

# Selecting Successful Salespersons with the 16PF®: Form A Validity Studies

Stephen J. Guastello, Ph.D.; Mark L. Rieke, Ph.D.

## Executive Summary

This report summarizes a series of studies that were undertaken to evaluate the validity of 16PF traits for predicting success in sales occupations. There is strong evidence that the successful salesperson is warm and outgoing (A+), intelligent (B+), emotionally stable (C+), assertive (E+), cheerful and optimistic (F+), socially bold (H+), conscientious (G+), and self-assured (O-).

In some of the studies, traits that were known to affect sales performance were compiled into a weighted composite variable and tested for cross-validity with several types of criteria: sales performance measured by ratings and objective indices, membership in a group of employed successful salespersons, length of stay, and job satisfaction. In other studies, multiple regression strategies were used to identify situational specific traits that were related to performance.

The average cross-validity coefficient for the composite of sales traits with membership in a successful sales group was .71. Cross-validity of the composite and situational multiple regression coefficients for the other criteria ranged from .30 to .31. The average value across all criteria, based on 683 salespersons was .41. In addition to the "standard" constellation of sales traits, successful salespersons are often sensitive (I+) and empathetic. It appears that the 16PF can produce an increase in the relative efficiency of the selection process by 17% to 32%, depending on job difficulty and the criteria important to the particular job.

## Introduction

Successful sales work is the result of several types of economic and social forces. Individual salespersons contribute their talents and ambition. Markets and products, at the very least, can make the job easy or difficult. Management sets the motivational climate and structures the work. Because talent for sales involves a certain quality of social interaction, it would follow that normal range personality traits would go a great distance toward explaining success in sales.

The goal of this report is to summarize the findings of studies undertaken to identify the most pervasive traits of salespeople. Such information is often vital for personnel selection, training, and career guidance. This report is organized into five basic sections and a summary. The first section provides an overview of the personnel selection process with particular attention to salesperson issues. The second section elaborates on the role of trait variables within the broad array of forces affecting sales performance. The third section describes the structure of personality traits. The fourth section describes the known validity studies for the 16PF traits with sales performance. The results of the 16PF studies are compiled in a form that allows for a straight forward comparison with results from selection studies prepared by independent sources. The fifth section evaluates the utility of selection with the 16PF of improving the relative efficiency of a personnel selection procedure.

## 1. How to Compose a Selections System for Salespersons

The purpose of this section is to describe the essentials of personnel selection for developing a valid battery of predictive measures. Entire books have been written on this subject, and standard texts in industrial and organizational psychology typically spend about a third of their page space on personnel selection issues. It is necessary, therefore, to limit the discussion to the most fundamental ideas.

Before continuing, however, it is worthwhile to reference two short documents concerning standards for personnel selection systems: *Uniform Guidelines on Employee Selection Procedures* (Miner & Miner, 1980), and *Principles for the Validation and Use of Personnel Selection Procedures* (Society for Industrial and Organizational Psychologists, [SIOP], 1987). The *Uniform Guidelines* refers to standards for personnel selection validity studies as they might be required by Equal Employment Opportunity law. The *Principles* interpret the *Uniform Guidelines* further as they apply to new ideas in personnel selection that have come into practice since the publication of the *Uniform Guidelines*.

There are four basic steps to composing a selection battery:

- job analysis,
- performance appraisals,
- identification of plausible predictors of success and testing their validity empirically
- assessment of the utility of the selection system in terms of how well it improves the selection system

## 1.1 Job Analysis

Job analysis is perhaps the most fundamental activity for selecting any occupational group. The objectives are to determine what is required by a job and, in many cases, to compare requirements among jobs. While there are many approaches one might take to job analysis, the range of options can be summarized into three basic categories: functional job analysis, task-based analysis, and psychological requirements analysis. Each type of job analysis has its assets and limitations.

### Job Analysis Methods

Functional job analysis is the type of job analysis found in the *Dictionary of Occupational Titles*, published by the United States Department of Labor. For each job the goal is to prepare a descriptive paragraph detailing what behaviors are necessary to

perform the job. The verbal descriptions should address the goals of the job in terms of what the worker must accomplish, the tools or other resources needed to accomplish those objectives, the performance standards associated with those objectives, and the training requirements for the job. Functional job analysis is usually accompanied by numerical scales that index the job in terms of how much of the work involves use of tools and manipulation of objects, use of information and decision making skills, interaction with people, use of reasoning, and use of math and language skills. A full description of the functional job analysis technique is found in Fine (1989).

Task-based job analysis begins with a large group of identifiable tasks that are completed by a heterogeneous group of employees. The task-based job analysis procedure would proceed as follows. First, the researcher would compile a nearly exhaustive list of tasks for the entire work group that is under study. The list of tasks is usually compiled from available work records. The tasks would be itemized on a rating form whereby, for each task, the incumbent would rate the task for importance, frequency with which it occurs, and possibly its difficult and time requirement. The ratings would then be factor-analyzed to produce a small number of task groups (or task *factors*). Each worker would be responsible for some work in each task group, although different workers may perform tasks in some groups but not in others.

A more psychological approach to job analysis is typified by the Position Analysis Questionnaire (PAQ; McCormick, Jeanneret, & Mecham, 1972), which is a standardized set of 194 questions about a person's job. The PAQ was originally developed to provide a taxonomy of Federal jobs, but has since been widely adopted by industry. The analysis is from the vantage point of the worker's activities, rather than from the vantage of task functions. The final job analysis takes the form of 27 numerical scores. These scores are combined into five broader level scores which give information about the type of information used by an incumbent, mental processes that are required, type of work output, relationships

with other people, and how the job might be related to work performed by other people.

No one method of job analysis is universally preferable to another. Rather, the method of choice depends on the requirements of the situation. The primary benefit of functional job analysis lies in the verbal descriptions that make communication about jobs easier than what would be accomplished by a string of numbers. The definition of performance objectives and training requirements is not easy or direct, however. The task-based method is attractive in situations where the assignment of tasks to people is not consistent. The method is useful for clarifying the allocation of tasks to people and the definition of performance standards. It says little about the psychological resources needed to perform those jobs effectively. The PAQ method focuses on psychological demands and lends itself to the identification of psychological tests that could be used to select people. Its connection to performance standards is not direct, and the numbers that it produces does not lend itself to easy communication about the jobs.

### Sales Jobs

Before leaving the topic of job analysis, it is important to call attention to some aspects of sales work that could distinguish one job from another, or distinguish sales jobs found in one setting from those found in another (Lambert, 1979). Whatever method of job analysis is used, some basic questions should be addressed.

- Does the job involve telephone sales, or face-to-face contact, or some of each?
- Does the salesperson develop his or her own sales leads, or is the customer base generated by the employer in some way?
- Does the job involve travel to visit customers, or do customers visit the store? Is travel local, regional, national, or international?
- Is the salesperson responsible for one product line, or for a territory in which many types of products can be sold?

- To what extent are sales, marketing, and advertising separate jobs or interwoven in one job?
- Is the product line tangible or intangible, simple or complicated (e.g. cosmetics versus mathematical software)?
- Is physical delivery of the product part of the job?
- To what extent is paperwork and summary reporting part of the job?

## 1.2 Performance Appraisals

Having determined what a job requires, the second step for the human resource manager is to establish standards of performance, along with a measurement system that captures those standards. The following is a summary of only the most important points. Interested readers should consult Carroll & Schneier (1982), Landy and Farr (1980, 1983), and Landy, Zedeck, and Cleveland (1983) for elaborations on these and many other aspects of performance appraisal. There are two broad categories of criteria: those that involve "objective" work outcomes, and more subjective rating methods. Each has its own assets and limitations.

### Objective Criteria

Examples of "objective" or "hard" criteria would include the number of items produced, sold, or scrapped. In the case of sales, further examples would include the number of new customers or clients, dollar volume of sales, number of real estate listings obtained, and so forth. The major assets of these types of criteria are that they are observable directly and would appear to involve little interpretation.

The limitation of so-called objective criteria is that a certain amount of the outcome lies beyond the control of the individual being assessed. That is, many such outcomes are partially the result of environmental and economic influences, or the result of the work of a team, rather than of a single individual.

### Subjective Criteria

Subjective criteria, such as ratings of work behavior, are more flexible, in principle, for purposes of isolating an individual's contribution to work outcomes. The inherent subjectivity in these criteria, however, lends itself to possible inaccuracies in evaluation. Such inaccuracies may be inherent in the rating scheme itself or a result of errors in judgment on the part of raters using the systems.

### Cautionary Notes

Two cautionary notes should be made regarding performance criteria for salespersons. First, success in a training program might be meaningful criterion to the organization's training staff, but it does not by itself meet the standards in the *Uniform Guidelines* for job-relatedness. According to the *Guidelines*, criteria should represent the job into which the applicant is hired. Success in future jobs in a possible career sequence is not a fair criterion, nor is success in a training program that transpires prior to the actual job.

The second cautionary note is that the statements of job performance in the rating scales or in the dimension descriptions should describe *behaviors*, not attributes, personality traits, or other personal characteristics. For instance, Lambert (1979) listed 15 aspects of salesperson performance (p. 42), all of which were accompanied by some compelling rationale. Consider five of them, however: (a) quality of field work being performed, (b) degree of preparation in pre-call planning, (c) skill in new product presentations, (d) high degree of empathy with customers, and (e) ability to build on brand and product loyalty. The first three items can be legitimately viewed as behaviors. Critical incidents analysis could produce examples of good and poor field work and pre-call planning, and could characterize an attractive product presentation.

Empathy, on the other hand, is a personality characteristic. A behavioral criterion could describe behaviors in which empathy might be involved, but the focus of attention should be on the presence or absence of that behavior, not the presence or

absence of a trait. Similarly, ability is possibly inferred from behaviors, but the criterion scale should be focused on the behavior, not the inferred ability.

### How Many Criteria?

Personnel psychology recognizes no upper limit to the number of possible rating scales or total number of criteria that could be used in a validity study or performance appraisal system. A small number of measures is simpler to use for administrative purposes, and would greatly simplify a test validation study. Larger numbers of more detailed scales, however, are preferable for employee counseling purposes where specific feedback to the employee is greatly beneficial to improving performance.

Cognitive psychologists have known for years that human information processing capabilities are generally limited in the number of categories, channels, or dimensions of information that can be cogently handled. The limitation has been known as "the magical number 7 plus or minus 2" (Miller, 1956). This robust finding suggests that a small number of carefully developed ratings are preferred to a large set of ratings. When too many scales are used, raters will have serious difficulty in keeping the meanings of the scales separate, and a large amount of overlap will occur from one rating to another.

Ultimately, one should use as many criteria as it takes to describe the relevant behaviors of the job. Factor analysis of a large number of scales should make obvious the more pervasive themes underlying performance. On a conceptual level, we have identified five aspects of performance that should be at least considered in the development of any criteria for sales work.

- **Customer service and field work** – How well has the incumbent developed a positive working relationship with the client?
- **Sales skill** – To what extent has the incumbent demonstrated competency in product knowledge and skill in present company products?

- **Knowledge of company products** – How well has the incumbent communicated correct information about the company’s products or policies?
- **Commitment to the organization** – How committed is the incumbent to organizational goals and policies as demonstrated by actions?
- **Development of new business** – How successful has the incumbent been in staking out new territory, or starting up new accounts? Has the incumbent contributed to any collective efforts in this area?

The foregoing list of performance dimensions lends itself to rating methods. Other criteria of sales performance encountered in this report include dollar value or quantity of goods and services sold, length of stay on the job, survival in a sales career, and job satisfaction. Job satisfaction is not an example of a performance criterion, but is included here for the purpose of substantiating the link between personality measures and career success from the individual’s point of view.

Similarly, organization commitment is usually viewed as a useful criterion, although it can be thought of as an attitude. When considered as a criterion, the question the researcher is asking is, “What personality traits (or other characteristics) of a person predict whether a person can become committed to a particular organization?” An individual’s commitment is something that must be earned by the organization through proper treatment of the employees.

#### **How Many Raters?**

Where possible, at least two raters should be used to evaluate a person’s performance. Different raters have different opportunities to observe a particular worker and often have different evaluations of the same behaviors. King, Hunter, and Schmidt (1980) reported that the average inter-rater reliability obtained from many studies in which at least two raters observed workers on several scales was .60. This rather low reliability coefficient should not be interpreted as a flaw in the set of rating scales, but

largely a result of difference in situations in which performance is observed.

### 1.3 Test Validity

Having established job requirements and standards of performance, the next steps in the process are to identify some plausible predictors of success, and to test whether those test measurements are indeed related to the behaviors of interest. A discussion of variables known to affect sales performance appears in a subsequent section of this report. The remainder of this section is devoted to the assessment of validity for a possible test battery.

It should be noted that validity is not a characteristic inherent in the test itself. Rather, validity lies in the use of a test for a particular purpose. Indeed, as more research is conducted with a test, and more about its construct and other properties is learned, the concept of validity for any purpose evolves (SIOP, 1987). There are three aspects to test validity: content, construct, and criterion-relatedness. These principles are described below as they apply to personnel selection objectives, and they are followed by a brief mention of validity generalization techniques for pooling information from many validity studies.

#### **Content Validity**

Content validity is the extent to which the test items or other measurements represent actual work behaviors. One might ask whether the test poses questions, problems or situations that are similar in form to questions, problems or situations the examinee would face on the job.

#### **Construct Validity**

Construct validity is the extent to which the test measures what it is supposed to measure. Construct validity is determined by past research where the test in question was found to correlate significantly with other measures of the same construct, or perhaps with behaviors that are indicative of the hypothetical construct. In personnel selection,

construct validity goes a step further to address the questions, "To what extent does the construct measured by the test underlie or explain success on the job?"

### Criterion-Related Validity

Criterion-related validity studies are those in which the test scores have been compared against measures of job performance. There are two basic types of criterion-related validity studies: predictive and concurrent. The former is ideal. The latter is more common.

Predictive designs establish whether a causal relationship exists between the test scores taken at one point in time and behavior at a future date. In predictive studies, all job applicants would be tested and hired. The new employees would be observed for performance measures after a period of time on the job has transpired. Unfortunately, predictive studies are difficult and time consuming.

In concurrent designs, incumbent employees are tested and measured on performance at a time close to the time of testing, and a correlation between test scores and performance is established. There are two major problems with concurrent designs. First, training effects need to be taken into consideration. If incumbents are unequally trained, which is often the case, it is necessary to isolate the effect of training from the effect of the test construct.

A second threat to a concurrent validity study is restriction of the sample's range on the criterion. In the classic situations, all or most of the poor performers either quit their jobs or are fired before the study of incumbents takes place. Thus low scores on the criterion are not represented in the sample. This type of restriction of range would have the effects of (a) rendering a lower correlation between the predictor (test under consideration) and criterion than what would have been obtained in an unrestricted sample, and (b) over predicting performance of new examinees (applicants) in 99% of the cases (Olson & Becker, 1983).

### Validity Generalization

Validity generalizations (or meta-analyses) are a set of techniques for determining a population correlation between predictors (selection tests) and work criteria by compiling the correlation coefficients from many different companies with employees in the same type of job. In principle, validity generalization allows for stronger conclusions about the relevance of a class of tests with similar constructs to similar jobs. These techniques often involve making corrections for restriction of range and criterion unreliability. However, it is often valuable to interpret generalized data without those corrections in order to assess the effect size one would obtain in a realistic situation where all studies being compared are based on the same test.

Validity generalization is mentioned in this report for two reasons. First, we did use a basic validity generalization procedure on 16PF variables, particularly the salesperson composite, with a variety of performance measures. Second, we discuss the validity generalization concept with the alternative viewpoint that a test should be validated for each situation where it is used. *Uniform Guidelines* and selection experts (SIOP, 1987) recommend that validity studies be conducted whenever it is feasible to do so.

Both the *Uniform Guidelines* and SIOP encourage organizations to pool their samples to create a large enough sample to establish validity for a given type of job. Furthermore, if two jobs are similar, as determined by critical features of job analysis, those incumbents can be pooled into one study.

### 1.4. Utility

Utility is a calculation using the validity coefficient to estimate the economic impact for the organization of using tests for selecting successful job candidates, versus continuing to hire without testing. While current theories about utility estimation are beyond the scope of this report, a brief synopsis of fundamental utility ideas is valuable here. According to Taylor and Russell (1939), utility can be thought

of as *relative efficiency* expressed in percentages of successful people on the job. In this report we use the Taylor-Russell method of utility estimation to assess the relative impact of selection with the 16PF on an organization's selection efficiency.

Relative efficiency is a function of three parameters; base rate of success, selection ratio, and the validity coefficient. The base rate is the percentage of job applicants who are successful on the job when the selection device is not used. The selection ratio is the ratio of the number of job openings to the number of applicants. For instance, if there are five job openings and 100 applicants, the selection ratio is .05. For a given level of validity, the test will have greater utility if selection ratios are small (i.e., the odds of success on the job are low when the test is not used).

## 2. Personality Variables in Context

It might be desirable to make statements such as, "Our top performers are worth ten times as much to the company as the worst performers." It was not until recently, however, that the data and techniques for making accurate statements of this type became available. It is now recognized that a greater amount of variability in quality of job performance is observed for more complex jobs compared to simple ones. Furthermore, the greater the variability in performance between the top and bottom performers, the greater the utility (i.e., benefit to the company) of a good personnel selection technique. For life insurance salespersons, calculations show that the top 1% of employees can outsell the average employee by 4.2 times; for noninsurance sales, the top 1% of employees can outsell the average employee by 2.3 times, (Hunter, Schmidt, & Judeisch, 1990).

A comprehensive review of the sales performance literature (Churchill, Ford, Hartley, & Walker, 1985) showed that six types of variables have successfully explained performance among a wide variety of salespersons. The six groups of variables are: role

variables, skill, motivation, personal factors, aptitude, and organizational factors.

"Role variables" is a general term for differences in job assignments, responsibilities, and interpersonal interaction patterns. Many of these could be discerned from job analysis.

"Skill" would refer to the ability to present products, listen, find opportunities to sell, and close the sales. It is typically assessed through job application or resume information, sales knowledge tests such as the Sales Comprehension Test (Bruce, 1984, 1988), or interview questions where applicants explain what they would do in hypothetical situations (Latham & Saari, 1984; Weekley & Gier, 1987).

"Motivation" in work psychology takes on two basic forms. "Extrinsic" motivation is typified by concrete rewards such as money, benefits, commission programs, and so forth. "Intrinsic" motivation is more abstract, and includes "liking the work" and achievement motivation. Motivation can be measured by a TAT-like projective technique (McClelland, 1961) or structured questionnaires (Harrell & Stahl, 1981; Maehr & Braskamp, 1986).

"Liking the job" actually has two kinds of meaning. On the one hand, "liking the job" is synonymous with "job satisfaction." Job satisfaction has traditionally involved the measurement of job attitudes on five basic topics: pay, promotion opportunities, supervision, coworkers, and job content. Satisfaction with job content may be related to salesperson's personality characteristics or abilities, but it is often a result of the way the employer defines the job and divides the work among staff members. Indeed, the other aspects of job satisfaction are under management control rather than the incumbent's, with the possible exception of satisfaction with one's coworkers.

On the other hand, "liking the job" can refer to the broader matter of enjoying a career in sales. A person may be perfectly suited to a career in sales, but may work for an employer who engages in stress producing or demotivating policies and directives. Liking the job in the sense of a person-job fit with a

specific job and in the sense of compatibility of the person to a career in sales is best measured through a personality assessment instrument, such as the 16PF, or on interest inventories such as the Strong-Campbell. The following sections of this report describe studies showing that salespersons who possess certain combinations of personality characteristics are more likely to gravitate toward careers in sales, perform their work better, like their jobs better, show lower levels of absenteeism, and have a lesser likelihood of quitting those jobs.

“Personal factors” explaining sales job success would include variables unique to a given situation. Examples would include knowledge of the geographic sales area, experience with the product line or customer base, whether a person is working full or part time, or whether the incumbent’s career goals are consistent with what the job and the employer have to offer.

“Aptitude” is the “raw material” that a person needs for success in an occupation before suitable training and experience. Aptitude test scores usually show substantial correlation with skill measures taken at a later date. Aptitude is generally measured in place of skill when the desired tests of skills are not applicable, such as when the applicant population consists primarily of entry level people. Aptitude is more closely connected to basic psychological traits and cognitive ability (e.g. intelligence measures); skill is more closely related to occupational knowledge.

“Organizational and environmental” factors refer to the company’s management style, organizational culture, and person-organization fit. While the latter can be assessed through personality profiles, some organizational and environmental factors are not tractable through selection procedures. Organizational and environmental factors are enhanced more effectively through organizational development and structured change procedures.

### 3. 16PF Trait Structure

The 16PF is composed of 16 basic personality traits known as primary or first-order factors. Because of their statistical properties, they are thought to comprise a finite set of traits. All 16 scales are bipolar.

A factor analysis of the first-order personality factors showed that the 16 traits can be aggregated into five broader themes, known as the second-order factors (Cattell, Eber, & Tatsuoka, 1970). They are: (a) introversion *v.* extraversion, (b) high anxiety *v.* low anxiety, (c) tough poise *v.* sensitivity, (d) independence *v.* subduedness, and (e) behavior control *v.* impulsivity. These second-order factors correspond to the Big Five personality traits (McCrae & Costa, 1985) that are widely recognized respectively as: extraversion, anxiety, agreeableness (reversed), openness to experience, and conscientiousness. [note: The Big Five trait Conscientiousness should not be confused with 16PF primary trait Factor G which has the same name in most 16PF publications.]

There are some special scales that are commonly reported with 16PF interpretations such as the Human Resource Development Report (HRDR), Narrative Score Report (NSR), or Personal Career Development Profile (PCDP). The special scales are composites of primary 16PF traits that were developed in empirical research (Cattell et al., 1970); they often have some occupational relevance. For instance, the Leadership Potential index is particularly relevant to the selection of managers and leaders, and a history of its development and validity is available in a separate report (Guastello & Rieke, 1993). The Creativity index is particularly relevant to the identification of creative talent among artists, musicians, writers, and research scientists; a history of its development and validity is in preparation. The Adjustment index has clinical applications but little pertinence to personnel selection.

NSR and PCDP interpretive reports also provide indicators of how closely an examinee matches any of a variety of occupational profiles on file.

Occupational profiles are built from the characteristic profile equation (CPE) method (Cattell et al., 1970). A CPE is a weighted combination of 16PF primary traits that was developed by comparing scores from a target occupational group (such as salespersons) with scores from the general population. The CPE is a convenient method for testing the hypothesis that a combination of variables thought to characterize one sample generalizes to other samples from the same occupation.

## 4. Personality Traits and Sales Performance

In this section of the report we review studies of the validity of personality trait measures with sales performance. Subsection 4.1 is a summary of a validity generalization study that produced results relevant to the selection of salespersons and was based on tests other than the 16PF. The magnitude of the relationships between those personality variables and sales performance serves as a benchmark value with which to compare the results of 16PF salesperson validity studies in subsection 4.2.

### 4.1. Big Five Generalization Study

The Big Five taxonomy of traits has proven to be a valuable heuristic for classifying personality traits and for assessing their validity with respect to job performance. Barrick and Mount (1991) conducted such a review, and a synopsis of their results for salesperson performance appears in Table 1. The table entries are average correlations obtained from many studies between a personality variable and a subjective or objective measure of sales performance.

Also appearing in Table 1 are population values of the average corrected validity coefficients, which were obtained by correcting for range restriction and by correcting both predictors and criteria for unreliability. Those correction techniques are commonplace in validity generalization studies and serve to place the results from different research situations on a common footing with each other for

comparison. It should be noted, however, that correcting the predictor test scores for unreliability is unrealistic in practice because the examiner must use the test in its available form, which has less than perfect reliability. Overall, the corrected estimates of the population correlation coefficient ( $\rho$ ) are larger than values researchers could replicate with real situations and data with their inherent limitations.

According to Barrick and Mount (1991), the average correlation between a personality variable in the one of the Big Five categories with sales performance ranged from  $-.01$  to  $.09$ . Corrected values that appear in the second column of Table 1 represent estimates of population correlation coefficients. The population correlation coefficients were larger in size, but Conscientiousness was the only trait that was consistently related to performance, as evidenced by a positive  $\rho$  and a 90% confidence value (CV) that was also greater than zero. The observed mean correlation of  $.09$  and population mean correlation of  $.23$  serve as benchmark values with which to compare the 16PF validity coefficients summarized in the following sections of this report.

**Table 1: Summary of Generalized Validity Coefficients for Big Five Tests and Salesperson Performance**

Big Five Trait	Obs. r	Pop. ( $\rho$ )	.90% CV	No. of Corrs.	Total - N
<b>Extraversion</b>	0.09	0.15	-0.05	22	2,316
<b>Anxiety</b>	0.04	0.07	-0.18	19	2,486
<b>Agreeableness</b>	0.00	0.00	-0.31	16	2,344
<b>Conscientiousness</b>	0.09	0.23	0.23	21	2,263
<b>Openness to Experience</b>	-0.01	-0.02	0.18	12	1,566

## 4.2. 16PF and Sales Performance

Successful salespersons in a variety of settings seem to show eight common characteristics (Rieke & Russell, 1987): A+: Warm and outgoing *v.* reserved and aloof; B+: Capable of abstract thought *v.* prone to concrete thinking; C+: Emotionally stable *v.* easily upset; E+: Assertive, aggressive or competitive *v.* docile or easily led; F+: Cheerful, optimistic, and enthusiastic *v.* sober and serious; G+: Conscientious and persistent *v.* expedient and inattentive to rules and details; H+: Venturesome and socially bold *v.* shy and timid; O-: Self-assured and secure *v.* self-reproaching and insecure.

Traits A+, B+, H+, and O- capture the entrepreneurial aspects of sales, that is, the preparatory part of the job involving the development of new sales targets, formulation of a sales pitch, and organizing and planning a sequence of potentially successful sales encounters. The combination of A+, C+, G+, E+, and F+ are more germane to making actual sales. For a sales manager, we would want to see additional strength in A+, B+, and E+.

Experts on sales performance appear to have recognized empathy as an important variable in sales success (Lambert, 1979). Unfortunately, clear empirical support for that premise is lacking in the published literature. The following studies go some distance in filling that information gap. It is now

known that empathy is a composite of 16PF traits that are significantly correlated with the Hogan Empathy Scale: H+, M+, Q2-, I+, B+, Q4-, N-, F+, and Q3-. The nine traits were identified through stepwise regression and cross-validated (Guastello, Choi, Rieke, & Billings, 1992). The correlation between the nine traits and the external measure of empathy was .65. In terms of second-order factors, an empathetic person would be extraverted, sensitive, less anxious, and somewhat impulsive.

Three traits in the empathy group also appear in the sales group of traits: H+, B+, and F+. The 16PF validity studies, considered below, illustrate the relative importance of the core sales traits, the empathy group of traits, and other situationally specific traits.

### Sales for Technical Products

Professionals have found that the qualities desirable when selecting technical salespersons are different from those suitable for non-technical salespersons. Here, technical sales involve product lines which require specialized engineering knowledge. For technical sales, the optimal strategy would be to select an engineer with a high extraversion score. This idea makes sense: Three variables important to sales (A+, F+, H+) are part of the extraversion second-order factor. Engineers are typically B+. A fourth element of extraversion, Q2-, would cancel or

counteract the Q2+ often found in engineers' profiles (Cattell et al., 1970).

Some 16PF profiles for the 470 Australian engineers employed in eight different job categories have recently become available. Dowling and DeCieri (1992) reported that in that analysis, the eight engineer profiles were not significantly different from each other except for two variables, Q2 and I. The marketing and sales subgroup scored lower than most of the others on Q2 (mean = 6.0), and higher than most others on I (mean = 6.0 also). What Dowling and DeCieri did not discuss was whether any cumulative effects emerge when the groups are compared on composite variables, such as extraversion, or when the groups are aggregated, such as marketing and sales versus other types of engineers. Those comparisons would be valuable for substantiating the "engineer plus extravert" rule discussed above, and we examine those comparisons in our treatment of 16PF validity studies below.

### Validation Strategies for 16PF Studies

Several of the empirical studies discussed below test the validity of a Sales CPE with respect to measures of performance, turnover, job satisfaction, and occupational group membership in other samples. Other studies evaluate the validity of 16PF variables directly using stepwise multiple regression. In a third type of validity study, a Sales CPE score is calculated from the means for groups of salespersons then compared against general population values.

Krug (1980) first developed the CPE for sales and tested it for cross-validity with nine samples and four criteria: job satisfaction, tenure (length of stay with the company), occupational group membership, and performance. His results are summarized in Table 2, Samples 1 through 9.

Other samples from IPAT files (unpublished) provided validity coefficients with performance using multiple regression (Samples 10 through 12). Sample 13 provided group membership data for experienced salespersons, and represented a subset of Sample 12. Sample 14 is British and results for that group were published in Handyside (1988). Sample 15 is

the group of Australian engineers working in marketing and sales capacities (Dowling & DeCieri, 1992). When all the data are pooled, it is possible to draw some clear conclusions about the generalizability of the Sales CPE or other combinations of 16PF traits across a variety of sales situations.

### Group Membership

The most striking effect is the comparison between profiles for people who are successful salesperson and those who are not. In the typical study, a test on the mean CPE scores was used for the comparison (Krug, 1980). We converted the  $t$  values to corresponding correlation coefficients by first converting  $t$  to omega-squared ( $w^2$ ), then taking the square root to obtain an estimate of  $r$  (see Technical Appendix). No corrections for restriction of range or unreliability of the criterion were used for any of the data summarized in Table 2. The resulting average validity coefficient for occupational group membership was .72.

Sample 15 was treated separately in light of the specialized hypothesis for engineering sales. The marketing and sales group's extraversion score was computed as 6.04 based on the published profile of means. A one-sample  $z$  test indicated that the group's extraversion score was significantly greater than the population mean ( $p < .05$ , two-tailed). The value of  $z$  converted to an estimated  $r$  of .23, using the  $w^2$  procedure. In contrast, the mean extraversion score for the remaining 403 cases (5.44) was trivially below the population mean (5.5). The results of this comparison indicated that the "engineer plus extravert" rule is a valid selection strategy.

### Performance

Performance measures were available for eight samples. The average value of  $r$  was .17 for the five CPE studies, and .31 across all eight samples. The values of  $r$  in Table 2 for Samples 11 and 12 are multiple  $R$  coefficients that were adjusted for the number of variables in the model (Wherry's formula), thereby making them comparable to the other five entries for work performance.

In Sample 11, there were five significant bivariate correlations between primary traits and performance. Two variables from the sales profile (A+ and C+) were unique predictors of performance in the multiple regression model. The remaining three variables (I+, L-, N-,) were part of the empathy trait group, and appeared to be redundant with the first two in the amount of performance they could account for.

In Sample 12, there were seven traits that made unique contributions to the multiple regression model. Four out of the seven were part of the sales composite (F+, B+, A+, H+), one was part of the empathy composite (I+), and two were unique to the sample (Q1+ and Q4+). Samples 11 and 12 demonstrated the importance of the empathy factor in sales performance.

Sample 14 (Handyside, 1988) produced a profile that was somewhat different from the others discussed so far. The sample consisted of 21 home construction salespersons, and the criterion was an overall performance rating score. The significant predictors of performance were A+, C-, E-, H+, and I- (Multiple  $R = .64$ ). The stability of the multiple regression coefficient was tested using the "jack-knife," or "leave-out-one" technique (Darlington, 1990);  $R$  was found to fluctuate between .55 and .73. The predictors of performance for this group were consistent with the Rieke and Russell (1987) profile on the basis of A+ and H+, opposite of their profile with respect to E- and C-, and inconsistent with the Empathy Composite on the basis of I-. Because mean values on 16PF traits were not provided by Handyside, it was not possible to calculate a CPE score for this group.

**Table 2: Summary of Validity Coefficients for 16PF with Salespersons with Weighted Averages by Sample Size**

Sample	Group Membership	Performance	Job Satisfaction	Tenure
(1) 75 recently hired salespersons			0.31	0.24
(2) 41 salespersons on job more than one year	0.76	0.11 <sup>a</sup>	0.21	0.42
(3) 29 real estate mgrs 6 months or more		0.30 <sup>b</sup>		
(4) 24 retail salespersons		0.31	0.35	
(5) 46 farm equipment salespersons		0.09		
(6) 42 orthopedic equipment salespersons		0.13		
(7) 31 department store salespersons	0.36			
(8) 157 national corporation salespersons	0.94			
(9) 40 successful salespersons	0.87			
(10) 31 salespersons	0.80			
(11) 24 sales representatives		0.59		
(12) 55 salespersons		0.53		
(13) 31 experienced salespersons	0.65			
(14) 21 home construction salespersons		0.64		
(15) 67 engineering salespersons	0.23			
TOTAL N all samples	398	282	140	116
Average $r$ weighted by sample size	0.72*	0.31*	0.31*	0.30*

**Note:** For Samples 1-9 and 13, the predictor is the Sales CPE. For Samples 10-12 and 14, the predictors are 16PF primary traits identified through multiple regression. For Sample 15, the predictor is Extraversion. <sup>a</sup>Earnings, <sup>b</sup>pooled earnings of subordinates; all others are combinations of sales and ratings. \*  $p < .01$ .

### Other Criteria

Data for job satisfaction and tenure were fairly straightforward. The average  $r$  for the CPE with job satisfaction was .31 ( $p < .01$ ). The average  $r$  for the CPE with tenure was .30 ( $p < .01$ ).

Finally the (unweighted) average correlation between 16PF sales scores and four aspects of performance was .43. The value for group membership was most distinct from the other three weighted averages. The latter observation meant that salespersons were more different from the general population than they were from each other in performance, tenure or satisfaction. Corrections for range restriction or criterion unreliability were not made here because the numerical values necessary to make those corrections were unavailable. The uncorrected values, however, represent the type of results a researcher might observe from groups of salespersons similar to those whose data are included in Table 2.

## 5. Utility

The results of the 16PF salesperson selection studies can be translated into projections of how well the selection system can be improved by using the 16PF. The calculation of utility requires a validity coefficient: a base rate of success, and a selection ratio. We used two estimates for the validity coefficient; the average of the weighted averages for group membership and performance (.52) and the weighted average validity coefficient for performance criteria only (.31). The engineering sales sample was not included in these averages; the utility projections provided here, therefore, generalize to non-technical sales situations only. For the selection ratio, it was assumed that the organization had a sufficiently large applicant pool and favorable economic conditions so as to allow itself to select persons who would score in the top 20% of the predictor variable (composite score).

We used a base rate of success of .40 because, according to Hunter et al. (1990), the distribution of sales performance scores is positively skewed such

that the median level of performance is only one standard deviation above the bottom of the performance score distribution. A substantial percentage of salespersons perform at the near-zero level. A base rate of .40 would be about 0.5 standard deviations above the median. Utility was also figured with a base rate of .20, which would represent the more difficult sales assignments.

If an organization experiences a base rate of success of .40 and selects from the top 20%, an average  $r$  of .31 (for performance criteria) would translate into a relative efficiency of .57. Thus, the 16PF would provide an increase in relative efficiency of 17% compared to selection by conventional methods, which typically consists of job applications and interviews.

If an organization is hiring for an entry level position such that a substantial proportion of the applicants have never held sales positions, then the average validity coefficient for group membership and performance would be applicable. A validity of .52 combined with a base rate of success of .40 and selection from the top 20% of application would translate into a relative efficiency of .70. Thus, the 16PF would provide an increase in relative efficiency of 30%. If an organization experiences the base rate of success of .30 and selects from the top 20%, the average  $r$  of .31 for performance would translate into a relative efficiency of 47%, and an increase in relative efficiency of 17%. Similarly, if the base rate is .30, the validity coefficient is .52, and the organization selected from the top 20% of applicants, the relative efficiency is 59% and the increase in relative efficiency is 29%.

On the basis of the research results accumulated in this report, it is possible to draw some conclusions regarding the usefulness of the 16PF for the prediction of success in sales jobs. First, the sales CPE generalizes to a variety of sales situations. The average validity coefficient obtained for all criteria and analytic strategies was .42. Overall, a 24% increase in relative efficiency of selection can be obtained, compared to selection without the 16PF, with a range between 17% and 32% depending on

the job difficulty and the specific criteria that are relevant to the job.

Second, deviations from a simple sales profile were found mostly when factors associated with empathy were significantly related to performance. At the present time, not enough research has been accumulated with the 16PF or other tests to discern which situations would require greater than average empathy levels from its incumbents. It would be reasonable to suggest, however, that empathy would be most relevant where: (a) the products or services being offered are complex or involve many intangibles, (b) the customers or clients could have many different motives to do business, or (c) the customers or clients are not sure of what they need or what form the final result should take. A thorough job analysis should address the levels of product complexity, purchase motivation, and clarity inherent in a selling situation.

Third, there were occasional findings for single personality factors that were not part of the sales CPE or empathy composite being relevant to work outcomes. Situationally specific outcomes will occur, and their importance should not be discounted. The standard recommendation to organizations would be to conduct situationally specific validation studies for all personnel selection instruments rather than to rely on generalization studies alone. It would appear that situationally specific validity studies could produce twice the relative efficiency of selection compared to what one would obtain by selection with the most generalizable measures only.

Fourth, the range of psychological variables that are potentially relevant to sales extends beyond the realm of personality characteristics. Several other categories of predictors were summarized in the introduction to this report. A selection battery should, therefore, augment the 16PF with other types of measures; selection decisions should not be made on the basis of personality results alone.

## 6.1. Simplified Personality Patterns

The results of the validity studies can be simplified into two basic approaches to salesperson selection. The general rules, which are summarized in Table 3, would be particularly valuable in applications where there is only a small number of incumbent salespersons available to conduct an independent validity study. Next we present two approaches for selecting salespersons. Approach A, for general sales circumstances, is closely tied to the results of the validity studies already reported. The distinction between trait subgroups for the Sales CPE for preparation and customer contact was drawn from the trait constructs. Future research with multiple criteria of sales performance may clarify further which traits are uniquely tied to specific aspects of sales performance.

Variables related to the Empathy composite often appear in profiles of successful salespeople. Of those variables, I+ is the most notable in terms of its frequency of appearance and centrality to the empathy construct. Several other traits appear in both the Empathy composites and the Sales CPE; this pattern additionally reinforces the role of empathy in sales. Although further research is needed on the role of empathy in sales, it would appear that empathy is of greatest importance when customer needs and products are complex and considerable efforts are necessary to match the needs with the products. Furthermore, the Empathy composite was found to be particularly relevant to success in social service occupations (Guastello et al., 1992); it would follow that empathy would be of particular value to the sales function in service industries and in customer service jobs which are often part of an organization's sales function.

Approach B for technical sales follows a couple of rules that have yet to be explored more fully. In technical sales, the likely candidate would need an engineering background to grasp the technical aspects of the product line and to understand customer questions. Extraversion (second-order) captures three of the key sales traits which should then be added to the overall profile of the successful

candidate. Thus the people making selection decisions should be looking for, in essence, extraverted engineers.

The traits listed in Table 3 were drawn from profiles published in Cattell et al. (1970), and represent 25 female and 24 male engineers. Traits that were significantly less than or greater than the mean

( $p < .05$ ) were selected for Table 3. Factor B+ is common to both engineering and sales. Several traits are unique to engineers. Two traits common among engineers (Q2+ and F-) are canceled by a high extraversion score. Thus when selecting engineering salespeople, the professionals making the selection decision should be looking for the combination of A+, B+, F+, G-, H+, I-, L+, Q1+, Q2-.

**Table 3: Summary of Strategies for Selecting Salespeople**

Composite Equation	Traits Contained	Specific Function
<b>Approach A: General Circumstances</b>		
Sales CPE	A+ B+ H+ O- A+ O- C+ G+ E+ F+	Preparatory Work, Customer Contact
Empathy	H+ B+ F+ I+ M+ Q2- Q4- N- Q3-	Same as Sales CPE, Perspective Taking
<b>Approach B: Technical Sales</b>		
Extraversion	A+ F+ H+ Q2-	Extraverted Engineers
Engineer's Profile	B+ G- I- L+ Q1+ F- Q2+	Same as Sales CPE, Unique to Engineers, Canceled by Extraversion

## Technical Appendix

The omega-squared formula ( $w^2$ ) was developed as a measure of percentage of variance accounted for by a set of nominally scaled groups such as those commonly found in analysis of variance (ANOVA) research designs. It is comparable in meaning to eta-squared, which is in turn comparable to  $r^2$  when there is no nonlinear effect. (There is no nonlinear effect possible when the independent variable consists of two groups.) The conversion of a  $t$  test to  $w^2$  is a special case of the main principle:

$$w^2 = (t^2 - 1) / (t^2 + df) \sim r^2, \text{ or}$$

$$r = [(t^2 - 1) / (t^2 + df)]^{1/2}$$

Where  $df$  represents the total degrees of freedom for the problem. Note that the  $w^2$  formula is similar

to the more common transformation of  $t$  to the point-biserial correlation:

$R_{pb} = [t^2 / (t^2 + df)]^{1/2}$ , but differs with respect to the -1 term in the numerator. Thus,  $w^2$  would render lower values of  $r$  than the more common method. The difference in estimated  $r$  would be slight when  $t$  and  $df$  are large, but the difference in estimates would be greater when  $t$  and  $df$  are small.

The differences in estimated  $r$  produced from the two formulae are most pronounced for  $t < 1$ , where all estimated values of  $r$  would be zero for the  $w^2$  formula and nonzero for the common formula. We preferred to use the  $w^2$  formula for two reasons. First, because it produces smaller values of  $r$ , we felt it minimized the risk of exaggeration of the  $r$  estimates for our sample-population comparisons. Second, the  $w^2$  formula recognizes  $t$  as a ratio between an observed difference between means and

the standard error of those differences; it is only when that ratio substantially exceeds 1.00 that there is any correlation taking place. (Note that asymptotically minimum significant value of  $t$  is 1.96 at  $p < .05$ .) Therefore, a value of  $t < 1.00$  represents a mean difference that is less than what could occur by chance, and therefore must represent a population estimate of  $r = .00$ .

## Acknowledgements

The authors would like to gratefully acknowledge Dody Bullerman, Dr. Heather E. P. Cattell, Mrs. A. Karen Cattell, Casey Jones, Darcie Karol, and Dr. David Watterson in their respective roles of typesetting, proofreading, editing and providing valuable input.

## References

- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1-26.
- Bruce, M. M. (1984). *Manuals supplement*. Larchmont, NY: Martin M. Bruce, Publishers.
- Bruce, M. M. (1988). *Examiner's manual: Sales Comprehension Test Revised*. Larchmont, NY: Martin M. Bruce, Publishers.
- Carroll, S. J., & Schneier, C. E. (1982). *Performance appraisal and review systems*. Glenview, IL: Scott Foresman.
- Cattell, R.B., Eber, H.W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire [16PF]*. Champaign, IL: Institute for Personality and Ability Testing.
- Churchill, G. A., Ford, N. M., Hartley, S. M., & Walker, O. C. (1985). The determinants of salesperson performance: A meta-analysis. *Journal of Marketing Research*, *22*, 103-118.
- Darlington, R.B (1990). *Regression and Linear Models*. New York: Wiley.
- Dowling, A. J., & DeCieri, H. (1992). Human resource management for engineers. In D. Samson (Ed.), *Management for engineers* (pp. 179-207). Melbourne, Australia: Longman Cheshire.
- Fine, S. A. (1989) *Functional job analysis scales* (Revised Edition). Milwaukee, WI: Author.
- Guastello, S. J., Choi, J., Rieke, M. L., & Billings, S. W. (1992). *Personality and volunteer service orientation*. Technical report. Milwaukee, WI: Marquette University.
- Guastello, S. J., & Rieke, M. L. (1993). *The 16PF and leadership: Summary of research findings 1954-1992*. Technical Report. Champaign, IL: IPAT.
- Handyside, J. D. (1988). Occupational validity studies in the UK. In Independent Assessment and Research Centre (IARC; Eds.), *The analysis of personality in research and assessment: In tribute to Raymond B. Cattell* (pp. 125-130). London, UK: IARC.
- Harrell, A. M., & Stahl, M. J. (1981). A behavioral decision approach to measuring McClelland's trichotomy of needs. *Journal of Applied Psychology*, *66*, 242-247.
- Hunter, J. E., Schmidt, F. L. & Judeisch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*, *75*, 28-42.
- King, K. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in a multidimensional forced choice performance evaluation scale. *Journal of Applied Psychology*, *65*, 507-516.
- Krug, S. E. (1980). *Reliability and validity of the sales CPE*. Technical report. Champaign, IL: IPAT.
- Lambert, C. (1979). *Field sales performance appraisal*. New York: Wiley.

Landy, F.J., & Farr, J.L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

Landy, F. J., & Farr, J. L. (1983). The measurement of work performance: Methods, theory, and applications. New York: Academic Press.

Landy, F., Zedeck, S., & Cleveland, J. (Eds. 1983). Performance measurement and theory. Hillsdale, NJ: Lawrence Erlbaum.

Latham, G. P., & Saari, L. (1984). Do people do what they say? Further studies on the situational interview. Journal of Applied Psychology, 69, 569-573.

Maehr, M. L., & Braskamp, L. A. (1986). The motivation factor: A theory of personal investment. Lexington, MA: Lexington.

McClelland, D. C. (1961). The achieving society. New York, Van Nostrand.

McClelland, D. C. (1975). Power: The inner experience. New York: Wiley.

McCormick, E. J., Jeanneret, P., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire [Monograph]. Journal of Applied Psychology, 36, 347-368.

McCrae, R. R., & Costa, P. T. Jr. (1985). Updating Norman's "adequate taxonomy:" Intelligence and personality dimensions in natural language and in questionnaires. Journal of Personality & Social Psychology, 49, 710-721.

Mendoza, J. L., & Reinhardt, R. N. (1991). Validity generalization procedures using sample-based estimates: A comparison of six procedures. Psychological Bulletin, 110, 596-610.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Miner, M. G., & Miner, J. B. (1980). Uniform guidelines on employee selection procedures. Washington, DC: Bureau of National Affairs.

Olson, C. A., Becker, B. E. (1983). A proposed technique for the treatment of restriction of range in selection validation. Psychological Bulletin, 93, 137-148.

Rieke, M. L., & Russell, M. T. (1987). The Narrative Score Report user's guide. Champaign, IL: IPAT.

Society for Industrial and Organizational Psychology (1987). Principles for the validation and use of personnel selection procedures (Third Edition). College Park, MD: Author.

Taylor, H. C., & Russell, J. T. (1939). The relationships of validity coefficients to the practical effectiveness of test in selection: Discussion and tables. Journal of Applied Psychology, 23, 565-578.

Weekley, J. A., & Grier, J. A. (1987). Reliability and validity of the situational interview for a sales position. Journal of Applied Psychology, 72, 484-487.

© Copyright 1993 IPAT, Inc.

® 16PF is a registered trademark of the Institute for Personality and Ability Testing, Inc (IPAT). IPAT is a wholly owned subsidiary of OPP Ltd.

® OPP is a registered trademark of OPP.