A META-ANALYTIC REVIEW OF THE EFFECTIVENESS OF PERSONNEL SELECTION PROCEDURES AND TRAINING INTERVENTIONS IN SALES OCCUPATIONS

by

SEONAID FARRELL

B.A. (Honours), Queen's University, 1994

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1999

© Seonaid Farrell, 1999
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date 31st August 1999
Abstract

Applied research on the effectiveness of personnel selection procedures and training interventions in sales occupations was examined by meta-analytic techniques. The literature review included 170 predictor-criterion effect sizes of selection procedures, and 12 effect sizes of training interventions. On average, composite-domain selection assessments predicted salesperson performance (multiple $R = .59$, $p < .05$), while the effects of single domain selection assessments and training interventions ($r = .27$, $p > .05$; $r = .16$, $p > .05$, respectively) were not statistically significant. Among the personnel selection procedures, subjective ratings of performance were predicted by personality assessments ($r = .20$, $p < .05$), biographical information inventories ($r = .20$, $p < .05$), and cognitive ability ($r = .32$, $p < .05$). While objective measures of performance were predicted by special purpose sales assessments ($r = .49$, $p < .05$). An analysis of the utility of selection procedures and training interventions based on the average effect sizes found in the present study revealed substantial dollar value gains to the company from the use of personnel selection procedures and training interventions in sales occupations.
# Table Of Contents

Abstract .................................................................................................................. ii  
List of Tables .......................................................................................................... iv  
Acknowledgements ............................................................................................... v

Introduction ........................................................................................................... 1

The Sales Occupation ......................................................................................... 1
Approaches to the Improvement of Sales Performance ....................................... 3
   Personnel Selection ......................................................................................... 3
   Training ......................................................................................................... 5

Purposes of the Present Study ............................................................................. 6

Measurement of Sales Performance ................................................................. 6
   Objective measurement of performance ...................................................... 6
   Subjective measurement of performance .................................................... 7
   Comparison of objective and subjective measures of performance ........... 8

Effectiveness of Selection Procedures and Training Interventions for Salespeople.. 8
   Effectiveness of predictors of salesperson performance ................................ 8
   Effectiveness of training interventions for salespeople ............................... 11

Method ............................................................................................................... 13

   Literature Search ......................................................................................... 13
   Criteria for Inclusion of Results .................................................................. 13
   Coding ......................................................................................................... 14
   Meta-Analytic Procedures .......................................................................... 15

Results and Discussion ..................................................................................... 20

   Sample Description .................................................................................... 20
   Meta-Analytic Results and Discussion ....................................................... 21
   Analysis of the Utility of Selection Procedures and Training Interventions .... 23

Summary ............................................................................................................ 27

References ......................................................................................................... 30
List of Tables

Table 1  Correlation Coefficients by Single Domain Predictor, Composite Domain Predictor, or Training Intervention for the Combined Criterion ................................................................. 41

Table 2  Validity Coefficients by Content of Single Domain Selection Assessment for the Combined Criterion ........................................................................................................ 42

Table 3  Validity Coefficients by Content of Single Domain Selection Assessment for the Subjective Criterion ........................................................................................................ 43

Table 4  Validity Coefficients by Content of Single Domain Selection Assessment for the Objective Criterion ........................................................................................................ 44

Table 5  Relative Utility Estimates of Valid Selection Procedures and Effective Training Interventions ................................................................................................................. 45
Acknowledgements

I am grateful to many people who supported the achievement of this research. I am indebted to my committee members for their encouragement and valuable insights. In particular Dr. Ralph Hakstian, who developed my enjoyment of research by challenging and inspiring me. I appreciate my close friends who gave their support and encouragement generously. Finally, I am grateful for a good family who support my endeavors unselfishly, and continually demonstrate the value of determination and personal achievement.
Personnel Selection and Training in Sales

A Meta-analytic Review of the Effectiveness of Personnel Selection Procedures and Training Interventions in Sales Occupations

The Sales Occupation

The sales occupation is an important area of study because of the strong influence it exerts on the economy, the wide range in job performance of salespeople, and the potential gains that can be achieved by sales management techniques. There were 14,633,000 salespeople in the United States in 1996, making up 11% of all jobs (Bureau of Labor Statistics, 1998). The number of sales and marketing jobs is expected to rise by 15.5% by 2006; a rate of increase which is greater than the average of all occupations (14%). The total compensation of salespeople that includes salary, commission, and incentives is substantial. In the United States, the average compensation of a sales representative was $63,200 in 1998 (Pricewaterhouse Coopers, 1998). The predominance of salespeople in the workforce, the expected rate of growth of sales jobs, and their level of compensation suggest that the job performance of salespeople is an important influence on the economy.

The great variability in the job performance among salespeople makes it a compelling area of study. Hunter, Schmidt, and Judiesch (1990) examined the standard deviation of employee output as a percentage of mean output for a variety of occupations and found that the variability was greatest for salespeople. These authors found that the average standard deviation of output, as a percentage of mean output, was 120% for life insurance, and 48% for non-insurance sales, compared with 48% for high complexity non-sales jobs, 32% for medium complexity non-sales jobs, and 19% for low complexity non-sales jobs. The increased variability in output for sales jobs can be expected to result in potentially greater gains in utility by human resource management activities, because the greater the variability (in dollar value to the organization) the greater the average gain in dollars from successful interventions. Hunter et al. (1990) suggested that the large variability in sales output may be due to the multiplicative effect of specific traits and abilities required of salespeople. These authors suggested that the average percentage increase in output from improved selection is approximately 2.5 times greater in sales jobs than in low complexity non-sales jobs.
Another unique characteristic of the sales occupation is the unusual features that are required to perform the job. Salespeople must work autonomously (Churchill, Ford, & Walker, 1997), relying on their own initiative and determination. They must also endure a high amount of rejection in proportion to successes, requiring persistence and achievement striving. Salespeople must solve the different needs of many clients in a frequently changing marketplace, requiring adaptability and problem solving skills.

A meta-analysis of the predictors of sales performance from a marketing perspective (Churchill, Ford, Hartley, & Walker, 1985) delineated six variables to account for individual differences in job performance. The six variables include: (1) aptitude, which are abilities and personal traits related to the performance of the job; (2) personal characteristics, which are physical traits, and family, education, and work background experiences; (3) skill level, which refers to the extent of learned proficiencies at performing job activities; (4) role perceptions, which are the employee's perceptions of job demands and expectations from others; (5) motivation, which is the desire to expend effort on specific job activities; and (6) organizational and environmental factors, which include sales potential of territory, salesperson's autonomy, and the company's competitive strength (Churchill et al., 1997, pp. 327-335). Churchill et al. (p. 367) suggested that among the variety of human resource management activities, recruitment and selection influence aptitude and personal characteristics, whereas training and supervision influence skill levels and role perceptions.

There are a variety of jobs included in the sales occupation. The various sales jobs can be classified as involving retail or industrial selling (Churchill et al., 1997). Retail selling refers to selling goods and services directly to the consumer of these products. The products are intended for the consumer's personal use. Industrial selling pertains to the sale of goods and services to wholesale customers. Wholesale customers can be: (1) resellers, who sell the product to a retailer, (2) business users, who use the product to produce another product, or (3) institutions, who use the product in their business operation. There are similarities and differences in the selling function of retail and industrial sales jobs. While retail sales requires many of the same skills and personal characteristics as industrial sales, industrial sales is often viewed as unique because the products are often more expensive and technically complex.
Approaches to the Improvement of Sales Performance

The importance of sales to the economy, great variability in job performance of sales people, and unique characteristics of the occupation are compelling reasons for the study of salespeople. One aspect of the study of this occupation is the factors that influence improvement in salesperson performance, two of which are personnel selection and training interventions.

Personnel Selection

Personnel selection involves the use of systematic, rational, valid assessment methods to select the best candidates for employment. The objective is to predict future job performance by assessing the applicant’s characteristics, capacity to function, and interest in the job. The use of assessment procedures for selection assumes that employees who are selected by valid procedures will be more productive than employees selected by inferior procedures. This is especially true in sales occupations if we can assume that individual differences account for much of the large variability in job performance.

The validity of assessment procedures for selection serves the needs of the salesperson as well. A salesperson who is selected to work at a company should have the requisite abilities to perform the job, or s/he will suffer frustration and failure. Alternatively, the applicant who is not hired for the sales job, but who does indeed have potential, is unfairly treated if this potential is not recognized and the decision is not to hire.

There are many types of selection procedures in use for the selection of salespeople. In the present study their content is categorized into six distinct groups that include:

1. Job skill/knowledge assessments, which measure an applicant’s knowledge about sales-related skills and information, such as personal selling, planning, time management, and market potential. In most cases, job skill/knowledge is gained as the result of specific work experience in sales occupations.

2. Personality inventories, which measure enduring personal dispositions that reflect an individual’s consistent reactions to situations. The personality dimensions most often studied among salespeople include dominance, empathy, sociability, responsibility, self-esteem, need for achievement, and need for power.
3. Biographical information, which is demographic, educational, and occupational information about an applicant that is obtained by asking questions related to the applicant's family, work, and lifestyle background experiences. Biographical information most often obtained in the selection of salespeople includes level of education, work experience, activities and preferences, and marital status.

4. Cognitive ability tests, which measure general mental ability and specific mental processes and abilities related to sales ability, such as verbal comprehension, numerical ability, and visual speed and accuracy.

5. Job specific aptitude assessments, which measure personal characteristics and abilities thought to be related to the performance of specific sales tasks, such as ease of interpersonal interactions and business judgment. Job specific aptitude differs from job skill/knowledge in that aptitude is not necessarily gained by specific sales experience and it does not measure concrete knowledge about sales.

6. Special purpose sales assessments are paper and pencil tests that measure aptitudes and personal qualities related specifically to success in sales occupations. They differ from the previously mentioned selection procedures in that special purpose tests are developed only for the prediction of sales performance dimensions, rather than general traits.

The term “selection procedure” (sometimes “assessment procedure” in what follows) will be used to refer to any of the systematic personnel assessment-selection methods identified in this categorization. More specifically, “single domain predictor” will be used to refer to any one of the selection procedures when used individually to predict sales performance, and “composite domain predictor” will be used to refer to any optimally- or unit-weighted combination of two or more of the selection procedures.

Personnel selection procedures are widely used in organizations. In a review of the use of selection procedures for sales occupations among a sample of 121 firms (Dubinsky & Barry, 1982), 90% of the firms reported use of interviews, 66% reported use of application blanks and reference checks, 33% of larger firms reported use of psychological tests, and 22% of smaller firms reported use of psychological tests.

Historically, the dominant belief has been that personnel selection validity is specific to particular
situations and individual jobs. A review of the raw validation results across many occupations found high variability in results among studies even when jobs and tests appeared to be similar (Ghiselli, 1966, 1973). Findings similar to this supported the theory of situational specificity which viewed the generalization of validity results across similar jobs and personnel selection procedures as essentially impossible (Albright, Glennon, & Smith, 1963, p. 18; Ghiselli, 1966, p. 28). The theory of situational specificity remained dominant until evidence mounted to suggest that the variability in validity study results within job-test combinations was due to statistical and measurement artifacts, rather than real differences between jobs (Schmidt, Berner, & Hunter, 1973; Schmidt, & Hunter, 1977; Schmidt, Hunter, Pearlman, & Shane, 1979; Schmidt, Hunter, & Urry, 1976). The variability among validity results was largely accounted for by error variance such as sampling error caused by small sample sizes (Schmidt, & Hunter, 1998), computational and typographical errors, differences between studies in criterion reliability, differences between studies in amount and kind of criterion contamination and deficiency, and differences between studies in degree of range restriction (Schmidt & Hunter, 1977). Quantitative techniques were developed to combine validity estimates across studies and correct for the statistical and measurement artifacts (Hunter & Schmidt, 1980; Hunter, Schmidt, & Jackson, 1982).

Training

Training can be defined as a planned program within the organization which endeavors to bring about relatively permanent changes in employee knowledge, skills, attitudes, and behavior (Wexley & Latham, 1981). The objective of training programs is to improve the performance of salespeople. It is assumed that improved performance of salespeople will lead to increased profits to the company. Training may improve performance by increasing productivity, improving morale, reducing turnover, improving customer relations, and improving management of time and territory (Churchill et al., 1997).

The content of training for salespeople has been categorized into four areas that are described by Hopkins (1978). These areas include training in (a) product knowledge, (b) company orientation, (c) market/industry orientation, and (d) selling techniques and related topics. It is the fourth area of training, selling techniques and related topics, that is of interest in this thesis. The related topics that are considered
include: human relations, self-awareness, problem solving and decision making, and motivation/values.

When the term "training intervention" is used in what follows, it should be taken to mean training in selling techniques or a related topic identified in the previous sentence.

In the majority of companies, most salespeople will receive some form of company or product training to enable them to perform their job duties at a minimally acceptable level. This training, however, is specific to the company, and its effectiveness may be irrelevant to training issues in other companies. Assuming that salespeople have a base of job-, company-, and product-knowledge, training of employees in selling techniques and related topics can be undertaken. Training in this area is interesting because the knowledge and skills that a salesperson learns at this level can be generalized to different sales situations at different companies. The effectiveness of the sales training can be compared across companies because the training is not specific only to the company that sponsors the training program.

**Purposes of the Present Study**

There were two purposes of the present study. The first was to describe—from a reading of the applied research literature—the extent to which selection procedures and training interventions have improved the performance of salespeople. The second purpose was to compare whether it is selection procedures or training interventions that has contributed more to the improvement of the performance of salespeople.

**Measurement of Sales Performance**

Salesperson performance is more amenable to measurement than is the performance of other occupations because of the concrete and quantifiable nature of the output. Campbell (1991) defined performance as the behavior that is relevant to the organization’s goals. Where behavior is unobservable, it must be inferred from its results, such as sales volume. 

*Objective measurement of performance*

Sales volume is an objective index of salesperson performance. It is the most common outcome measure of salesperson performance (Weitz, 1981) and that most frequently used in academic research (Comer & Dubinsky, 1985). Objective measures of performance are a function of the individual
Objective measures are unique in that they require no abstraction or synthesis by the evaluator (Landy & Farr, 1993).

The drawbacks to using an objective measure of salesperson performance such as sales volume include: (1) the extent to which it is influenced by external factors beyond the salesperson's control, such as market fluctuations and type of product sold, and (2) the information about performance that is not measured, such as the strength of the relationship with customers and the commitment to group performance over individual performance. Objective measures may be adjusted for contamination by external factors, by comparing output with a normative reference for the sample, or by reporting sales volume as a proportion of sales quota (Borman, 1994), since sales quota is an index that is established with a consideration of external factors. Borman notes, however, that this practice may be suboptimal because information about particular subgroups of salespeople is lost when large groups of salespeople are used as a normative reference.

In a study examining the stability of salesperson performance measures over time, Levy and Sharma (1993) found that objective performance measured over six months was significantly related to objective performance over one year; however, objective performance measured over three months was not significantly related. This suggests that reliable measurement of objective performance requires at least a six-month period of assessment.

Subjective measurement of performance

Subjective measures of performance, such as a supervisor's ratings, require an individual to make a judgment about the performance level of another person or of him/herself. The judgment may be based on knowledge of results, such as sales volume, as well as inferences about less quantifiable aspects of the salesperson's performance inferred from his/her behavior, such as communication skills. In this way, subjective measures may include more information about performance and include information that is under the control of the salesperson. Levy and Sharma (1993) found that subjective measures (managerial- and self-ratings) of performance, were upwardly biased and restricted in range. These authors also found that, between the subjective measures examined, managerial ratings were positively related to the length of
employment of the salespeople, and self-ratings were related to objective performance, whereas managerial ratings were not.

A major drawback of subjective measures is the error that results when the supervisor makes judgments about performance. These judgments, usually recorded on some form of rating instrument, may be influenced by irrelevant factors, such as selective attention to behavior, impression management by the salesperson, inadequate sampling of behaviors on the rating instrument, and biases of the supervisor when rating performance (Bush, Bush, Ortinau, & Hair, 1990).

Comparison of objective and subjective measures of performance

Across all occupations there is evidence of only a moderate relationship between subjective and objective measures of performance (Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995; Heneman, 1986). In a meta-analysis of salesperson performance, in particular, Bommer et al. (1995) found that the average correlation coefficient between subjective and objective measures of performance was $r = .41$. It may be that managerial ratings are measuring non-selling aspects of the sales job and should be used in combination with objective ratings to give the most information about job performance. Overall, there is agreement that subjective and objective measures of performance should not be used interchangeably.

Effectiveness of Selection Procedures and Training Interventions for Salespeople

A disparity exists between the amount of information available that describes the effectiveness of selection procedures and that available dealing with training interventions for salespeople. Individual studies and meta-analyses have examined single and composite domain predictors of salesperson performance—predictors that can be applied as selection procedures. On the other hand, few individual studies of the effectiveness of training interventions for salespeople are available.

Effectiveness of predictors of salesperson performance

Several meta-analyses have examined the effects of predictors of salesperson performance. Vinchur, Schippmann, Switzer, and Roth (1998) examined the relationship between a variety of predictors and performance in 129 studies and found biographical information and sales ability (i.e. job skill/knowledge assessments) were relatively strong predictors of subjective ratings ($r = .52$ and $r = .45$;
correlations for this study were corrected for range restriction and criterion unreliability). They also found special purpose sales assessments predicted subjective ratings \( (r = .45) \) and objective sales \( (r = .37) \).

Among the personality dimensions, they found that Potency (a subdimension of Extraversion) predicted subjective ratings \( (r = .28) \) and objective sales \( (r = .26) \), and that Achievement (a subdimension of Conscientiousness) predicted subjective ratings \( (r = .25) \) and objective sales \( (r = .41) \). Cognitive ability was identified as a unique predictor in that, although it bore a moderate relationship to subjective ratings \( (r = .31) \), it was unrelated to objective sales \( (r = .04) \).

A meta-analysis of predictors of sales performance by Ford, Walker, Churchill, and Hartley (1987) from a sales and marketing perspective, examined the relationships between biographical variables (demographic, background and work experiences, and family status and lifestyle) and psychological variables (aptitudes, mental abilities, personality traits, and learned skills) with sales performance. These authors found that biographical variables bore the strongest relationship to objective performance \( (r = .46) \).

Ford et al. (1987) found that among the aptitude dimensions examined, cognitive ability was the best predictor of performance \( (r = .26) \) followed by sales aptitude \( (r = .19) \). Ford et al. suggested that dominance \( (r = .16) \) and self-esteem \( (r = .14) \) were the most important personality dimensions. They suggested that the variation among other personality dimensions was due to sampling error.

There is some evidence that a difference in validity appears when comparing higher-level sales representatives with lower-level sales clerks. Ghiselli (1973) found intelligence tests predicted performance of salespeople \( (r = .30) \), but not sales clerks \( (r = -.10) \). Personality tests were found to predict performance of salespeople \( (r = .40) \) and sales clerks \( (r = .30) \).

The individual predictors of salesperson performance differ in the extent of their intercorrelation. Cognitive ability is moderately related to biographical information (Gray & Rosen, 1956; Johnson, 1940 Vinchur et al., 1998). Cognitive ability is also moderately related to sales ability (Cotham, 1968; Gray & Rosen, 1956; Vinchur et al.).

Individual predictors of salesperson performance also differ in the extent to which they predict either subjective or objective measures of performance. Ford et al. (1987) found that cognitive ability,
sales aptitude and self-esteem were better predictors of subjective data; whereas biographical information was a better predictor of objective sales data. These authors suggested that objective sales data may be influenced by external uncontrollable factors, and fail to take into account the strength of the relationship between salesperson and customer.

The single domain predictors of performance vary among themselves with respect to their cost and ease of development and administration, as well as their validity. In a meta-analysis of alternative predictors of performance, Schmidt and Hunter (1998) noted that cognitive ability has the lowest cost of administration and was the best predictor of job-related learning, acquisition, and transfer to actual job performance (Schmidt & Hunter, 1982; Schmidt, Hunter, & Outerbridge, 1986). Alternatively, job knowledge tests, work sample measures, and some structured interviews require more time to develop and assume a base of specific knowledge of the job. Biographical information inventories were found to be difficult and time consuming to construct.

Where there is great variability among effect sizes of the same predictor-criterion relationship, moderator variables may account for some of the variability. Churchill et al. (1985) suggested that customer and product type moderate the size of effect of the determinants of salesperson performance. The performance measure used may also moderate the effect (Ford et al., 1987). Ford et al. found that biographical information was a stronger predictor for salespeople who sell or provide services to individual customers when measured objectively. Alternatively, Ford et al. found that aptitudes were stronger predictors for salespeople who sell to institutions when measured subjectively. The personality dimension of dominance was found to be a stronger predictor for salespeople selling services, rather than products, to individuals. Overall, Ford et al. concluded that the best predictors for salespeople selling industrial goods to institutional customers included: self-esteem, sales aptitude, and cognitive ability. The best predictors for salespeople selling consumer goods to individual customers included: biographical information (personal history, marital and family status) and interpersonal skills.

Predictors of salesperson performance have also been studied in combination. Schmidt and Hunter (1998) found the best combinations of predictors of performance in all occupations to be general mental
ability (GMA) and work sample \((multiple R = .63)\), GMA and integrity test \((multiple R = .65)\), and GMA and structured interview \((multiple R = .63)\). Considering only the sales occupation, Vinchur et al. (1998) found the best combination of predictors of subjective performance to be sales ability and cognitive ability \((multiple R = .36)\). They found the best predictor of objective data to be the personality dimension, achievement, \((r = .23)\) and noted that other predictors added little to the prediction.

Effectiveness of training interventions for salespeople

Training is well accepted as an effective means of enhancing productivity at work (Burke & Day, 1986; Guzzo, Jette, & Katzell, 1985; Russell, Terborg, & Powers, 1985). However, the evaluation of training interventions in applied settings is deficient (Bunker & Cohen, 1978; Kerr, 1975; Sackett & Mullen, 1993). Often the only evaluation that is completed is the subjective response of employees to the training program.

There are several reasons that account for limited evaluation of training interventions (Dipboye, 1997). Even when resources are made available for evaluation it can be difficult to establish comparison groups. The perceptions of unfairness by those who remain untrained as a comparison group could threaten the internal validity of the research design. Also, trainers within the organization may oppose the objective, empirical evaluation of training programs. Political support for the training intervention from senior executives may reduce the motivation to objectively evaluate the program.

Despite some attempts to evaluate training interventions, scant empirical evidence exists reporting their effectiveness. Meyer and Raich (1983) found behavior modelling training had a direct effect on the number of sales achieved by the participants. Ivancevich (1974) found that goal setting training had an effect on performance for up to 18 months, after which time the effect disappeared. He suggested that post-training improvements would be observed only after three months. Ralis and O'Brien (1987) found similar positive effects on sales performance from goal-setting training.

Meta-analytic Review of Personnel Selection Procedures and Training Interventions

The goal of the present study is to improve the job performance of salespeople by increasing our understanding of the extent to which personnel selection procedures and training interventions influence
sales performance. Across the variety of personnel selection procedures, training interventions, and measures of sales performance, meta-analytic techniques were used to reduce the effect of statistical and measurement artifacts on the true effects of these human resource management activities. This meta-analytic review advances our understanding of how to improve sales performance. First, it examines the effects of both personnel selection and training interventions in the same study. Second, it refines previous meta-analyses of the effectiveness of personnel selection procedures by including only independent effect sizes, disaggregating effect sizes by type of performance measure used, using more conservative corrections for statistical artifacts than in a previous meta-analysis of predictors of sales performance (Vinchur et al., 1998), and examining the homogeneity of the effects of personnel selection and training.
Method

Literture Search

Several approaches were used to obtain published and unpublished articles, dissertations, and book chapters that examined the effectiveness of selection procedures and training interventions. Computer-based searches of published empirical studies and dissertations were conducted by accessing PsycINFO (1967-1999), ABI/Inform (1971-1990), Canadian Business and Current Affairs (1982-1998), and ERIC (1966-1998). Manual searches of the literature were accomplished through reference lists accompanying several previous reviews of the literature (Vinchur et al., 1998; Churchill et al., 1985; Ford et al., 1987), and through individual studies found in the process of the search, along with relevant books and book chapters.

These approaches yielded a paucity of published studies of the effectiveness of training interventions in sales. An extended search was, therefore, directed to procuring more studies of training effectiveness. Three industrial psychologists were contacted personally, and asked for unpublished applied research. A group of 45 salespeople and sales managers in an executive sales management diploma course at the University of British Columbia were approached and asked to provide the names of companies with “best practices” in salesforce training. Based on the information supplied by 28 respondents, the training departments of five companies were contacted for studies examining the effectiveness of their sales training interventions.

The “file drawer” problem (Rosenthal, 1979) influences meta-analyses, in that unpublished studies not yielding statistically significant results remain in the “file-drawers” of researchers, and are not included in the sample of studies examined by meta-analysis. An attempt was made to address this problem through the approaches previously described.

Criteria for the Inclusion of Results

Studies examining selection procedures. To be included in the present investigation, individual studies of selection procedures had to meet two criteria. First, studies had to present an effect size (correlation coefficient between predictor, or composite of predictors, and criterion measure of salesperson
performance) of the relationship between an individual or set of personnel selection procedures and one or more measures of sales performance. Studies with dependent variables of training proficiency, turnover, or absenteeism were not included, to ensure consistency in the criterion. Second, studies had to use either job applicants or incumbents (i.e., not students) as subjects.

To be included in the meta-analysis, the effect sizes of the individual studies had to be independent of each other, where subjects did not contribute to more than one of the same predictor-criterion relationship. Where studies reported more than one correlation coefficient for the same predictor-criterion relationship with the same subjects, the effect sizes were averaged.

*Studies examining training interventions.* There were four criteria for studies of training interventions. First, the study had to examine a training intervention related to selling effectiveness. Such training programs included topics in human relations, self-awareness, problem solving, motivation, and the sales process. Training programs related to product knowledge, company orientation, or market/industry orientation were excluded. Second, the study had to include a measure of performance indexing a salesperson's behavior or business results such as sales productivity. Thus, measures of salespeople's subjective responses to the training program, or measurements of amount of information learned were excluded. Third, the study had to include a comparison group that did not receive the specific training intervention. The comparison group could receive another intervention, such as a placebo, or no training at all. Fourth, the subjects of the study had to be job incumbents rather than student subjects.

*Coding*

Coding of specific study variables (customer type, product type, job level, content of training or selection procedure, criterion type, and effect size information) was done independently by the author. The most widely used estimate of reliability of coding was achieved by calculating the agreement rate (Orwin, 1994, p.147) between codes assigned by the author and a second coder on three studies of selection procedures and two studies of training interventions. The agreement rate, a ratio of number of observations agreed upon to the total number of observations, was found to be 90% on the studies coded.
Meta-Analytic Procedures

The effect sizes in the individual studies of selection procedures were standard Pearson correlation coefficients between selection assessment and criterion scores—validity coefficients. The results of individual studies of training interventions were reported as either standardized mean differences on the criterion performance variable, or as t-test statistics for comparing the mean difference between the experimental group who received the training program and the control group who did not receive the training program. In order to compare the effect sizes with selection procedures, standardized mean differences for each \( i \)th study, \( d_i \), were converted to Pearson correlation coefficients by the following equation given by Rosenthal (1994, p. 239):

\[
    r = \frac{d_i^2}{d_i^2 + 4}
\]

This formula is appropriate for use when the populations of those given the treatment (in this case the training program) and those not given the treatment are regarded as essentially equal in size (Rosenthal, 1994, p. 239). T-test statistics were converted to Pearson correlation coefficients by the following equation given by Rosenthal (1994, p. 239):

\[
    r = \frac{t^2}{t^2 + df}
\]

where \( df \) is the degrees of freedom, \( (n_e + n_c - 2) \), where \( n_e \) is the number of subjects in the experimental training condition, and where \( n_c \) is the number of subjects in the control condition.

Corrections for statistical artifacts. The correlation coefficients for selection procedures were individually corrected for range restriction in the independent variable and criterion unreliability of subjective performance ratings prior to being averaged. The correlation coefficients for training interventions were individually corrected for criterion unreliability of subjective performance ratings only. In each personnel-selection study, correlation coefficients were corrected for range restriction because low scores on the independent variable were systematically excluded from the analysis; low scores would be
relevant in the intended population of job applicants. In most selection studies, information necessary to correct for range restriction—ratio of standard deviation of population of scores to standard deviation of sample of scores—is unavailable. Given the lack of information in the studies necessary to correct for range restriction, analyses were completed with an average estimate of range restriction provided by Schmidt and Hunter (1977) that has been accepted as accurate (Alexander, Carson, Alliger, & Cronshaw, 1989; Vinchur et al., 1998). The correlation coefficient corrected for range restriction was found by the following equation given by Glass and Hopkins (1996, p. 122):

\[
r_{rr} = \frac{r^2 \left( \frac{\sigma_{pop}}{\sigma_{sample}} \right)^2}{1 + r^2 \left( \frac{\sigma_{pop}}{\sigma_{sample}} \right)^2 - r^2},
\]

where \( r_{rr} \) is the correlation coefficient corrected for range restriction, \( r \) is the correlation coefficient uncorrected for range restriction, \( \sigma_{pop} \) is the standard deviation of the scores on the independent variable in the population, and \( \sigma_{sample} \) is the standard deviation of scores on the independent variable in the sample.

Corrections for criterion unreliability were made in the range-restriction-corrected correlation coefficients in selection studies and uncorrected correlation coefficients in training intervention studies, based on reliability estimates of the criterion given in the individual studies. Where criterion reliability information was unavailable in the study, an estimate of the average coefficient alpha—an estimate of intra-rater reliability—for supervisory ratings of overall job performance was used. This average was drawn from the distribution of intra-rater reliability coefficients provided by Viswesvaran, Ones, and Schmidt (1996). The average coefficient alpha for supervisory ratings of overall job performance provided by this distribution is .86 (\( SD = .1433 \)). Correlation coefficients for studies in which the criterion was a subjective measure of performance were corrected by standard partial disattenuation, i.e. dividing the correlation coefficient by the square root of the criterion reliability coefficient. Objective measures of performance were not corrected, because it was expected that there would be no systematic measurement error in the chosen criterion. The present study differs from the previous meta-analysis of predictors of salesperson performance (Vinchur et al., 1998), in that the previous meta-analysis reported the inter-rater
reliability estimate of .52, rather than the intra-rater reliability estimate of .86 in the present study, in correction for criterion unreliability. It was expected that this difference would lead to considerably lower disattenuated validity coefficients in the present study because of much lesser correction for criterion unreliability.

The only effect size corrections made in the present study were for unreliability in subjective measures of the criterion and range restriction of the independent variable in personnel selection studies. Corrections for predictor unreliability were not made because the influence of this unreliability is relevant to the use of these tools in personnel selection. Hunter and Schmidt (1994) provided means of correcting for the following additional statistical artifacts: artificial dichotomization of the continuous dependent and independent variables split into proportions, imperfect construct validity of the independent variable, range restriction in the dependent variable, bias in the correlation coefficient, and study-caused variation. These additional corrections were not made in the present study because of the lack of evidence that they systematically influenced the results of the individual studies, lack of information on artifacts in the individual studies, and also, because of Rosenthal’s (1994, p. 240) caution against over-correcting for the statistical artifacts.

**Combining correlation coefficients.** In order to assess the relationship between various selection procedures and performance, and to compare selection procedures with training interventions, the weighted mean correlation coefficient for all predictor-criterion relationships, and for all training effects were calculated. Corrected correlation coefficients for each study were first converted to a standard metric by Fisher’s Z-transformation, \( Z_r \). The Fisher’s Z-transformation was used because it has a normal sampling distribution irrespective of the population correlation coefficient or the sample size of each effect size (Glass & Hopkins, 1996). The weighted mean transformed correlation coefficient was found by assigning weights to the studies corresponding to their degrees of freedom and calculating the weighted average \( \bar{Z}_r \), by standard formulas (see, e.g., Shadish & Haddock, 1994), as:
where the weight \( w_i \) assigned to each study is \((n_i - 3)\). The weighted mean \( \overline{Z}_r' \) was then converted back to \( r \) by referring to standard tabled values of \( Z_r \) and \( r \).

The sample estimate of the variance component of effect sizes, \( s^2(Z_r) \), was estimated by the standard formula for the sample variance, which in the present context is:

\[
s^2(Z_r) = \frac{\sum_{i=1}^{k} (Z_{i}^2) - \left( \frac{\sum_{i=1}^{k} Z_i}{k} \right)^2}{k-1},
\]

where \( Z_i \) is the mean standardized effect size for the \( i \)th study, and \( k \) is the number of studies contributing to the overall effect size.

The expected value of \( s^2(Z_r) \) was found by the following equation given by Shadish and Haddock (1994):

\[
\hat{\sigma}_{Z_r}^2 = \left( s^2(Z_r) - \left( \frac{1}{k} \sum_v \right) \right)
\]

where \( v_i \) is the conditional variance of \( Z_i \) estimated by:

\[
v_i = \frac{1}{(n_i - 3)},
\]

and \( n_i \) is the sample size of the \( i \)th study.

The expected value of the sample estimate of the variance component, \( \hat{\sigma}_{Z_r}^2 \), was corrected for sampling error due to small sample sizes, consistent with the literature in integrating results in personnel selection (Hedges, 1983; Hunter, Schmidt, & Jackson, 1982; Schmidt & Hunter, 1977). The corrected variance component was found by the following equation given by Hunter et al. (1982, p. 44):
\[ \hat{\sigma}^2_{z_p} = \hat{\sigma}^2_{z_r} - \hat{\sigma}^2_e, \]

where \( \hat{\sigma}^2_{z_p} \) is the variance component corrected for sampling error, \( \hat{\sigma}^2_{z_r} \) is the expected value of the variance component uncorrected for sampling error, and \( \hat{\sigma}^2_e \) is the variance component of sampling error.

The value of \( \hat{\sigma}^2_e \) is estimated by the following equation given by Hunter et al. (1982, p. 70):

\[ \hat{\sigma}^2_e = \bar{a}^2 \frac{k(1 - \bar{\rho}^2)^2}{N}, \]

where \( k \) is the total number of studies, \( N \) is the total number of subjects, \( \bar{\rho} \) is the average correlation coefficient across studies, and \( \bar{a} \) is the average ratio of corrected correlation coefficients (corrected for range restriction and criterion unreliability) to uncorrected coefficients across all of the studies, estimated by the following equation:

\[ \bar{a} = 1/k \sum_{r} \frac{r_c}{r}, \]

where \( r_c \) is the correlation coefficient corrected for all the relevant statistical artifacts, and \( r \) is the uncorrected correlation coefficient.

The upper and lower limits of the 95% confidence interval for each weighted mean of the correlation coefficient, \( \bar{Z}' \), were found by multiplying the standard error of \( Z_r \) corrected for sampling error, \( \hat{\sigma}^2_{Z_p} \), by the critical value of 1.96 (for \( \alpha = .05 \)), adding and subtracting the resulting product from \( \bar{Z}', \) and converting the endpoints of the confidence interval back to \( rs \) by referring to standard tabled values of \( Z \) and \( r \). There is a probability of .95 that intervals so constructed will capture the true mean correlation coefficient. The confidence interval also illustrates the degree of precision in the estimate of the true mean correlation coefficient.

*Statistical test of the assumption of homogeneity of effect.* The equation for the weighted mean
correlation coefficient assumes that the studies share a common population effect size (Shadish & Haddock, 1994, p. 266). This assumption was tested by the following equation for the homogeneity test statistic, $Q$, given by Shadish & Haddock (1994, p. 266):

$$Q = \sum [(Z_r - \bar{Z}_r)^2 / (v_i)],$$

where $Z_r$ is the $Z$-transformed correlation coefficient and $v_i$ is the conditional variance of $Z_r$ defined above. When $Q$ exceeds the critical value of chi-square with $k - 1$ degrees of freedom, the conclusion is that the variability in effect sizes is significantly greater than would be expected by chance alone, if all studies shared a common population effect size. In this case, the weighted mean correlation coefficient would not be interpreted, but rather it would be taken as a description of the mean of observed effect sizes.

Results and Discussion

Sample Description

The exhaustive search of the literature produced 170 independent predictor-criterion effect sizes of selection procedures from 86 different samples in 59 studies, and 12 independent effect sizes of training interventions from 11 studies. Independent effect sizes were achieved by ensuring that no subject appeared in more than one of the same predictor-criterion relationship.

Among the studies of selection procedures, 58 of the 59 were published articles, with one unpublished doctoral dissertation. Among the studies of training interventions, 7 of the 12 were published articles; the remaining 5 were unpublished doctoral dissertations. The studies were performed during the period 1953 to 1997, with the exclusion of 1971-1976 during which time no study in the sample appeared.

Among the selection studies, the research was carried out with 34 samples of industrial salespeople, 26 samples of retail salespeople, and 26 with insufficient information to code for this variable. These salespeople worked in 26 companies that provide consumer goods, 20 companies that provide industrial goods, 37 that provide services, and 3 whose product is unknown. Among the training studies, the research was carried out with 3 samples of salespeople who sell to institutions, 7 who sell to individuals, 1 selling to both institutions and individuals, and 1 with insufficient information on this
variable. The research was done in 7 companies that provide consumer goods, 1 that provides industrial goods, 3 that provide services, and 1 with insufficient information. The sample size of the studies ranged from 16 subjects to 16,230 in selection procedures, with a median sample size of 88. The studies of training interventions ranged in size from 17 subjects to 1,017 with a median sample size of 105.

**Meta-Analytic Results and Discussion**

The overall results of the meta-analysis are presented in Table 1 for the combined criterion (average effect size across subjective and/or objective performance assessments for each independent predictor-criterion relationship). The average corrected correlation coefficients are presented for the single domain predictors of salesperson performance, composite domain predictors of salesperson performance, and training interventions in sales effectiveness. With reference to personnel-selection studies, in what follows criterion-related validity coefficients have been corrected for range restriction in the independent variable and attenuation in the criterion where the criterion was subjective ratings. With reference to training intervention studies, in what follows correlation coefficients for training interventions have been corrected for criterion unreliability where the criterion was subjective ratings. On average, composite-domain selection assessments predicted salesperson performance with a corrected multiple correlation of .59 (p < .05). The average corrected correlation coefficients of single-domain selection assessments (r = .27, p > .05) and training interventions (r = .16, p > .05) were not statistically significant. These results suggest that the job performance of a salesforce can be improved by the use of a composite domain assessment for personnel selection. The non-significant results for single-domain selection assessment and training interventions suggest that we cannot say with 95% confidence that these approaches are effective in salesforce improvement.

In all of the analyses in the present study the homogeneity test statistic, $Q$, exceeded the critical value. This result suggests that the observed variance in effect sizes is greater than what would be expected by chance if all studies shared a common population effect size. When $Q$ is rejected, Shadish and Haddock (1994, p. 266) suggested disaggregating the study effect sizes by breaking the studies into
smaller groups by appropriate categories until a homogeneous effect is detected. Due to very small sample sizes found when breaking the studies into smaller groups, a further disaggregation of the study effect sizes was not possible in the present study. The results of the weighted mean correlation coefficients are presented in what follows as a description of the mean of observed effect sizes, with limited interpretation because of the heterogeneity of effects.

The results of the meta-analysis by type of content of the selection procedure for the combined criterion are presented in Table 2. The corrected mean validity coefficients of the single domain selection procedures range between $r = .22$ and $r = .44$ ($p > .05$ for all coefficients). Although the corrected validity coefficients are positive and moderate, the variability of the effect does not allow us to conclude that the results are statistically significant. The results are similar to those obtained by Ford et al. (1987) in a previous meta-analysis of dimensions used in the prediction of salesperson performance. The results of the previous study were disaggregated into more narrowly focussed dimensions than the present study, and uncorrected for statistical artifacts other than sampling error. The range of effect sizes among the set of narrow dimensions in the previous study that are comparable to the broad dimensions in the present study contained the uncorrected effect sizes of the present study. This finding lends support to the estimates of the effect sizes of the various predictors of sales performance, because the studies included in the previous meta-analysis were drawn from journals of marketing and consumer research, whereas the present meta-analysis included studies drawn from the psychology journals.

The results of the meta-analysis by type of content of the selection procedure for the subjective criterion and objective criterion are presented in Tables 3 and 4, respectively.
Subjective ratings of performance were predicted by the single domains of personality dimension \( (r = .20, p < .05) \), biographical information \( (r = .20, p < .05) \), and cognitive ability \( (r = .32, p < .05) \). The results of a previous meta-analysis of sales predictors (Vinchur et al., 1998) found a similar validity coefficient for the prediction of subjective ratings of performance by overall cognitive ability \( (r = .31) \), but a larger corrected correlation coefficient for biographical information \( (r = .52) \). The discrepancy between the correlation coefficients of biographical information and subjective ratings in the Vinchur et al. study and in the present study may be due to both fewer studies used in the present study \( (k = 5) \) than in the previous study \( (k = 8) \) and lesser correction for statistical artifacts in the present study. In general, the results suggest that subjective ratings of salesperson performance can be predicted by personality dimensions, biographical information, and cognitive ability.

Objective performance assessments were predicted by the single domain of special purpose sales assessments with an average corrected correlation of \( .49 (p < .05) \). This result is similar to the average corrected correlation of \( .45 \) in the Vinchur et al. (1998) study. It is interesting to note that cognitive ability demonstrated no relationship with objective performance \( (r = -.09, p > .05) \), whereas it was moderately related to subjective ratings of performance \( (r = .32, p < .05) \). This finding is consistent with previous investigations which similarly found a positive relationship between cognitive ability and subjective ratings of performance, but no relationship with objective criteria (Bommer et al., 1995; Vinchur et al., 1998). That cognitive ability predicts supervisory ratings of performance and not objective performance suggests that some important aspects of a salesperson’s job not reflected in their sales productivity may be related to cognitive ability.

Analysis of the Utility of Selection Procedures and Training Interventions

To evaluate the cost-benefit implications of the meta-analytic results, a utility analysis was conducted. The analysis of utility of human resource management (HRM) activities, such as personnel selection and training, is frequently used in industrial psychology to predict, describe, and/or explain their
usefulness or desirability (Boudreau, 1983; Cascio, 1980). The purpose of utility analysis is to express the substantial value of particular HRM activities to the organization to allow management to evaluate what is gained from the use of information from industrial and organizational psychology, economics, and information theory. It also allows the company to evaluate the relative contribution of its HRM investments, and establish their value in comparison to investments in other functional areas (Cascio, 1989).

Utility analysis models were first developed in response to the deficits of traditional measurement and test theories in expressing the usefulness of tests. Utility analysis has come to reflect the influence of an HRM activity on: (1) the number of employees affected and the duration of the effect, (2) the average effect of the program, and (3) the resources required to implement and maintain the activity.

Utility of Selection Procedures. Originally, utility analysis models were developed specifically for selection assessment procedures. One of the first utility analysis models defined the benefit of assessment as the success ratio, which is the proportion of selected employees whose job performance is found to be successful (Taylor & Russell, 1939). This model was found to be insufficient because it lacked information about the number of employees affected and the duration of effect. It was also criticized because it expressed the criterion as a dichotomous variable (success or failure), whereas job performance is usually considered as a continuous variable. A second model of utility defined the benefit of assessment as the expected mean overall job performance of selected applicants (Naylor & Shine, 1965). It expressed the utility of a selection procedure as the difference in average standardized criterion score between those selected by a test and those selected without it. This model was criticized because the difference in average standardized criterion score is difficult to interpret in an applied situation. A third model of utility, and the one employed in the present study, defined the benefit of selection assessment by the dollar-value gained by use of the assessment (Brogden, 1946a, 1946b, 1949; Cronbach & Gleser 1965, pp. 308-309). This model, which will be referred to as the Brogden-Cronbach-Gleser selection utility model, estimates the overall net utility of a selection procedure, $\Delta U$, based on the number of employees hired by the selection assessment procedure, $N$, the average tenure of the employees, $T$, the validity coefficient of the selection
assessments, \( r_{xy} \), the standard deviation of job performance in the applicant pool (preferably in dollars), \( SD_y \), the standardized average score on the predictor assessment, \( z_x \), and the overall cost of implementing and maintaining the selection assessment procedure, \( C \), by the following equation,

\[
\Delta U = N T r_{xy} z_x SD_y - C
\]

The utility of selection procedures in the present study was estimated by the Brogden-Cronbach-Gleser model with the estimated validity coefficients of .27 for the single domain selection procedure and .59 for the composite domain selection procedure. The average tenure of salespeople was estimated to be 2.4 years, based on data from the Bureau of Labor Statistics (1998). The estimation of the standard deviation of job performance (\( SD_y \)) was obtained from a meta-analysis by Hunter et al. (1990) that examined empirical estimates of the variability of employees' productivity as a percentage of their average salary. Hunter et al. reported separate estimates of \( SD_y \) for incumbent employees and applicants. Where selection procedures were examined in the present study, the estimate of \( SD_y \) for applicants was employed because applicants are the population of interest. The average standard score on the predictor, \( z_x \), reflects the extent to which an organization can be selective, because the more applicants considered, the greater the difference between those selected and rejected. The estimates of \( z_x \) used in the present study were obtained from average standard scores on the predictor variable for those selected at various selection ratios (assuming a normal distribution on the predictor), given by Hunter and Hunter (1984). The selection ratio is the ratio of the number of applicants selected for employment to the number of applicants in the applicant pool. The estimate of \( z_x \) used, based on a selection ratio of .20, was 1.40. The estimate of \( z_x \) used, based on a selection ratio of .40, was .97. Estimates of the cost, \( C \), of selection assessment procedures were not included because they vary greatly depending on which procedure is used, how it is developed, and how its use is maintained in the organization. The estimates of utility in the present study represent the expected dollar-value gain to the organization from the use of single domain or composite domain selection assessments to hire one sales representative. Therefore, we set \( N \) to 1.0 in the equation for \( \Delta U \).
The results in Table 5 illustrate the dollar value gain for each applicant selected by single or composite domain assessment over random selection, for two levels of sales representative at two selection ratios, per year ($T = 1$) and per average tenure ($T = 2.4$). Selection by a composite domain assessment procedure when selecting the top 20% of applicants for senior sales representative positions would yield a return to the organization of approximately $24,901 per employee per year of the employee’s tenure. Selection by a composite domain assessment procedure when selecting the top 40% of applicants would yield a return to the organization of approximately $17,253 for each sales representative hired per year. Table 5 also includes the total dollar value gain for each applicant selected for the duration of the effect; in the case of personnel selection, the duration of effect is the length of the employee’s tenure. With an estimate of the average tenure of a salesperson in the United States of 2.4 years (Bureau of Labor Statistics, 1998), the total return to the company of using a composite domain assessment procedure to select a sales representative from among the top 20% of applicants would be $59,762, and when selecting from among the top 40% of applicants would be $41,407.

The cost of development, implementation, and maintenance of the selection assessment has not been taken into account in the present analyses; thus, these are gross utilities. The calculation of the cost-benefit implications of selection can be easily made by subtracting the proposed costs from the corresponding dollar-value benefit given in Table 5 for each sales representative hired.

*Utility of Training Interventions.* Although utility analysis has been applied primarily to personnel selection procedures it has also been developed to express the outcomes of training and development programs (Cascio, 1989; Landy, Farr, & Jacobs, 1982; Schmidt, Hunter, & Pearlman, 1982). Computational formulas exist which estimate the utility of training interventions as a function of their cost, effectiveness, and duration of effect. In the studies examined in the present thesis observed values of cost and duration of effect were unavailable. Likewise, empirical distributions of these values were unavailable. The annual utility of training interventions was estimated by the following equation given by Cascio (1989,
\[ \Delta U = N \cdot d \cdot SD_y, \]

where \( \Delta U \) is the expected annual net benefit of training, \( N \) is the number of employees trained, \( d \) is the standardized mean difference of performance between those trained and those untrained, and \( SD_y \) is the standard deviation of job performance. The standard deviation of job performance was the same value that was taken from empirical estimates provided by Hunter et al. (1990) in the calculation of utility of selection procedures.

The empirical estimates of the utility of training interventions for an average sales representative and senior sales representative are presented in Table 5 along with utility estimates for personnel selection. Training an average sales representative would yield a return to the organization of approximately $9,773 assuming that the duration of the effect of training was one year. The assumption that the duration of the effect of training endures for at least one year is not established by the empirical research reviewed in the present study. Therefore, the utility estimates of training interventions must be interpreted with caution.

The comparison of the relative utilities of single domain selection procedures, composite domain selection procedures, and training interventions are based on the weighted mean effect sizes found by meta-analytic procedures in the present study. The utility estimates refer to the dollar value gain per year for each employee selected by valid selection procedures or trained with an effective training program for the duration of the effect. The duration of the effect of selection procedures is assumed to be the employee’s tenure, \( T \); however, the duration of the effect of training is unknown from the evidence in the present study. The cost of developing, implementing and maintaining personnel selection procedures and training interventions has not been taken into account in these utility estimates. On average, the dollar value gain per year from personnel selection procedures is greater than the gain from training interventions.

**Summary**

On average, composite domain selection assessments predicted salesperson performance, while the effects of single domain selection assessments and training interventions were not statistically significant.
The disaggregation of results of personnel selection procedures by type of content and criterion, however, clarified the results. Personality assessments, biographical information inventories, and cognitive ability tests were found to be promising predictors of subjective ratings of performance. Special purpose sales assessments were good predictors of objective ratings of performance. It is worthwhile to consider subjective and objective performance criteria separately in sales occupations, as they appeared to be predicted by different individual difference dimensions.

The disaggregation of the results of single-predictor personnel selection procedures by type of predictor content and criterion measurement reduced the variability among the effect sizes of single predictors overall. This reduction in the variability among effect sizes was demonstrated where the effect size of single predictors overall was not significantly different from zero, whereas some of the individual single predictors were significantly different from zero when either a subjective or objective criterion was considered. This finding suggests that the type of content of the predictor and the type of criterion measurement moderate the relationship between single dimension personnel selection methods and sales performance.

The applied research of training interventions is limited to a relatively small number of studies that produced small effects with a high degree of variability. This latter finding argues for further research into the effectiveness of training interventions in sales occupations, with a particular focus on the duration of effect.

There are two limitations in this study. First, the comparison between selection procedures and training interventions was limited by a disproportionately small number of studies available in training, greater proportion of unpublished studies in training, and greater variability among effect sizes in training. Second, the sales occupation is extremely broad, representing substantial differences in job complexity which range from sales clerks in department stores, to representatives selling industrial products to international customers.

The comparison among single- and multiple dimension selection procedures and training interventions suggests that selection by multiple dimensions is the only approach examined in the present
study that leads to a valid improvement in the job performance of salespeople. Training interventions were marked by a greater amount of variability than selection procedures. It was not possible in this study to examine possible moderators of the effect of training, because of the small number of training studies.
References

References marked with an asterisk indicate studies included in the meta-analysis.


psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.


Table 1

**Correlation Coefficients by Single Domain Predictor, Composite Domain Predictor, or Training Intervention for the Combined Criterion (average of subjective and objective criteria)**

<table>
<thead>
<tr>
<th>Approach to Improve Job Performance</th>
<th>k</th>
<th>N</th>
<th>Mean r</th>
<th>SE</th>
<th>lower</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>single predictor</td>
<td>144</td>
<td>58,310</td>
<td>.27</td>
<td>.29</td>
<td>-.30</td>
<td>.69</td>
</tr>
<tr>
<td>composite predictor</td>
<td>10</td>
<td>1,246</td>
<td>.59</td>
<td>.20</td>
<td>.27</td>
<td>.79</td>
</tr>
<tr>
<td>training intervention</td>
<td>12</td>
<td>2,378</td>
<td>.16</td>
<td>.24</td>
<td>-.30</td>
<td>.57</td>
</tr>
</tbody>
</table>

*Note: k = number of studies; N = number of subjects; Mean r = average correlation coefficient corrected for criterion unreliability, and range restriction in the independent variable for single and composite predictors; SE = standard error corrected for sampling error; 95% CI = 95% confidence interval.*
Table 2

*Validity Coefficients by Content of Single Domain Selection Assessment for the Combined Criterion*  
*(average of subjective and objective criteria)*

<table>
<thead>
<tr>
<th>Content of Selection Assessment</th>
<th>k</th>
<th>N</th>
<th>Mean $r_{cr}$</th>
<th>Mean $r_{cr, CR}$</th>
<th>SE</th>
<th>95% CI for $r_{cr, CR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>job skill/knowledge</td>
<td>15</td>
<td>3,171</td>
<td>.22</td>
<td>.33</td>
<td>.31</td>
<td>-.26, .74</td>
</tr>
<tr>
<td>personality dimension</td>
<td>41</td>
<td>4,325</td>
<td>.14</td>
<td>.24</td>
<td>.17</td>
<td>-.09, .51</td>
</tr>
<tr>
<td>biographical information</td>
<td>24</td>
<td>44,509</td>
<td>.16</td>
<td>.26</td>
<td>.18</td>
<td>-.09, .54</td>
</tr>
<tr>
<td>cognitive ability</td>
<td>22</td>
<td>2,784</td>
<td>.13</td>
<td>.22</td>
<td>.22</td>
<td>-.22, .57</td>
</tr>
<tr>
<td>job specific aptitude</td>
<td>16</td>
<td>1,856</td>
<td>.19</td>
<td>.30</td>
<td>.39</td>
<td>-.44, .80</td>
</tr>
<tr>
<td>special purpose</td>
<td>24</td>
<td>1,481</td>
<td>.28</td>
<td>.44</td>
<td>.32</td>
<td>-.16, .80</td>
</tr>
</tbody>
</table>

*Note:* $k =$ number of studies; $N =$ number of subjects; Mean $r_{cr} =$ average validity coefficient corrected for criterion unreliability of subjective ratings; Mean $r_{cr, CR} =$ average validity coefficient corrected for criterion unreliability of subjective ratings and range restriction in the independent variable; $SE =$ standard error corrected for sampling error; 95% CI = 95% confidence interval.
Table 3

*Validity Coefficients by Content of Single Domain Selection Assessment for the Subjective Criterion*

<table>
<thead>
<tr>
<th>Content of Selection Assessment</th>
<th>k</th>
<th>N</th>
<th>Mean $r_c$</th>
<th>Mean $r_{crr}$</th>
<th>SE</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>job skill/knowledge</td>
<td>11</td>
<td>1,626</td>
<td>.24</td>
<td>.36</td>
<td>.28</td>
<td>-.17</td>
<td>.74</td>
</tr>
<tr>
<td>personality dimension</td>
<td>24</td>
<td>1,919</td>
<td>.12</td>
<td>.20</td>
<td>.03</td>
<td>.13</td>
<td>.27</td>
</tr>
<tr>
<td>biographical information</td>
<td>5</td>
<td>670</td>
<td>.12</td>
<td>.20</td>
<td>.08</td>
<td>.03</td>
<td>.36</td>
</tr>
<tr>
<td>cognitive ability</td>
<td>17</td>
<td>2,190</td>
<td>.20</td>
<td>.32</td>
<td>.00</td>
<td>.27</td>
<td>.35</td>
</tr>
<tr>
<td>job specific aptitude</td>
<td>9</td>
<td>1,350</td>
<td>.17</td>
<td>.28</td>
<td>.29</td>
<td>-.29</td>
<td>.70</td>
</tr>
<tr>
<td>special purpose</td>
<td>16</td>
<td>1,069</td>
<td>.27</td>
<td>.42</td>
<td>.38</td>
<td>-.29</td>
<td>.81</td>
</tr>
</tbody>
</table>

*Note:* $k =$ number of studies; $N =$ number of subjects; Mean $r_c =$ average validity coefficient corrected for criterion unreliability of subjective ratings; Mean $r_{crr} =$ average validity coefficient corrected for criterion unreliability of subjective ratings and range restriction in the independent variable; $SE =$ standard error corrected for sampling error; 95% CI = 95% confidence interval.
Table 4

Validity Coefficients by Content of Single Domain Selection Assessment for the Objective Criterion

<table>
<thead>
<tr>
<th>Content of Selection Assessment</th>
<th>k</th>
<th>N</th>
<th>Mean r</th>
<th>Mean $r_{\tau}$</th>
<th>SE</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>job skill/knowledge</td>
<td>6</td>
<td>1,666</td>
<td>.20</td>
<td>.30</td>
<td>.40</td>
<td>-.43</td>
<td>.80</td>
</tr>
<tr>
<td>personality dimension</td>
<td>18</td>
<td>2,539</td>
<td>.15</td>
<td>.26</td>
<td>.28</td>
<td>-.26</td>
<td>.67</td>
</tr>
<tr>
<td>biographical information</td>
<td>17</td>
<td>13,665</td>
<td>.20</td>
<td>.32</td>
<td>.18</td>
<td>-.03</td>
<td>.60</td>
</tr>
<tr>
<td>cognitive ability</td>
<td>6</td>
<td>679</td>
<td>-.06</td>
<td>-.09</td>
<td>.38</td>
<td>-.69</td>
<td>.59</td>
</tr>
<tr>
<td>job specific aptitude</td>
<td>4</td>
<td>452</td>
<td>.21</td>
<td>.34</td>
<td>.21</td>
<td>-.08</td>
<td>.65</td>
</tr>
<tr>
<td>special purpose</td>
<td>10</td>
<td>558</td>
<td>.32</td>
<td>.49</td>
<td>.14</td>
<td>.25</td>
<td>.68</td>
</tr>
</tbody>
</table>

Note: $k =$ number of studies; $N =$ number of subjects; Mean $r =$ average validity coefficient; Mean $r_{\tau} =$ average validity coefficient corrected for range restriction in the independent variable; $SE =$ standard error corrected for sampling error; 95% CI = 95% confidence interval.
Table 5

*Relative Utility Estimates of Valid Selection Procedures and Effective Training Interventions*

<table>
<thead>
<tr>
<th>Human Resource Management Activity</th>
<th>Dollar value gain per year for each employee</th>
<th>Dollar value gain for the duration of tenure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selection Assessment Procedure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Selection Ratio = .20</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single dimension ((r = .27))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior sales representative</td>
<td>$16,119</td>
<td>$38,686</td>
</tr>
<tr>
<td>Sales representative</td>
<td>11,395</td>
<td>27,349</td>
</tr>
<tr>
<td>Multiple dimension ((R = .59))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior sales representative</td>
<td>35,224</td>
<td>84,537</td>
</tr>
<tr>
<td>Sales representative</td>
<td>24,901</td>
<td>59,762</td>
</tr>
<tr>
<td><em>Selection Ratio = .40</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single dimension ((r = .27))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior sales representative</td>
<td>11,168</td>
<td>26,804</td>
</tr>
<tr>
<td>Sales representative</td>
<td>7,895</td>
<td>18,949</td>
</tr>
<tr>
<td>Multiple dimension ((R = .59))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior sales representative</td>
<td>24,405</td>
<td>58,572</td>
</tr>
<tr>
<td>Sales representative</td>
<td>17,253</td>
<td>41,407</td>
</tr>
<tr>
<td><strong>Training Intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training effectiveness ((d = .32))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior sales representative</td>
<td>$13,824</td>
<td></td>
</tr>
<tr>
<td>Sales representative</td>
<td>9,773</td>
<td></td>
</tr>
</tbody>
</table>