utilizing both the Minnesota Multiphasic Personality Inventory and the Guilford-Zimmerman Temperament Survey) from 532 commercial real estate sales representatives in seven geographical regions. Utilizing traditional interpretational procedures led to the conclusion of situational specificity. However, when two sources of artifacts were removed statistically (range restriction and sampling error) and three held constant (criterion contamination and deficiency, predictor unreliability, and slight differences in factor structures between measures of similar constructs), nonsignificant overall variation in validities remained--thereby leaving little room for the operation of situational moderation. Support for generalizability (transportability) was strong in 12 test distributions, which had validities between -.02 and -.49. Results also indicated that a battery composed of the individual scales: Schizophrenia, Ego-Strength, L, K, and Anxiety, predicted 19% of the criterion variance. Moreover, the latter two scales acted as suppressor variables. Although

limitations with the Hunter et al. procedure might exist when applied to personality-turnover relationships in a field setting, the procedure helps to explain the conflicting results across studies and shows the importance of not totally ignoring personality variables.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Personality Test as Predictors	3
	Positive findings	3
	No relationship	7
	Summary	9
	Meta-Analysis as a Method of Determining Validity	10
	Purpose of the Study and Hypotheses	15
II.	METHOD	17
	Subjects	17
	Short tenure group	17
	Long tenure group	18
	Procedure	18
	Temperament and Personality Predictors	20
	Minnesota Multiphasic Personality Inventory	20
	Guilford-Zimmerman Temperament Survey	21
	Meta-Analysis Procedure	22
III.	RESULTS	27
	Situational Specificity	31
	Validity Generalization	36
	Determing a Predictor Battery	40
IV.	DISCUSSION	44
	Situational Specificity/Generalizability of Personality Predictors	44
	Individual Scales and Turnover	47
	Psychopathic Deviate	48
	Schizophrenia	49
	Ego-Strength	49
	Control	49
	L scale	50
	Thoughtfulness	50
	Predictor Battery	51
	Future Research	53
	LIST OF REFERENCES	55
	VITA	66

LIST OF TABLES

TABLE PAGE

1. Demographic Characteristics for Employees in Long Tenure and Short Tenure Groups 19

2. Point Biserial Correlations and Correlations Corrected for Range Restriction Between Personality and Turnover. 28

3. Situational Specificity Hypothesis Results for both MMPI and GZTS Predictors 32

4. Validity Generalization Results for Personality Predictors Where the Criterion is Turnover. 38

5. Correlations Between Variables. 41

6. Summary of Multiple Regression Analysis Using Turnover as Criteria and Personality Variables. 42

CHAPTER I

INTRODUCTION

Much of the research in the field of Organizational Psychology has tended to minimize the importance of personality constructs and variables. In fact, much of the research has used personality traits as moderator variables. This situation is typified by the following statement:

> We will find throughout this review that personality traits appear as predictors of attitudes (e.g., involvement), motivation (e.g., expectancies), and leadership (e.g., behavioral styles), but the central focus of that research is usually motivation, attitudes, or leadership and not personality . . . .
>
> This secondary role seems justified and necessary. If Mischel's arguments are correct, then we will be better served by continuing in the direction we are heading. Personality variables probably control only a minor percentage of variance in behavior when compared to situational factors, and there is little consistency over time and settings (Mitchell, 1979, p. 247).

The same state of affairs holds true for the topic area of job turnover. Reviews of the literature on job turnover have definitely established that characteristics of the external environment of the organization (e.g., economic conditions and employment perceptions), structural or work related factors (e.g., pay, job performance, job satisfaction, job autonomy, and satisfaction with supervision), and individual employee characteristics (e.g., age, tenure, gender, number of dependents, and education) are related to turnover in organizations (Baysinger & Mobley, 1983; Bluedorn, 1982a; Brayfield & Crockett, 1955; Cotton & Tuttle, 1986; Forrest, Cummings, & Johnson, 1977; Herzberg, Mausner, Peterson, & Capwell, 1957; Lefkowitz, 1971; March & Simon,

1958; Mobley, 1982; Mobley, Griffeth, Hand, & Meglino, 1979; Mowday, Porter, & Steers 1982; Muchinsky & Tuttle, 1979; Pettman, 1973; Porter & Steers, 1973; Price, 1977; Schuh, 1967; Vroom, 1964). However, the relative emphasis in previous research appears to have been placed on investigating structural and work-related and extra-organizational determinants of withdrawal behavior. As a result, far less is known about the influence of personal characteristics/traits of the individual employee with respect to the turnover decision. This seems unusual because many personal variables and characteristics, as contrasted with most other predictors of withdrawal, can be assessed prior to one's employment (Porter & Steers, 1973).

Several conceptual models have been developed to explain why people leave a job. In the participation model of March and Simon (1958), the "intermediate linkages" version by Mobley (1977) and Mobley et al. (1979), the investment model of Farrell and Rusbult (1981), the met expectations model by Steers and Mowday (1981), the matching model of Wanous (1980), or the models by Price (1977) and Price and Mueller (1981), as well as by Bluedorn (1982a, 1982b), the individual difference variables (if mentioned at all) have played a minor role (Tenopyr & Oeltjen, 1982). For example, in both the Steers and Mobley models of turnover, the absence of personality variables would not significantly alter the processes involved. An exception is the model developed by Muchinsky and Morrow (1980). In their model, individual factors can not only be moderated by economic factors but they can also function as direct determinants of employee turnover.

Personality Tests as Predictors

One set of personal factors that has not received much attention lately are the personality characteristics of employees. Several reasons may account for this state of affairs. Schmitt, Gooding, Noe, and Kirsh (1984) conducted a meta-analysis on validation studies published in either the Journal of Applied Psychology or Personnel Psychology between 1964 and 1982. Results indicated that personality measures are associated with lower validity coefficients when compared to other predictors (e.g., biodata). This is partially a function of the fact that the latter predictors are conceptually simpler and better defined than complex and ambiguous personality constructs, such as Psychological Mindedness as measured by the California Psychological Inventory (Lent, Aurbach, & Levin, 1971). They also found that the criterion used may produce lower or higher validity coefficients. For example, performance ratings yielded better coefficients than turnover criteria. Finally, personality measures tended to be less valid predictors of turnover than other predictor variables. These reasons have contributed to the perception by reviewers of the literature that personality variables seemingly do not account for much of the variance in turnover criteria. As a result, researchers have tended to place greater emphasis on structural/work, organizational, and other personal variables. This recent neglect of the role of personality variables in turnover is reflective of the mixed and inconclusive findings by researchers.

Positive findings. In an early study, Hanna (1935), using the Thurstone Personality Schedule to classify individuals with respect to

emotional stability, found that the "adjusted" group demonstrated greater job stability than those classified as "maladjusted." However the relationship might have been confounded since the latter group was older by a median age of approximately two years than the "adjusted" group. In another study, Sinaiko (1949) administered the Rosenzweig Picture-Frustration Study to 53 department store section managers. He found that length of service was positively related to impunitiveness.

Two studies of turnover were conducted using samples of hospital aides. The first of these showed that long-term employees (10 years or more on the job) scored significantly higher on the California F scale (authoritarianism) than short-term employees with six or fewer months on the job (Cleland & Peck, 1959). Butterfield and Warren (1963) demonstrated that aides who were discharged during the first six months of employment exhibited higher levels of defensiveness, irresponsibility, and dominance, as measured by the Minnesota Multiphasic Personality Inventory (MMPI) than those who were retained.

MacKinney and Wolins (1960) found that the job tenure of male production foremen was positively related to the ascendance and sociability scales of the Guilford-Zimmerman Temperament Survey (GZTS) and negatively related to the neurotism scale of the Bernreuter Personality Inventory (BPI). Hakkinen and Toivainen (1960) also found that those who left their work scored significantly higher on a test of neuroticism than those that stayed on the job. In another study of 1360 engineers, Meyer and Cuomo (1962) showed that those who terminated

employment had higher levels of achievement motivation, aggression, independence, self-confidence, and sociability. By comparison, "stayers" exhibited more emotional stability, maturity, sincerity, stronger job orientation, and a moderate level of achievement orientation. These results are consistent with findings by Mowday and Spencer (1981) that high needs for achievement and autonomy are related to turnover. Farris (1971) also found high achievement orientation to be predictive of turnover among scientific and engineering personnel. Consistent with these findings, Hines (1973) using a nonprojective inventory, the Lynn Achievement Motivation Questionnaire, found that engineers, accountants, and middle managers who were high in need achievement displayed greater labor turnover than those who were low in need achievement. On the other hand, he found need for achievement was negatively related to turnover for entrepreneurs. In subsequent research, Mowday, Porter, and Stone (1978) used Jackson's Personality Research Form (JPRF) in a study of 314 female clerical employees in two regional offices of a large insurance company. The individuals who voluntarily withdrew from the job exhibited a higher need for autonomy and a lower need for harm avoidance than those who remained with the organization. The results held up under double cross-validation. In another validational study for a large utility, the Humm-Wadsworth Temperament Scale was found to be a valid predictor of job tenure.[1]

The studies cited thus far are generally supportive of Porter and Steers (1973) "polar" hypothesis. They posited that employees who

---

[1]Ruch, W. W. Personal communication, June 1, 1986.

withdraw from an organization tend to exhibit personality characteristics at extreme positions along certain trait continua than employees with more moderate levels of these traits. In a test of the "polar" hypothesis, Bernardin (1977) administered the Sixteen Personality Factor Questionnaire to both non-supervisory and non-secretarial employees who had voluntarily or involuntarily terminated employment in a midwestern sales operation. Results indicated that four factors were associated with turnover: anxiety, conscientiousness, ego strength, and superego strength. The analysis also indicated that the first two factors accounted for the largest proportion of the variance, and the relationships survived a double cross-validation. In addition, the "polar" hypothesis was not supported except for the anxiety factor. Mowday, Stone, and Porter (1979) administered the JPRF to employees in either low-scope (machine operators) and high-scope (scientific/ technical) jobs. For employees on high-scope jobs, turnover was negatively related to the need for achievement, and positively related to the need for affiliation. In contrast, turnover was negatively related to the need for affiliation among employees on low-scope jobs. Finally, the need for autonomy was unrelated among employees on either high or low-scope positions. In a study of a large national oil company using the GZTS, Dreher (1982) found significant differences between stayers and leavers among managerial, professional, and technical employees. Finally, Lang and Krug (1978) administered the Executive Profile Survey to a large sample of bank presidents. Those more likely to remain with the organization tended to score as less ambitious, self-assertive, enthusiastic, creative, innovative,

self-directed, and more adaptable, perceptive, and systematic compared with those who did not remain.

No relationship. A body of evidence exists that suggests there is no relationship between personality measures and turnover. Wickert (1951) administered a 60-item measure of "neurotic tendency" to both female telephone operators and service representatives and found no difference between those that stayed or those that left the company. Similarly, MacKinney and Wolins (1960) reported no relationship between turnover and the personality measures of the Guilford-Martin Personal Inventory with male production foremen. Hedberg and Baxter (1962) classified life insurance salesmen as either high or low on ten sub-scales of the GZTS, and did not find any reliable differences. An examination of the raw score distribution did, however, reveal differences on three sub-tests among the high-low criterion groups, but these differences vanished under cross-validation.

Butterfield and Warren (1962) reported that none of the scales on the MMPI differentiated between psychiatric aides who either stayed, were dismissed, or resigned from an institution. A variety of configurations did (Psychopathic Deviate-Hypomania, Psychopathic Deviate-Psychasthenia, Psychopathic Deviate-Social Introversion, Hypomania-Social Introversion, Hypomania-Psychasthenia) distinguish between stayers from leavers, but these were not found in a cross-validation sample. These results are inconsistent with results reported in a later study by Butterfield and Warren (1963). Several reasons were suggested for these findings. First, personality trait/patterns may not be predictive of turnover in various settings, since different

personality patterns may be necessary in different settings. Also, personality measures designed for clinical use might not be sensitive enough to differentiate between individuals from "normal" populations. Cuadra and Reed (1957) developed a motivational subscale consisting of 43 of the 472 items of the California Psychological Inventory. The subscale appeared to differentiate between long-term (i.e., those that remained on the job more than ten months) from short-term hospital aides (i.e., those that remained on the job less than three months). The results, however, did not hold up under cross-validation. Robbins and King (1961) found that sales managers who were terminated scored higher on only one (Sociability) of the six scales comprising the BPI. Unfortunately, the effect disappeared when extraneous variables were controlled for. In another study where the original results were not supported under cross-validation, Vincent and Dugan (1962) found the "Original Thinking" scale of the Gordon Personal Inventory (GPI) was predictive of turnover among insurance salesmen. Schwab and Packard (1973) reported the results of a study in which the GPI and the Gordon Personal Profile (GPP) were administered in a manufacturing setting to one group of female employees prior to hiring and to another group immediately after they were hired (i.e., replication group). There was a significant negative relationship between turnover and the Vigor scale of the GPI in the group prior to hiring. On the GPP, ascendancy and sociability scores were positively related to turnover in this group as well. However, no significant relationships were found in the replication group. Finally, in a reanalysis of an earlier study

Atchison and Lefferts (1972) found no differences in needs for achievement, affiliation, and power for air force pilots who stayed or left.

Summary. Overall, in studies in which replication or cross-validation have been attempted, the results have been rather disappointing (e.g., Butterfield & Warren, 1962; Cuadra & Reed, 1957; Hedberg & Baxter, 1957; Robbins & King, 1961; Schwab & Packard, 1973; Vincent & Dugan, 1962). In only two studies did the relationship between personality characteristics and turnover hold up under cross-validation (Bernardin, 1977; Mowday, Porter, & Stone, 1978). This state of affairs has made it rather difficult to draw inferences regarding the effects of personality on withdrawal behavior. There are several other reasons for this as well. First, the multitude of personality tests and subscales utilized make comparisons among studies invalid (Muchinsky & Tuttle, 1979). Secondly, the inconclusive findings are partly a function of the differences in the factor structure of the personality measures employed. Thirdly, situational specific moderator variables (e.g., age, marital status, and sex) are thought to cause a personality measure to appear valid in one setting, but not in other circumstances. Fourthly, most studies of turnover have operationally partitioned the criterion into long and short tenure groups. Meaningful comparisons among studies are difficult because a common criterion-cutoff or base rate as to what length of employment constitutes "long tenure" has not been established (Muchinsky & Tuttle, 1979). The criterion-cutoff is often determined arbitrarily, and it is partially a function of the position in consideration. For example, Vincent and Dugan (1962) defined long tenure as being on the job for

more than one year. Whereas, Robbins and King (1961) described their short-term (tenure) group as those that remained on the job less than nine years. Thus, some reviewers have concluded that under certain conditions/circumstances, some personality measures yield good predictive results (e.g., Porter & Steers, 1973; Schuh, 1967). These conclusions agree with those of Ghiselli and Barthol (1953) and Guion and Gottier (1965) when they surveyed the validities of personality tests for industrial use. Moreover, Guion and Gottier posited that it is difficult to know in advance what personality measure should be utilized in which situation. They also asserted that any personality measure must be developed for a specific setting rather than utilizing a standard personality measure. Yet, this has not prevented other reviewers from continuing to argue that although much research shows some significant relationships between person variables and turnover, it was best to conclude that personality differences seem to have a marginal impact on turnover (Muchinsky & Tuttle, 1979; Mobley, 1982; Pettman, 1973).

Meta-Analysis as a Method of Determining Validity

The traditional narrative literature review method used by previous investigators can be informative, but is not the optimal method of integrating research findings since it may lead one to erroneous conclusions. Glass, McGaw, and Glass (1981) cite various reasons for this state of affairs. For example, some reviewers do not include a representative sample of studies since their theoretical orientation and possible bias may have influenced them to include only studies that affirmed their position, and to exclude studies that might have

contradicted their beliefs. In addition, reviewers frequently misrepresent the findings of the study by utilizing the "vote counting" or "box score" method in which studies are grouped into sets according to whether the findings were statistically significant or not. Finally, reviewers also frequently fail to consider or recognize that random sampling error can create inconsistent findings among studies.

Within the last decade, a collection of methods have been developed that allows quantitative aggregation of results across studies. Glass (1976, 1977) coined the term "Meta-analysis" for these methods. Bangert-Drowns (1986) has identified several major meta-analytic approaches that vary in purpose and sophistication: Glassian meta-analysis, study effect meta-analysis (e.g., Mansfield & Busse, 1977), combined probability method (e.g., Rosenthal, 1978, 1984), and approximate data pooling with tests of homogeneity (e.g., Hedges, 1982). Several of these methods (particularly the Glassian meta approach) have a serious weakness in that they take the variance across studies at face value. Thus, the problem of sampling error is not dealt with adequately (Hedges, 1983).

A fifth approach developed by Schmidt, Hunter, and their associates (Hirsh, Northrup, & Schmidt, 1986; Hunter, Schmidt, & Jackson, 1982; McDaniel, Hirsh, Schmidt, Raju, & Hunter, 1986; Pearlman, Schmidt, & Hunter, 1980; Schmidt, Gast-Rosenberg, & Hunter, 1980; Schmidt & Hunter, 1977, 1978, 1984; Schmidt, Hunter, & Pearlman, 1981; Schmidt, Hunter, Pearlman, & Hirsh, 1985; Schmidt, Hunter, Pearlman, & Shane, 1979; Schmidt, Hunter, & Urry, 1976; Schmidt, Ocasio, Hillery, & Hunter, 1985) was used in the project for several

reasons. First, it provides specific formulas for statistically aggregating effect sizes across studies. Second, Schmidt and Hunter have specifically applied their meta-analytic method to bivariate predictor-criterion (i.e., correlational) relationships. Finally, it is based on the idea that much of the variation in results across studies may be accounted for by statistical, measurement, and other artifacts present in most studies rather than substantive differences in underlying population relationships. Moreover, the effect of these artifactual mediators is to reduce the correlations below their true (e.g., population) values.

Several sources of spurious between study variance have been identified by Schmidt and Hunter (1977) that can contribute to the appearance of variability in validity coefficients. The predictive validity of personality test scores may be susceptible to the same artifacts. The first source of variance results from differences between studies in criterion reliability. A second source of variance is accounted for by differences in predictor reliability. While some studies utilized well established measures with published reliability and validity data (e.g., Bernardin, 1977), others used measures of debatable value (e.g., Farris, 1971; Hakkinen & Tiovainen, 1960; MacKinney & Wolins, 1960; Robbins & King, 1961). In both studies by Robbins and King and MacKinney and Wolins the BPI was used. In a review of the test, Becker (1965) stated:

> The test is easily faked to give an emotionally stable, somewhat outgoing profile, and while successful and unsuccessful employees of various sorts are occasionally differentiated, the differences are so small as to preclude their usefulness in making decisions about individuals (p. 157).

Another source of artifactual variance is sampling error due to small sample size. While most psychologists recognized the existence of sampling error at the abstract level, they have traditionally ignored it at the operational level (Schmidt et al., 1976). As believers in the "law of small numbers" (i.e., that small random samples are about as representative of their corresponding populations as are large random samples), they overlooked the problem of inadequate statistical power (Tversky & Kahneman, 1971). As a result of this belief, variation in results among studies is often attributed to moderator variables rather than sampling variability (Schmidt & Hunter, 1978, 1984). A fourth source of variation results from restriction in range. Numerous studies have examined the effects of range restriction on the magnitude and sign of observed validity coefficients (e.g., Alexander, Carson, Alliger, & Barrett, 1984; Brewer & Hills, 1969; Forsyth, 1971; Greener & Osburn, 1979, 1980; Gross, 1982; Gross & Fleishman, 1983; Lee, Miller, & Graham, 1982; Linn, 1968, 1983). In most personnel selection settings, range restriction most often occurs in the predictor, and can be severe. This restriction results from applicants being selected (or promoted) solely on the basis of some predictor (i.e., test). Direct range restriction results when a group of applicants are tested, and only those who score the highest are selected. Since complete test-criterion information is not available for those who are not hired, the validity coefficients are attenuated by the restricted variance of the test.

Schmidt and Hunter (1977) demonstrated that these four sources alone account for most of the observed variance between studies for single predictor-criterion relationships. They also delineated several other sources of artifactual, between-study variance: computational and typographical errors (Wolins, 1962); and slight differences in factor structures between tests thought to measure similar constructs. The final source of variance was due to the amount and kind of criterion contamination and deficiency (Brogden & Taylor, 1950). It is well documented in the turnover literature that lack of comparability of turnover measures across studies is a major problem (Mobley et al., 1979; Muchinsky & Tuttle, 1979; Schuh, 1967). For example, some studies used measures of voluntary turnover (e.g., Schwab & Packard, 1973; Hedbert & Baxter, 1962), others of involuntary turnover (e.g., Butterfield & Warren, 1962, 1963), and still others failed to distinguish voluntary from involuntary turnover (e.g., MacKinney & Wolins, 1960; Vincent & Dugan, 1962). Bluedorn (1978) and Price (1977) stated that the variables and causal processes that produce involuntary separations are possibly quite different from the variables and processes which produce voluntary separations. MacKinney and Wolins (1959) found that an aptitude test could predict the turnover of discharged, but not voluntary separators. Moreover, the classification of the reasons for turnover are frequently taken from company personnel records. Yet, Lefkowitz and Katz (1969) reported significant differences between reasons for termination found in company records and those reported by employees following termination of their employment.

The Hunter et al. (1982) method computes the observed variance of a particular validity distribution, and the variance due to each of the sources of error variance is subtracted from this total variance. The situational specificity hypothesis is rejected if the remaining residual variance is near zero. In other words, the overall variation in validity coefficients can be accounted for by the identified sources of error variance. Schmidt and Hunter (1977) further stated that validity is generalizable even where some situational variation appears to exist as long as most of the values (e.g., 90%) in the distribution of estimated true validities exceed some minimally useful level of validity.

Purpose of the Study and Hypotheses

The purpose of the present study was to extend and compliment previous reviews of personality measures by (a) using the Hunter and Schmidt technique to provide a more accurate picture of the relationship between personality measures and employee turnover, and (b) discover whether apparently inconsistent results across studies might possibly reflect the existence of statistical artifacts (e.g., sampling error) or situational specificity (i.e., situational moderators).

It was hypothesized that 75% or more of the variance in the overall personality test score/turnover distribution will be accounted for by artifactual variance. The implications of a positive finding would mean that in general the variance left within which situational specificity (i.e., situational moderators) can operate is very limited.

It was also hypothesized that the personality predictors will generalize across the various regions; specifically, that none of the 90%

credibility values will include zero (i.e., the value at or above which 90% of all estimated true validities are expected to lie after accounting for sources of artifactual error), but will be substantial for most of the tests.

CHAPTER II

METHOD

Subjects

The subjects in this study were sales representatives in seven geographically distinct regions of a large national commercial real estate sales organization. Consistent with the majority of personality-turnover research, employees were designated in either of two criterion groups: long tenure (i.e., those that remained with the organization) or short tenure (i.e., those that withdrew from the organization.

Short tenure group. The short tenure group was constructed by including all employees who permanently left the organization from 1981-1985 (N = 738). From this group, only those that voluntarily resigned (i.e., the employee initiated the action) were included. Also, employees whose reasons for voluntarily leaving the organization that were vague and possibly contaminated by involuntary considerations (e.g., "personal") were excluded. An examination of the archival personnel records indicated that the majority of voluntary turnover occured during the first five years of employment--approximately 65%. Therefore, the short tenure group was operationally defined as consisting of individuals who had terminated with five years or less on the job. Another reason for deciding on the five-year cutoff was that employees are paid a salary the first year during training, and on a commission basis starting the second year. Additionally, it takes several years past the training period to build contacts and

establish oneself in the commercial real estate business. These procedures resulted in a final short tenure sample of 284 employees. This short tenure group had the following demographic characteristics: 89% of them were male; 73% had a four year college degree; the average age at hire was 31 years; and the average length of tenure was approximately three years.

Long tenure group. The long tenure group was made up of employees with six years or more of service. Individuals were randomly selected from the organization who were hired between the years 1970-1979 and were still employed as of December, 1985 (N = 248). The long tenure group had an average age at hire of 32 years and an average tenure of 10 years. A breakdown of the group by sex indicated that 95% were males. Also 72% had received a college degree.

A summary comparing the long and short-term (tenure) groups in terms of demographic characteristics are shown in Table 1. There were no significant differences between the groups with respect to age, male-female ratio, or educational background ($p > .01$). Each region was composed of employees from both long and short tenure groups. The number of employees from region one through region seven was 95, 69, 90, 108, 47, 88, and 35, respectively. The number of short (and long) tenure employees per region was 32(63), 24(45), 65(25), 47(61), 36(11), 69(19), and 11(24), respectively.

Procedure

Biographical data for employees in short and long tenure groups were obtained from two sources: the application blank and a computerized company data base. The biographical items of gender, age at

TABLE 1

DEMOGRAPHIC CHARACTERISTICS FOR EMPLOYEES IN LONG TENURE AND SHORT TENURE GROUPS

Employee Characteristic	Short Tenure	Long Tenure
Age at hire	30.87[a] (6.32)[b]	31.54[a] (6.05)[b]
Gender (percentage male)	89.08	94.76
Years of education	16.00[a] (.76)[b]	16.10[a] (.72)[b]
Tenure	2.65[a] (1.61)[b]	10.32[a] (2.93)[b]
Sample sizes	284	248

Note. Tenure was measured in years.

[a] Mean

[b] Standard deviation

hire, date of hire, age at termination, and reason for termination were obtained from the latter source. The sales branch and region of employment were also obtained from the data base. The level of education was recorded from the application blank. As part of the organization's hiring/selection process, each prospective employee was given a battery of tests by an aptitude testing company in order to determine mental abilities, interests, personality, and motivations. From this battery, the personality test scores of employees classified in either short or long tenure groups, for the purposes this project, were obtained. The raw test data was gathered from the aptitude testing company's evaluative report.

Temperament and Personality Predictors

Two instruments were utilized to assess the personality traits of employees that might be related to turnover.

Minnesota Multiphasic Personality Inventory. The MMPI consists of 560 statements covering physical and psychological functioning (Hathaway & McKinley, 1967). There are 10 basic clinical scales and four validity scales. The clinical scales were developed by contrasting the responses of normal groups with the responses of clinical cases. A revised and shortened version of the MMPI containing 505 items was developed and employed by the aptitude testing company for use in industrial settings. The "Cannot Say" validity scale was eliminated. In addition, the clinical scales of Masculinity-Femininity and Social Introversion were omitted.

The following standard scales of the MMPI were used in this study: L, K, Hypochondriasis (Hs), Depression (D), Hysteria (Hy),

Psychopathic Deviate (Pd), Paranoia (Pa), Psychasthenia (Pt), Schizophrenia (Sc), and Hypomania (Ma). In addition, the following research scales were included: Anxiety (A; Welsh, 1956), Ego-Strength (Es; Barren, 1953), Dependency (Dy; Navran, 1954), Dominance (Do; Gough, McClosky, & Meehl, 1951), Social Responsibility (Re; Gough, McClosky, & Meehl, 1952), Prejudice (Pr; Gough, 1951), Control (Cn; Cuadra, 1953), Work Attitude (Wa; Tydlaska & Mengel, 1953), Emotional Immaturity (Em; Pearson, 1954), and Impulsivity (Im; Gough, 1957).

Guilford-Zimmerman Temperament Survey. The GZTS was designed primarily for use with "normal" populations (Guilford & Zimmerman, 1949). It consists of 300 items scored for 10 factor traits. The developement and psychometric properties of the GZTS are described in depth by Guilford, Zimmerman, and Guilford (1976). The Restraint (R), Friendliness (F), Thoughtfulness (T), and Personal Relations (P) scales were examined in this project.

Butterfield and Warren (1962) suggested that one reason for the failure of personality measures to predict turnover is that there has been no theoretical rationale in their selection. The following scales were selected because each was expected to be predictive for some specified reason. For example, high scorers on the Psychopathic Deviate scale tend to be rebellious toward authority figures, impatient, as well as having poor work histories. To illustrate further, the Hypomania scale is able to predict individuals that are easily bored, dissatisfied, and have difficulties at work (Graham, 1977).

Meta-Analysis Procedure

The Hunter et al. (1982) procedure allows for corrections of artifactual variance from four sources: (a) sampling error, (b) range restriction, (c) predictor unreliability, and (d) criterion unreliability. The correction for variance due to sampling error was a weighted average of the sampling error across geographical regions (Hunter et al., 1982, chap. 3).

Range restriction due to explicit selection was corrected utilizing the method developed by Allexander, Alliger, and Hanges (1984), since information was only available on the selected sample, and the usual correction formulas require an estimate of the unrestricted population variance. The method provides an estimate of the unrestricted variance of a variable from the restricted sample itself, which can then be used in Thorndike's (1949) standard Case II (i.e., direct univariate range restriction) formula: $\underline{r}_c = \underline{r} / \sqrt{\underline{r}^2 + (\hat{\underline{SD}}^2/\underline{SD}^2_{obs}) - (\hat{\underline{SD}}^2/\underline{SD}^2_{obs})(\underline{r}^2)}$, where $\underline{r}$ is the observed correlation, $\underline{SD}_{obs}$ is the observed standard deviation, and $\hat{\underline{SD}}$ is the corrected standard deviation. Moreover, the assumption can be made that the test scores were in all liklihood not the sole basis for selection (Linn, 1983; Roe, 1979). This suggests that the estimates of variance attributable to range restriction will be attenuated. The result will be a slight overestimation of the residual (or true) variance and an underestimation of the population correlation (Linn, 1968; Linn, Harnish, & Dunbar, 1981). Nonetheless, the usual corrections for univariate range restriction are more accurate than the corresponding uncorrected estimates of the population correlation (Greener & Osburn, 1979, 1980; Linn et al., 1981).

Most meta-analytic studies have corrected the population correlation using assumed artifact distributions. In the present study, the adjustment for range restriction was applied at the level of the individual validity coefficient, and the adjustment for sampling error was at the level of the meta-analytic set of studies.

Also, since each validity distribution was based on a single test, test reliability differences between regions could be eliminated as a source of artifactual variance. In other words, the estimate of variance due to this source was essentially zero. In addition, although the reliability of the "objective" criterion measure of turnover has probably a high degree of reliability (since once an employee leaves he cannot contribute to turnover again), no correction was made. The absence of this correction is probably not a serious shortcoming, because previous meta-analyses have found that sampling error alone is the most serious artifact--accounting for 75-90% of the variance (Hunter et al., 1982; Pearlman et al., 1980; Schmidt & Hunter, 1984; Schmidt, Hunter, & Caplan, 1981; Schmidt et al., 1980).

Pearlman et al. (1980) have asserted that it is hard to control for differences across studies in amount and kind of criterion contamination and deficiency, as well as for computational and typographical errors. No correction was made for the latter since it is difficult to estimate the frequency or magnitude of such errors. In regards to the former, it is well documented in the turnover literature that the use of different conceptual and operational measures of turnover across studies is a problem (Mobley et al., 1979; Muchinsky & Tuttle, 1979; Schuh, 1967; Wolpin & Burke, 1985). In consideration of this, the meta-analysis was confined to seven replications

(i.e., regions) where greater control of criterion measures could be obtained than if published personality/turnover research was utilized. Although the use of seven regions might appear to be low, there is no theoretical basis from a Bayesian perspective as to the number of coefficients that are required for a validity generalization analysis, because the procedure utilized here weights by the information value (which is a function of distributional variance) rather than the absolute size of the distribution (Pearlman et al., 1980). Use of such data also will eliminate another source of error. Because the same personality measures were utilized, any variance between factor structures would be slight compared to differences that arise when nonidentical personality tests are used to measure the same personality construct. In their meta-analytic studies, Schmidt and Hunter have ascertained that it is quite difficult to control for these two sources of error. The present design utilizing data collected from seven regions controls these artifacts and affords a stronger test of of the situational specificity hypothesis than is possible if published research were employed as a data source. Another advantage of using data from seven regions is that it avoids potential problems in meta-analysis research that utilizes published research. Bullock and Svyantek (1985) and Orwin and Cordray (1985) found that deficient reporting (e.g., resulting from problems in coding study characteristics as well as problems in sampling the entire domain of studies available) can introduce a great deal of "noise" into this kind of meta-analysis and that can lead to spurious conclusions.

After the variance due to each of the sources of error variance are subtracted from the total variance, the remaining (residual)

variance may be insignificant in size and likely due to other statistical artifacts, or it may be nontrivial and suggest the possible presence of one or more moderators in the data. Two tests of the situational specificity hypothesis were performed. The first was the Schmidt-Hunter 75% decision rule which states that if 75% or more of the observed variance is accounted for by variance due to statistical artifacts, then the situational hypothesis is rejected. The rationale for this decision rule is that the remaining artifacts (i.e., criterion unreliability, and computational, typographical, and transcriptional, errors) for which one cannot correct are likely to account for at least 25% of the observed between study variance.

In addition, it is important to examine the actual amount of residual variation (after removal of sampling error), as well as the percentage of observed variance. Since sampling error is primarily a function of sample size, correlations based on large sample sizes have smaller sampling error variances than those made up of smaller sample studies--regardless of whether the residual variance is large or small. This arises from the fact that other sources of variance (e.g., range restriction) are unaffected by sample size. For instance, if validity coefficients based on large samples are unaffected by situational or other moderators, the percentage of observed variance attributable to sampling error will be small, and since neither sampling error nor moderators significantly affect the coefficients, both the observed and residual variances will be small (McDaniel, Hirsh, Schmidt, Raju, & Hunter, 1986). Another rationale for investigating the residual variance is because it and not the observed

variance is what determines the transportability of validity through its impact on the standard deviation of true validities (McDaniel et al.).

To assess whether the observed residual variation across studies was greater than that expected by chance, an approximation to a chi-square test was done following the formula given in Hunter et al. (1982) as $\chi^2_{(K-1)} = K(\sigma^2/\sigma^2_e)$, where K = the number of studies, σ^2 = the sample size weighted variance, and σ^2_e = the effect size variance due to sampling error. The Schmidt-Hunter 75% rule has greater power (to detect the existence of moderators) than the chi-square procedure. Whereas, the latter provides better protection against Type I errors (Sackett, Harris, & Orr, 1986). Thus, moderators were concluded to exist when there was a statistically significant amount of unexplained variance and the residual variation was a greater then 25%.

CHAPTER III

RESULTS

The analysis will be presented in two sections. The first part will demonstrate that using traditional data analytic interpretation procedures, significant differences across the seven regions did in fact exist--thereby supporting the validity of the situational specificity hypothesis. The second section will present the results when the Hunter et al. (1982) validity generalization procedures were applied to the data.

Situational effects were investigated through the computation of pearson product-moment (point biserial) correlation coefficients between test scales and the turnover criterion for each region. Situational specificity is present when a predictor that is valid in one situation is not valid in another situation (Ghiselli, 1966, p. 28). Further manifestations of situational specificity would be a positive correlation in one situation and a negative correlation in a different situation.

Inspection of Table 2 reveals a pattern of results consistent with Ghiselli's description. For region one, it is possible to conclude that the only reliable predictors of turnover were the MMPI scales: L, K, Hysteria, Psychasthenia, Schizophrenia, Anxiety, Ego-Strength, and Prejudice; and the GZTS Friendliness scale. Results for region two indicated that only the MMPI scales Schizophrenia and Prejudice predicted turnover. The subscale Schizophrenia was the only MMPI predictor of turnover in region three, whereas the following

TABLE 2

POINT BISERIAL CORRELATIONS AND CORRELATIONS CORRECTED FOR RANGE RESTRICTION BETWEEN PERSONALITY AND TURNOVER

Predictor	Region 1 (N = 95)	2 (N = 69)	3 (N = 90)	4 (N = 108)	5 (N = 47)	6 (N = 88)	7 (N = 35)
MMPI							
L	-.28(-.55)**	-.16(-.27)	-.10(-.19)	-.20(-.38)*	-.32(-.56)*	-.03(-.06)	-.22(-.21)
K	-.26(-.50)**	.06(.11)	-.10(-.20)	-.14(-.27)	-.20(-.35)	.06(.12)	.02(.03)
Hs	-.09(-.17)	.00(.00)	.00(.00)	-.04(-.08)	-.31(-.54)*	-.01(-.02)	.18(.23)
D	-.03(-.06)	.19(.35)	-.10(-.21)	.01(.02)	-.41(-.36)***	.07(.12)	-.10(-.17)
Hy	-.26(-.50)**	.20(.39)	-.11(-.22)	.00(.00)	-.09(-.19)	.05(.10)	-.21(-.36)
Pd	-.06(-.11)	-.08(-.14)	-.12(-.24)	-.15(-.31)	-.41(-.80)***	-.12(-.22)	-.08(-.10)
Pa	-.12(-.21)	.03(.04)	-.20(-.40)	-.04(-.07)	-.20(-.40)	-.09(-.17)	.26(.46)
Pt	-.21(-.40)*	-.06(-.11)	-.04(-.07)	-.06(-.10)	-.26(-.49)	.05(.09)	.00(.00)
Sc	-.31(-.59)***	-.35(-.60)***	-.40(-.76)***	-.31(-.57)***	-.42(-.44)***	-.03(-.07)	-.14(-.22)
Ma	.06(.11)	-.13(-.23)	-.08(-.15)	-.04(-.07)	.06(.11)	.01(.02)	-.10(-.14)
A	.22(.44)*	-.16(-.31)	.05(.10)	.26(.51)**	-.01(-.03)	.03(.06)	.16(.30)
Es	.21(.39)*	.16(.31)	.11(.24)	.23(.45)*	.23(.41)	.24(.23)*	-.10(-.14)
Dy	.17(.25)	-.23(-.36)	.02(.02)	-.01(-.01)	.02(.02)	-.05(-.08)	.32(.49)
Do	-.02(-.03)	.03(.06)	.02(.01)	.18(.36)	.28(.41)*	.13(.25)	-.31(-.49)
Re	-.07(-.13)	.11(.21)	.17(.34)	-.05(-.08)	.09(.14)	.22(.43)*	-.05(-.09)
Pr	.21(.42)*	-.24(-.43)*	-.08(-.16)	.05(.10)	-.12(-.23)	-.18(-.35)	.04(.07)
Cn	.14(.28)	.07(.13)	-.07(-.13)	.28(.54)***	.18(.34)	.07(.13)	-.01(-.02)
Wa	.14(.18)	-.19(-.30)	-.04(-.05)	.08(.15)	.01(.01)	-.11(-.20)	-.01(-.03)
Em	.19(.30)	.05(.10)	.00(.00)	.12(.20)	-.09(-.10)	-.19(-.33)	.02(.03)
Im	.11(.19)	-.11(-.16)	-.16(-.27)	.12(.21)	.17(.31)	-.22(-.38)*	-.06(-.11)

TABLE 2 (Continued)

Predictor	Region 1 (N = 95)	2 (N = 69)	3 (N = 90)	4 (N = 108)	5 (N = 47)	6 (N = 88)	7 (N = 35)
GZTS							
R	-.14(-.27)	-.12(-.22)	.08(.16)	-.10(-.20)	.04(.08)	.23(.46)*	-.14(-.27)
F	-.23(-.43)*	.17(.29)	.03(.06)	.14(.28)	-.08(-.11)	.02(.03)	.18(.32)
T	-.09(-.19)	-.11(-.22)	-.15(-.29)	-.26(-.50)**	.05(.10)	.13(.26)	.06(.09)
P	-.13(-.23)	.15(.26)	.06(.10)	.26(.51)**	-.02(-.03)	-.09(-.18)	-.19(-.38)

Note. Corrected values are in parentheses. MMPI = Minnesota Multiphasic Personality Inventory; Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; A = Anxiety; Es = Ego-Strength; Dy = Dependency; Do = Dominance; Re = Social Responsibility; Pr = Prejudice; Cn = Control; Wa = Work Attitude; Em = Emotional Immaturity; Im = Impulsivity; GZTS = Guilford Zimmerman Temperament Survey; R = Restraint; F = Friendliness; T = Thoughtfulness; P = Personal Relations.

*$p < .05$.

**$p < .01$.

***$p < .005$.

scales had significant point biserial correlations with the turnover criterion in region five: L, Hypochondriasis, Depression, Psychopathic Deviate, Schizophrenia, and Dominance. Two GZTS scales, Thoughtfulness and Personal Relations, were the best predictors of turnover in region four; along with the MMPI scales L, Schizophrenia, Anxiety, Ego-Strength, and Control. In region six, the Ego-Strength, Social Responsibility, and Impulsivity MMPI scales, with the GZTS trait of Restraint were most predictive of turnover. None of the predictors were significantly related to turnover in region seven. Forty-four percent of the validity coefficients were positive, supporting the contention of a slightly negative relationship between personality traits and turnover. A selective examination of validity coefficients illustrates these differences in results; Hypochondriasis correlated -.31 with turnover in region five and .18 in region seven; Depression correlated .19 in region two and -.41 in region five; Hysteria significantly correlated -.26 with turnover in region one and .20 in region two; Anxiety correlated .26 in region four, while -.16 in region two; and the GZTS Personal Relations scale correlated .26 and -.19 with turnover in regions four and seven, respectively.

A more rigourous criterion for determining situational specificity than that utilized by Ghiselli (1966) would be evidence of statistically significant differences in validity coefficients across situations. A chi square test was conducted for each of the 24 predictors following procedures outlined by Cohen and Cohen (1975, pp. 51-52). The GZTS Personal Relations scale was the only predictor

whose validity coefficients were heterogenous with an alpha level of .05 (χ^2 = 12.688). In addition, validity differences for Depression, Hysteria, Prejudice, and Impulsivity were statistically significant when the alpha was .10, and the chi-square values were 11.680, 10.859, 12.267, and 11.112, respectively. Also, the Anxiety, Dependency, Emotional Immaturity, Restraint, Friendliness, and Thoughtfulness scales approached significance ($p < .25$).

Thus, using traditional data analysis, the pattern of results in Table 2 leads one to conclude that different predictive relationships did exist across regions. In other words, the personality/turnover validities were not transportable across the different regions in the same national commercial real estate sales organization. This further implies that empirical validation of each predictor is required in each situation/setting.

Situational Specificity

The results to assess the hypothesis of situational specificity utilizing the Hunter et al. (1982) procedure are presented in Table 3. The first two columns of the table contain, respectively, the uncorrected sample size weighted mean ($\bar{r}$) of each distribution, and the observed standard deviation (SD) of each range-restricted corrected distribution. It should be recalled that each individual validity coefficient in a distribution was subjected to a univariate correction for range restriction prior to the meta-analysis. The predicted SD is the square root of the sum of the variance due to sampling error. This computation correctly reflected the fact that the sampling error variance of a range restricted corrected r_c is different

TABLE 3

SITUATIONAL SPECIFICITY HYPOTHESIS RESULTS FOR
BOTH MMPI AND GZTS PREDICTORS

Predictor	$\bar{r}$ [ab]	Observed SD	Predicted SD	Percentage of Variance Accounted For[c]	Residual SD	Chi-square
MMPI						
L	-.177	.177	.198	100	.000	5.567
K	-.091	.227	.210	86	.086	8.181
Hypochondriasis	-.041	.171	.205	100	.000	4.867
Depression	-.028	.198	.200	100	.000	6.876
Hysteria	-.053	.276	.219	63	.117	11.059
Psychopathic Deviate	-.135	.315	.212	45	.233	15.447**
Paranoia	-.075	.216	.210	94	.053	7.455
Psychasthenia	-.080	.186	.202	100	.000	5.917
Schizophrenia	-.284	.226	.214	90	.073	7.815
Hypomania	-.026	.119	.204	100	.000	2.400
Anxiety	.095	.271	.222	67	.155	10.429
Ego-Strength	.176	.145	.198	100	.000	3.799
Dependency	.017	.209	.146	49	.150	14.417**
Dominance	.068	.227	.188	68	.128	10.228
Social Responsibility	.063	.219	.211	93	.058	7.522
Prejudice	-.033	.288	.217	57	.189	12.349
Control	.104	.223	.218	96	.044	7.286
Work Attitude	-.008	.169	.185	100	.000	5.854
Emotional Immaturity	.027	.205	.174	72	.108	9.705
Impulsivity	-.020	.250	.196	62	.155	11.382

TABLE 3 (Continued)

Predictor	$\bar{r}$ [ab]	Observed SD	Predicted SD	Percentage of Variance Accounted For [c]	Residual SD	Chi-square
GZTS						
Restraint	-.014	.267	.222	70	.147	10.039
Friendliness	.022	.258	.200	60	.163	11.645
Thoughtfulness	-.078	.257	.218	72	.136	9.707
Personal Relations	.030	.291	.216	55	.195	12.728*
Total/Means	-.023	.225	.204	79	.091	29.196

Note. For each predictor, N = 532. MMPI = Minnesota Multiphasic Personality Inventory; GZTS = Guilford-Zimmerman Temperament Survey.

[a] For each predictor, total number of r's = 7.

[b] Sample size weighted estimates of the mean validity coefficients uncorrected for range restriction.

[c] Where predicted variance exceeded the observed variance, the percentage of variance accounted for is shown as 100%

*$p < .05$.

**$p < .025$

(i.e., larger) than that of an uncorrected $\underline{r}$ by a factor approximately equal to the squared ratio of the corrected to the uncorrected $\underline{r}$ (Hunter et al., 1982; Pearlman et al., 1980). Sampling error variance was estimated in this analysis by multiplying each distribution's mean range restricted corrected validity-mean uncorrected validity ratio squared $(\bar{r}_c/\bar{r})^2$ by the sampling error of the observed uncorrected correlations, $K(1 - \bar{r}^2)^2/N$, where K is the number of studies (regions), $\bar{r}$ = sample size weighted mean correlation of the distribution, and $\underline{N}$ is the total sample size. The next column reports the percentage of variance that is accounted for by statistical artifacts (i.e., sampling error). The column headed residual $\underline{SD}$ presents the square root of the variance that remains after the sampling error variance is subtracted from the observed variance. Consequently, the residual variance can be attributable to differences among regions in criterion unreliability, computational and typographical errors, and to true situation effects.

Using general guidelines suggested by Pearlman et al. (1980) and Schmidt, Hunter, and Pearlman (1982) the situational specificity hypothesis can be rejected for 12 of the 24 distributions, because 75% or more of the observed variance can be accounted for by artifacts. For the following scales: L, K, Paranoia, Psychasthenia, Schizophrenia, Ego-Strength, Social Responsibility, Control, Hypochondriasis, Depression, Hypomania, and Work Attitude the amount of residual variance was nonsignificant. Thus no evidence for differences in underlying subpopulation correlations exist. In other words, the variance left within which situational moderators can operate is

very limited. Moreover, for the L, Ego-Strength, Psychasthenia, Hypochondriasis, Depression, Hypomania, and Work Attitude scales, it appears that no variation is left at all.

The remaining 12 distributions (i.e., Hysteria, Dependency, Prejudice, Emotional Immaturity, Impulsivity, Restraint, Thoughtfulness, Friendliness, Personal Relations, Psychopathic Deviate, Anxiety, and Dominance) did not meet the 75% criterion--thereby not allowing rejection of the situational specificity hypothesis. Situational variance accounted for an average of 38% of the variation in observed validity coefficients across regions, with a range from 28% to 51%. Because two sources of error variance could not be estimated or controlled, the estimate of the situation-specific variance may be inflated (and underestimate the true degree of validity transportability). To put it another way, if the two unassessed error sources could be accounted for and subtracted, there would be even less room (or perhaps none at all) for situational factors (i.e., variance) to operate. Although a fairly large proportion of the observed variance is unaccounted for, nine of the 12 distributions had insignificant amounts of unexplained residual variance, and validity generalization is still possible. In the remaining three distributions (Dependency, Personal Relations, and Psychopathic Deviate), the chi-squares were significant ($p < .05$) indicating fair amounts of unexplained residual variance. Thus, in these instances situational moderators are apparently operating and validity generalization is not likely.

It is also observed in Table 3 that for seven of the 24 distributions (L, Ego-Strength, Psychasthenia, Depression, Hypochondriasis,

Hypomania, and Work Attitude) the predicted SD exceeds the observed SD. This is probably due to second order sampling error in the estimate of observed variation stemming from the small number of validity coefficients in each distribution (Schmidt, Hunter, Pearlman, & Hirsh, 1985). In these seven distributions when the percentage of variance accounted for is rounded down to 100%, the average percentage of variance accounted for across all distributions is 79.10% (with a mean residual SD of .091). A more accurate estimate of the percentage of variance accounted for can be obtained by correcting for second order sampling error. If the actual computed percentages are used and weighted by sample size, the mean rises to 95.21%.

Validity Generalization

Several criteria may be utilized in assessing the transportability of validities. The most conservative criterion to determine the validity generalization of a measure is when the true or corrected variance of the distribution of its validities is essentially zero. In this case there is no room for factors to moderate the relationship. Another criterion is when most of the values in the distribution of estimated true validities exceed some minimum practical level of validity. For example, Hunter and Schmidt (1982) and Schmidt, Hunter, McKenzie, and Muldrow (1979) have adopted .10 as this minimum level. The criterion for evaluating validity generalizability used in this study is the one proposed by Callender and Osburn (1981). Their approach is directly analogous to procedures for the traditional testing of statistical significance. They have suggested that a conclusion of validity generalization be reached whenever the 90%

credibility value is either positive or negative--depending upon whether one is expecting positive or negative validities, respectively. A significant result using customary significance tests means only that the confidence interval does not include zero.

Table 4 presents the validity generalization results. The first two columns of the table show the mean ($\bar{\hat{\rho}}$) and standard deviation ($SD_{\hat{\rho}}$) of the prior Bayesian validity distributions. The mean of the prior distribution is the mean validity coefficient corrected for range restriction. The $SD_{\hat{\rho}}$ is the square root of the variance of estimated true validities corrected for sampling error variance. In this study, the $SD_{\hat{\rho}}$ is the same as the residual SD. The last column shows the 90% credibility values (i.e., the value at or above which 90% of all estimated true values are expected to lie). These values were computed by multiplying the $SD_{\hat{\rho}}$ by 1.2816 (the standard score corresponding to the tenth percentile of the normal distribution) and either adding or subtracting this value from $\bar{\hat{\rho}}$--depending upon whether $\bar{\hat{\rho}}$ is negative or positive, respectively. The 90% credibility values tend to be slight underestimates of the values of the true validity estimates at the 10th percentile, since the $SD_{\hat{\rho}}$ is inflated as a result of not correcting for all sources of error variance.

In the case of the Schizophrenia, L, Ego-Strength, and Control distributions, the mean validities and credibility values range from moderate to substantial--thus demonstrating that validities are generalizable across regions. These scales can confidently be used to predict turnover. The K, Paranoia, Psychasthenia, and Social Responsibility predictors show low mean validity. While three of

TABLE 4

VALIDITY GENERALIZATION RESULTS FOR PERSONALITY PREDICTORS WHERE THE CRITERION IS TURNOVER

Predictor	Bayesian Prior Distribution		90% C. V.
	$\bar{\hat{\rho}}$ [a]	$SD\hat{\rho}$ [b]	
MMPI			
L	-.316	.000	-.316
K	-.174	.086	-.063
Hypochondriasis	-.080	.000	-.080
Depression	-.020	.000	-.020
Hysteria	-.099	.117	.051
Psychopathic Deviate	-.254	.233	.044
Paranoia	-.147	.053	-.078
Psychasthenia	-.148	.000	-.148
Schizophrenia	-.490	.073	-.397
Hypomania	-.044	.000	-.044
Anxiety	.186	.155	-.013
Ego-Strength	.307	.000	.307
Dependency	.021	.150	-.171
Dominance	.122	.128	-.042
Social Responsibility	.122	.058	.048
Prejudice	-.060	.189	.183
Control	.204	.044	.148
Work Attitude	-.021	.000	-.021
Emotional Immaturity	.045	.108	-.094
Impulsivity	-.035	.155	.164
GZTS			
Restraint	-.025	.147	.163
Friendliness	.044	.163	-.165
Thoughtfulness	-.155	.136	.019
Personal Relations	.057	.195	-.194

Note. C. V. = credibility value; MMPI = Minnesota Multiphasic Personality Inventory; GZTS = Guilford-Zimmerman Temperament Survey.

[a] Estimated mean true validities. These figures are corrected only for range restriction.

[b] $SD\hat{\rho}$ = Residual SD.

the four distributions 90% credibility values are not that far above zero, there is still considerable evidence for generalizability according to the Callender and Osburn decision rule.

The means of the Work Attitude, Hypochondriasis, Depression, and Hypomania distributions are very near zero, and the confidence interval for each distribution does not include zero. It will be recalled that the situational specificity hypothesis was rejected as well. Thus, these results indicate that it is the lack of validity that can be generalized. In other words, in a new region these scales will have, at best, a negligible validity for predicting turnover. In regards to the distributions discussed thus far, the mean $\hat{\rho}$ of each distribution is the best estimate of the validity in a new setting and not the 90% credibility values (Pearlman et al., 1980).

A further examination of Table 4 indicates that the Hysteria, Dependency, Prejudice, Emotional Immaturity, Impulsivity, Restraint, Friendliness, and Personal Relations scales (where the hypothesis of of situational specificity could not be rejected) lack validity as well as generalizability (since the confidence intervals include zero). The Anxiety, Dominance, and Thoughtfulness scales demonstrate somewhat higher validity, but again do not generalize. Finally, the Psychopathic Deviate distribution displays moderate negative validity, yet the 90% credibility value is positive. The lack of generalizability of this scale (as well as both the Personal Relations and Dependency scales) is consistent with earlier findings that the unexplained residual variance was significant.

Determining a Predictor Battery

The results presented above apply to individual tests/scales. In practice, several tests are frequently used in decision making. The same can be done in developing a battery to predict voluntary turnover.

Table 5 shows the matrix of intercorrelations among all the variables of interest in this study. The correlations that are of primary concern here are those that address the relationship between turnover and the test scales. As can be seen in Table 5, turnover and Schizophrenia were correlated (r = -.30). Turnover was also correlated significantly with both L and Ego-Strength (r = -.20 and .20, respectively) as well as with Psychopathic Deviate (r = -.14), Control (r = .12), and Thoughtfulness (r = -.11)--although to a smaller degree. In addition, for the most part there were significant correlations between the majority of the predictor variables. This partially reflects the fact that for the MMPI, many of the items are included in more than one scale (Grant, Welsh, & Dahlstrom, 1972). The risk associated with common method variance is that it would result in distortions in the estimation of the magnitude, as well as sign, of the regression coefficients (Kerlinger & Pedhazur, 1973). An examination of each predictor's squared multiple correlation with all the other predictors indicated that multicollinearity was not a problem.

In order to determine which combination of predictors provided the optimum battery in predicting turnover, a SAS regression program was performed (SAS Institute Inc., 1985). In this analysis, the

TABLE 5

CORRELATIONS BETWEEN VARIABLES

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. L																								
2. K	46**																							
3. Hs	26**	61**																						
4. D	21**	18**	27**																					
5. Hy	13*	42**	55**	19**																				
6. Pd	02	42**	45**	20**	39**																			
7. Pa	09	24**	25**	04	28**	26**																		
8. Pt	02	48**	46**	27**	39**	47**	30**																	
9. Sc	14*	55**	50**	10	37**	50**	27**	62**																
10. Ma	-15**	-26**	-13*	-38**	-09	02	02	-02	10															
11. A	-39**	-68**	-28**	02	-19**	-13*	-07	00	-16**	20**														
12. Es	-07	11*	-08	-18**	-14**	-07	-12*	-12*	-15**	08	-23**													
13. Dy	-45**	-55**	-23**	-06	-21**	-03	01	12*	-06	20**	71**	-31**												
14. Do	-10	09	05	-12*	00	-01	-01	-11*	-06	00	-18**	22**	-22**											
15. Re	37**	41**	23**	17**	22**	-07	-04	02	00	-30**	-37**	01	-38**	19**										
16. Pr	-26**	-63**	-33**	-09	-32**	-16**	-31**	-22**	-15**	24**	53**	-15**	46**	-20**	-48**									
17. Cn	-52**	-52**	-30**	-15**	-25**	-14*	-07	-14**	-25**	-25**	41**	19**	32**	18**	-49**	28**								
18. Wa	-23**	-59**	-26**	-07	-25**	-16**	-15**	-14**	-15**	-15**	56**	-07	40**	-18**	-34**	59**	30**							
19. Em	-39**	-51**	-10	-02	-02	07	01	12*	-03	-03	51**	-12*	52**	-09	-39**	36**	40**	43**						
20. Im	-61**	-59**	-31**	-29**	-19**	01	-03	-07	-06	-06	53**	-07	49**	01	-55**	43**	54**	41**	49**					
21. R	14**	10	06	17**	08	00	00	04	-03	-17**	-06	-17**	-02	05	27**	-15**	-19**	-13*	-12	-20**				
22. F	27**	48**	33**	22**	32**	16**	14*	22**	17**	-37**	-35**	-10	-26**	05	40**	-43**	-45**	-46**	-25**	-45**	27**			
23. T	-03	-23**	-13*	-08	-12*	-10	-04	-04	-07	18**	17**	-14**	24**	-04	-02	16**	12*	16**	12*	16**	31**	-15**		
24. P	19**	52**	21**	05	24**	20**	22**	28**	20**	-24**	-43**	06	-31**	03	27**	-55**	-33**	-43**	-32**	-34**	08	45**	-22**	
25. Turnover	-20**	-06	-04	-05	-05	-14**	-08	-08	-30**	-05	05	20**	-03	09	05	-06	12*	-04	-03	-01	-05	03	-11**	05

Note. N = 532. Decimals have been omitted. Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; A = Anxiety; Es = Ego-Strength; Dy = Dependency; Do = Dominance; Re = Social Responsibility; Pr = Prejudice; Cn = Control; Wa = Work Attitude; Em = Emotional Immaturity; Im = Impulsivity; R = Restraint; F = Friendliness; T = Thoughtfulness: P = Personal Relations.

*$p < .01$.

**$p < .001$.

criterion for variable entry and retention in the model was the .01 level of significance or better. As can be seen in Table 6, the Schizophrenia scale accounted for a large proportion (eight percent) of the variance in voluntary turnover.

TABLE 6

SUMMARY OF MULTIPLE REGRESSION ANALYSIS USING TURNOVER AS CRITERIA AND PERSONALITY VARIABLES

Variable	Regression Weights	R^2	R^2 Change
Schizophrenia	-.06	.08	.08*
Ego-Strength	.02	.11	.03*
L	-.05	.14	.03*
K	.06	.17	.04*
Anxiety	.03	.19	.02*

Note. N = 521. The only variables shown are those entering equation with p .01. The variables are listed in the order of inclusion in the regression equation. R^2 Change is the increase in the R^2 when the variable is added to the equation.

*p < .001.

An additional six percent of this is explained by both the Ego-Strength and L scales. It is also interesting to note that while the K and Anxiety scales of the MMPI were not correlated with the criterion, they contributed significant proportions (four and two percent, respectively) of variance to the prediction equation. This suggests that both the K and Anxiety scales might act as suppressor variables.

Conger (1974) has defined a suppressor variable as one which increases the predictive validity of another variable (or group of variables) by its entry into a regression equation. In other words, Conger asserts that only those variables whose regression weights are increased should a variable be considered a suppressor. Following procedures outlined by Conger for identifying suppressor variables in multiple regression equations, an examination of the weighted validities (i.e., the product of the regression weight and zero order validity) indicated that the K scale acted to suppress (or remove the irrelevant predictive variance) of both the Schizophrenia and L scales. Moreover, the Anxiety scale tended to supress the variance of the K scale--thereby contributing to the overall prediction. Thus, 19% of voluntary turnover variance was accounted for by the five variable set (shrunken $\underline{R}^2$ = .18).

CHAPTER IV

DISCUSSION

Situational Specificity/Generalizability of Personality Predictors

The present study is an attempt to understand why past research has produced conflicting and unclear results with regard to the nature of the relationship between personality and employee turnover. The Hunter et al. (1982) validity generalization procedure was applied to a data set that would allow for a test of the situational hypothesis. The overall results are consistent with findings in other validity generalization studies regarding the percentage of artifactual variation accounted for (Pearlman et al., 1980; Schmidt et al., 1980; Schmidt et al., 1981; Schmidt et al., 1979). In 22 of the 24 distributions representing individual test scale-turnover combinations, more than 50% of the variance in validities across regions was due to statistical artifacts. In 12 of these cases this percentage exceeded the 75% decision criterion--thereby refuting the situational hypothesis. For nine of the remaining 12 distributions (Hysteria, Prejudice, Emotional Immaturity, Impulsivity, Anxiety, Dominance, Restraint, Friendliness, and Thoughtfulness) the situational specificity hypothesis could not be rejected. But, even in these instances the residual variances were small and insignificant--thereby leaving little room for situational moderators of any type to operate. On the other hand, the residual variances for the Dependency, Psychopathic Deviate, and Personal Relations scales were significant--indicating that situational differences exist. Several potential moderators for investigation were unavailable due to

methodological concerns and data unavailability. It is possible to rule out sex as a moderator variable since the majority of employees were males (92%), and female employees were fairly equally distributed throughout the seven regions. Any effects would be negligible. The problem of looking at age as a moderator is that it is confounded with other variables like tenure (Mobley et al., 1979; Price, 1977). Level of performance has also been suggested as a possible moderator. However, the only "objective" measure available was the actual amount of dollar sales. No data were available for the first two years of employment for the short tenure group. Additional research looking at potential moderators needs to be done for the future. Overall, it would appear that one explanation for past inconsistent results in the personality-turnover literature is the failure to consider statistical artifacts.

The findings of the meta-analysis also support the conclusions of some researchers that personality measures, overall, have low validity regarding turnover (Mobley, 1982; Muchinsky & Tuttle, 1979; Pettman, 1973). Personality test validities corrected for range restriction yielded a mean true validity coefficient of -.04. Moreover, the average amount of variance accounted for in the 24 distributions was 79%--thereby providing support for the first hypothesis. And, the overall residual variance was insignificant ($p > .05$)--thereby allowing one to conclude that the lack of validity does generalize. Thus, this finding that "personality" overall is a weak determinant of turnover, in most situations, tends to provide support for investigators (e.g., Mobley, 1977; Steers & Mowday, 1981) who have

previously relegated personality to a minor role in models of voluntary turnover. But, personality traits come in all shapes and sizes. One must be cautious in making inductive statements about an entire class of constructs when a multitude of individual traits have not been looked at. More specifically, in examining the results of this project, the conclusion that there is a weak relation between "personality" and turnover must be interpreted with caution, since only selected scales from the MMPI and GZTS were utilized, and did not include other personality measures.

Moreover, 12 individual scales did generalize (using the criterion proposed by Callender & Osburn and selected for application in this study), across regions (thereby providing partial support for the second hypothesis), with four of the 12 measures (Schizophrenia, L, Ego-Strength, and Control) showing moderate to high mean true validity (-.49, -.32, .31, and .20, respectively). The K ($\underline{r}$ = -.17), Paranoia ($\underline{r}$ = -.15), Psychasthenia ($\underline{r}$ = -.15), and Social Responsibility ($\underline{r}$ = .12) scales showed low mean true validity. The remaining four scales (Hypochondraisis, Depression, Hypomania, and Work Attitude) had mean true validities near zero.

Nine other scales (which did not exceed the 75% rule) failed to meet Callender and Osburns (1981) generalizability criterion. However, another criterion for transsituational generalizability that has been utilized by other researchers is when the actual amount of residual (true) variance is zero. In these scales the residual variation was insignificant, thereby leaving little room for situational moderation. Thus, using this criterion these measures can be said

to be transportable, but the evidence is weaker since they failed to meet the 75% decision rule. This conclusion is consistent with the findings by Pearlman et al. (1980) and Schmidt et al. (1980) that validity generalization is still feasible even when the situational specificity hypothesis cannot be rejected. The majority of these predictors had mean true validities near zero, with the exceptions being the Anxiety ($\underline{r}$ = .19) and Dominance ($\underline{r}$ = .12) scales.

In addition to being generalizable, these results overall support the conclusions by Porter and Steers (1973) and Schuh (1967) that some individual personality measures yield good predictive results. It also suggests that one personality pattern (or trait) might not be generalizable nor predictive of turnover using different occupational groups. Moreover, research needs to be done to discover for a particular occupational group which traits are key to identifying the "hobo factor" for individuals in that group. In addition, these 21 scales (12 applying the Callender & Osburn criterion) do not provide support for Guion and Gottier's (1965) contention that personality measures must be developed and validated for a specific setting and use.

Individual Scales and Turnover

After completing this meta-analysis, an examination of the uncorrected scales and turnover was conducted. Only five MMPI scales (L, Schizophrenia, Ego-Strength, Psychopathic Deviate, and Control) were significantly correlated with voluntary turnover. This is in contrast to Butterfield and Warren's (1962) early finding that no single subscale was predictive of turnover. The present results are

partially in agreement with a later study by Butterfield and Warren (1963) in which only the Psychopathic Deviate scale, of the four scales (Hypomania, K, and L) they examined, was significant using another occupational group. Another reason to account for the incongruity in results, besides that resulting from statistical artifacts, is that the criterion Butterfield and Warren (1963) used was involuntary and not voluntary turnover. It is possible that some personality measures are able to predict only one or the other type of turnover criterion. Clearly, this issue deserves further research. Another explanation for these findings is that the MMPI might not be sensitive enough in predicting turnover using a "normal" population (which would include different occupational groupings), since it was primarily designed for use in clinical samples.

An examination of each of the five MMPI scales might also shed some light on their significant relationship with voluntary turnover.

Psychopathic Deviate. The finding that employees who left the organization exhibited high scores on the Psychopathic Deviate scale can be explained by considering the nature of the commercial sales job. It generally takes a considerable amount of time from initial client contact to final contract. In other words, it requires a great deal of persistence. In contrast, employees with high Psychopathic Deviate tendencies tend to be quite impatient, display poor work histories, and are rebellious toward authority (Larchar, 1974). Consequently, they are much more likely to engage in risk taking behavior (i.e., more willing to leave the job) as a result of their im-

patience, and greater likelihood of their inability to get along with superiors.

Schizophrenia. Turnover was also related to higher scores on the Schizophrenia scale of the MMPI. Dahlstrom, Welsh, and Dahlstrom (1972) characterized persons with high levels of this factor as being alienated from their social environment--including avoiding dealing with people and new situations. They are further characterized as being resentful, shy, hostle, and aggressive toward others. Those employees who left apparently found it difficult to handle the interpersonal relationships that are essential in a sales position.

Ego-Strength. Previous research utilizing the Sixteen Personality Factor Questionnaire found that low ego-strength is associated with voluntary turnover (Bernardin, 1977). This study confirms such a relationship using a similar sample of sales representatives. It is possible that employees in this sales position (which is on a strict commision basis) found it quite stressful, and thus had a greater propensity to leave.

Control. Additionally, withdrawal was related to lower scores on the Control scale. Graham (1977) described persons with low levels of this trait as being overly conventional and rigid. In other words, they were less likely to experiment with or explore their environment. Those employees who left the organization apparently perceived their sales jobs as interfering with manifestations of this personality trait. They subsequently terminated their employment to seek a job that accomodated their inflexibility.

L scale. Finally, voluntary turnover was related to a greater need to appear in a favorable light as measured by the L scale of the MMPI. Those employees with high levels of this trait were probably much more unwilling to admit to shortcomings. When their performance on the job exposed their weaknesses, the persons sought employment elsewhere.

The analysis of the uncorrected scales also revealed that the only GZTS scale significantly related to turnover was Thoughtfulness.

Thoughtfulness. Employees with high levels of this trait are more likely to be characterized by reflectiveness, going off on tangents, and overcomplicating matters (Guilford et al., 1976). The result is that these individuals apparently do not manage their time effectively. The nature of the commercial sales position requires one to be fairly well organized in order to be effective. Employees with a high level of this trait would find it difficult to be satisfied in this environment, and thereby would be less likely to remain with the organization.

Other researchers have found no relation between the GZTS scales used in this study and turnover (Hedberg & Baxter, 1957; MacKinney & Wolins, 1960). A possible explanation for this apparent lack of results with the GZTS scales in this study is that some confounding might have possibly occurred. The sales managers in the organization received a profile of only the GZTS subscales along with a copy of the evaluative report. It is possible that managers are basing their hiring decisions on one or two scales that in their subjective

judgement they feel predict top performers.

An extension of the above argument can be made as to why certain scales of both the MMPI and GZTS were significantly related to turnover. The short and long tenure groups utilized in this project were selected from the population of all those individuals hired as sales representatives at the organization. The MMPI and GZTS were often used as one criteria in making the recommendation to hire. Moreover, certain scales of both instruments were given greater weight than others. It is possible that employing both instruments in the initial selection of the employees reduced the variability within the population from which the long and short tenure groups were chosen--thereby making it difficult for certain scales to differentiate between long and short tenure employees.

Predictor Battery

This study also demonstrated a screening battery of good predictive power. A substantial proportion of the criterion variance (19%) was explained by five predictors (Schizophrenia, Ego-Strength, L, K, and Anxiety). The Schizophrenia scale was, without doubt, the most powerful among the predictors evaluated, accounting for 8% of the variance. The insignificant correlations between the K and Anxiety scales with the criterion, as well as the fact that both added significantly to the proportion of explained criterion variance, point to their operation as suppressor variables. The finding that the former scale acts as a suppressor is in agreement with research by Meehl and Hathaway (1946).

However, many employers have questioned the utility of including

suppressor variables in a predictive equation for a multitude of reasons, including the possibility that applicants could learn how scores were utilized and then purposely attempt to get low scores on the test. Even if the K and Anxiety scales were dropped from the equation, one would still account for a good proportion of criterion variance (14%). Moreover, since the three remaining scales generalize, one would be able to use the equation across regions with a high degree of confidence. Even if one did not reject the use of suppressor variables, the original five variable model equation could be utilized, or a four variable model which excluded the Anxiety scale. An advantage of the latter model is that the evidence for its' generalizability across regions is stronger (depending upon the criterion utilized) than with the former model.

A major theme throughout the literature is that utilizing certain employee selection procedures can result in a significant reduction in turnover. The screening battery appears to be an effective tool for management in attempting to control voluntary turnover. The reduction of turnover using such a screening procedure would have financial implications for managers.

One method of estimating the economic impact of utilizing the five variable battery in reducing turnover would be to calculate the number of voluntary turnovers the test battery saves and multiply that by an estimate of the turnover cost per employee. An estimate of the turnover cost for a sales position in the organization studied was approximately $210 per employee (the cost of testing one applicant). This estimate is very conservative since it does not

include other direct costs such as employment advertising, employment staff compensation, and relocation expenses which were not available. Moreover, it does not include the indirect costs involved such as training and productivity losses. In order to compute the expected reduction in turnover rates from the test battery, the "binomial effect-size display" (BESD) procedure was employed (Rosenthal & Rosnow, 1984). The BESD addresses the question of what is the effect on the turnover rate of a new predictor battery. In other words, it displays the change in the voluntary turnover rate attributable to the new predictor battery. The results suggested that managers implementing the test battery can expect to reduce voluntary turnover approximately 43 percentage points over the rates they are experiencing without using the battery. Therefore, for every 100 sales employees that are in the organization, a savings of $9,030 per year (which is a very conservative estimate) can be expected utilizing the test battery.

Future Research

Several directions for future research are suggested from the results of this project. First the findings are descriptive of only one organization and one job type. There is a definite need for replication in different organizations before the results can be generalized with greater confidence.

Secondly, sampling error has been shown to account for most of the variation in validities. However, the correction for sampling error assumes random sampling. Yet, in this field study the sample was not random. Thus, application of the Hunter et al. (1982)

procedure might be compromised by lack of random samples. This suggests that additional research might need to be done utilizing random samples.

Also, overall, validity generalization was shown in this project to be a fairly robust phenomenon. Yet, the variation in three distributions indicated the possible presence of situational moderators. Since data were unavailable to assess which situational moderators were accounting for the variation, further research needs to be done in this area.

Future research also needs to examine the relationship between personality and turnover employing additional instruments. This project examined a narrow set of predictors. In order to develop a more accurate profile of long and short tenure employees, other personality inventories need to be assessed for their effectiveness.

Finally, more research needs to be done in regards to the type of criterion used. Whereas, this project used voluntary turnover (i.e., the employee initiated the action), it would be interesting to see whether the same personality traits would be able to predict involuntary turnover as well. In addition, the criterion employed in this study for the short tenure group was defined as being on the job five years or less. It would be interesting to see if different personality test scores would be correlated with turnover if the short tenure criterion was defined at six months, twelve months, or longer periods of time. Kriedt and Gadel (1953) found higher validity coefficients for predicting three month separators than twelve month separators. In other words, certain test scales might have a stronger relationship with turnover at different periods of tenure.

LIST OF REFERENCES

REFERENCES

Alexander, R. A., Alliger, G. M., & Hanges, P. J. Correcting for range restriction when the population variance is unknown. Applied Psychological Measurement, 1984, 8, 431-437.

Atchison, T. J., & Lefferts, E. A. The prediction of turnover using Herzberg's job satisfaction technique. Personnel Psychology, 1972, 25, 53-64.

Bangert-Drowns, R. L. Review of developments in meta-analytic method. Psychological Bulletin, 1986, 99, 388-399.

Barron, F. An ego strength scale which predicts response to psychotherapy. Journal of Consulting Psychology, 1953, 17, 327-323.

Baysinger, B. D., & Mobley, W. H. Employee turnover: Individual and organizational analysis. In K. M. Rowland & G. R. Ferris (Eds.), Research in Personnel and Human Resources Management (Vol. 1). Greenwich, CT.: JAI Press Inc., 1983.

Becker, W. C. The Personality Inventory. In O. K. Buros (Ed.), Sixth Mental Measurements Yearbook. Highland Park, N. J.: Gryphon Press, 1965.

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

Bluedorn, A. C. A taxonomy of turnover. Academy of Management Review, 1978, 3, 647-651.

Bluedorn, A. C. The theories of turnover: Causes, effects and meaning. In S. Bacharach (Ed.), Research in the Sociology of Organizations (Vol. 1). Greenwich, CT.: JAI Press, 1982. (a)

Bluedorn, A. C. A unified model of turnover from organizations. Human Relations, 1982, 35, 135-153. (b)

Brayfield, A. H. & Crockett, W. H. Employee attitudes and employee performance. Psychological Bulletin, 1955, 52, 396-424.

Brewer, J. K. & Hills, J. R. Univariate selection: The effects of size of correlation, degree of skew, and degree of restriciton. Psychometrika, 1969, 26, 347-372.

Brogden, H. E., & Taylor, E. K. A theory and classification of criterion bias. Educational and Psychological Measurement, 1950, 10, 159-186.

Bullock, R. J., & Svyantek, D. J. Analyzing meta-analysis: Potential problems, an unsuccessful replication, and evaluation criteria. Journal of Applied Psychology, 1985, 70, 108-115.

Butterfield, E. C., & Warren, S. A. The use of the MMPI in the selection of hospital aides. Journal of Applied Psychology, 1962, 46, 34-40.

Butterfield, E. C., & Warren, S. A. Prediction of attendant tenure. Journal of Applied Psychology, 1963, 47, 101-103.

Callender, J. C., & Osburn, H. G. Testing the constancy of validity with computer-generated sampling distributions of the multiplicative model variance estimate: Results for petroleum industry validation research. Journal of Applied Psychology, 1981, 66, 274-281.

Cleland, C. C., & Peck, R. F. Psychological determinants of tenure in institutional personnel. American Journal of Mental Deficiency, 1959, 64, 876-888.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. New York: Lawrence Erlbaum Associates, Inc., 1975.

Conger, A. J. A revised definition for suppressor variables: A guide to their identification and interpretation. Educational and Psychological Measurement, 1974, 34, 35-46.

Cotton, J. L., & Tuttle, J. M. Employee turnover: A meta-analysis and review with implications for research. Academy of Management Review, 1986, 11, 55-70.

Cuadra, C. A. A scale for control in psychological adjustment (Cn). In G. W. Welsh & W. G. Dahlstrom (Eds.), Basic readings on the MMPI in psychology and medicine. Minneapolis: University of Minnesota Press, 1953.

Cuadra, C. A., & Reed, C. F. Prediction of psychiatric aide performance. Journal of Applied Psychology, 1957, 41, 195-197.

Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. An MMPI handbook. Volume I. Clinical interpretation. Minneapolis: University of Minnesota Press, 1972.

Dreher, G. F. The role of performance in the turnover process. Academy of Management Journal, 1982, 25, 137-147.

Farrell, D., & Rusbuilt, C. E. Exchange variables as predictors of job satisfaction, job commitment, and turnover: The impact of rewards, costs, alternatives, and investments. Organizational Behavior and Human Performance, 1981, 27, 78-95.

Farris, G. F. A predictive study of turnover. Personnel Psychology, 1971, 24, 311-328.

Forest, C. R., Cummings, L. L., & Johnson, A. C. Organizational participation: A critique and model. Academy of Management Review, 1977, 2, 586-601.

Forsyth, R. A. An empirical note on correlation coefficients corrected for restriction in range. Educational and Psychological Measurement, 1971, 31, 115-123.

Ghiselli, E. E. The validity of occupational aptitude tests. New York: John Wiley & Sons, Inc., 1966.

Ghiselli, E. E., & Barthol, R. P. The validity of personality inventories in the selection of employees. Journal of Applied Psychology, 1953, 37, 18-20.

Glass, G. V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5, 3-8.

Glass, G. V. Integrating findings: The meta-analysis of research. Review of Research in Education, 1977, 5, 351-379.

Glass, G. V., McGaw, B., & Smith, M. L. Meta-analysis in social research. Beverly Hills, CA.: Sage Publications, 1981.

Gough, H. G. Studies for social intolerance: II. A personality scale for anti-Semitism. Journal of Social Psychology, 1951, 33, 247-255.

Gough, H. G. California Psychological Inventory manual. Palo Alto: Consulting Psychologists Press, 1957.

Gough, H. G., McClosky, H., & Meehl, P. E. A personality scale for dominance. Journal of Abnormal and Social Psychology, 1951, 46, 360-366.

Gough, H. G., McClosky, H., & Meehl, P. E. A personality scale for social responsibility. Journal of Abnormal and Social Psychology, 1952, 47, 73-80.

Graham, J. R. The MMPI: A practical guide. New York: Oxford University Press, 1977.

Greener, J. M., & Osburn, H. G. An empirical study of the accuracy of corrections for restriction in range due to explicit selection. Applied Psychological Measurement, 1979, 3, 31-41.

Greener, J. M., & Osburn, H. G. Accuracy of corrections for restriction in range due to explicit selection in heteroscedastic and nonlinear distributions. Educational and Psychological Measurement, 1980, 40, 337-346.

Gross, A. L. Relaxing assumptions underlying corrections for restriction of range. Educational and Psychological Measurement, 1982, 42, 795-801.

Gross, A. L., & Fleischman, L. Restriction of range corrections when both distributions and selection assumptions are violated. Applied Psychological Measurement, 1983, 7, 227-237.

Guilford, J. P., & Zimmerman, W. S. The Guilford-Zimmerman Temperament Survey: manual. Beverly Hills, CA.: Sheridan Supply Company, 1949,

Guilford, J. P., Zimmerman, W. S., & Guilford, J. P. The Guilford-Zimmerman Temperament Survey handbook: Twenty-five years of research and application. San Diego, CA.: EdITS Publishers, 1976.

Guion, R. M., & Gottier, R. F. Validity of personality measures in personnel selection. Personnel Psychology, 1965, 18, 135-164.

Hanna, J. V. Job stability and earning power of emotionally maladjusted as compared with emotionally adjusted workers. Journal of Abnormal and Social Psychology, 1935, 30, 155-163.

Hakkinen, S., & Toivainen, Y. Psychological factors causing labour turnover among underground workers. Occupational Psychology, 1960, 34, 15-30.

Hathaway, S. R., & McKinley, J. C. The Minnesota Multiphasic Personality Inventory manual. New York: Psychological Corporation, 1967.

Hedberg, R., & Baxter, B. A second look at personality test validation. Personnel Psychology, 1962, 15, 223-225.

Hedges, L. V. Estimation of effect size from a series of independent experiments. Psychological Bulletin, 1982, 92, 490-499.

Hedges, L. V. A random effects model for effect sizes. Psychological Bulletin, 1983, 93, 388-395.

Herzberg, F., Mausner, B., Peterson, R. O., & capwell, R. Job attitudes: Review of research and opinions. Pittsburgh: Pittsburg Psychological Services, 1957.

Hines, G. H. Achievement motivation, occupations, and labor turnover in New Zealand. Journal of Applied Psychology, 1973, 58, 313-317.

Hirsh, H. R., Northrop, L. C., & Schmidt, F. L. Validity generalization results for law enforcement occupations. Personnel Psychology, 1986, 39, 399-420.

Hunter, J. E., & Schmidt, F. L. Fitting people to jobs: Implications of personnel selection for national productivity. In M. D. Dunnette & E. A. Fleishman (Eds.), Human performance and productivity, Vol. 1: Human capability assessment. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1982.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA.: Sage Publications, 1982.

Kerlinger, F. N., & Pedhazur, E. Multiple regression in behavioral research. New York: Holt, Rinehart, & Winston, 1973.

Kriedt, P. H., & Gadel, M. S. Prediction of turnover among clerical workers. Journal of Applied Psychology, 1953, 37, 338-340.

Lang, V. R., & Krug, S. E. Perspectives on the executive personality: A manual for the Executive Profile Survey. Champaign, Ill.: Institute for Personality and Ability Testing, 1978.

Larchar, D. The MMPI: Clinical assessment and automated interpretation. Los Angeles, CA.: Western Psychological Services, 1974.

Lee, R., Miller, K. J., & Graham, W. Y. Corrections for restriction of range and attenuation in criterion-related studies. Journal of Applied Psychology, 1982, 67, 637-639.

Lefkowitz, J. Personnel turnover. Progress in Clinical Psychology, 1971, 69-90.

Lefkowitz, J., & Katz, M. L. Validity of exit interviews. Personnel Psychology, 1969, 22, 445-455.

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

Linn, R. L. Range restriction problems in the use of self-selected groups for test validation. Psychological Bulletin, 1968, 69, 69-73.

Linn, R. L. Pearson selection formulas: Implications for studies of predictive bias and estimates of educational effects in selected samples. Journal of Educational Measurement, 1983, 20, 1-15.

Linn, R. L., Harnish, D. L., & Dunbar, S. B. Corrections for range restriction: An empirical investigation of conditions resulting in conservative corrections. Journal of Applied Psychology, 1981, 66, 655-663.

MacKinney, A. C., & Wolins, L. Validity information exchange. Personnel Psychology, 1959, 12, 482-483.

MacKinney, A. C., & Wolins, L. Validity information exchange. Personnel Psychology, 1960, 13, 443-447.

Mansfield, R. S., & Busse, T. V. Meta-analysis of research: A rejoinder to Glass. Educational Researcher, 1977, 6, 3.

March, J. G., & Simon, H. A. Organizations. New York: John Wiley & Sons, Inc., 1958.

McDaniel, M. A., Hirsh, H. R., Schmidt, F. L., Raju, N. S., & Hunter, J. E. Interpreting the results of the meta-analytic research: A comment on Schmidt, Gooding, Noe, and Kirsch (1984). Personnel Psychology, 1986, 39, 141-148.

Meehl, P. E., & Hathaway, S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. Journal of Applied Psychology, 1946, 30, 525-564.

Meyer, H., & Cuomo, S. Who leaves? A study of background characteristics of engineers associated with turnover. Crotonville, N. Y.: General Electric Company, Behavioral Science Research, 1962.

Mitchell, T. R. Organizational behavior. In M. R. Rosenzweig & L. W. Porter (Eds.), Annual Review of Psychology (Vol. 30). Palo Alto: Annual Reviews Inc., 1979.

Mobley, W. H. Intermediate linkages in the relationship between job satisfaction and employee turnover. Journal of Applied Psychology, 1977, 62, 237-240.

Mobley, W. H. Employee turnover: Causes, consequences, and control. Reading, MA.: Addison-Wesley, 1982.

Mobley, W. H., Griffeth, R. W., Hand, H. H., & Meglino, B. M. Review and conceptual analysis of the employee turnover process. Psychological Bulletin, 1979, 86, 493-522.

Mowday, R. T., Porter, L. M., & Steers, R. M. Employee-organizational linkages: The psychology of commitment, absenteeism, and turnover. New York: Academic Press, 1982.

Mowday, R. T., Porter, L. M., & Stone, E. F. Employee characteristics as predictors of turnover among female clerical employees in two organizations. Journal of Vocational Behavior, 1978, 12, 321-332.

Mowday, R. T., & Spencer, D. G. The influence of task and personality characteristics on employee turnover and absenteeism incidents. Academy of Management Journal, 1981, 24, 634-642.

Mowday, R. T., Stone, E. F., & Porter, L. W. The interaction of personality and job scope in predicting turnover. Journal of Vocational Behavior, 1979, 15, 78-89.

Muchinsky, P. M., & Morrow, P. C. A multidisciplinary model of voluntary turnover. Journal of Vocational Behavior, 1980, 17, 263-290.

Muchinsky, P. M., & Tuttle, M. L. Employee turnover: An empirical and methodological assessment. Journal of Vocational Behavior, 1979, 14, 43-77.

Navran, L. A. A rationally derived MMPI scale to measure dependence. Journal of Consulting Psychology, 1954, 18, 192.

Orwin, R. G., & Cordray, D. S. Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. Psychological Bulletin, 1985, 97, 134-147.

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

Pearson, J. S. Psychometric correlates of emotional immaturity (Doctoral dissertation, University of Minnesota, 1954). Dissertation Abstracts, 1954, 14, 2129.

Pettman, B. O. Some factors influencing labour turnover: A review of the literature. Industrial Relations Journal, 1973, 4, 43-61.

Porter, L. W., & Steers, R. M. Organizational, work, and personal factors in employee turnover and absenteeism. Psychological Bulletin, 1973, 80, 151-176.

Price, J. L. The study of turnover. Ames: Iowa State University Press, 1977.

Price, J. L., & Mueller, C. W. A causal model of turnover for nurses. Academy of Management Journal, 1981, 24, 543-565.

Robbins, J. E., & King, D. C. Validity information exchange. Personnel Psychology, 1961, 14, 217-219.

Roe, R. A. The correction for restriction of range and the difference between intended and actual selection. Educational and Psychological Measurement, 1979, 39, 551-559.

Rosenthal, R. Combining results of independent studies. Psychological Bulletin, 1978, 85, 185-193.

Rosenthal, R. Meta-analytic procedures for social research. Beverly Hills, CA.: Sage Publications, 1984.

Rosenthal, R., & Rosnow, R. L. Essentials of behavioral research. New York: McGraw-Hill, 1984.

Sackett, P. R., Harris, M. H., & Orr, J. M. On seeking moderator variables in the meta-analysis of correlational data: A monte carlo investigation of statistical power and resistance to type I error. Journal of Applied Psychology, 1986, 71, 302-310.

SAS Institute Inc. SAS/STAT guide for personal computers, version 6 edition. Computer program manual Cary, NC.: SAS Institute Inc., 1985.

Schmidt, F. L., Gast-Rosenberg, I., & Hunter, J. E. Validity generalization results for computer programmers. Journal of Applied Psychology, 1980, 65, 643-661.

Schmidt, F. L., & Hunter, J. E. Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 1977, 62, 529-540.

Schmidt, F. L., & Hunter, J. E. Moderator research and the law of small numbers. Personnel Psychology, 1978, 31, 215-231.

Schmidt, F. L., & Hunter, J. E. A within setting empirical test of the situational specificity hypothesis in personnel selection. Personnel Psychology, 1984, 37, 317-326.

Schmidt, F. L., Hunter, J. E., & Caplan, J. R. Validity generalization results for two job groups in the petroleum industry. Journal of Applied Psychology, 1981, 66, 261-272.

Schmidt, F. L., Hunter, J. E., McKenzie, R., & Muldrow, T. The impact of valid selection procedures on workforce productivity. Journal of Applied Psychology, 1979, 64, 609-626.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. Task differences and validity of aptitude tests in selection: A red herring. Journal of Applied Psychology, 1981, 66, 161-185.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. Progress in validity generalization: Comments on Callender and Osburn and further developments. Journal of Applied Psychology, 1982, 67, 835-845.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Hirsh, H. R. Forty questions about validity generalization and meta-analysis. Personnel Psychology, 1985, 38, 697-798.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. Further tests of the Schmidt-Hunter Bayesian validity generalization procedure. Personnel Psychology, 1979, 32, 257-281.

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

Schmidt, F. L., Ocasio, B. P., Hillery, J. M., & Hunter, J. E. Further within-setting empirical tests of the situational specificity hypothesis in personnel selection. Personnel Psychology, 1985, 38, 509-524.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 1984, 37, 407-422.

Schuh, A. J. The predictability of employee tenure: A review of the literature. Personnel Psychology, 1967, 20, 133-152.

Schwab, D. P., & Packard, G. L. Response distortion on the Gordon Personal Inventory and the Gordon Personal Profile in a selection context: Some implications for predicting employee tenure. Journal of Applied Psychology, 1973, 58, 372-374.

Sinaiko, H. W. The Rosenzweig Picture-Frustration Study in the selection of department store section managers. Journal of Applied Psychology, 1949, 33, 36-42.

Steers, R. M., & Mowday, R. T. Employee turnover and post desision accommodation processes. In L. Cummings & Staw (Eds.), Research in organizational behavior (Vol. 3). Greenwich, CT.: JAI Press Inc., 1981.

Tenopyr, M. L., & Oeltjen, P. D. Personnel selection and classification. In M. R. Rosenzweig & L. W. Porter (Eds.), Annual Review of Psychology (Vol. 33). Palo Alto: Annual Reviews Inc., 1982.

Thorndike, R. L. Personnel selection: Test and measurement techniques. New York: John Wiley & Sons, Inc., 1949

Tversky, A., & Kahneman, D. Belief in the law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

Tydlaska, M., & Mengel, R. A scale for measuring work attitude for the MMPI. Journal of Applied Psychology, 1953, 37, 474-477.

Vincent, N. L., & Dugan, R. D. Validity information exchange. Personnel Psychology, 1951, 15, 223-225.

Vroom, V. Work and motivation. New York: John Wiley & Sons, Inc., 1964.

Wanous, J. P. Organizational entry: Recruitment, selection, and socialization of newcomers. Reading, MA.: Addison-Wesley, 1980.

Welsh, G. S. Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), Basic readings on the MMPI in psychology and medicine. Minneapolis: University of Minnesota Press, 1956.

Wickert, F. R. Turnover and employees' feelings of ego-involvement in the day-to-day operations of a company. Personnel Psychology, 1951, 4, 185-197.

Wolins, L. Responsibility for raw data. American Psychologist, 1962, 17, 657-658.

Wolpin, J., & Burke, R. J. Relationships between absenteeism and turnover: A function of the measures? Personnel Psychology, 1985, 38, 57-74.

VITA

John R. Burnett-Doering was born in Summit, New Jersey on June 20, 1953. He attended elementary schools in Wisconsin and was graduated from Brookfield East High School in June 1971. He subsequently attended the University of Wisconsin and received a Bachelor of Arts degree in 1975. Following the award of a Master of Science degree from the same university in Industrial/Experimental Psychology in 1977, he entered The Graduate School of The University of Tennessee, Knoxville. He received the Doctor of Philosophy degree with a major in Psychology in December 1987.

The author is currently employed by a consulting company in California specializing in personnel selection.