

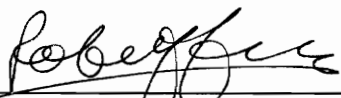
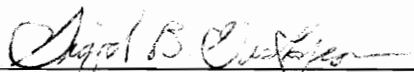
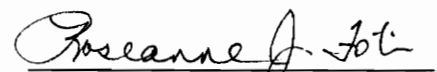
**An Examination of the Structure and Predictability of
Myers-Briggs Type Indicator Preferences Using a
Job Component Validity Strategy Based on the
Common-Metric Questionnaire**

by
Reagan Brown

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
MASTERS OF SCIENCE

IN
PSYCHOLOGY

APPROVED:


Robert L. Harvey, chair
Sigrid B. Gustafson
Roseanne J. Foti

November, 1995

Blacksburg, Virginia

Key words: Psychology, Selection, Validation, Synthetic validity, Job component validity, Personality testing

2

LD
S655
V855
1995
B772
C.2

**AN EXAMINATION OF THE STRUCTURE AND PREDICTABILITY OF
MYERS-BRIGGS TYPE INDICATOR PREFERENCES USING A
JOB COMPONENT VALIDITY STRATEGY BASED ON THE
COMMON-METRIC QUESTIONNAIRE**

by

Reagan Brown

Robert J. Harvey, Chairman

Psychology

(ABSTRACT)

When criterion related validity studies are not feasible due to statistical power problems, a number of alternate methods for validation are available. Indirect validity offers a few of these strategies. One of these is synthetic validity which, in its most generalized form, is the linking of predictor items or tests directly to a job's elements. Three main operational definitions exist for synthetic validity, but only one, job component validity (JCV), actually specifies minimum scores for a given job. In an era of minimum cutoff scores, having an empirical procedure to determine requisite scores can be quite valuable. JCV can be used to predict personality test scores for jobs, scores which can be used for selection as well as other personnel decisions. While much is known about the process of predicting cognitive ability test scores as well as the efficacy of such predictions, little is known about the use of JCV with personality test profiles. This study

sought to examine the efficacy of the JCV model when applied to a personality test, namely the Myers-Briggs Type Indicator, using the work dimension scores of the Common-Metric Questionnaire (CMQ) as predictors. Results indicated that the JCV model is applicable to non-ability tests such as the MBTI. Moreover, prediction of personality test scores is not moderated by the type of work performed. Finally, results supported the MBTI's claim that its scales synergistically interact with each other.

Acknowledgments

Without my advisor, R. J. Harvey, there wouldn't be a place for an acknowledgments section, so the first round of thanks must go to him. His re-direction of my wayward conclusions spared me from public embarrassment on multiple occasions. In essence, without R. J. there would be no thesis.

Thanks also go out to my committee members, Sigrid Gustafson and Roseanne Foti. Their helpful comments and suggestions served to spur me in new directions with much of this study's logic. Much of this paper's polish and insight belongs to them.

I can proceed no further without mentioning my wife, Amy Lynn Hill Brown, whose support and willingness to wait to do "anything as long as it's outside of the house" until a "later" that would never arrive was invaluable. Now, we can do those fun things like long walks together or the pumpkin festival, if only it wasn't so cold.

Finally, I must acknowledge the support provided by my parents all through the years. Anyone who received the amount of love and support that they offered could have finished a thesis, probably in less time.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iv
OVERVIEW OF VALIDATION	1
SYNTHETIC VALIDITY	6
The J-Coefficient	8
Criterion Oriented Synthetic Validity	12
Job Component Validity	18
THE GRAVITATIONAL HYPOTHESIS	27
PERSONALITY TESTS AND INTEREST INVENTORIES IN SELECTION...	32
THE MYERS-BRIGGS TYPE INDICATOR	35
THE COMMON-METRIC QUESTIONNAIRE	41
HYPOTHESES IN THE PRESENT STUDY	48
METHOD	49
Subjects	49
Unit of Analysis	49
Data Manipulation	50
Analyses	51
RESULTS	55
Factor Analyses	55
Job Component Validity Analyses	56

Moderation Analyses	58
Synergistic Interaction Analysis	61
DISCUSSION	62
REFERENCES	71
Figure 1. The J-Coefficient Concept	79
Figure 2. The Hollenbeck and Whitener (1988) Synthetic Validity Model	80
Figure 3. Pictorial View of the Analytical Strategy for the Planned Comparison	81
Figure 4. Pictorial View of the Analytical Strategy for the Unplanned Comp.	82
Figure 5. Scree Plot of Eigenvalues from the Reduced Correlation Matrix of the 80 CMQ First-order Scales	83
Figure 6. Scree Plot of Eigenvalues 2-80 from the Reduced Correlation Matrix of the CMQ First-order Scales	83
Figure 7. Scatterplot of Predicted and Actual Extraversion/Introversion Percentages from the 21 Predictor Model	84
Figure 8. Scatterplot of Predicted and actual Sensing/Intuition Percentages from the 21 Predictor Model	84
Figure 9. Scatterplot of Predicted and actual Thinking/Feeling Percentages from the 21 Predictor Model	85
Figure 10. Scatterplot of Predicted and actual Judgment/Perception Percentages from the 21 Predictor Model	85
Figure 11. Graph of Marketing/Sales and Percent of Extraverts	86

Table 1. PAQ Attribute Ratings to Derive Job Requirements 87

Table 2. Mean Test Scores of GATB Intelligence 88

Table 3. Gravitation to Jobs Commensurate with Ability 89

Table 4. Loadings of the 6 Factor Higher-order Solution 90

Table 5. Intercorrelations of the 6 Second-order Factors 95

Table 6. Loadings of the 21 factor Higher-order Solution 96

Table 7. Intercorrelations of the 21 Second-order Factors 103

Table 8. Multiple Correlations, R²'s, and Associated p-values for MBTI Scales . . . 104

Table 9. Regression Weights and Associated p-values for the 21 Predictor Model
for all Four MBTI Scales 105

Table 10. Regression Weights and Associated p-values for the 6 Predictor
Model for all Four MBTI Scales 110

Table 11. Regression Results for Interaction Analyses for 6 & 21 Pred. Models . . . 113

Table 12. Regression Results for Analysis of Gender Within 6 & 21 Pred. Models. 114

Table 13. Residual Group Means and ANOVA Results for Each Scale 115

Table 14. Actual and Predicted MBTI Types for Each Job Title 119

Table 15. Agreement Rates Between Predicted and Observed MBTI Types 125

VITA 126

Overview of Validation

In the field of personnel selection, validation is the process of determining the accuracy of inferences made from selection test scores (Binning & Barret, 1989). This can be accomplished using a variety of techniques. The most common method involves the use of a criterion oriented strategy, whereby scores on a predictor are correlated with scores on a criterion. While this strategy may be both powerful and simple, it suffers from a variety of limitations and problems. In particular, it is not a very powerful method for organizations with small sample sizes due to the vagaries of sampling error. In a power analysis, Schmidt, Hunter, and Urry (1976) showed that under typical and realistic conditions (sample size of 68), the practitioner has only a 50% chance of detecting a significant validity coefficient, given that the test is a valid predictor ("true" validity is .45). As the sample size drops, chances for detection also fall. By computing sample sizes needed for minimum power (.50) across a variety of selection ratios and criterion reliabilities, they demonstrated that criterion related validity studies are "technically feasible" much less often than is commonly assumed (p. 480).

EEOC guidelines and good business practice dictate that as a job changes in its content, old validation efforts should not be used as evidence for validity of a selection procedure. Due to this, for jobs that change rapidly, selection procedures would necessarily have to be validated with every change in the content of the job. A similar problem exists when an organization has, in total, a moderate to large number of people employed. These people, however, are distributed into many job titles with only moderate overlap in job activities and a relatively small number of positions per job.

A number of solutions exist to combat these deficiencies. For openers, an organization could eschew the criterion oriented strategy altogether, in favor of a content validity strategy. In a content validity effort, no correlations are computed between predictor performance and job performance. Rather, the link is made by using tests that have an isomorphic relationship to behaviors on the job. Experts are used to determine the degree of relation between a test item and actual job behaviors. Problems arise when the cutoff scores are challenged. Because no traditional validity coefficient is computed, no one knows with certainty that an x point difference on the test leads to a y point difference on the criterion of job performance.

Construct validation provides another option to the criterion validity model. The general thrust of the construct validation effort is “doing good science in the sense of understanding both predictor and performance domains” (Barret, 1992, p.14). In the construct validation paradigm, a thorough understanding of the job is needed to identify the domains of work with the goal of specifying underlying constructs needed for successful performance. Tests are then selected or devised to tap each one of the constructs, with the goal of minimizing construct irrelevant variance (i.e., predictor contamination). Naturally, the same amount of care and caution is used when constructing the performance appraisal instrument, with specific dimensions for each construct. Using an information processing approach to construct validation, Barret (1992) found a .80 correlation between predictor and criteria designed to tap word processing performance. To date, very few construct validation efforts have been undertaken, mainly due to their costly, time consuming nature. Additionally, correlations

between the predictor and criterion scores still have the same power problems as are present with criterion related validity.

Under the general rubric of indirect validity, even more alternate methods exist for validation. They can be classified as either synthetic validity or validity generalization (VG) techniques. In both cases, they seek to infer the validity of a selection procedure without following the traditional criterion or content related strategies. Their methodology and theoretical thrust, however, are quite dissimilar. Validity generalization, in its simplest form, seeks to minimize and gloss over differences among jobs. The process begins with the identification of relevant criterion related validation efforts previously conducted. Drawing on these studies, VG makes a statistical determination of the presence of moderating variables (Schmidt & Hunter, 1977). These moderators can be differences in the structure of the job or differences in the test used to select applicants for the job. Ideally, one would like for all of the tests used to be the same, or at least load onto the same constructs. The technique then entails corrections of the observed variance among the validity coefficients for four main statistical artifacts (sampling error, unreliability in the predictor, unreliability in the criterion, and range restriction in the dependent variable). Traditionally, if one could eliminate 75% of the observed variance among the validity coefficients, then one can conclude that no moderators are present and “safely” generalize the evidence of validity from the reference studies (Schmidt, Hunter, & Pearlman, 1981). The estimate for the validity coefficient is taken to be the mean of the distribution of validity coefficients and is usually corrected for attenuation due to range restriction and unreliability in the criterion (Hunter & Schmidt, 1990).

Problems with the VG procedure exist on technical, theoretical, and legal grounds.

The technical qualms relate to the statistical artifact corrections. Specifically, the procedure has a number of stages where assumed distributions are employed. This occurs when corrections are made for criterion unreliability and restriction of range. Few validation efforts measure criterion reliability and range restriction, and even fewer report it. Due to this, distributions of these artifacts must be constructed based on what past research indicates might be the correct pattern. In both of these cases, it is possible to overcorrect for unreliability and range restriction, thus predicting more artifactual variance than actually exists. This could lead to two incorrect conclusions. First, one might obtain the recommended level of at least 75% of the distributional variance and conclude that situational moderators are not present and validity can be safely generalized. Instances of this are on display in almost every VG study in which roughly 20% of the results show a case where the predicted standard deviation (predicted on the basis of the big four experimental artifacts) exceeds the actual standard deviation (for an example, see Pearlman, Schmidt, & Hunter, 1980). That is, VG has clearly overcorrected for artifactual variance among validity coefficients. It is not illogical to assume that this same pattern of overcorrection exists in other, less obvious cases as well. Second, this overcorrection of artifactual variance via the assumed distributions of criterion reliability and range restriction also leads to inflated estimates of the population (or overall) validity for a selection procedure. That is, when the mean validity coefficient is corrected for range restriction and criterion unreliability, these corrections could be too large, resulting in a biased estimate of the “true” validity.

James, Demaree, Mulaik, and Ladd (1992) raised the theoretical concern that the reliability of the criterion and the validity coefficient were not independent. In what they term the “restrictiveness of climate model”, situations exist where the variability in performance is low due to organizational constraints (high standardization of work procedures, non-merit based reward systems, programmed approaches to problem solving, low levels of individual autonomy, low job involvement, low achievement motivation, etc.). In cases such as this, individual levels of achievement are not likely to widely vary, meaning the hypothetical utility of a selection system cannot be high. Whenever variance in the range of performance is restricted, reliability in the measurement of performance is also necessarily low. The ramifications of this scenario are the same as before. Not only will one derive an overall validity coefficient that is greater than the “true” validity, but more importantly, because the criterion reliability is so low, corrections made on the variability of the distribution of validity coefficients will again be too large, precluding the chance of concluding that situational moderators are indeed present. This results in situations in which one incorrectly generalizes validity from the reference populations.

As mentioned, VG has not fared well in the courts (Van Aken vs. Young, 1982; EEOC vs. Atlas Paper Co., 1989). This is a sizable problem, as any validation procedure that cannot withstand legal scrutiny is functionally useless. In both of its direct challenges the courts wanted the VG effort to focus its analyses on validation studies from organizations that are similar to the one in question, with similar predictors and similar populations. In addition, the courts also wanted an on-site job analysis. In essence, the tie

between the reference validation efforts and the target organization must be firmly established.

Synthetic Validity

Synthetic validity, another form of indirect validity, faces none of these problems. Regardless of the specific technique, synthetic validity is rooted in the details and structure of the individual job. That is, instead of ignoring differences among jobs (and using a global appraisal of job similarity), synthetic validity measures the differences via a detailed examination of the work situation. In its general form, synthetic validity refers to the process of “inferring of validity in a specific situation from a logical analysis of jobs into their elements, a determination of test validity for these elements, and a combination of elemental validities into a whole” (Balma, 1959, p.395). Even more generally, synthetic validity is the quantitative linking of the predictor(s) to a given job’s elements.

To reiterate, synthetic validity is ideal for situations in which: there are too few employees in the organization to conduct a traditional criterion related validation effort with decent statistical power; an organization which has a large number of employees in total scattered into a number of highly dissimilar jobs; and jobs of a dynamic nature whose content changes significantly over time.

Synthetic validity capitalizes on a common sense logic of work behavior and selection. That is, since job behaviors are multi-dimensional, one should be able to validate predictors against each one of these dimensions. Moreover, many jobs within an organization share common elements, so the individual predictors should hold across

different jobs sharing that element. Synthetic validity bears an interesting relationship with criterion development. If one is using multi-dimensional criteria, it naturally makes sense to employ a predictor battery, with a predictor designed to tap each criterion. This allows for synthetic validity to be a useful strategy whenever multi-dimensional criteria are employed, regardless of the organization's size (Balma, 1959; Guion, 1965). Synthetic validity can actually be very similar to standard criterion related validation of a battery of tests. The key difference is that in synthetic validity test selection and validation is done at the individual element level of a job, not a global assessment of the job.

The term synthetic validity is actually a misnomer. As Guion (1965) clarifies, "validity cannot be synthesized or created, it can only be observed and reported. It is not, however, validity which is synthetic; it is the valid test battery which is synthesized. 'Synthetic validity' should then be considered a convenient shorthand statement for 'synthesis of a valid test battery'" (p. 62).

The origins of synthetic validity date to C. H. Lawshe (1952), who conceptualized the process as the psychological equivalent of the industrial engineering time and motion studies of the halcyon days of Taylorism. In these studies, one of the goals was to break basic human motions down to their smallest levels in order to develop standard times for these individual motions. This allowed one to combine the individual standard times into a synthesized whole whenever a new behavior appeared (Lawshe, 1952; Guion, 1965; Jeanneret, 1992). The development of the psychological equivalent of this process was slow, mainly due to the fact that much of the original work was for the federal government, and never published in psychological journals.

Three main operational definitions exist for the synthetic validity. Specifically they are the J-coefficient, criterion oriented synthetic validity (no specific name exists), and the job component validity (JCV) model.

The J-coefficient

Introduced by Ernest Primoff in 1955, the J-coefficient is a procedure that can accommodate each one of the traditional validation strategies: criterion, content, and construct. However, due to the fact that using this procedure requires large sample sizes for the construct or criterion strategies, they are seldom used, meaning the J-coefficient is in actuality another method for content validation. The J-coefficient “provides a numerical estimate of the correlation between a selection test and job performance” (Hamilton, 1981, p. 807) and is generated by the formula $J = \sum (B_{xe} R_{ye})$, where B_{xe} represents the test-element weights and R_{ye} the job performance-element correlations (Mossholder & Arvey, 1984). For each element, the J-coefficient multiplies the degree to which a test taps that element (B_{xe} , which can also be expressed as a correlation; R_{xe}) by that element’s importance to overall job performance (R_{ye}). One then simply sums across elements to derive the coefficient. A cursory glance at this formula shows that at no point are applicants entered into the process; one is simply rating the importance of job elements and the relevance of tests purporting to measure them.

The first step naturally involves the identification of the job elements. Estimation of B_{xe} then can proceed via a variety of methods, including expert, supervisor, or self ratings. Since self ratings were found to be weak in the results area (Hamilton &

Dickinson, 1978, cited in Mossholder & Arvey, 1988), supervisory ratings were used. Supervisors were eventually found to be not completely knowledgeable about job content, leading to the use of experts, namely psychologists. The use of expert ratings appears to be the most common technique for the estimation of B_{xe} . The experts first rate the relevance of each test item to each job element; the proportion of items judged relevant then becomes the estimate of test relevance for each element (Snedecor & Cochran, 1967, cited in Hamilton, 1981). It is worth mentioning that this particular expert rating procedure allows for the situation where one test is used for selection, or different tests are used for each job element (consistent with content validity methodology). When one test is used, the expert ratings signify what portion or how much of the test is relevant to the job element. However, when a separate test is used for each element, the expert ratings indicate how well the test taps the element.

As mentioned earlier, facets of the J-coefficient procedure can be executed with criterion oriented data. The estimation of the job performance-element (R_{ye}) correlations can indeed be done this way by collecting detailed performance data for employees' elemental and overall performance (never an easy task), then correlating the elemental performance with overall performance. If one has the sample size to do this, then perhaps one should proceed with the full criterion strategy, not a synthetic one. Given that the criterion strategy for estimation of R_{ye} is not feasible, two judgmental methods for estimating R_{ye} exist. The first involves a panel of judges who rate the salience of each job element to overall performance (Mossholder & Arvey, 1984). The ratings are combined and scaled on a percentage agreement basis where 100% indicates that all raters judged

the element to be of maximal importance to the job. Classic policy capturing methods may be used to determine element-job relationships (Hamilton & Dickinson, 1978, cited in Mossholder & Arvey, 1988). Policy capturing entails having raters assign overall ratings to carefully manipulated profiles of employee performance. Multiple regression techniques are then used to determine the weight raters have been placing on various aspects of job elements. Figure 1 provides a pictorial representation of the J-coefficient concept (Hamilton, 1981, p. 807).

The J-coefficient model makes a number of assumptions. The most important assumptions involve the linear relationship between job elements and performance and a complete job analysis to identify all of the crucial elements of the job. To the extent that these assumptions are violated, the J-coefficient becomes an imperfect approximation of the validity coefficient (Hamilton, 1981). Although it is unrealistic to expect that all of the elements of a job will be identified, alternate formulae for the calculation of the J-coefficient exist to correct for imperfect identification (similar to correcting a validity coefficient for unreliability in the criterion and predictor). Hamilton's formula (p.807), implementing this correction for incomplete identification, is:

$$J_{xy} = \frac{R_{xe} R_{ee}^{-1} R_{ye}}{(R_{xe}' R_{ee}^{-1} R_{xe})^{1/2} (R_{ye}' R_{ee}^{-1} R_{ye})^{1/2}}$$

R_{xe} and R_{ye} are both the same as before (test-element weightings and element-job performance weightings, respectively) and can be estimated using the same methods as before. R_{ee} (a holdover from Primoff's, 1959, original formula) represents the matrix of intercorrelations among job elements and can be constructed from judgments of interrelationships by experts or from the actual correlations of elements. Hamilton's J-coefficient formula is analogous to the classic formula used to correct for attenuation due to unreliability in the predictor and criterion. As a point of reference, the formula to correct for attenuation is:

$$r_{xy} = \frac{r_{xy}}{(r_{xx})^{1/2} (r_{yy})^{1/2}}.$$

In terms of statistical accuracy, when its pieces are calculated with empirical methods (i.e., using the correlations between the employees' elemental and overall performance to estimate R_{ye}), the J-coefficient will be identical to the classic r_{xy} (Dickinson, 1979, cited in Hamilton, 1981). One might wonder in a theoretical sense what it is exactly that the J-coefficient is doing. If the job elements are identified in terms of knowledge, skills, and abilities, then the J-coefficient quantifies the evidence for content validity (Hamilton, 1981). However, even when the elements are identified in terms of behaviors and observables, a case can be made that the J-coefficient still represents the same quantified estimate of content validity since all one is doing is moving the inferential leap (translating behaviors and activities into abilities) from the job element identification phase to the B_{xe} (test-element relations) portion of the formula.

Empirical tests of the J-coefficient procedure have been encouraging. Hamilton and Dickinson (1978, cited in Mossholder & Arvey, 1984) demonstrated that when supervisory ratings or policy capturing methods were used for the estimation of the test-element (B_{xe}) and element-job performance (R_{ye}) links subsequent J-coefficients reproduced patterns of test-criterion validities. As an example (Primoff, 1959), validities estimated by the J-coefficient procedure ranged between .23 and .62.

Finally, Trattner (1982) introduced another method to estimate the calculation of the J-coefficient. Not only did the actual formula depart from the previous two, this method also departed from the traditional focus of the J-coefficient (content validity) and adopted a criterion oriented approach to the estimation of the test-element links. Trattner's method involves using individual criterion related validity coefficients for each element, then aggregating them. All of this would be fine if not for the fact that "this requires large numbers of research participants in occupations that have detailed performance requirements" (p.395). As previously mentioned, a synthetic validity method that requires large numbers of participants defeats the main reason for employing this validation strategy: applicability to situations with small sample sizes. This method does, however, foreshadow much of the theoretical aim (i.e., to make the process less subjective and more criterion oriented) of the other synthetic validity strategies.

Criterion Oriented Synthetic Validity

Robert Guion (1965) provided the first example of a criterion oriented synthetic validity strategy with his example of its implementation in a small business setting. The

procedure was quite simple, starting with a job analysis that identified seven elements across nine job titles. Clearly, there was overlap of elements across jobs. Great care was taken to collect criterion ratings for each element relevant to a given job. Tests were then chosen *a priori* for their relationship to the given job elements. Finally, employees were administered the relevant tests (a concurrent design in a firm of only 48 employees).

For analysis, Guion used a simple procedure where the criterion and test scores were divided at the median. For each test, the percentage of the employees above the median who were also above the median in performance ratings (for that job element) were identified as well as the percentage in the lower scoring group who were designated as in the top half of the ratings. The significance of the difference was tested using a simple monograph. “Quite arbitrarily” (p.57), he retained the two best predictors for each element and computed their multiple correlation (ranging from .31 to .82; median = .59). The tests were then cross validated on a sample of 13 new employees, yielding 76% correct classifications.

Guion’s example of synthetic validity was the first to empirically link the predictor scores to a criterion. In effect, it serves to validate a selection test for each element. While many multiple correlations are computed between the two best tests and performance ratings for job element, no overall measure of effectiveness is computed.

In 1988, John Hollenbeck and Ellen Whitener published a refinement of Guion’s procedure. The chief difference between this and any other procedure is the order of the aggregation and validation steps. With Guion’s procedure, one validates each test, then aggregates across tests within element or job. The Hollenbeck and Whitener procedure

requires one to aggregate first, then validate, allowing for greater statistical power.

Comparing this method to the standard criterion validation procedure, they provide this example: in order to obtain a statistical power of .80 to detect a correlation of .30 at the .05 significance level, one would need 84 employees for that job (either hire 84 employees for the same job title over a given time period, or, concurrently, have 84 employees employed within the same job title). By using a synthetic validity procedure and aggregating before validating, one only needs 84 employees within the organization whose jobs have some overlap of elements.

The procedure is almost as straightforward as the Guion method. One begins with a job analysis with the goal of identifying the elements comprising work behavior. Hollenbeck and Whitener recommend using a standardized worker oriented instrument that produces dimensional scores for each job title. While dimension scores are not conceptually the same for all worker oriented job analysis instruments, they are usually based on the frequency and importance, or criticality, of the behaviors encompassed by the dimension. For each dimension present in a given job title, weights are computed on a percentage basis. That is, if a job title is rated as having a .67 score on dimension one, a .50 score on dimension two, a .25 score on dimension three, and no other dimension with any score above zero, then the weights for job title are .47 (due to the fact that its element's score was 47% of the total element scores) for dimension one, .35 (35%) for dimension two, and .18 (18%) for dimension three. These weights, specific to each job title, will be used for all of the dimensional criterion ratings and test scores for each employee within the same job title.

Based on the job analysis results, tests are chosen to tap each job dimension. Most proprietary job analysis instruments suggest tests that can be used for their various dimensions. Incumbents or candidates are administered all the tests relevant to their specific job. As with any validation effort, the collection of accurate criterion data is essential. Appraisal forms must be constructed for each job title that include all of its relevant dimensions. As with the Guion method, ratings must be made for each dimension of every employee's job. One potential aid in collecting useful criterion data is the classic technique of instructing the raters that the ratings they will provide will be used strictly for the purpose of the validation effort.

Once all of the data are collected, two person by element matrices are constructed, one for each person's test scores and one for their performance ratings. In both matrices, the data are standardized on national norms (in the case of the selection tests) or local, within company, norms (for the criterion data). The z-scores for both are multiplied by the weight of each dimension. That is, for dimension one of job x , both the test score measuring that dimension and the performance rating for that dimension will be multiplied by the same coefficient (elemental weight). Once these matrices are constructed, one simply sums across dimensions for each person in the predictor matrix, resulting in a predicted criterion value (in z-score form) based on the test results. Similarly, the criterion matrix is also summed across dimensions resulting in an actual composite criterion score for each person in z-score form. Figure 2 offers a matrix algebraic representation of the aggregation-process (Hollenbeck and Whitener, 1988; p. 539). Finally, the predicted criterion values are correlated with the actual criterion composite

value using the traditional Pearson formula. In a practical demonstration of this procedure in a chemical manufacturing company with 83 employees, Hollenbeck and Whitener calculated a .22 validity coefficient ($p = .05$).

The first issue one may raise with this procedure is whether each dimension's contribution to the composite value (predictor or criterion) is actually unique. As long as one uses a job analysis instrument that has orthogonal dimensions, then each contribution will be unique. This is the advantage of using a well researched job analysis instrument that publishes data indicating the correlations among its dimensions (such as the CMQ). A second query with the procedure is whether it produces a biased estimate. Hollenbeck and Whitener tested this by inserting data with known properties. For example, when the true correlation between predictor and criterion for all of the dimensions is zero, one inserts random numbers into the matrices (since the weights remain constant for each dimension, they change nothing). In this case, the procedure yields a correlation of zero. In the other extreme, where all of the dimensions are correlated .30, one inserts two lists of numbers correlated at .30 into the matrices. The resulting correlation will also be .30. Finally, consider the case where some of the dimensions are uncorrelated while the rest are correlated at .30. The resulting correlation will be bounded by zero and .30 and will be determined by the weights of the dimensions. That is, if dimensions with the .30 correlation are weighted heavier than the dimensions with the zero correlation, then the resultant correlation will be closer to .30 than zero. The reverse also holds true.

One potential alternative explanation deals with the possibility that the correlation reflects job placement differences and not job performance differences. That is, the

average level of performance by incumbents within a given job could systematically vary among job titles along with the average test scores of those incumbents. This possibility was tested by aggregating across subjects within a given job, resulting in a mean score on the predictor and the criterion for each job title ($n = 51$). The resulting coefficient was .06. If in fact it was the case that the relatedness was occurring at the job title level and not the individual level, we would not only expect a significant correlation, but a very sizable one since we are correlating mean scores, and thus capitalizing on the stability of the mean (James, 1982).

A final issue with the Hollenbeck and Whitener (1988) procedure concerns the degree of overlap among elements and jobs. Clearly, not all jobs will have many elements in common with other jobs. They may only have one or two out of ten, resulting in dimension scores that also have little or nothing in common. This would not be out of the question when blue collar jobs and a sole white collar job are included in the same validation effort (other, less obvious, more realistic cases may apply). A situation such as this challenges some of the logic of the procedure because one will have a situation in which a job has a large number of unique elements causing one to lose out on the statistical power of validation using a collection of jobs that share elements. In a sense, one will be validating an almost separate and distinct job with other jobs who share a higher degree of association and similarity. Hollenbeck and Whitener do not place a restriction the identification of job dimensions (by saying, for example, that an dimension must be shared with another job to “count”), for doing so would ignore a facet of the job and alter the true validity of the procedure. The solution to this situation remains

unanswered other than by decreasing the level of specificity at which the job dimensions are defined (fewer potential dimensions results in a greater chance of shared dimensions). This, however, is not an advisable solution since one will be creating one problem (by perverting the job analysis procedure) to fix another.

Job Component Validity

The use of the JCV procedure avoids many of the previously discussed problems with other approaches to synthetic validity. First, it is an empirically based procedure, not a judgment oriented content validity technique (as is the J-coefficient). In addition one does not have the potential problem of non-overlap among the work dimensions that the Hollenbeck and Whitener (1988) procedure faces. Finally, since one is not computing an overall validity coefficient, one does not face statistical power problems. Of course, the lack of an overall validity coefficient can also be viewed as a shortcoming.

Perhaps the main advantage of the JCV model is that it establishes minimum competency standards. In a legal era of the minimum cutoff score, having an empirical determination of minimum levels of a given aptitude, free from judgmental biases, can be valuable. In essence, the JCV model seeks to empirically link the knowledge, skill, ability domains to job characteristics.

While credit for the JCV model is typically given to Ernest McCormick (1959), one of the developers of the Position Analysis Questionnaire (PAQ; McCormick, Jeanneret, and Mecham, 1972), a reanalysis of Lawshe and Steinberg's 1955 article on synthetic validity shows it to be the first JCV procedure. While the Lawshe and Steinberg

procedure is more judgmental in nature than the state of the art in JCV, its logic is the same. Lawshe and Steinberg attempted to find out whether jobs with larger amounts of similar critical operations (i.e., a job where at more than one point in the successful exercise of job duties an employee must tap the domain of “spelling” versus a job where correct spelling is only needed for one operation) are staffed by people who, in fact, are better at that critical operation. They started with a job analysis, in order to identify the functions of each job in an organization. For each job element, judges then rated the criticality of a given aptitude for the successful exercise of that element (dichotomously rated: “critical” or “not critical”). The domain of aptitude was simply assessed by the subtests of a clerical test battery (naturally, these were all clerical jobs). Finally, the test was administered to all of the incumbents. Expectancy charts were constructed showing the probability of exceeding the median score on a given subset of the test for jobs broken down by number of critical operations. For example, incumbents in jobs judged to have four or five critical operations involving spelling had an 84% chance of exceeding the median score of the spelling subtest. However, incumbents in jobs with only two or three critical spelling operations had only a 58% chance of exceeding the median score. This trend continued for each of the job elements and their associated critical abilities (usually more than one for each element).

This study, as with all JCV techniques (regardless of the level of sophistication), rests on one fundamental assumption known as the gravitational hypothesis. The assumption is that “in general, those incumbents who are currently employed in positions, by and large, possess those qualities which the positions demand of them. The rationale is

that if an individual cannot or does not ‘deliver’ on the job, either the employee is replaced or the job assignment is modified” (Lawshe & Steinberg, 1955). The tenability of this assumption will be discussed later in detail.

As previously mentioned, the creators of the PAQ (McCormick, Jeanneret, and Mecham, 1972) refined and popularized the JCV model. The methodology behind the model is explained in their 1972 monograph.

It can conceivably be argued that if a given kind of work activity occurs in two or several different jobs, the human requirements for such jobs would be comparable insofar as that common activity is concerned. Let us consider, for example, the aptitude requirements of jobs. If by hook or by crook one can identify - in the case of a sample of jobs that do have a particular communality of work activities - the corresponding aptitude(s) that are so required, it would then seem reasonable to assume that the corresponding aptitude(s) would likewise be required in other jobs that have the same communality of work activities. This analogy, repeated for other types of activities and their corresponding aptitudes, would then presumably make it possible to “build up” the total aptitude requirements for any given job synthetically. (p.358)

Notice the hallmark of the JCV procedure. We are not synthesizing validity coefficients, but rather, aptitude requirements.

Clearly, the first step in a procedure such as this is a job analysis. Here, for the first time, we see the critical need for a job analytic technique that is standardized, with a common metric, that allows for meaningful comparisons across jobs. It was for these reasons that the PAQ was developed. The creators of the instrument decided to focus their items on the behaviors of the job. That is, they were interested in attempting to quantify what one does in order to accomplish the tasks of the job. In essence they focused on the generalized behavior of the workers (not the specific outcomes of the behaviors), giving rise to the euphemism “worker oriented.” The defining characteristic of a worker oriented instrument is its use of reasonably specific descriptions of many kinds of behaviors. McCormick et al. drew upon their vast experience as job analyst experts to insure that the range of behaviors covered in the PAQ sufficiently spanned all potential job behaviors. Their final version of the instrument contains 194 behavioral elements and work context factors (working conditions, work schedule factors, etc.) grouped into six general categories (information input, mediation processes, work output, interpersonal activities, work situation and job context, and miscellaneous factors). It should be noted that in worker oriented job analysis, the term element has a different meaning than it does in the context of task analysis. Here, the term element is used to describe the individual general work behaviors rated in an instrument. Each element includes scales rating factors such as importance, amount of time spent, applicability, and extent of use.

Factor analysis of a sample including over 500 jobs resulted in a 27 factor solution (later analyses found factor solutions ranging between 27 and 32 factors). These factors represent the job dimensions. That is, they are groupings of similar elements. The

weighting of the elements, combined with the scores on them, results in job dimension scores. It is these job dimension scores, comparable across jobs drastically heterogeneous in content, that are necessary for the JCV technique.

Once the correct tool was developed, the work involved with the JCV technique lost much of its complexity. PAQ results collected from a variety of jobs are matched with standardized test scores (in the case of the McCormick et al., 1972, study the test was the GATB). In cases where more than one profile (collection of job dimension scores) is collected for the same job (not often the case with the PAQ as its post-graduate reading level precludes many incumbents from completing the instrument), the mean profile is used. Similarly, the mean profile of GATB subtest results are also used. It is important to note that the PAQ and GATB results typically do not originate from the same job. Rather, they are matched on the basis of similarity, operationally defined by job title. Once the mean scores are matched, a multiple regression analysis is executed in which the PAQ dimension scores constitute the independent variable and the GATB scores become the dependent variable (separate analyses are conducted for each subtest of the GATB). The resulting regression equations allow for future prediction of GATB scores based on the PAQ results for a given job.

Because one has no validity coefficient to examine as a measure of effectiveness, the efficacy of the JCV procedure is determined by the correlation of the predicted mean test scores with the actual mean test scores for a wide range of jobs. Using the GATB tests, McCormick et al. (1972) reported correlations ranging between .59 and .80 (median $r = .71$) for the nine GATB subtests. Mecham (1977, cited in McCormick, DeNisi, &

Shaw, 1979) replicated the JCV procedure with the GATB data and a new set of 163 jobs, resulting in a wider range of correlations (.30 to .83 for the nine subtests) but a similar median correlation (.73). McCormick, DeNisi, and Shaw (1979) repeated the procedure, substituting commercially available tests for the GATB. Correlations between the predicted mean values and the actual mean scores for each job ranged between .52 and .75 with a median correlation of .71. Consistent with previous results, McCormick, Mecham, and Jeanneret (1989, cited in Jeanneret, 1992) reported GATB correlations in the range of .24 to .78 with a median of .69, in an analysis of 460 jobs.

That there are strong correlations between predicted scores and actual scores is encouraging, however not surprising. Median correlations around .70 should be viewed with a skeptical eye. As mentioned before, whenever one correlates mean scores (in this case, mean scores are correlated with predicted mean scores), one can potentially capitalize on the stability of the mean, allowing for spuriously high correlations (James, 1982).

The JCV strategy as a method of validation has successfully faced a court challenge. In the case of *Taylor vs. James River Corporation* (1989) the court affirmed the use of the technique. Moreover, the plaintiff's expert witness "conceded" that it was sound. However, because the case was argued in 1989, the court used the Ward's Cove rules which require a less stringent standard for business necessity and do not shift the burden of proof onto the defendant once the prima facie case has been made by the plaintiff. That is, the defendant (James River Corporation) did not have the burden of proving that the selection procedure was based on a legitimate business reason; it was the

plaintiff's job to show that it was not. In addition, the standard that the defendant had to meet was much lower than it was for cases previous to the Ward's Cove ruling (and since the 1991 Civil Rights Act). A legitimate business reason (or concern) for the selection procedure is much easier to show than a business necessity. Due to the fact that the 1991 Civil Rights Act restored the shifting burden of proof model, this legal affirmation of the JCV model should be considered tentative.

In developing the PAQ, McCormick et al. (1972) set out to describe jobs in terms other than general worker behaviors. They also attempted to describe each element in terms of abilities needed to perform the behavior. After an exhaustive search, they created a list of 68 worker attributes that could be applied to 165 of the 176 items in the original PAQ (the list was later expanded to 76 by Marquardt & McCormick, 1972). The "attributes" consisted of aptitudes (verbal comprehension, perceptual speed) or situational descriptors (variety of duties, personal risk). Approximately ten experts (some of whom were graduate students) rated the relevance for each attribute to each element. Reliabilities of the ratings were generally acceptable (.80's). Median ratings were used for each attribute for each element. Two matrices could then be constructed with these attribute ratings. One is a profile of all attributes for each element, while the other is a profile of a given attribute across elements. This second matrix is valuable since it allows one to see the amounts of every attribute for a given job. Table 1 provides an illustration of these two PAQ attribute dimension profiles (Jeanneret, 1992, p. 85). Factor analysis of the attribute data revealed a 21 factor solution with easily interpretable factors (visual input, general body handling activities, information processing).

The use of the attribute profile data bears a resemblance to the work of Fleishman's "ability requirements approach." The Fleishman approach involves a similar profile of abilities which account for job performance. One seeks to find jobs with similar ability requirements, allowing one to predict performance based on the amount of requisite abilities possessed (Fleishman, 1967).

As with the job dimension profiles, McCormick et al. (1972) used the attribute profile data to predict mean GATB scores. Their results were almost exactly the same (median $r = .70$ vs. $.71$ for job dimension scores). One examination of the attribute profile concept was offered by Sparrow (1989) who compared alternative methods of computing the profile scores for a given job. The study demonstrated that the additive approach, which consists of summing the weights of the attributes for each element checked (regardless of the element's weight), was judged superior (more correct classifications) to the critical behavior method (only using attributes for elements that had been assigned the highest rating in the job analysis phase) and cross-product method (score of each job element is multiplied by the weighting of the corresponding attribute) for generating attribute profiles.

One potential problem with using the attribute profile scores is that one has doubled the amount of human judgment in the validation process as well as further removed the job analysis process from the observable stage. As the behavioral specificity of a job analysis procedure increases, it becomes more easily verifiable. This also pleases the courts. Using the attribute profile scores as the basis for validation is a step in the less

verifiable direction. This, in conjunction with the serious questions of holistic ratings on both professional and legal grounds, most likely results in the rare use of attribute profiles.

In a test of JCV using the attribute profile scores, Sparrow, Patrick, and Barwell (1982) used the JCV procedure to identify tests that should be related to performance. Using the additive procedure to calculate attribute profile scores for the job of “setter” (in a plastics injection molding company), they selected commercially available tests to tap the top seven attributes (abilities were ranked from highest to lowest in terms of their weightings) all of which were of an ability nature (long term memory, near visual acuity, etc.). Due to overlap of the test domains and ability domains, only five tests were needed to tap the seven abilities. After collecting criterion data, they correlated test scores with the overall ratings for the fourteen employees. All correlations were large (absolute values ranging from .29 to .89) but only two were significant due to the small sample size involved. Multiple regression analysis showed that the top two predictors accounted for 91% (adjusted for the number of variables in the equation) of the variance in criterion ratings. The other five predictors failed to add variance. The results were undoubtedly inflated by the small sample size of the study ($n=14$). Moreover, the criterion data were not of a true performance quality (or quantity) type, but actually consisted of supervisors’ ratings of employees’ ability in several specific areas. These same areas (mechanical) were tapped by one of the predictors (a test of mechanical comprehension) which had the highest single correlation with ratings. However, even after discounting the magnitude of the results, it appears that the JCV procedure is an effective strategy for selection.

The Gravitational Hypothesis

This “test” offered by Sparrow et al. (1982) seems to hint at the veracity of the fundamental assumption of the JCV model, the gravitational hypothesis. To reiterate, when using this form of synthetic validity one is assuming that jobs differ in their requisite amounts of ability and that people somehow gravitate towards jobs commensurate with their ability. By targeting the mean level scores for a job, one is setting a cutoff at the average level of performance, and will be hiring candidates at least as good as the average employee. Logically, it follows that in the long run, one will be raising performance for the organization. McCormick et al. (1979) further clarify the gravitational hypothesis:

To put it another way, “survival” in a given job might be thought of as implying that the incumbents have performed at a reasonably satisfactory level and have found the job to be reasonably acceptable. Following this rationale, it is logical to consider significant mean differences in test scores of incumbents in various jobs as reflecting differences in the nature and level of the human characteristics or constructs that contribute to success on the jobs in question. For this purpose, “success” is considered the ability and willingness to survive in the jobs in question, presumably performing at an acceptable level. (p. 52)

Empirical support for the gravitational hypothesis and the JCV model was traditionally offered by the data set presented in Table 2 (Jeanneret, 1992). The pattern of between job

differences in GATB intelligence test scores is clear. PAQ predictions of these scores using the JCV method are generally close to the actual score. Moreover, Jeanneret (1992) states the standard deviations observed for the scores within each occupation are “relatively small, indicating that the distribution of ability of incumbents in that occupation is narrow and consistent across a range of performance and tenure” (p. 88). Data to support this claim are not provided.

Perhaps the main problem with the gravitational hypothesis is that it represents only a nominal level of measurement (Mossholder & Arvey, 1984). Furthermore, Mossholder and Arvey caution that if the hypothesis is true, average test scores may indicate only the minimal level of competency needed for survival; if false, the process may be very misleading, leading to large amounts of incorrect hires. This raises the concern that at no point does one have information indicating the preponderance of false positives and false negatives. One could argue that those not hired due to a lower test score would add variability to the average test scores of the job, but also be successful on the job. Similarly, one could meet the minimum test score criteria and still fail on the job (McCormick et al., 1979). There is no mechanism (i.e., the validity coefficient) to give us an indication of the preponderance of incorrect classifications.

Trattner (1982) states that the JCV procedure establishes job requirements (or job specifications) not job performance. Mossholder and Arvey (1984) raise the possibility that the reciprocal influence between job experience and test performance could affect the relevance of using incumbent test scores as the benchmark for hiring decisions, a criticism common to all concurrent designs. Trattner further notes that the link between test scores

and actual job performance is indirect and does not meet the Uniform Guidelines selection validation requirements. Keep in mind that if one could establish a direct relationship between scores and job performance, then one would not need to use a synthetic validity procedure (all of which fall under the rubric of indirect validity).

Perhaps the best empirical demonstration of the gravitational hypothesis was offered by Wilk, Desmarais, and Sackett (1995). They outlined the often unspoken logic of the gravitational hypothesis, then tested it with both longitudinal and cross sectional designs. In essence, the rationale is that individuals overqualified for a job are likely to search for a job that provides challenges lacking in their current job. In addition, overqualified individuals can be promoted to jobs with more responsibilities and challenges. Alternatively, underqualified individuals will most likely be terminated or voluntarily change to another occupation where they are less likely to feel the sting of daily failure.

Wilk et al.'s (1995) longitudinal test tracked the movement of almost 4,000 people (ages between 23 and 27) across jobs rated by cognitive complexity. The complexity ratings of jobs were based on a system devised by Linda Gottfredson (1986) called the Occupational Aptitudes Patterns Map. Jobs can be classified into ten levels based on their cognitive complexity with occupations such as chemist, physician, and engineer defining the top bracket. Middle range jobs include carpenter, nurse, and mail clerk, whereas occupations at the bottom are generally of a pure labor nature (laborer, baker helper, etc.).

Subjects were a subset of the National Longitudinal Survey of Youth database. For each subject, ASVAB general intelligence scores were recorded in 1982. In addition,

the levels of job complexity were recorded in 1982 and 1987. After controlling for age, the general cognitive ability scores were segregated by job complexity level at 1982 and further broken down by those people who moved up in complexity level, stayed at the same level, or moved down in level. Table 3 presents the results, showing the cognitive ability test scores for each category.

In almost every case, the pattern of results was the same. People who moved to jobs lower in cognitive complexity, had lower intelligence scores than did those who remained at the same level of job complexity. Furthermore, people who moved up in job complexity had higher scores than the other two groups. Hierarchical regression analysis provided the formal test of this hypothesis. Cognitive ability was a significant predictor of 1987 job complexity level after controlling for age and 1982 complexity level (age and 1982 level explained 10% of variance in 1987 level; adding intelligence explained another 5%). The pattern is clear: over time people in this sample (almost 4,000) tended to gravitate towards jobs commensurate with their ability.

The (Wilk et al., 1995) cross-sectional study examined the variability of cognitive ability test scores as a function of experience. If people are truly leaving jobs where they do not possess the requisite amount of cognitive ability and gravitate to jobs where they do meet the current standards, then as tenure on a job increases, so will the amounts of similarity in cognitive ability among incumbents. That is, groups within a given job with longer tenure should be more homogeneous (in terms of cognitive ability) than those with a shorter tenure, controlling for the cognitive complexity of the job. The dependent variable in the test of this was the natural log of the variance of test scores plus a

correction for bias due to different sample sizes. This correction allows for a measure of variability to be used as a dependent variable without violating the homoscedasticity of variance assumption (Bryk & Raudenbush, 1988).

The results supported the assertion that the longer an individual remains in a given job (or organization), the more likely it is that he or she is similar to coworkers in terms of cognitive ability (Wilk et al., 1995). The variance indices were larger for less experienced workers than for more experienced workers (41.98 vs. 40.13) when experience is restricted to current job tenure (firm experience). When experience includes all jobs within a given category (job experience, which allows for greater experience since it includes everything relevant), the results were similar (42.45 vs. 39.32).

The Wilk et al. (1995) study appears to offer convincing evidence that people do indeed gravitate to jobs commensurate with their abilities. Both of their studies employed large sample sizes (the cross sectional design had over 5,500 people and over 50 job-firm combinations for both the job experience and firm experience categories). The results indicate that using the JCV model may indeed be theoretically sound for validation purposes.

There is a facet of JCV involving ability tests that must be considered. The use of ability tests where one always selects from the top down poses something of a logical problem for JCV. In a tight job market where there is an excess labor supply looking for an available job, one will necessarily be hiring candidates who are overqualified. That is, selection from top down in a tight job market will result in only those at the very top of the distribution being hired. Over time these people will leave for jobs that are more

challenging, and additionally, better paying (as suggested by the higher scores of those who those who 'moved up'). This seems to indicate that selection via ability tests and JCV will doom an organization to high turnover of its best employees.

Personality Tests and Interest Inventories in Selection

If ability tests are a potential problem for JCV, then perhaps interest and personality inventories offer a better platform for JCV methodology. This change from ability tests to personality and interest inventories emphasizes the similarity factor of the JCV process. Both McCormick et al. (1979) and Wilk et al. (1995) make the point that the gravitational hypothesis applies to more than just abilities. Furthermore, in a second longitudinal study, Wilk and Sackett (1995) replicated their earlier findings but also concluded that “cognitive ability is not the sole determinant of job complexity change” (p. 22). The gravitational hypothesis also applies to interests and values. To accommodate this, the JCV procedure remains the same; only the emphasis changes. We are changing our focus from “selection of those with the highest score as long it is above the cutoff point” to “selection of those who are most similar to incumbents.” This process of establishing and predicting mean levels of a personality or interest profile by using job analysis information is exactly what the JCV model entails.

Not all of personality selection research has embraced the philosophy of selection by similarity. Much of it still concerns itself with ability-type selection where the highest score “wins.” Many studies have focused on trying to find the most valid selection instrument. To date, the best meta-analytic findings indicate that Conscientiousness (a big

five dimension) correlates .22 with performance (Barrick & Mount, 1991). A search such as this may be a bit misguided and doomed to mediocre results. It stands to reason that the necessary personality aspects of work will be moderated by the nature of a given job (abilities required of an employee as well as contextual factors). The most fruitful search probably will not be one that looks for the personality factor that predicts across all jobs, but rather a system that correctly identifies the best type of personality profile for each different job. This tailoring of a multi-dimensional instrument for every job has the best chance of maximizing predictability within and among all jobs. This potential avenue for better results is echoed by Tett, Jackson, and Rothstein (1991) who, in a meta-analysis of personality inventories and selection, state that:

...The full potential of personality traits in personnel selection will be realized only when confirmatory research strategies employing personality-oriented job analysis become the standard practice for determining which traits are relevant for predicting performance on a given job... (p.732)

Their own meta-analytic results echoed some of this optimism, but were later attacked by Ones, Mount, Barrick, and Hunter (1994) on technical grounds. This setback notwithstanding, the potential for obtaining better results using job analysis, specifically in the context of JCV, still exists.

Related to this aim is the use of interest inventories for selection. In the case of interest inventories, one is tailoring the desired responses to the job in question. In

actuality, the observed responses are compared to most similar profiles of responses for each job, however the end result is the same. Much of the work in this area has been done with the Strong-Campbell Interest Inventory (Campbell, 1977). The Strong-Campbell is predominantly used for career guidance. The instrument's validity at predicting occupational choice has been tested in a predictive design, with a 59% hit rate for males (43% for females) as well as a concurrent design with hit rates around 60% (Spokane, 1979).

Selection using the Strong-Campbell has generally been considered to be ineffective due to the fact that one is matching interests with the average of all job incumbents, not the high performers (a common criticism of using occupational membership as a criterion). Meta-analytic findings echo this sentiment. Hunter and Hunter (1984) estimated the population validity coefficient for the Strong-Campbell to be .10. When examining the correlation between Strong-Campbell results and tenure and promotion, the findings were not quite as bleak (.22 and .25, respectively). Overall, the results (which were based on an average of only three studies each) suggest that selection using the Strong-Campbell results in employees who are not terribly productive for long periods of time.

The use of the Strong-Campbell in conjunction with a cognitive ability test provides very different results from the overall perspective than a meta-analysis provides. In a firefighter selection system, Frank and Wilcox (1978) use a subset of the Strong-Campbell keyed specifically for firefighter interest in conjunction with a non-verbal intelligence test. By dividing the Strong-Campbell results at the median, they created two

groups of candidates, high similarity (above median) and low similarity (below median). Their results were striking. For the low similarity group, intelligence test results correlated with supervisor ratings $-.32$. The high similarity group's results correlated $.47$ with performance ratings (total $n=25$). Clearly using the intelligence test without the Strong-Campbell data would have produced an overall validity coefficient in the $.20$ range.

Frank and Wilcox (1978) repeated this procedure in a predictive design with similar results. Using training proficiency scores as criteria, high similarity candidates' cognitive ability test scores correlated $.83$ with training success while low similarity candidates correlated only $.10$ (total $n = 22$). While these correlations were based on sample sizes of slightly more than 20 employees (and must be taken with the appropriate statistical grain of salt), the pattern that they reveal is intriguing. In addition, it is worth noting that the supervisors' ratings for training success (second study) could not have been contaminated by test score knowledge as all scores were sealed for six months by court order as soon as they were collected.

While the Frank and Wilcox (1978) results do not directly relate to Job Component Validity, they do drive home the point that interest inventories (and probably personality inventories as well) can be useful for selection when they are keyed on individual job data and used in conjunction with other tests.

The Myers-Briggs Type Indicator

Another type of personality inventory extensively used in the workplace is the Myers-Briggs Type Indicator (MBTI; Myers & McCauley, 1985). The MBTI consists of

95 scored forced choice questions (alternate versions of the instrument contain extra questions for research purposes) that tap behavioral preferences and preferred self-descriptions. Items are scored along four bipolar scales. The scales encompass attitudes (Extraversion/Introversion scale), orientations (Judgment/Perception scale), and functions (Sensing perception/Intuitive perception and Thinking/Feeling scales). It is believed that although every person possesses both processes in each dimension, they intrinsically prefer one over the other (McCauley, 1990). Moreover, the four scales are hypothesized to synergistically interact with each other to produce 16 distinct personality types, which are considered qualitatively unique in terms of cognitive and affective preferences (Murray, 1990; McCauley, 1990).

Some characteristics of the instrument are worth mentioning. First, it is not a trait instrument. Jung's theories are based on the psychological types that one favors. Traits are continuous, while types are dichotomous. While continuous preference scores for each of the four dimensions are reported, suggesting that they are continuous traits, strict followers of the instrument deal only in the type (i.e., the letter indicating which pole one favors).

The second aspect relates to the scoring of the MBTI. If the instrument actually does measure psychological types, then we would expect to see a bimodal distribution of the frequency of preference scores. The bimodal distribution would place two large segments of the test-takers safely into each pole for each dimension, with very few people in the middle. Bimodality is considered to be a necessary condition for the existence of a "type" (Weiss, Mendelsohn, & Feimer, 1982). Research along these lines using the

traditional scoring procedure has produced dismal results. Both Hicks (1984) and McCrae and Costa (1989) found unimodal distributions for all four scales. These results painted a bleak picture of the MBTI's status as a type indicator until Harvey and Murry (1994) used an Item Response Theory (IRT) latent trait model to score the instrument. The scores using a latent trait model converged with the continuous preference scores (traditional method) strongly ($r = .97$) but also produced distributions that were indeed bimodal. In addition, the IRT method produced its highest measurement precision at or near the midpoint between poles, a factor that serves to increase the type stability (i.e., test-retest reliability).

While the Harvey and Murry (1994) results using IRT offer a great deal of hope for the MBTI, their scoring methods were not used to generate the vast databases of MBTI information. The use of the traditional preference score model is likely at the root of the instrument's reliability problems. That is, with so many scores clustered around the midpoint (or zero point of the scale), small differences in observed scores will lead to changes in classifications. As an example, over a five week period, 50% of subjects changed poles on at least one scale (McCarley & Carskadon, 1983).

In spite of its problems and limitations, the MBTI is still widely used, selling more than two million copies annually (Pittenger, 1993). Business accounted for 40% of its 1.5 million sales in 1986 (Moore, 1987). Its primary use in the corporate world relates to team building. Undoubtedly it is also used for screening decisions, much to the chagrin of the test's developers. Coe (1992) delineated the rationale for not using the MBTI for selection purposes. Primarily, the instrument is susceptible to faking. It does not

incorporate a 'Lie' scale as do other instruments such as the MMPI. The other main inadequacy of the MBTI is its assignment of people to various types without measuring how well one functions as a member of that type. Although these are definite limitations, the MBTI does have one attractive asset that most personality inventories lack, the wording style of its questions. As Harvey, Murry, and Markham (1995) note, the MBTI's items are worded in a "positive, affirming" (p.8) manner that is unlikely to be legally challenged for violating the applicant's right to privacy. The extent to which MBTI is actually used for screening and selection is open for debate, but its sales volume as well as its "cult" following in the business world suggest that personnel selection (or screening) may indeed be commonplace.

Further opportunity and justification for using the MBTI for selection purposes is provided by the PAQ. All recent PAQ reports predict MBTI profiles for each job. That is, on every job analysis report, the PAQ provides the estimated percent of incumbents in each pole of the four dimensions. Seeing the prediction that a given maintenance job (for example) should be staffed by an ESTJ person (i.e., a person who favors the Extraversion, Sensing, Thinking, and Judgment poles) might be very tempting for a manager who believes in MBTI dogma. What the PAQ people do not like to tell us is how they derived these profiles (presumably JCV methodology) or how well they can predict the profiles. That is, they do not widely report the degree to which the JCV model (if that is how they are making their predictions) accurately predicts the profile types. At least when we are dealing with aptitude tests, we have some indication of success (correspondence between predicted scores and actual scores). The only data the PAQ people provide to support the

predictions of MBTI profiles are the multiple correlations between percentages of people in a given occupation with a specified MBTI preference for each scale (as in 73% favor Extraversion over Introversion). These multiple correlations are only reported in the PAQ manual and range from .60 for Extraversion/Introversion to .65 for Thinking/Feeling (median multiple $R=.62$; Mecham, 1988, cited in McCormick, Mecham, & Jeanneret; 1989). Clearly, there is room for improvement.

Prediction of MBTI types can be aided by using a better job analysis instrument. When the PAQ was introduced in the early 1970's, it was clearly the best standardized job analysis instrument. However, in the 20 years that followed, a number of inadequacies have been identified. Foremost among the PAQ's deficiencies is its coverage. Many elements of higher-level jobs (i.e. managerial) are not tapped by its items, prompting McCormick and friends to develop a new instrument specifically aimed at white collar jobs. This, however, does not fix the PAQ's problem. It is still not an instrument that can describe almost all jobs well with a common pool of items, a necessary component for an instrument purporting to make meaningful cross-job comparisons on a quantitative level. Given that this is the goal of worker oriented job analysis (Harvey, 1991), the instrument must be considered inadequate. Related to JCV, the poor content coverage of the higher-level work behaviors results in less accurate predictions (of cognitive ability test scores, MBTI profiles, etc.).

One of the great benefits thought to arrive with the development of a standardized job analysis instrument is the ability to have each incumbent rate his or her own job. This allows for one to average the individual position profiles at the job level, resulting in a

quantitative description of the job that is not affected by individual raters' errors. The assumption underlying this logic is that an average of many descriptions of a job is the best description. That is, most errors are random and will cancel out upon aggregation. The PAQ would have served this end well if it had not been for the fact that its reading level was estimated to be at the post-graduate level (Ash & Edgell, 1975). The high reading level virtually insures that the instrument must be completed by job analysts at the rate of one per job, potentially resulting in lower quality ratings.

Further handicapping the PAQ's chances of accuracy is its behaviorally abstract item pool. The items were originally intended to be more general than task statements in order to allow for comparisons of task dissimilar jobs. The level of abstraction in the PAQ (as well as other worker oriented instruments) is so general that many have described its items as "behaviorally nebulous" (Harvey, 1993). Overgeneralized items suffer from accuracy problems as the connection with actual job behaviors becomes blurred: as Harvey, 1993 noted, "when raters attempt to make accurate ratings, the cognitive task facing them when judging highly abstract, unobservable items is daunting" (p. 3).

Finally, the rating scales for 60% of the PAQ's items suffer from serious problems. Nearly 60% of the items are rated on an importance scale that is arguably within-job-relative (Harvey, 1991). That is, the PAQ's instructions direct one to rate the importance of behaviors to the job in question. Importance is anchored only by vague descriptions (minor, high, extreme, etc.) This results in ipsative measurement where the meaning of various levels of rating (i.e. a "5" versus a "3") differs across raters. Ipsative measurement insures that cross-job comparison of results using the PAQ lacks much of its purported

meaning. Anchors for other scales are similarly ambiguous. Moreover, the PAQ allows for unanchored ratings to be made at midpoints between two anchors. These item-level inadequacies serve to introduce more error in dimension scores and the JCV results based on them.

The Common-Metric Questionnaire

Designed to be free of these deficiencies, the Common-Metric Questionnaire (CMQ; Harvey, 1990) provides an improved instrument for synthetic validity efforts. Foremost among the improvements is the range of content coverage that includes the higher-level work behaviors needed to describe managerial jobs. Moreover, the items are written at a level more general than a task statement, but certainly more specific than the PAQ's items. For an example of the feasibility of this middle ground of behavioral specificity, see the Occupational Analysis Inventory (Cunningham, Boese, Neeb, & Pass, 1983).

In addition to its behaviorally specific wording, CMQ items are rated on absolute levels of frequency that allow for true cross-job comparisons to be made. The scales, similar to ones used in the Executive Checklist (EXCEL by Lozada-Larsen, 1988), feature anchors such as: daily, every few days to weekly, monthly, etc. Other types of scales feature similar levels of specificity. Finally, the CMQ employs an eighth grade reading level in order to not preclude a knowledgeable incumbent from completing the instrument. The combination of these features allows for the CMQ to have near-universal applicability,

greater accuracy, true cross-job comparative ability, and incumbent completion resulting in position-level data.

Another feature of the CMQ is dimension profile scores that convey “Does not apply” information. Typically, worker oriented instruments have dimensions scores resulting from factor analysis that are presented in z-score formats. It is difficult to determine from this information whether a dimension simply did not apply to the job in question or is only present in low levels (Harvey, 1993). To illustrate, it is impossible to determine from a score of -1.89 on dimension x whether the dimension is not tapped at all by the job or is simply present in extremely low levels. CMQ dimensions were initially identified by factor analysis, resulting in an 80 factor solution (Harvey, 1993). Scores for each item are simply the percentage of the total possible ratings for each item within each dimension. This percentage score allows for a cursory glance to reveal all dimensions that were completely inapplicable. In short, this true zero level is a key component of the absolute nature of CMQ dimension scores, and allows for greater utility than the relativistic scores on other instruments.

Because the CMQ does not suffer from many of the weaknesses common to other worker oriented job analysis instruments, it would be ideal for a JCV study. As previously mentioned, the PAQ provides an estimate of MBTI profiles for every job analyzed. Use of the CMQ to predict profiles will most likely result in better prediction, increasing the practical utility of selection using JCV methodology and the MBTI.

This leads to our first hypothesis. Given the improved precision and coverage of the CMQ and the extensive MBTI database, we predict that CMQ dimension scores will

have a significant multiple correlation with MBTI scores. That is, for each scale of the MBTI, the CMQ dimension scores should have a large R-squared. To clarify, the CMQ dimensions scores will function as the independent variables while the MBTI subscales will serve as the dependent variables (with a separate analysis for each scale). One reason that we expect significant prediction from the CMQ is that it predicts other external criteria very well. Adjusted multiple correlations for predicting pay are around .90. Hit rates for predicting exempt status exceed 90% (Harvey & Lozada-Larsen, 1993). In addition to mere significance, we predict that the CMQ's multiple correlations will be significantly higher than the PAQ's multiple R's for each of the four MBTI scales.

One may wonder why we should even bother using the JCV model to predict MBTI profiles from job analytic data since the MBTI publishers list the percentages of each type for a great number of occupations (the very database that will be used for this study). The answer is that we are not interested in simply a job title-level description of the MBTI profile. In many organizations, job titles can be misleading. We want to tie the MBTI profile to the behaviors and characteristics of the job. On a theoretical level, this can potentially help us understand the nature of work. Moreover, tying the MBTI profiles to the job's characteristics allows for prediction (or synthesis) of new profiles as the job's content changes (something that would not be reflected at the title-level).

Although we originally intended to predict the MBTI data from the full 80 dimension profile of the CMQ, the low subjects-to-predictors ratio (110 job titles to 80 predictors) renders this analysis meaningless. Because an analysis using all 80 dimensions would be problematic, we decided to employ a higher order factor solution. Based on a

previous analysis we know that a 22 factor solution exists that explains 99.6% of the common variance (Harvey & Lozada-Larsen, 1993). In addition, a 6 factor, highly generalized, solution explains 71.2% of the variance. While it is expected that the 6 factor solution will not account for as much variance in the MBTI scores as the 22 factor solution, it is quite reasonable that the 22 factor solution will predict as well as the full profile of CMQ dimensions had the subjects-to-predictors ratio been high enough to permit such an analysis. Because the CMQ database has grown by almost 10% since the higher-order factor solutions were originally generated, we decided to re-analyze the CMQ profiles in order to update the higher-order solutions.

One additional piece of information that is provided on PAQ reports is a predicted validity coefficient for all of the GATB test information (predicted mean scores) derived using the JCV model. These predicted validity coefficients are intended to provide an estimate of the test's validity if used for hiring. Data for these coefficients date back to the original article (McCormick et al., 1972), which developed regression equations to predict validity coefficients based on PAQ data. The regression coefficients were derived via multiple regression analysis between actual validity coefficients for a sample of 90 jobs and PAQ results for these jobs. In essence, McCormick et al. used the JCV methodology to predict validity coefficients. Results showed that the predicted validity coefficients correlated with actual coefficients at around the .5 level. While this is certainly a high correlation, it renders meaningless the interpretation of the predicted validity coefficients provided by the PAQ reports. If the predicted coefficient for a general intelligence test is .3 and we know that this correlates .5 with an actual validity, there is certainly much room

for error. While some systematic relationship does exist, the interpretation of these numbers is nearly meaningless and without practical utility.

What the concept of predicted validity coefficients does serve to do is remind everyone not in the VG camp that characteristics of the job (tasks, behaviors, information processing demands, etc.) moderate the utility of various predictors. Empirically, this was demonstrated by Gutenberg, Arvey, Osburn, and Jeanneret (1983). They tested the hypothesis that jobs higher in information-processing demands will also have higher validities for intelligence tests. Actual GATB validity coefficients for intelligence subtests (verbal, general, numerical) correlated significantly (median $r = .26$) with PAQ dimension scores for decision making, information-processing, and communication/general responsibility.

Gutenberg et al.'s (1983) results clearly demonstrate that the characteristics of the job moderate the validity of a selection test. This prompts the search for moderators of the MBTI's usefulness. That is, for certain jobs, are MBTI profiles systematically more predictable, and thus more useful than for others? We say "systematically more predictable" since sampling error dictates that there will be a distribution of correlations when examining all jobs. We will be specifically looking at the MBTI's predictability on two levels. First, the gender domination of the job will be examined. That is, as a job title becomes more female dominated (and thus less male dominated), does the relationship between CMQ scores and MBTI profiles change? Our rationale for this examination lies in the MBTI's sensitivity to gender. The instrument employs separate scoring keys for males and females on the Thinking/Feeling scale. This is a technique not used by Big-Five

measures of personality inventories and should play a role in the MBTI's relation to external variables. Second, we will examine holistic groupings of job titles into job families in order to determine whether significant differences exist by family in the multiple correlations between CMQ dimension scores and MBTI preferences. For if such differences do exist, then separate regression equations will be required for each level of job family. The job family groupings are provided by Harvey & Lozada-Larsen (1992). Specifically, jobs will be classified into one of these nine categories: executive, managerial, supervisory, professional/technical, sales, clerical, farming, precision, and labor. As suggested by Gutenberg et al. (1983) predictability is moderated by characteristics of the job. Use of these "big picture" groupings will enable us to examine predictability changes as a function of gross level differences in job characteristics.

When the PAQ provides predicted MBTI profiles for jobs its proprietors are assuming that there is no moderation. If moderation does indeed exist, then using the same regression equations for all jobs leads to greater amounts of incorrect predictions.

The MBTI authors claim that in addition to the existence of discrete personality types, each of the four dimensions interact synergistically with each other to produce 16 distinct psychological types (Murray, 1990; McCauley, 1990). Empirical examinations of this synergistic interaction claim are rare. To date, the debate has centered around whether an individual scale (such as any one of the four MBTI scales) produces distinct psychological types (for general debate on existence of psychological types see Mendelsohn, Weiss, & Feimer, 1982; Block & Ozer, 1982; Weiss et al., 1982; Miller & Thayer, 1989; Gangestad & Snyder, 1991). The validity of the MBTI's interaction claim

is generally considered to be weak (Pittenger, 1993) since the preponderance of the evidence for the scales indicates that bimodal distributions do not exist (Hicks, 1984; McCrae & Costa, 1989). The logic in these tests is that if an instrument is to have interactions among its types, then the lack of evidence for types within each scale precludes any interaction. McCrae and Costa (1989) present one of the few tests of the interactions among the four scales. They examined the effects of interactions among the scales on a correlated big five scale. Results indicated that only 1 of 55 potential interactions were significant and led them to conclude that the 16 types simply summarize the four independent main effects. This test offered by McCrae and Costa is only the first examination of the 16 distinct types and is just one method of examination of the personality type claim.

As mentioned, Harvey and Murry (1994) found convincing evidence for the existence of types within each scale with their use of IRT scoring methods. This offers hope for the interaction perspective of the MBTI's publishers. Further support could be garnered by examining the predictability of the total type profile versus the four individual scales. That is, if the four scales do synergistically interact within the person to produce a distinct psychological type (one of the 16 possible types) then this should manifest itself in one's behavior in a different manner from the four preferences considered separately. In our case the behavior in question is one's occupational choice and indirectly, job performance (yet another use for the gravitational hypothesis). Experimentally, this can be examined by determining whether the overall combined type profile (four letter profile) is more predictable from job analytic information than the four individual scales considered

separately. This draws on Mendelsohn et al.'s (1982) assertion that claims for the existence of distinct types must be supported by prediction of external criteria.

To reiterate our hypotheses regarding JCV between the CMQ and the MBTI:

Hypothesis 1a: There will be a significant multiple correlation between CMQ dimension scores and MBTI preferences for each of the four scales (Extraversion/Introversion, Sensing perception/Intuitive perception, Thinking/Feeling, Judgment/Perception).

Hypothesis 1b: The multiple correlations between the CMQ job dimension profiles and MBTI preferences will be significantly higher than the multiple correlations between PAQ scores and MBTI preferences (PAQ results: .60 for Extraversion/Introversion, .63 for Sensing/Intuition, .65 for Thinking/Feeling, and .61 for Judgment/Perception).

Hypothesis 1c: The higher-order 22 factor solution of the CMQ will have a significantly higher multiple correlation with the MBTI data than the 6 factor solution.

Hypothesis 2a: The gender domination of a job will interact with the CMQ dimension scores in the prediction of MBTI profiles.

Hypothesis 2b: There will be significant differences in predictability among the levels of job family.

Hypothesis 3: The predictability for the overall MBTI profile will be better than the predictability for the four individual scales considered separately.

Method

Subjects

Data for the analyses were provided by two sources. The CMQ data consists of results from field tests for the standardization edition. This resulted in a database of over 5,000 job incumbents for over 900 jobs. MBTI profiles and preferences for each of the four scales as well as the overall type came from Appendix D of Manual: a Guide to the Development and Use of the Myers-Briggs Type Indicator (Myers & McCauley, 1985). The appendix lists MBTI preferences for each scale and overall types for approximately 180 jobs. While this in no way was a random sample, it did span much of the economy including the white-collar professions not measured well by the PAQ. Data for the gender domination of the jobs were obtained from the U. S. Bureau of Labor Statistics ("Employed Civilians," 1990).

Unit of Analysis

The unit of analysis for all hypotheses was the job title level. MBTI scores reflect the percent of members of a given occupation that prefer each dimension of the four scales (e.g., 62% of receptionists are Extraverts, meaning that only 38% are introverts). The percentages were expressed in the E, S, T, and J directions. In addition to the percentage of incumbents who favored each pole of the four scales, data were also obtained indicating the one MBTI type favored by each job's incumbents (i.e., modal type). The CMQ data reflect the average profile of responses for each job title. That is, for the job of "truck driver" the scores reflect the median values of all position-level profiles.

For the factor analyses of the CMQ data, the individual position served as the unit of analysis. That is, the factor analyses were executed on all available 5311 CMQ profiles.

Data Manipulation

Because many job titles in the CMQ database were worded in slightly different manners (e.g., “secretary legal” versus “legal secretary”) or were simply misspelled, the CMQ data had to be manually grouped into similar job titles. Once the data were grouped, the median profiles were calculated for each job title.

When this final version of the CMQ data was obtained, the CMQ data and MBTI data had to be manually matched by job title. It was originally thought that almost every MBTI title would have a match in the CMQ database, but for many titles, no match could be located. In addition, many components of the MBTI database represented higher groupings of other titles, offering no useful data. The matching process yielded 110 job titles (less than our goal of 180 job titles). The average number of responses per title was 29.3 for the CMQ data (for a total sample size of 3223 out of the 5311 possible CMQ profiles) and 474.3 for the MBTI data. It is important to note that the minimum number of responses for MBTI job titles is 50. This was necessary due to the fact that the MBTI data are in the form of percentages of a favored type (or pole) per job title. Having an average as high as 474 people with a minimum of 50 per title lends stability to the percentages of people per group.

Two additional variables were then added to the dataset. These were the job family and the gender domination of the job. The job families were assigned using the

nine-level structure listed in Directions for Administering the CMQ (Harvey & Lozada-Larsen, 1992). The manual gives specific descriptions as well as examples for each level of job family hierarchy. The gender domination of each job, expressed as the percent of incumbents who are women, was added to the dataset. The U. S. Department of Labor report (1990) had statistics for only 96 of the 110 job titles, leaving 14 titles with no information on gender domination.

Analyses

For the factor analysis of the CMQ data, we used the common factor model with squared multiple correlation estimates of communality. The analysis was executed using all 5311 available CMQ profiles.

For hypotheses 1a, 1b, 1c, and 2a multiple regression analysis was used between CMQ dimension scores and percentages of MBTI preferences for each of the four subscales. To reiterate, the CMQ scores were entered as the independent variables; the MBTI preferences constituted the dependent variable. Significance of the differences in predictability were determined from the formula:

$$F = \frac{(R_f^2 - R_r^2) / (K_f - K_r)}{(1 - R_f^2) / (N - K - 1)}$$

where: R_f refers to the full model

R_r refers to the reduced model

K_f refers to the number of predictors in the full model

K_r refers to the number of predictors in the reduced model

N refers to the number of subjects

For the hypothesis that there will be significant differences in the predictability of MBTI preferences across job families (Hypothesis 2c), we examined the average size of the absolute value of the residuals ($y - y'$) resulting from multiple regression analysis for each level of job family. That is, the absolute values of the residuals calculated from the multiple regression analysis of all 110 job titles were grouped into the job family levels and averaged. A one-way Analysis of Variance was used to determine whether significant differences existed in the average size of the residuals. If indeed the MBTI profiles can be predicted better for some types of jobs than others, these differences will be manifested as smaller residuals of prediction for that group.

For Hypothesis 3 (synergistic interaction of the MBTI's four scales into a type that is qualitatively different than its components), the discriminant analyses were executed in two manners, resulting in a planned and unplanned analysis.

Planned Analysis

The planned analysis involved a comparison of agreement rates of predicted and actual MBTI types for two different methods of generating the predicted type. Thus, we have two sets of agreement rates. One agreement rate was the agreement between the modal MBTI type and the type predicted from the multiple regression analysis of the individual scales. For each scale, the predicted pole was obtained by examining the

continuous predicted value for each job title. If the predicted value was greater than 50.00 (the variables are expressed as percent of a job title's incumbents), then the job title was given the pole E, S, T, or J, depending on the scale. If the predicted value was less than 50.00, then the job title was assigned the value I, N, F, or P. This dichotomization was done for all four scales across every job title, resulting in a predicted (from multiple regression analysis) four-letter MBTI type for each job title (referred to as regtype). These types generated from multiple regression analysis were then compared with the modal type to determine agreement rates. It is this agreement rate that is compared with a second agreement rate that determines the efficacy of the procedure.

The second agreement rate referred to above is the agreement rate that is generated by examining the agreement of the MBTI type predicted from discriminant analysis of the modal MBTI type and the actual model type. The discriminant analysis functions as follows. The modal type is inserted as the dependent variable with the CMQ factors constituting the predictors (similar to the regression analysis except that the dependent variable is categorical). This results in a predicted MBTI type based on discriminant analysis (referred to disctype).

The agreement between the type generated from the discriminant analysis and the modal type was then computed. Figure 3 displays a pictorial view of the planned comparison. The procedure (regression of the four scales considered separately or the discriminant analysis of the modal type) more effective at predicting the modal type will offer support to the nature of its data (i.e., considering the scales separately versus considering the overall type). It is expected that this analytical strategy will offer a slight

advantage to the regression analysis (and thus the position that the scales do not synergistically interact) because the regression analysis has the advantage of tailoring its prediction to a continuous dependent variable.

Unplanned Analysis

Upon a close examination of the MBTI data, an interesting phenomenon surfaced. Dichotomization of the continuous preference scores for each scale of the MBTI yielded a preference that was not always a member of the modal type of that job title. For example, examination of the separate scales (observed scales, not predicted scales) for the job title of “teacher” yielded pole preferences of “E”, “S”, “T”, and “J” (which would be combined to form ESTJ) while the modal type was actually “ISTJ.” This lack of agreement between the MBTI type formed as a composite of the individual scale preferences (referred to as composite type) and the modal type occurred for roughly 45% of the job titles. This indicates that there is a distinct difference between the favored individual preferences and the favored type. Although this difference by itself does not indicate superiority of one form of data over the other, we can determine it by their relationships with external variables.

Whether or not the MBTI’s scales synergistically interact with each other in a meaningful manner that is associated with actual behaviors (i.e., career choice) was tested with the JCV model. For this test, two discriminant analyses were conducted and agreement rates were calculated. For both analyses, the CMQ’s higher-order dimension scores served as the predictors and a four-letter MBTI type served as the dependent variable. The only difference between the two analyses was the type of MBTI data used.

For one discriminant analysis, the MBTI type was the modal type (most frequently occurring type) while the other discriminant analysis used the composite MBTI type (type constructed by dichotomizing the four individual continuous scales and combining into a four-letter type) as the dependent variable. Figure 4 offers a pictorial view of the unplanned analysis.

Results

Factor Analyses

Scree plots for the eigenvalues are presented in Figures 3-4. Figure 3 shows all 80 eigenvalues, whereas in Figure 4 the first eigenvalue is dropped in order to make discontinuities more visible. As the results in Figures 5-6 indicate, a number of breaks can be seen at 3 (accounting for 57% of the total common variance), 6 (74%), 9 (83%), 11 (88%), 14 (93%), 18 (98%), and 21 (100%) factors. Our original goal of using a highly generalized solution of 6 factors and larger solution accounting for roughly all of the common variance was accomplished. The sole difference lies in the number of factors in the higher-order solution. Rather than using a 22 factor solution (derived from the Harvey & Lozada-Larsen, 1993, analysis), we are employing a 21 factor solution that explains almost the exact same percent of common variance (99.4 % to 100%) with one less factor.

Table 4 displays the rotated factor solution (Harris-Kaiser rotation with $p=.05$) for the 6 factor model. The 6 factors were easily interpretable in terms of their individual components. Table 5 displays the factor intercorrelations for the 6 factor solution. As can be seen, out of a possible 15 correlations, only 2 exceeded .25 and none exceeded .35.

This low level of item factor intercorrelations is desirable because these factor solutions constitute the predictors in our multiple regression analysis, an analytical technique in which high intercorrelation among predictors is problematic (i.e., multicollinearity).

The factor names and rotated factor loadings for the 21 factor solution are displayed in Table 6. As with the 6 factor solution, the factors are easily interpretable in terms of their latent dimensions of work activity. The factor intercorrelations are shown in Table 7. Once again, the degree of intercorrelation is low with only 13 out of a possible 210 correlations exceeding .25. Only one correlation was greater than .35 (.39).

The scores on the higher-order factors were generated using the score function of the SAS PROC FACTOR program. Simply stated, this function multiplies the standardized CMQ score by its scoring coefficient, then sums across the elements for each factor, resulting in a dimension score for each factor. Thus, two data sets were constructed, one with 21 dimensions (for all 110 job titles, the other with 6 dimensions for each job title). These dimension scores served as the predictors in the JCV (multiple regression) analyses.

JCV Analyses

Hypothesis 1a simply stated that we expected significant multiple correlations between the CMQ dimensions scores and MBTI preferences. Significance was achieved for both the 21 predictor model (R^2 's from .51 to .73) and 6 predictor model (R^2 's from .32 to .60) across 3 of the 4 scales (all except for the JP scale). Multiple correlations approached significance for the JP scale for both the 21 predictor model and 6 predictor

models (p-values .12 and .08 respectively). These multiple correlations would have undoubtedly achieved significance had the subjects-to-predictors ratio been higher. Evidence can be seen by the fact that the p-value came closer to significance when using a model with fewer predictors. While the subjects-to-predictors ratio for the 21 factor model (5.2 to 1) is lower than the conventional level of 10 to 1, the stability gained by using the median profile of predictor scores should more than compensate. Table 8 displays the R-squared, multiple correlations and associated p-values for the 21 and 6 predictor models for all four scales. Figures 7 - 10 show scatterplots of the predicted and actual scale values for the 21 predictor model. Tables 9 - 10 display the raw-score regression weights and associated p-values across all four scales for the 21 and 6 predictor models, respectively. In order to determine whether certain assumptions of regression were violated (homoscedasticity and linearity), the residuals of each scale were plotted against their predicted value. An examination of the scatterplots for signs of uneven variance (heteroscedasticity) and nonlinearity revealed no violations of these assumptions.

An examination of the regression weights for the 6 predictor model offers an intuitive confirmation of the results. For the Extraversion/Introversion scale, jobs high in “marketing, sales and management” activities and low in “tool/machine use” were high in terms of Extraversion. In essence, as contacts with people rise, so do levels of Extraversion in incumbents. Jobs high in “meetings” and “machine use” were associated with higher levels of Sensing preferences. Similarly, jobs high in “management” and low in “external contacts” had higher levels of Thinking than Feeling. Finally, Judgment preferences were favored in jobs high in “management” activities and “external contacts.”

Hypothesis 1b involved a comparison of the multiple correlations between the CMQ JCV results (21 predictor model) and the PAQ's JCV results. As previously mentioned, this unpublished study (Mecham, 1988, cited in McCormick, Mecham, & Jeanneret, 1989) is the only example of JCV methodology used with personality inventories. The CMQ's multiple correlations were significantly higher than the PAQ's for the Extraversion/Introversion scale only (.7321 to .60, $p < .05$). The Sensing/Intuition scale was non-significantly higher for the CMQ (.6566 to .632, $p > .05$). Both the Thinking/Feeling (.555 to .647) and Judgment/Perception (.609 to .506) scales were non-significantly higher ($p > .05$) for the PAQ.

The results from the JCV analyses for the 21 predictor model and 6 predictor models were compared with each other for Hypothesis 1c. Tests for the significance of the difference between the multiple correlations for the two models yielded significant results in only one case, the Extraversion/Introversion scale ($p < .05$). For the other scales, the differences approached significance ($F_{crit} = 1.82$; F_{obt} for S/N, T/F, & J/P were 1.41, 1.58, & 1.22, respectively). Had the sample size been higher (more job titles), these values clearly would have been significant.

Moderation Analyses

Hypothesis 2a involved more comparisons of multiple correlations for each scale of the MBTI for the 21 predictor model. In this case we compared a model with all 21 CMQ predictors and the gender domination of the job against a model that includes the above predictors plus interactions of all of the CMQ predictors with the gender

domination. Table 11 lists the multiple correlations and R-squared's for each comparison. As can be seen for the 21 CMQ predictor model, moderation of the predictability of MBTI types from gender domination was significant only for the Extraversion/Introversion scale.

Because examination of gender-based interaction within the context of the 21 predictor model burns up so many precious degrees of freedom, we examined the possible interaction within the 6 predictor model. Although we know the 6 predictor model does not predict as well as the 21 predictor model (Hypothesis 1c), we feel that this loss in predictability is offset by the gain in power via the extra degrees of freedom. Results for the gender-based interaction within the 6 predictor model paralleled those from the 21 predictor model. Significant interaction was achieved only for the Extroversion/Introversion scale. Significance was approached for the Sensing/Intuition and Thinking/Feeling scales ($F_{crit} = 2.21$; $F_{obt} = 1.95$ for S/N and 2.06 for T/F), but not achieved. See Table 11 for a full listing of the results.

The interaction (and non-interaction) of a job's gender domination with characteristics of the job poses some interesting questions and solutions on both a practical and theoretical level. Quite simply, the interaction of gender domination with a job's behaviors indicates that for some jobs with a high amount of a given work behavior, the female-dominated jobs might favor Extraversion and the male-dominated jobs Introversion. When the amount of that given work behavior is low, however, the relationship changes, with the female-dominated jobs favoring Introversion and the male-dominated jobs Extraversion. For the Extraversion/Introversion scale, significant interactions ($p < .05$) occurred in the 6 predictor model for "marketing/sales" and

“office/technical machine use.” As a concrete example, consider the dimension of “marketing/sales.” Figure 11 displays a graph of the amount of “marketing/sales” behavior (factor 5 from the 6 factor solution) and percent of incumbents who are Extraverts. Separate regression lines are drawn for male-dominated jobs (jobs where greater than 50% of incumbents are men) and female-dominated jobs. When the amount of “marketing/sales” activities are high in a job, both male-dominated and female-dominated jobs favor Extraversion (both groups are greater than 50% on E/I, thus Extraverts). For jobs with smaller amounts of “marketing/sales” behaviors, male-dominated jobs favor the Introversion pole, while female-dominated jobs continue to favor Extraversion (although only slightly). Thus, examination of the Extraversion/Introversion scale in isolation can offer insight into the dynamics and processes of work relating to the role of gender. Moreover, the level of specificity is in the control of the researcher. Rather than examining one of the factors from the 6 factor solution, we could have examined a factor from the 21 factor solution. Or, in cases in which the experimental power is better (i.e., more jobs), one could examine the interactions for a dimension from the 80 factor solution, allowing for even more specificity in our conclusions.

It is interesting to note the role played by gender domination as a main effect. As Table 12 shows, for both the 21 predictor and 6 predictor models, gender domination failed to add significant predictability to the base models of the Extraversion/Introversion, Sensing/Intuition, and Judgment/Perception scales. For the Thinking/Feeling scale, however, gender domination added substantial predictability to the model. Adding gender domination to the 21 predictor model accounted for another 25 percentage points of

variance ($p < .05$). In the 6 predictor model, gender domination added 42 percentage points of predictive variance ($p < .05$). Thus, simply knowing one's sex substantially increases the predictability of the Thinking/Feeling scale.

For our other moderation test, Hypothesis 2b, we examined the absolute values of the residuals of prediction for each scale of the MBTI. The residuals were generated by the 21 predictor model and grouped by job family. Table 13 lists the means, group labels, and significance of ANOVA results for each scale. The results were non-significant for all four scales (all p -values $> .05$). That is, the predictability of MBTI types does not differ significantly across all levels of job family groupings.

Synergistic Interaction Analysis

Table 14 lists the results of the analyses of Hypothesis 3. For each job title four MBTI types are listed: its actual MBTI modal type, the type predicted from regression analysis of the individual scales (regression type), the type predicted from discriminant analysis of the modal type (disc type), and the type generated by dichotomizing the four individual continuous scales (composite type). The middle two types were compared with the modal type to generate the agreement rates that constitute the planned analysis of Hypothesis 3. Table 15 lists the agreement rates for both groups. The discriminant analysis results (disc type) agreed with the modal MBTI type in 65.5 % of the cases on all four scales whereas the regression analysis of the individual scales (regression type) agreed with the modal type in only 29.1 % of the cases for all four scales ($p < .05$; using test for differences in proportion). Agreement on at least three of the four scales improved for

both disc type and regression type results (81.8 to 72.7) with disc type maintaining a non-significant advantage ($p > .05$). In the third case (at least two correct scales), the regtype performed slightly better than the disctype ($p > .05$). In the other cases (one correct scale, no scales correct), the disc type resulted in better prediction than the regression type (non-significant difference, $p > .05$). Thus in four out the five possible comparisons (including the most important comparison - all four scales), the predictions resulting from consideration of the overall type were superior to those of the individual scales considered separately. This superior prediction resulting from considering the overall type instead of a summary of scores on the four separate scales offers support to the notion that the MBTI's scales interact with each other. This support is bolstered by the fact that the regression type enjoyed a methodological advantage due to its use of continuous data for the dependent variable.

For the unplanned analysis of Hypothesis 3, we simply compared the hit rates for the discriminant analysis results on the modal MBTI type versus the composite type (type resulting by combining the favored pole in each of the individual scales). Hit rates were non-significantly ($p < .05$) higher for the discriminant analysis on the modal MBTI type (65.5% correct classifications) than the analysis on the composite type (60.0% correct classifications), again offering support for the interaction of the scales with each other.

Discussion

The successful results of the first hypothesis in our study confirm the notion that we can effectively predict non-ability requirements using JCV methodology. While the

previous work by Mecham (1988, cited in McCormick, Mecham, & Jeanneret; 1989) suggested that this may be the case, so very little information was known about the study that more research was necessary. Furthermore, the results presented clearly indicate that there is a systematic, predictable relationship between our personality profiles and the types of work behavior we choose.

Although the CMQ clearly has the advantage over the PAQ in the content coverage of its items, it predicted significantly better than the PAQ on just one (E/I) scale. This failure on the other scales of the MBTI can be attributed to a number of possible causes (with a likely suspect simply being the lack of information on exactly how Mecham proceeded with his JCV analysis), but the most probable cause simply lies in the relative sizes of the job analysis databases. Due to the fact the PAQ has had the advantage of collecting job analytic information since its inception in the early 1970's, the range of jobs for which it has information is relatively unbounded. This database allows its proprietors the luxury of matching profiles with MBTI data for a stratified selection of job titles (i.e., job titles covering a range of values across a number of variables). In contrast, the CMQ has a total database of just over 5,300 positions, meaning that job title matches with the MBTI database are on an availability basis. Resulting from this is the case of the Judgment/Perception scale (PAQ can boast an 11 percentage point advantage in R-squared) in which just 11 of the 110 (10%) job titles in our sample favored the "Perception" pole. This creates an analytical situation in which the range of possible values for Judgment/Perception is restricted, reducing variance and handicapping prediction. This restriction of range could be the cause of the decreased predictive ability

by the CMQ as compared with the PAQ on the Judgment/Perception scale. While there is no way to be certain that this is indeed the case, it is a viable possibility.

Our analysis of the interaction of gender domination with the type of work performed resulted in just one significant result, the Extraversion/Introversion scale. However, just as the range restriction problem in the Judgment/Perception scale hampered our initial analysis of predictability, it could also have hindered our chances of detecting moderation. In addition to the imbalance of type preferences in the Judgment/Perception scale, range restriction occurred for the Sensing/Intuition scale to a lesser extent (33 out of 110, 30%, favored Sensing), potentially accounting for its lack of interaction. Finally, the Thinking-Feeling scale's chances for significant interaction were affected by its sex-based scoring system. While not a certainty, it is entirely possible that gender moderation occurs for all four scales of the MBTI, but our sample's characteristics (as in the Sensing/Intuition and Judgment/Perception scales) or sex-based scoring schemes (for the Thinking/Feeling scale) serve to conceal this effect.

On a more practical level, practitioners who insist on using the individual scales of the MBTI must incorporate the interaction terms in their regression equation for the Extraversion/Introversion scale when making JCV-type predictions. Taking the interactions into account raised the multiple correlation for our 21 factor model (plus interactions) to .89. Failure to do so results in the loss of valuable predictive ability and ignores the role of the interaction between gender and the type of work performed.

Even more interesting than the interaction finding for the Extraversion/Introversion scale is the lack of a finding of interaction for the Thinking/Feeling scale. The gender

domination does add a significant main effect (almost doubling the percent of variance accounted for), but fails to interact with the type of work behavior performed. This potential for interaction may have been removed by the MBTI's use of separate scoring keys for each sex. The MBTI's manual (Myers & McCauley, 1985) reports that its separate scoring keys result in endorsement rates for "Thinking" for 60% of males and only 30% of females. In the context of JCV, the Thinking/Feeling scale serves as a proxy variable for sex (simply knowing one's sex increases the predictability enormously). It is because of this that the MBTI's publishers should consider abandoning its sex-based scoring schemes. This scoring system serves to dilute the scale's systematic relationships with external variables. Related to this is the use of the MBTI for selection purposes. Any organization that does so may be technically in violation of the 1991 Civil Rights Act which banned all group membership-based score adjustments, thus using one scoring scheme for both sexes is a legal necessity if one were to use the MBTI for selection purposes.

Our other examination of moderation, moderation by level of job family, proved to be completely at odds with the Gutenberg et al. (1983) article which found significant moderation of cognitive ability tests by holistic classification of jobs. In no case did our analysis even come close to finding significant differences in predictability among the groups (the lowest p-value was .24). This failure to find significant moderation by job family level essentially gives the practitioner the "green light" to use JCV methods with the MBTI across all types of jobs (e.g., high-level managerial and blue-collar heavy industry). Based upon these two studies, we may conclude that a job's general qualities

do moderate its predictability of cognitive ability test scores and do not moderate its predictability of MBTI types (and most likely, other personality tests).

The final hypothesis was related to the MBTI's claim that its scales synergistically interact with each other to offer information that cannot be expressed as simply the sum of its components. In both of our analyses, the "true" MBTI type had the more systematic relationship with the job analytic data. Had the MBTI's type been simply a summary of four main effects, consideration of the scales separately as opposed to a "true" type would offer no less (or more) of a systematic relationship with an external criterion. This linking of the various formats of MBTI data (separate scales versus overall type) to external criteria is exactly what Mendelsohn et al. (1982) decreed necessary to support the existence of a true type. Moreover, our choice of external criteria was not a trivial one. Few decisions in life affect us more than our choice of the type of work behavior we want to undertake.

When delineating the first analytical strategy for the test of interaction among the scales, we noted that the technique we employed offered an advantage to regression-based prediction (and thus the position that the types do not interact). In spite of this handicap, the "true" MBTI type had the more systematic relationship by a wide margin (36.4% more correct classifications). Consistency (although not significance) of the synergistic interaction was offered by the second analysis. The 5.5% advantage in predictability held by the "true" MBTI type in the second analysis, while not significant, was not trivial either.

This study focused on the MBTI at two levels of analysis: the overall type and the individual scale. There is a third level, occupying a middle ground, that is worthy of mention. Known as temperaments, these are carefully conceived relations between two scales. Four of these low-level interactions are hypothesized to exist by the MBTI's publishers. Specifically, they are Sensing-Thinkers, Sensing-Feelers, Intuitive-Feelers, and Intuitive-Thinkers. These interactions relate to the type of information we seek and how we process that information. While these temperaments ignore some valuable information (i.e., the other two scales), they do serve as a more efficient means of considering interactions among individual scales and are worthy of future research.

Clearly this study has more than a few weaknesses. The most obvious was noted when we first mentioned JCV. Specifically, it is the fact that we are correlating averages (median profiles), a technique known to inflate resulting correlation estimates (James, 1982). In an ideal paradigm, we would have MBTI results for each person administered the CMQ. Thus the analysis would be on the individual level, not the group level (at least with the CMQ we have this capability; the PAQ is generally completed at the job-level, rendering data matching by position impossible). Furthermore, we would employ the item response theory latent-trait scoring procedures used by Harvey and Murry (1994) to score all four of the MBTI's scales (allowing for better type classification). Inherent in the previous statement is the notion that we would not be using separate scoring keys for each sex on the Thinking/Feeling scale. This paradigm would have allowed us to report an overall estimate of the predictability of MBTI preferences, and would have eliminated an enormous amount of error variance. In this ideal world, we would also ensure that we

have adequate sampling of not only type preferences (in order to avoid range restriction), but also the domain of work (i.e., placing emphasis on collecting data from portions of the economy, such as farming, ignored by our current sample). Finally, an individual-level analysis would allow us to draw conclusions about the interaction of sex (not gender domination) with the type of work performed. Nevertheless, we must resort to what is feasible. It was because of this concession (and subsequent susceptibility to aggregation bias) that we limited our analysis and conclusions to situations where we compared the relative efficacy of two analyses using the same experimental design (such as the CMQ versus PAQ comparisons).

In addition to the specter of aggregation bias, we would have preferred a larger sample of job titles (i.e., larger sample size). Our original goal of 180 job titles would have offered substantially more power in our analyses, and would have had better coverage of the domain of work (more variability, leading to better prediction). As the CMQ database grows and the MBTI updates its listing of preferences by job title to include more titles, this goal may one day be realized. Under the current constraints, however, we maximized our resources.

Another limitation lies in the nature of the MBTI. The test-taker is simply assigned to a preference or type with no measure of how well one functions as a member of that type. Thus, we have no within-type variance to be measured. Borderline classifications (or incompetent performers within a type) serve to introduce an unsystematic relationship to our study. If it were possible to identify those who function

at a minimally competent level within a type, our results and conclusions would be stronger.

One final weakness with this study concerns the time at which the data was collected. The MBTI preferences are reported in the 1985 edition of the test manual (Myers & McCauley), meaning that the data were most likely collected in the early 1980's. The CMQ and government data (gender domination of each job), however, were collected from 1989-1994. This 7 to 10 year time-lapse in data collection must be considered a source of error. Although it is unlikely that this could systematically inflate the correlational estimates (as the aggregation bias does), it is possible. Most likely, this time lapse serves to introduce unsystematic variance, handicapping our prediction. This should not jeopardize the validity of our conclusions, as they are derived from relative comparisons of effectiveness within the same experimental paradigm (i.e., same rationale that aggregation bias does not jeopardize our conclusions).

This study, while not perfect from a design viewpoint, does allow for several clear conclusions to be drawn. First, there is a systematic relationship between the type of work performed on the job and the personality profile of the employee. That is, JCV is feasible for non-ability tests. Moreover, the efficacy of the JCV model with personality test data is not moderated by level of the job family. Finally, since the synergistic interaction was supported (albeit, less than overwhelmingly) in two separate analyses, we should no longer focus on the individual scales of the MBTI and acknowledge that their whole is greater than the sum of their parts. Future predictions and analyses based on and incorporating the overall type will most likely be superior to predictions made on the individual scales.

The MBTI's publishers caution against using the instrument for personnel selection. One final reason to refrain from doing so is the possibility that selection of a certain type (then using top-down selection on another predictor within types) for a given job might not lead to the most productive work force. Kroeger and Thuesen (1988) made the case that organizations suffering from type homogeneity are unlikely to have the multidimensional viewpoints necessary to see more than one side of a complicated problem. Thus, the best uses of JCV involving the MBTI are in career counseling and the research setting. Specifically, it is best suited for an understanding of the types of work behaviors that are favored by a given personality type. The improvement offered by this method over a simple declaration that "Introverts prefer desk-type jobs" (information garnered from an examination of MBTI preferences of basic job descriptions) is analogous to the improvement in job analysis gained by a thorough examination of the tasks and behaviors in a job as opposed to a holistic judgment of its title.

References

- Ash, R. A. & Edgell, S. L. (1975). A note on the readability of the Position Analysis Questionnaire. Journal of Applied Psychology, 60, 765-766.
- Balma, M. J. (1959). The concept of synthetic validity. Personnel Psychology, 12, 395-396.
- Barret, G. V. (1992). Clarifying construct validity: definitions, processes, and models. Human Performance, 5, 13-58.
- Barrick, M. R. & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. Personnel Psychology, 44, 1-26.
- Binning, J. F. & Barret, G. V. (1989). Validity of personnel decisions: a conceptual analysis of the inferential and evidential bases. Journal of Applied Psychology, 74, 478-494.
- Block, J. & Ozer, D. J. (1982). Two types of psychologists: remarks on the Mendelsohn, Weiss, & Feimer contribution. Journal of Personality and Applied Psychology, 42, 1171-1181.
- Bryk, A. S. & Raudenbush, S. W. (1988). Heterogeneity of variance in experimental studies: a challenge to conventional interpretations. Psychological Bulletin, 104, 396-404.
- Campbell, D. P. (1977). Manual for the Strong-Campbell Interest Inventory. Stanford, CA: Stanford University Press.

Coe, C. K. (1992). The MBTI: potential uses and misuses in personnel administration. Public Personnel Management, 21, 511-522.

Cunningham, J. W., Boese, R. R., Neeb, R. W. & Pass, J. J. (1983). Systematically derived work dimensions: factor analysis of the OAI. Journal of Applied Psychology, 68, 232-252.

EEOC v. Atlas Paper Co. (1989). 868 F. 2d. 1487.

Employed civilians by detailed occupation, sex, race, and Hispanic origin. (1990). Employment and Earnings, 183-188.

Fleishman, E. A. (1967). Development of a behavior taxonomy for describing human tasks: a correlational experimental approach. Journal of Applied Psychology, 51, 1-10.

Frank, H. & Wilcox, C. (1978). Development and preliminary cross-validation of a two-step procedure for firefighter selection. Psychological Reports, 43, 27-36.

Gangestad, S. W. & Snyder, M. (1991). Taxonomic analysis redux: some statistical considerations for testing a latent class model. Journal of Personality and Social Psychology, 61, 141-146.

Gottfredson, L. S. (1986). Occupational Aptitude Patterns Map: development and implications for a theory of job aptitude requirements [Monograph]. Journal of Vocational Behavior, 29, 254-291.

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

Gutemberg, R. L., Arvey, R. D., Osburn, H. G., Jeanneret, P. R. (1983).

Moderating effects of decision-making/information processing job dimensions on test validities. Journal of Applied Psychology, 68, 602-608.

Hamilton, J. W. (1981). Options for small sample sizes in validation: a case for the J-coefficient. Personnel Psychology, 34, 805-816.

Harvey, R. J. (1990). The common-metric questionnaire for the analysis and evaluation of jobs (field test version 1.12). San Antonio, TX: The Psychological Corporation.

Harvey, R. J. (1991). Job analysis. In M. D. Dunnette & L. Hough (Eds.), Handbook of Industrial and Organizational Psychology (2nd edition). Palo Alto, CA: Consulting Psychologists Press.

Harvey, R. J. (1993). Research monograph: the development of the Common-Metric Questionnaire. San Antonio, TX: The Psychological Corporation.

Harvey, R. J. & Lozada-Larsen, S. R. (1992). Directions for administering the CMQ. San Antonio, TX: The Psychological Corporation.

Harvey, R. J. & Lozada-Larsen, S. R. (1993, May). Predicting pay and exempt status using higher-order factors of the Common-Metric Questionnaire. Paper presented at annual conference of the Society for Industrial and Organizational Psychology, San Francisco.

Harvey, R. J. & Murry, W. D. (1994). Scoring the Myers-Briggs Type Indicator: empirical comparison of preference scores versus latent trait methods. Journal of Personality Assessment, 62, 116-129.

Harvey, R. J., Murry, W. D., Markham, S. E. (1995). A “big five” scoring system for the Myers-Briggs Type Indicator. Paper presented at annual conference of the Society for Industrial and Organizational Psychology, Orlando.

Hicks, L. E. (1984). Conceptual and empirical analysis of some assumptions of an explicit typological theory. Journal of Personality and Social Psychology, 46, 1118-1131.

Hollenbeck, J. R. & Whitener, E. M. (1988). Criterion-related validation for small sample contexts: an integrated approach to synthetic validity. Journal of Applied Psychology, 73, 536-544.

Hunter, J. E. & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. Psychological Bulletin, 96, 72- 98.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: correcting error and bias in research findings. Newbury Park, Cal.: Sage Publications.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. Journal of Applied Psychology, 67, 219-229.

James, L. R., Demaree, R. G., Mulaik, S. A., & Ladd, R. T. (1992). Validity generalization in the context of situational models. Journal of Applied Psychology, 77, 3-14.

Jeanneret, P. R. (1992). Applications of job component validity/synthetic validity to construct validity. Human Performance, 5, 81-96.

Kroeger, O., & Thuesen, J. M. (1988). Type Talk, New York: Bantam Doubleday Dell Publishing Group, Inc.

- Lawshe, C. H. (1952). Employee selection. Personnel Psychology, 5, 31-34.
- Lawshe, C. H. & Steinberg, M. D. (1955). Studies in synthetic validity I: an exploratory investigation of clerical jobs. Personnel Psychology, 8, 291-301.
- Lozada-Larsen, S. R. (1988). Going beyond criticism: management work theory and research. Unpublished doctoral dissertation, Rice University.
- Marquardt, L. D. & McCormick, E. J. (1972). Attribute ratings and profiles of the job elements of the Position Analysis Questionnaire (PAQ; Tech Rep. No. 1) West Lafayette, IN: Purdue University, Psychological Sciences Department.
- McCarley, N. G. & Carskadon, T. G. (1983). Test-retest reliabilities of scales and subscales of the Myers-Briggs Type Indicator and of criteria for clinical interpretive hypothesis involving them. Research in Psychological Type, 6, 24-36.
- McCarley, M. H. (1990). The Myers-Briggs Type Indicator and leadership. In K. E. Clark & M. B. Clark (Eds.), Measures of Leadership (pp. 381-418). West Orange, NJ: Leadership Library of America.
- McCormick, E. J., DeNisi, A. S., & Shaw, J. B. (1979). Use of the Position Analysis Questionnaire for establishing the job component validity of tests. Journal of Applied Psychology, 64, 51-56.
- McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire. Journal of Applied Psychology, 56, 347-368.

McCormick, E. J., Mecham, R. C., & Jeanneret, P. R. (1989). Technical Manual for the Position Analysis Questionnaire (PAQ; 2nd ed.). Logan, UT: PAQ Services. Palo Alto, CA: Consulting Psychologists Press Inc.

McCrae, R. R. & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. Journal of Personality, 60, 117-174.

Mendelsohn, G. A., Weiss, D. S., & Feimer, N. R. (1982). Conceptual and empirical analysis of the typological implications of patterns of socialization and femininity. Journal of Personality and Social Psychology, 42, 1157-1170.

Miller, M. L. & Thayer, J. F. (1989). On the existence of discrete classes in personality: is self-monitoring the correct joint to carve? Journal of Personality and Social Psychology, 57, 143-155.

Moore, T. (1987). Personality tests are back. Fortune, 115, 76-81.

Mossholder, K. W. & Arvey, R. D. (1984). Synthetic validity: a conceptual and comparative review. Journal of Applied Psychology, 69, 322-333.

Murray, J. B. (1990). Review of research on the Myers-Briggs Type Indicator. Perceptual and Motor Skills, 70, 1187-1202.

Myers, I. B. & McCauley, M. H. (1985). Manual: a guide to the development and use of the Myers-Briggs Type Indicator. Palo Alto, CA: Consulting Psychologists Press.

Ones, D. S., Mount, M. K., Barrick, M. R., & Hunter, J. E. (1994). Personality and job performance: a critique of the Tett, Jackson, and Rothstein (1991) meta-analysis. Personnel Psychology, 47, 147-156.

Pearlman, K., Schmidt, F. L., Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations.

Journal of Applied Psychology, 65, 373-406.

Pittenger, D. J. (1993). The utility of the Myers-Briggs Type Indicator. Review of Educational Research, 63, 467-488.

Primoff, E. S. (1955). Test selection by job analysis: the J-coefficient. Washington, D. C.: U. S. Civil Service Commission, Assembled Test Technical Edition.

Primoff, E. S. (1959). Empirical validation of the J-coefficient. Personnel Psychology, 12, 402-418.

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

Schmidt, F. L., & Hunter, J. E., & Pearlman, K. (1981). Task differences as moderators of aptitude test validity in selection: a red herring. Journal of Applied Psychology, 66, 166- 185.

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

Sparrow, J. (1989). The utility of the Position Analysis Questionnaire in relating job behaviors to traits. Journal of Occupational Psychology, 62, 151-162.

Sparrow, J., Patrick, J., Spurgeon, P., Barwell, F. (1982). The use of job component analysis and related aptitudes in personnel selection. Journal of Occupational Psychology, 55, 157-164.

Spokane, A. R. (1979). Occupational preference and the validity of the Strong-Campbell Interest Inventory for college women and men. Journal of Counseling Psychology, 26, 312-318.

Taylor v. James River Corporation. (1989). WL 165953, 1 (SD Ala.).

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: a meta-analytic review. Personnel Psychology, 44, 703-742.

Trattner, M. H. (1982). Synthetic validity and its application to the uniform guidelines validation requirements. Personnel Psychology, 35, 383-397.

Van Aken v. Young (1982). 541 F. Supp. 448.

Weiss, D. S., Mendelsohn, G. A., & Feimer, N. R. (1982). Reply to comments of Block & Ozer. Journal of Personality and Social Psychology, 42, 1182-1189.

Wilk, S. L., & Sackett, P. R. (1995). A longitudinal analysis of ability-job complexity fit and job change. Paper presented at annual conference of the Society for Industrial and Organizational Psychology, Orlando.

Wilk, S. L., Desmarais, L. B., & Sackett, P. R. (1995). Gravitation to jobs commensurate with ability: longitudinal and cross-sectional tests. Journal of Applied Psychology, 80, 79-85.

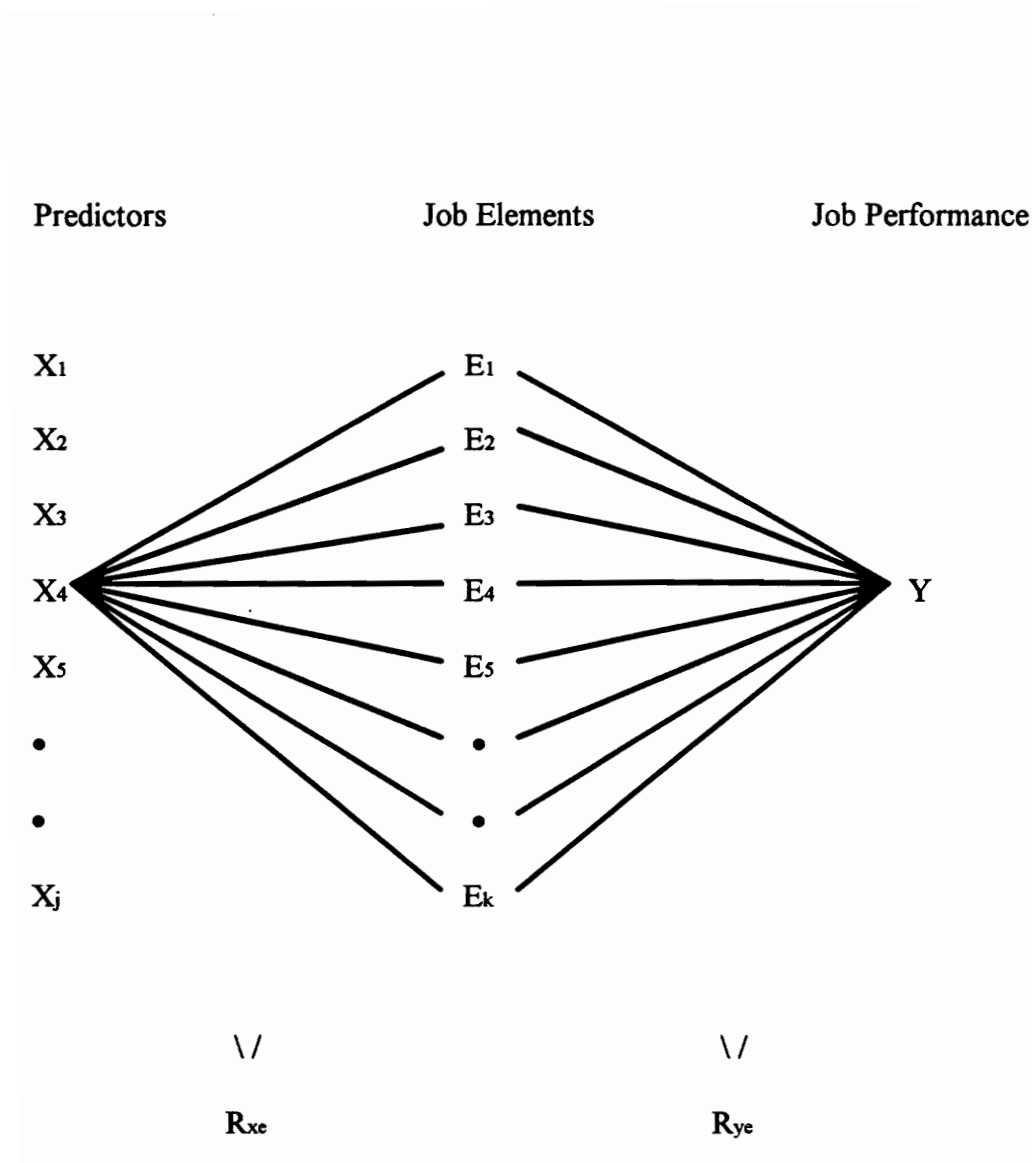


Figure 1. The J-coefficient concept.

Hamilton (1981, p. 807)

Predictor Score Components	Overall Predictor Score	Criterion Score Components	Overall Criterion Score
	$\{W_{y11}\}$	$\{W_{y11}\}$	
	$\{\bullet\}$	$\{\bullet\}$	
S1 $[Z_{x11} \dots Z_{xm1}]$	$\{\bullet\} = Z_{y1}$	$[Z_{y11} \dots Z_{ym1}]\{\bullet\} = Z_{y1}$	
	$\{W_{ym1}\}$	$\{W_{ym1}\}$	
	$\{W_{y12}\}$	$\{W_{y12}\}$	
	$\{\bullet\}$	$\{\bullet\}$	
S2 $[Z_{x12} \dots Z_{xm2}]$	$\{\bullet\} = Z_{y2}$	$[Z_{y12} \dots Z_{ym2}]\{\bullet\} = Z_{y2}$	
	$\{W_{ym2}\}$	$\{W_{ym2}\}$	
	$\bullet$	$\bullet$	
	$\bullet$	$\bullet$	
	$\bullet$	$\bullet$	
	$\{W_{yin}\}$	$\{W_{yin}\}$	
	$\{\bullet\}$	$\{\bullet\}$	
Sn $[Z_{x1n} \dots Z_{xmn}]$	$\{\bullet\} = Z_{yn}$	$[Z_{y1n} \dots Z_{ymn}]\{\bullet\} = Z_{yn}$	
	$\{W_{ymn}\}$	$\{W_{ymn}\}$	

Figure 2. The Hollenbeck and Whitener (1988) synthetic validity model.

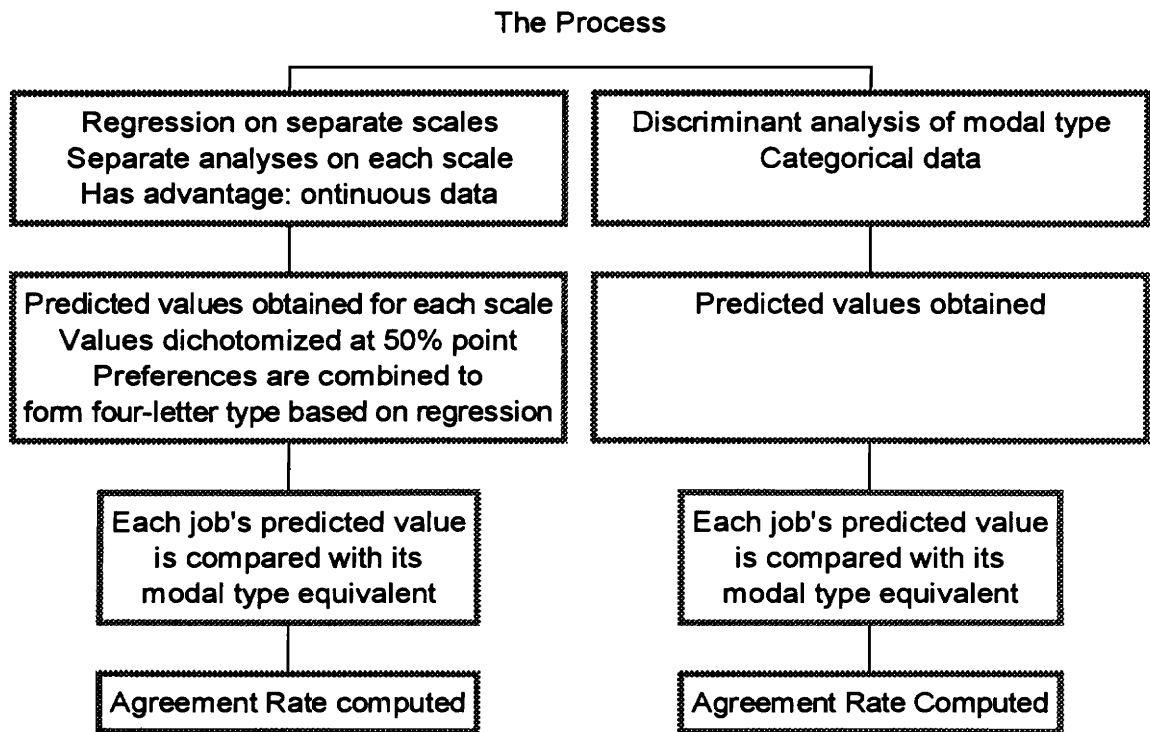


Figure 3. Pictorial view of the analytical strategy for the planned comparison.

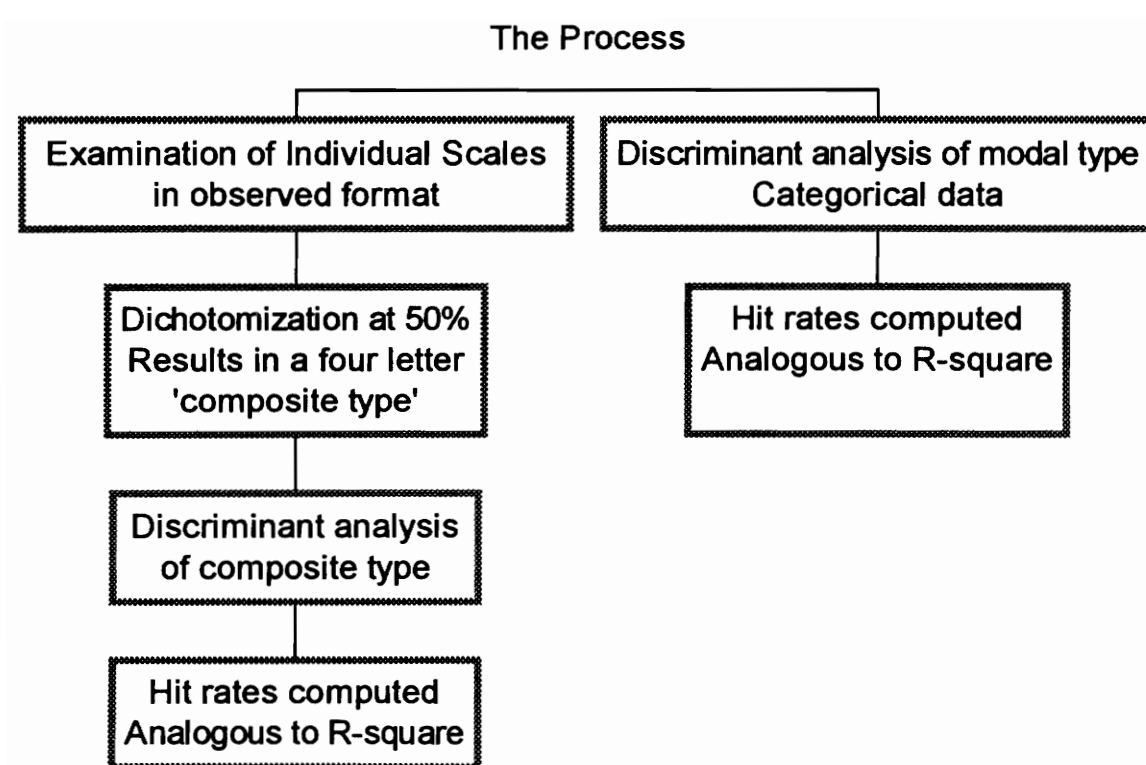


Figure 4. Pictorial view of the analytical strategy for the unplanned comparison.

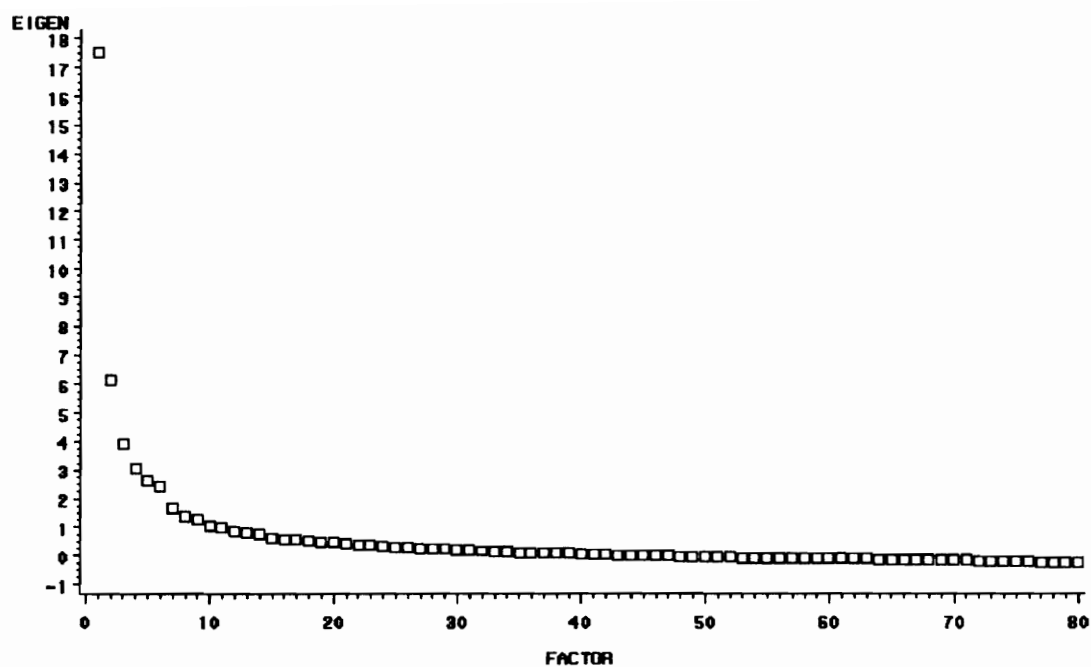


Figure 5. Scree plot of eigenvalues from the reduced correlation matrix of the 80 CMQ first-order scales.

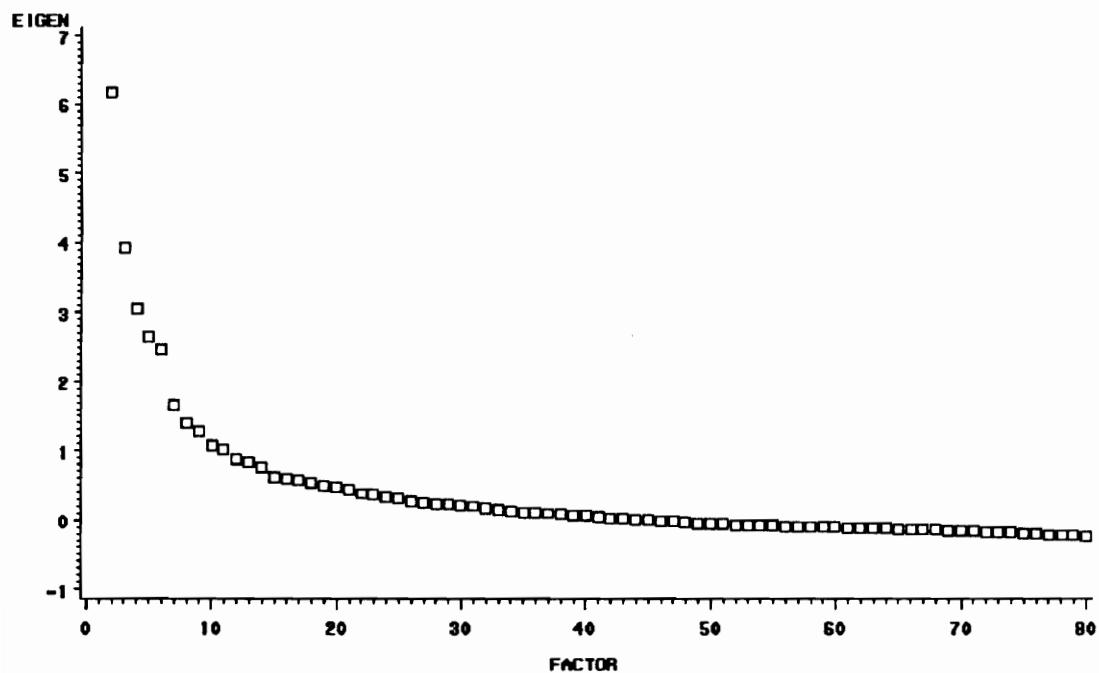


Figure 6. Scree plot of eigenvalues 2-80 from the reduced correlation matrix of the CMQ first-order scales

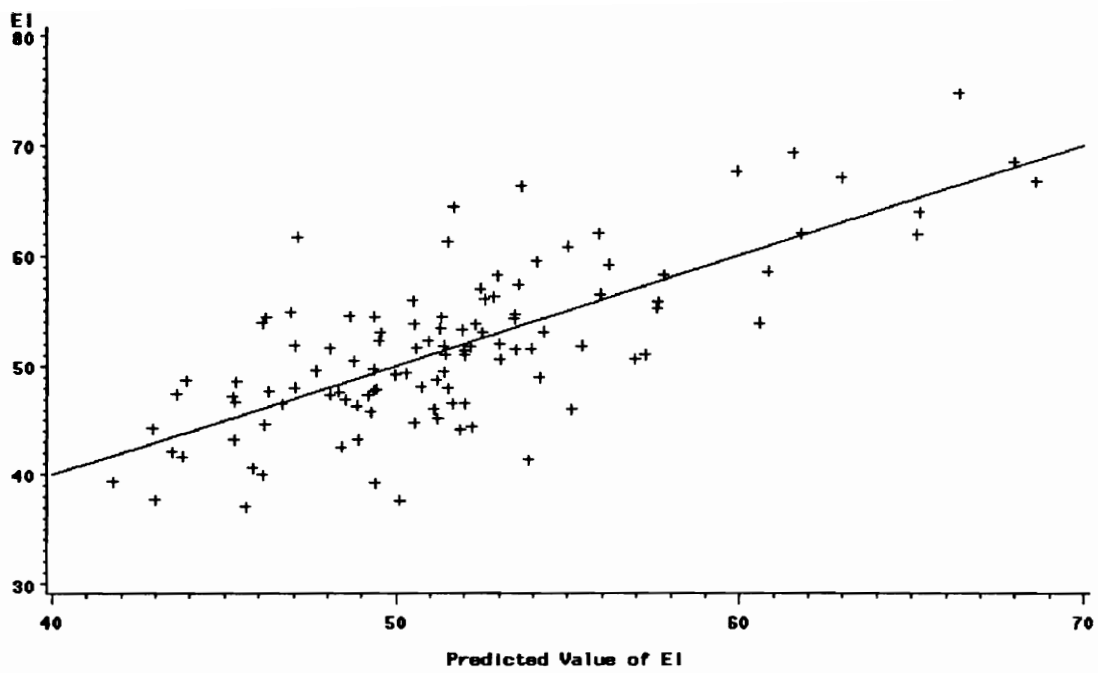


Figure 7. Scatterplot of predicted and actual Extraversion/Introversion percentages from the 21 Predictor model

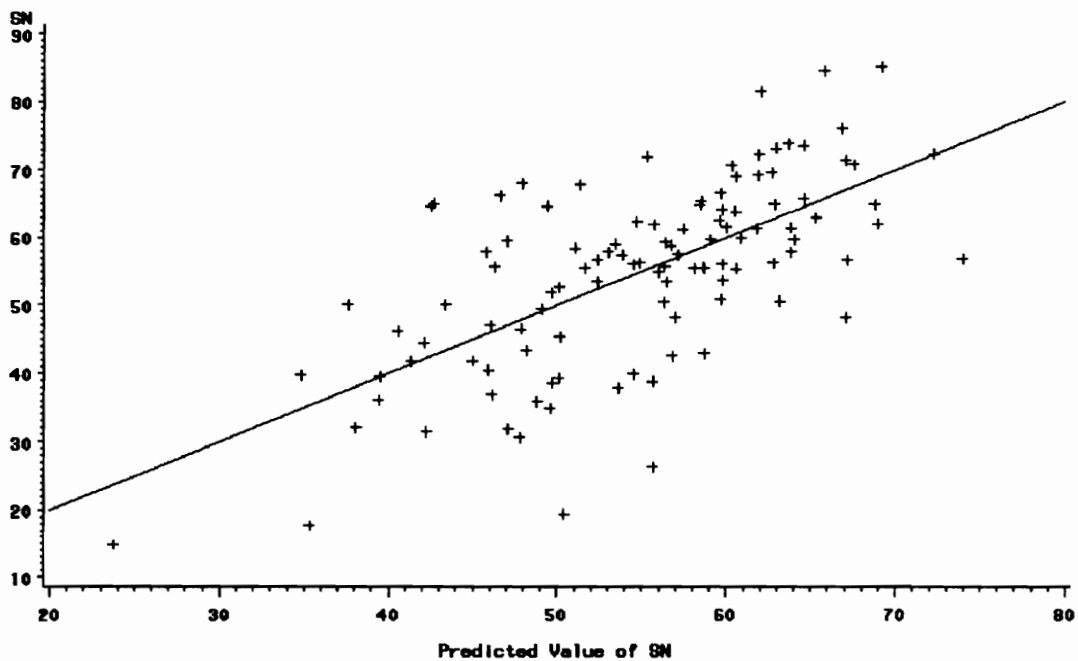


Figure 8. Scatterplot of predicted and actual Sensing/Intuition percentages from the 21 Predictor model.

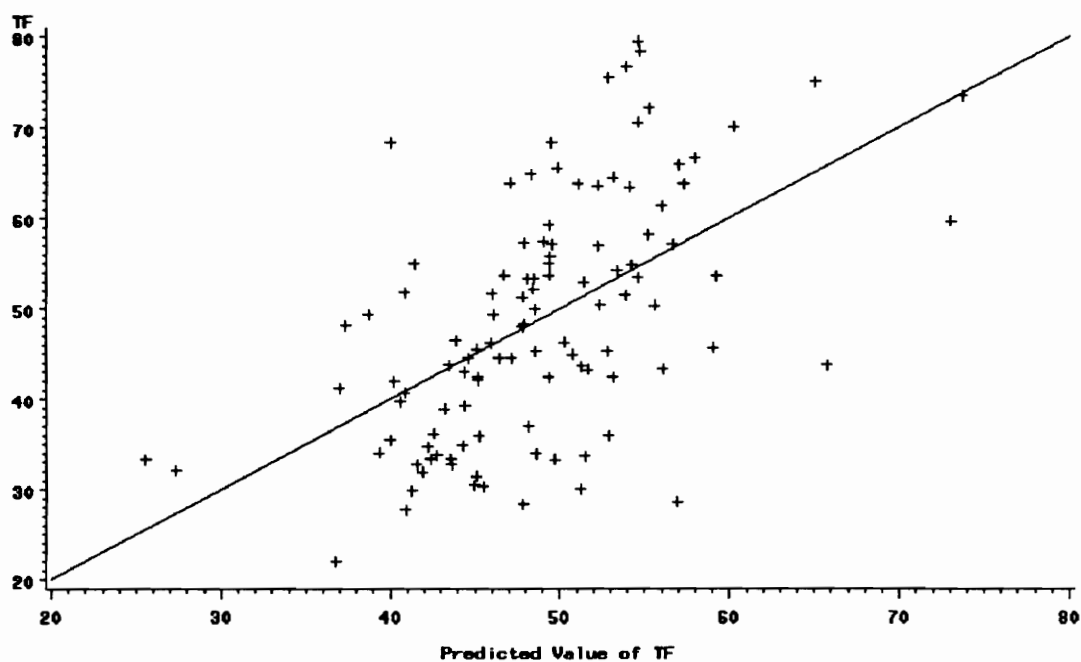


Figure 9. Scatterplot of predicted and actual Thinking/Feeling percentages from the 21 Predictor model.

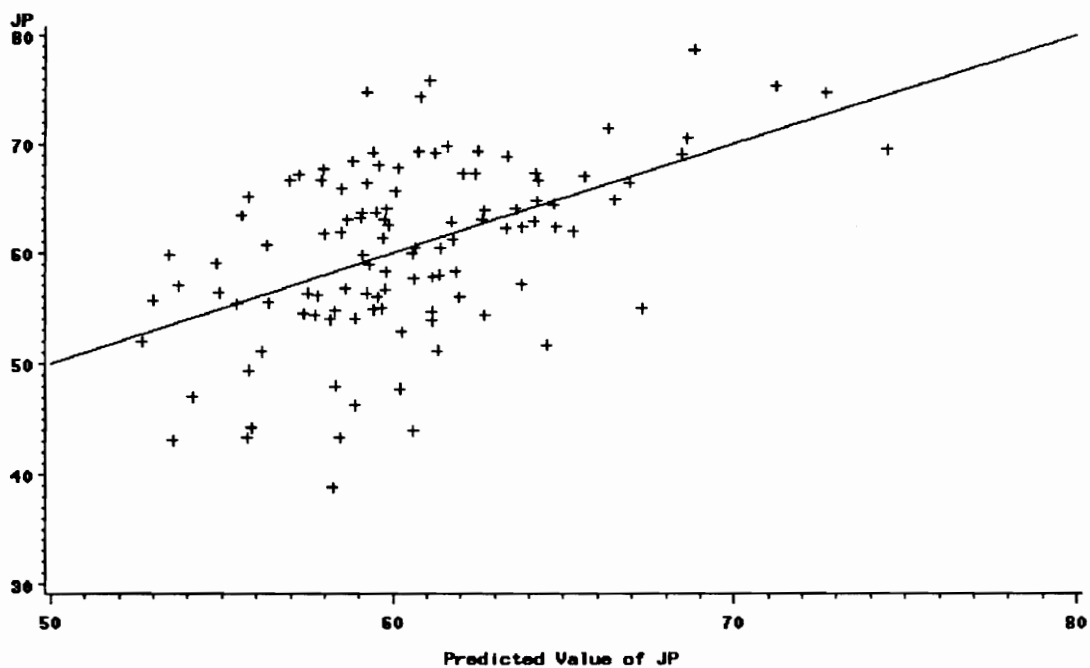
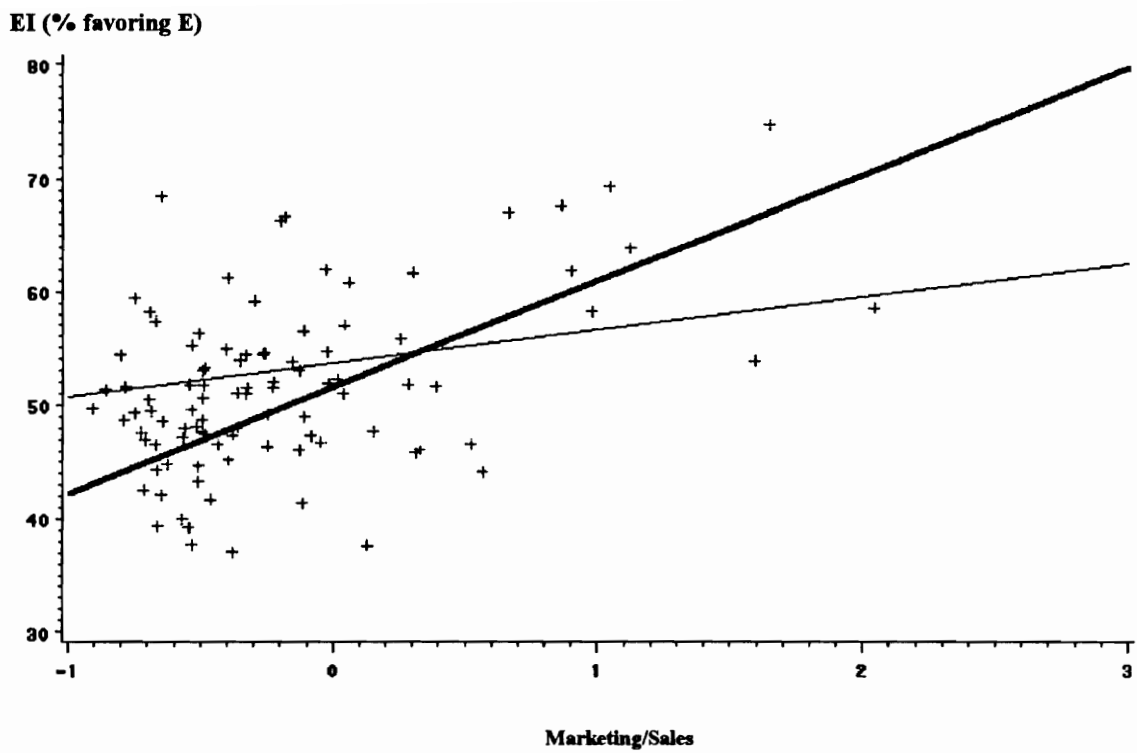


Figure 10. Scatterplot of predicted and actual Judgment/Perception percentages from the 21 Predictor model



Dark Line: Male-dominated jobs

Light Line: Female-dominated jobs

Figure 11. Graph of Marketing/Sales and percent of incumbents who are Extraverts.

Table 1. PAQ Attribute ratings to derive job requirements.

Matrix Compilations

PAQ Elements	Attributes				
	1	2	3	...	76
1	X1 ₁	X1 ₂	X1 ₃		X1 ₇₆
2	X2 ₁				
•					
•					
187	X187 ₁				

Applications of Attribute Ratings

PAQ Element	Element Rating	Attribute 1 (A1)	Attribute Profile
1	3	X1 ₁	(3)(X1 ₁)
2	2	X2 ₁	(2)(X2 ₁)
•	•	•	•
•	•	•	•
187	•	•	•
			<u> </u> ΣA ₁

Jeanneret (1992, p. 85)

Table 2. Mean test scores on the GATB intelligence (G) test for selected occupations based on actual Incumbents (I) and predictions from PAQ Job Dimension Scores (P).

<u>Occupations</u>	<u>Incumbents (I)</u>	<u>Predicted (P)</u>
General practitioner (M. D.)	135	130
Forester	124	117
Draftsman	116	122
Teller, bank	109	104
Stenographer	106	107
Power plant operator	105	108
Keypunch operator	101	105
Sheet metal worker	100	99
Barber	96	101
Plasterer	93	95
Bag machine operator	89	94
Custodian, school	88	93
Yarn winder	80	78

Jeanneret (1992, p.88)

Table 3. Gravitation to jobs commensurate with ability.

1982 job- complexity level	1987 job-complexity level		
	Lower	Same	Higher
10	1.07	1.42	NA
9	0.63	0.78	1.26
8	0.79	1.23	1.01
7	0.06	0.19	0.42
6	0.31	0.69	0.75
5	-0.12	0.08	0.43
4	-0.05	0.26	0.27
3	-0.47	-0.19	0.10
2	-0.28	-0.29	0.06
1	NA	-0.43	-0.04

total n = 3,887

Wilk, Desmarais, and Sackett (1995)

Table 4. Loadings of the 6 factor higher-order solution.

Dimension Name	Factor Name					
	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FAC6
Factor 1: Attending and Chairing Meetings						
Chair Mtgs: Outside sup/exec.	76	10	1	0	-1	5
Attend Mtgs: Outside/non-sup/exec.	72	5	3	9	-1	16
Attend Mtgs: Bargaining	68	9	3	1	23	0
Chair Mtgs: Regul/gov't/press	66	4	0	14	-4	-17
Chair Mtgs: Bargaining	66	9	3	-5	22	-10
Chair Mtgs: Public/customers	64	-10	-3	18	12	-9
Attend Mtgs: Regulators/gov't/press	63	3	0	23	-4	-11
Chair Mtgs: Own/other non-exec.	62	31	-3	-11	0	28
Chair Mtgs: Non-sup. empls.	62	12	1	1	-5	25
Attend Mtgs: Public/customers	61	-10	-2	25	14	-5
Chair Mtgs: Own/other,exec.	56	31	-2	-7	3	-3
Attend Mtgs: Own/other/non-sup.	51	-2	5	1	5	40
Attend Mtgs: Own/other exec.	51	34	-4	-3	6	10
Freq: Bargaining	49	13	1	2	30	-9
Freq: Chairing mtgs	46	42	-2	2	2	17
Freq: attending mtgs	39	31	2	9	5	35

Factor 2: Management

H. R. M.	9	86	-7	5	0	-5
Freq: H.R.M. benefits	2	79	-3	13	-8	3
Prod/operations mgt.	10	75	2	5	-3	13
Freq: H. R. M.	3	75	6	11	-9	16
Financial: Purchasing/budgeting	6	69	0	-7	4	8
Strategic planning: Products/services	18	65	0	0	5	-2
I.C.: Supervising mid-level emp	9	59	-8	8	21	2
Freq: Finance/budget	1	59	3	-6	12	12
Freq: Planning: Products/services	18	56	2	-2	2	4
I.C.: Supervising upper-level emp	13	43	-3	3	31	-9
I.C.: Resolving conflicts	12	42	-3	9	33	-6
Financial: Investments/cash	3	39	-6	-1	0	-8
I.C.: Supervising prof/technical	20	38	-3	-1	20	14
Strategic planning: Whole business	-1	37	3	1	6	-22
I.C.: Training	17	36	-7	-6	27	20
Freq: planning: Whole orgs.	-3	28	4	1	5	-17

Factor 3: Tool/Vehicle Use; Supervising Laborers

Freq: Tool/machine use	1	-3	85	-3	-4	-6
Stationary/manufacturing mach	6	-2	73	-19	0	-7

	<u>FACTOR1</u>	<u>FACTOR2</u>	<u>FACTOR3</u>	<u>FACTOR4</u>	<u>FACTOR5</u>	<u>FAC6</u>
Factor 3: Tool/Vehicle Use; Supervising Laborers (ct'd)						
Hand-held tools	0	-5	68	3	-1	21
Gross physical activity	-8	-5	65	35	-14	-2
Mobile tools/equip.	7	-3	65	-8	2	-2
Vehicle/machine repair	8	-1	63	-14	2	-2
Offroad/utility vehicles	5	-1	59	-13	2	-14
Direct other's tool/machine use	5	7	56	-5	4	-3
Hazardous work env.	-7	-6	56	35	-7	5
Freq: I.C.: Laborers	-17	9	48	1	25	5
Heavy highway vehicles	2	-1	48	-2	-1	-15
Sensory input from work env.	-8	1	48	44	-8	22
I.C.: Supervising laborers	-17	23	39	26	24	-1
Precise physical activity	-9	-5	39	29	-10	24
Water vehicles	6	4	22	5	-7	-1
Air vehicles	4	7	7	3	-4	-3
Factor 4: External Contacts						
Freq: E.C.: Pub, press, gov't	3	6	-9	83	-6	-4
E.C.: Pub, non-sales	6	1	1	80	14	-3
Demanding personal sits.	-6	7	8	69	-5	11
E.C.: Exchanging info	27	9	-7	61	26	-3

Factor 4: External Contacts (cont'd)

Firearms use	-2	4	-1	56	-25	6
Freq: E.C.	13	19	-8	48	23	12
E.C.: Resolving conflicts	19	-2	-5	47	27	-8
E.C.: Exchanging business info	24	12	-9	42	41	18
E.C.: Training	28	-1	3	41	31	-5
Light highway vehicles	10	2	26	33	-7	15
I.C.: Unions	1	7	9	26	12	0

Factor 5: Marketing/Sales, Service

E.C.: Customers: sales-related	15	-10	-2	15	72	6
I.C.: Marketing/sales employees	2	-1	-6	-10	65	12
Freq: Customers/mktg/sales	-9	-12	-7	5	59	11
E.C.: Supervising suppliers	19	14	6	6	50	16
I.C.: Selling/persuading	15	13	3	1	45	-1
E.C.: Bargaining/negot.	25	11	2	10	39	-14
I.C.: Bargaining/negot.	19	17	4	1	37	-11
I.C.: Supervising operatives	-7	16	30	-6	37	17
E.C.: Entertaining	7	-5	6	17	36	-9
I.C.: Entertaining	-5	3	7	9	33	-3
E.C.: Supervising: business related	21	5	7	24	30	-6

FACTOR1 FACTOR2 FACTOR3 FACTOR4 FACTOR5 FAC6

Factor 5: Marketing/Sales, Service

I.C.: Personal service employees	-9	10	3	16	23	-5
I.C.: Providing treatment, therapy	1	0	7	10	12	-5

Factor 6: Office/Technical Equipment Use

Office machines/equip.	2	2	-6	2	11	57
I.C.: Exchanging info	4	13	-7	-7	48	54
Freq: Office machines	-1	6	-26	-7	14	53
Technical/scientific equipment	-3	0	14	19	-18	52
Freq: I.C.: Mgrs, prof/tech	1	24	-12	-4	30	50
Attend Mtgs: Own/other/non-exec.	49	21	-2	-9	9	49
Numerical/verbal activities	0	7	25	10	9	48
Freq: Tech/scientific machs.	-3	1	7	20	-23	46

NOTE: Values greater than 25.4 have been marked by bold-faced print.

Table 5. Intercorrelations of the 6 second-order factors.

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6
FACTOR1	100	32	-2	16	29	16
FACTOR2	32	100	-2	9	23	20
FACTOR3	-2	-2	100	11	-2	2
FACTOR4	16	9	11	100	17	9
FACTOR5	29	23	-2	17	100	15
FACTOR6	16	20	2	9	15	100

NOTE: Printed values are multiplied by 100 and rounded to the nearest integer.

Values greater than 43.5 have marked by bold-faced print.

Table 6. Loadings of the 21 factor higher-order solution.

		Factor Name																				
Dimension name		Factor #																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
Factor 1: External Contacts, public, non-sales																						
Freq: E.C.: Pub, press, gov't																						
87	1	9	-2	-3	2	-3	0	2	-3	-5	0	5	2	-4	1	2	1	4	-2	-1		
E.C.: Pub, non-sales																						
69	3	10	1	5	-6	9	-2	-4	2	17	-1	2	4	-2	1	7	0	-9	1	18		
Freq: E.C.																						
58	5	-3	-2	-8	12	-7	6	18	5	23	14	6	-2	2	2	-1	1	25	3	-15		
Demanding personal sits.																						
56	8	37	-7	8	3	7	-1	-4	3	-16	11	-5	7	4	-1	4	2	-9	-3	-2		
E.C.: Exchanging info																						
54	4	-6	0	11	3	5	0	-3	4	34	-7	18	-3	5	2	8	2	4	3	11		
Firearms use																						
47	6	5	10	10	-1	-2	5	-9	-5	-7	-15	-4	14	10	-1	-2	-1	-17	26	-2		
E.C.: Resolving conflicts																						
33	0	4	2	22	8	7	-2	-3	-3	34	4	3	1	4	-5	10	0	-19	-4	0		
Factor 2: Management Decisions, Production Ops mgt,H.R.M.																						
Freq: H. R. M.																						
0	80	4	0	-3	3	4	3	-1	9	8	-5	0	4	7	14	1	-1	3	0	0		
Freq: H.R.M. benefits																						
9	79	-2	1	10	1	3	4	9	1	-7	5	2	0	-1	1	-2	2	-1	0	1		
Prod/operations mgt.																						
-5	62	4	-4	4	2	2	4	12	3	12	-6	5	5	4	22	1	3	6	2	2		
H.R.M.																						
4	60	-4	-1	20	2	2	4	32	-6	-5	4	8	-1	-6	8	-1	6	0	-1	4		

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

Factor 3: Physical Activity, Environmental Conditions

Gross physical activity

7	-2	71	11	-1	-3	-3	-1	-2	7	-3	-2	4	1	-18	-1	-2	0	-4	10	5
---	----	-----------	----	----	----	----	----	----	---	----	----	---	---	------------	----	----	---	----	----	---

Sensory input from work env.

15	3	71	1	0	2	3	1	1	12	0	-1	-7	8	3	0	6	0	-3	0	0
-----------	---	-----------	---	---	---	---	---	---	----	---	----	----	---	---	---	---	---	----	---	---

Precise physical activity

4	-1	71	3	-2	-4	1	0	-3	-2	-2	0	5	2	13	2	-5	0	2	-5	4
---	----	-----------	---	----	----	---	---	----	----	----	---	---	---	----	---	----	---	---	----	---

Hazardous work env.

17	-2	51	13	-3	2	-1	0	-4	12	-1	6	-3	8	-13	-3	0	3	-1	14	-5
-----------	----	-----------	----	----	---	----	---	----	----	----	---	----	---	-----	----	---	---	----	----	----

Numerical/verbal activities

-8	4	42	0	-1	3	11	0	10	4	14	4	-8	11	23	0	8	3	17	4	-7
----	---	-----------	---	----	---	----	---	----	---	----	---	----	----	-----------	---	---	---	-----------	---	----

Factor 4: Skilled Machine Use

Stationary/manufacturing mach

-4	-1	6	85	1	-2	-2	0	-2	3	1	-4	-1	1	0	1	2	0	-1	-4	-2
----	----	---	-----------	---	----	----	---	----	---	---	----	----	---	---	---	---	---	----	----	----

Vehicle/machine repair

2	-3	-1	75	1	0	3	0	2	-1	-5	0	0	4	4	1	0	1	-3	5	3
---	----	----	-----------	---	---	---	---	---	----	----	---	---	---	---	---	---	---	----	---	---

Direct other's tool/machine use

4	9	3	56	0	2	2	1	1	14	2	-4	-2	-1	0	-2	1	-3	-3	3	1
---	---	---	-----------	---	---	---	---	---	----	---	----	----	----	---	----	---	----	----	---	---

Freq: Tool/machine use

-13	0	37	47	-4	1	-5	-2	-1	10	10	-2	0	-2	-16	0	1	0	4	25	-7
-----	---	-----------	-----------	----	---	----	----	----	----	----	----	---	----	------------	---	---	---	---	-----------	----

Mobile tools/equip.

-6	-4	12	44	-3	2	2	-2	3	2	-4	3	7	0	-4	2	-5	0	2	32	10
----	----	----	-----------	----	---	---	----	---	---	----	---	---	---	----	---	----	---	---	-----------	----

Offroad/utility vehicles

-12	-2	11	38	-2	2	0	-3	-1	8	2	-2	7	-9	-12	-1	-4	4	2	23	2
-----	----	----	-----------	----	---	---	----	----	---	---	----	---	----	-----	----	----	---	---	-----------	---

Factor 5: Internal Contacts, mid-to upper-level employees

I.C.: Resolving conflicts

5	7	-2	0	75	6	3	3	2	8	3	2	2	-4	1	2	2	-1	0	0	-6
---	---	----	---	-----------	---	---	---	---	---	---	---	---	----	---	---	---	----	---	---	----

I.C.: Supervising upper-level emp

-1	5	1	2	69	4	0	1	4	-3	6	2	7	-2	-2	3	0	10	5	-2	0
----	---	---	---	-----------	---	---	---	---	----	---	---	---	----	----	---	---	----	---	----	---

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

Factor 5: Internal Contacts, mid-to upper-level employees (cont'd)

I.C.: Supervising mid-level emp

6	35	-6	-1	52	-1	1	4	10	5	0	3	4	-7	5	1	7	-3	7	1	-4
---	----	----	----	----	----	---	---	----	---	---	---	---	----	---	---	---	----	---	---	----

I.C.: Training

-5	2	-3	-5	48	2	9	10	1	-2	4	-7	1	3	6	14	-3	-2	27	4	16
----	---	----	----	----	---	---	----	---	----	---	----	---	---	---	----	----	----	----	---	----

I.C.: Supervising prof/technical

-2	6	2	-3	42	6	2	17	5	-6	11	-5	-1	11	-5	10	-5	0	22	-1	8
----	---	---	----	----	---	---	----	---	----	----	----	----	----	----	----	----	---	----	----	---

I.C.: Unions

21	5	-1	-1	23	6	7	-6	-7	17	-4	-3	0	0	4	-6	5	-4	-4	13	3
----	---	----	----	----	---	---	----	----	----	----	----	---	---	---	----	---	----	----	----	---

Factor 6: Bargaining/Negotiating

Freq: Bargaining

5	5	-3	-1	0	73	3	0	4	1	0	8	-6	-1	-1	6	18	6	5	2	3
---	---	----	----	---	----	---	---	---	---	---	---	----	----	----	---	----	---	---	---	---

Chair Mtgs: Bargaining

-2	0	1	0	6	64	-3	26	2	0	2	3	11	-3	4	6	4	0	-3	1	3
----	---	---	---	---	----	----	----	---	---	---	---	----	----	---	---	---	---	----	---	---

Attend Mtgs: Bargaining

0	2	1	-2	8	59	25	7	2	2	8	3	23	-2	3	3	1	0	-4	0	-2
---	---	---	----	---	----	----	---	---	---	---	---	----	----	---	---	---	---	----	---	----

E.C.: Bargaining/negot.

1	3	-1	3	12	35	2	-3	9	2	30	7	3	1	0	1	2	4	-14	-7	8
---	---	----	---	----	----	---	----	---	---	----	---	---	---	---	---	---	---	-----	----	---

I.C.: Selling/persuading

0	-1	-2	-3	28	31	5	0	-2	12	3	10	-1	0	1	5	-5	-5	3	2	24
---	----	----	----	----	----	---	---	----	----	---	----	----	---	---	---	----	----	---	---	----

I.C.: Bargaining/negot.

-2	-1	1	4	28	28	-2	5	4	5	9	8	3	5	-3	3	-6	3	-5	-8	17
----	----	---	---	----	----	----	---	---	---	---	---	---	---	----	---	----	---	----	----	----

Factor 7: Attending Meetings

Attend Mtgs: Own/other/non-sup.

1	-1	3	2	1	0	73	14	0	5	0	5	6	-1	11	1	2	-1	-6	1	3
---	----	---	---	---	---	----	----	---	---	---	---	---	----	----	---	---	----	----	---	---

Attend Mtgs: Own/other/non-exec.

-5	7	0	-3	4	10	59	7	6	-2	12	-3	7	8	3	9	-2	-2	27	1	-3
----	---	---	----	---	----	----	---	---	----	----	----	---	---	---	---	----	----	----	---	----

Freq: attending mtgs

13	20	1	2	4	19	49	-2	8	-1	0	-1	-8	7	-5	7	15	2	24	1	-6
----	----	---	---	---	----	----	----	---	----	---	----	----	---	----	---	----	---	----	---	----

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

Factor 8: Chairing Meetings

Chair Mtgs: Outside sup/exec.

-2	2	2	-2	2	12	-2	66	5	-2	13	2	18	2	1	6	9	-2	-1	2	-3
----	---	---	----	---	----	----	-----------	---	----	----	---	-----------	---	---	---	---	----	----	---	----

Chair Mtgs: Non-sup. empls.

4	8	-3	0	4	0	39	56	0	7	-5	-1	-8	-3	7	0	13	1	-8	2	4
---	---	----	---	---	---	-----------	-----------	---	---	----	----	----	----	---	---	----	---	----	---	---

Chair Mtgs: Own/other non-exec.

-2	12	-3	-1	10	15	25	50	5	-3	3	-7	-4	3	-3	8	4	0	24	-2	1
----	----	----	----	----	----	-----------	-----------	---	----	---	----	----	---	----	---	---	---	-----------	----	---

Attend Mtgs: Outside/non-sup/exec.

2	2	5	-4	-1	5	32	39	5	-2	18	2	29	2	0	6	2	-3	-6	5	-3
---	---	---	----	----	---	-----------	-----------	---	----	-----------	---	-----------	---	---	---	---	----	----	---	----

Chair Mtgs: Own/other, exec.

5	10	0	5	14	23	2	36	6	-8	-9	9	30	1	-9	7	-13	4	12	-10	-1
---	----	---	---	----	-----------	---	-----------	---	----	----	---	-----------	---	----	---	-----	---	----	-----	----

Freq: Chairing mtgs

11	30	-5	3	6	25	16	33	7	0	-6	-4	-22	1	-7	3	23	4	22	-1	1
----	-----------	----	---	---	-----------	-----------	-----------	---	---	----	----	------------	---	----	---	-----------	---	-----------	----	---

Factor 9: Decisions: Financial

Freq: Finance/budget

3	6	-1	2	-1	4	2	3	79	7	4	3	-1	0	6	7	0	0	3	-1	-2
---	---	----	---	----	---	---	---	-----------	---	---	---	----	---	---	---	---	---	---	----	----

Financial: Purchasing/budgeting

-3	17	-1	-2	7	2	3	3	73	-1	2	-5	3	2	2	9	1	2	3	3	4
----	-----------	----	----	---	---	---	---	-----------	----	---	----	---	---	---	---	---	---	---	---	---

Financial: Investments/cash

4	5	-1	0	5	0	3	2	41	-4	-1	-1	5	-6	2	7	-3	14	-9	-7	-1
---	---	----	---	---	---	---	---	-----------	----	----	----	---	----	---	---	----	----	----	----	----

Factor 10: Internal Contacts: low level supv.

Freq: I.C.: Laborers

-4	0	15	7	-5	3	-1	-3	3	70	0	2	4	0	-8	-1	-3	1	8	-1	-4
----	---	-----------	---	----	---	----	----	---	-----------	---	---	---	---	----	----	----	---	---	----	----

I.C.: Supervising operatives

-9	5	-2	8	7	5	6	4	1	62	11	0	0	7	7	0	-3	2	6	-6	4
----	---	----	---	---	---	---	---	---	-----------	----	---	---	---	---	---	----	---	---	----	---

I.C.: Supervising laborers

16	13	10	2	9	-5	1	0	6	60	0	4	-3	-1	-6	0	3	-2	-7	12	3
-----------	----	----	---	---	----	---	---	---	-----------	---	---	----	----	----	---	---	----	----	----	---

I.C.: Personal service employees

16	-8	-1	-5	13	0	-1	0	5	24	-6	9	0	2	-3	10	-3	0	-5	-3	10
-----------	----	----	----	----	---	----	---	---	-----------	----	---	---	---	----	----	----	---	----	----	----

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

Factor 11: External Contacts: Supervising

E.C.: Supervising suppliers

-2	0	-1	0	8	8	4	12	19	12	53	13	0	2	4	1	-2	-1	9	1	2
----	---	----	---	---	---	---	----	----	----	----	----	---	---	---	---	----	----	---	---	---

E.C.: Exchanging business info

39	11	-8	-2	3	-2	6	11	0	2	46	17	9	-3	7	3	2	-1	19	8	0
----	----	----	----	---	----	---	----	---	---	----	----	---	----	---	---	---	----	----	---	---

E.C.: Supervising: business related

10	1	1	6	13	-3	6	7	4	8	38	-1	10	2	-3	1	6	0	-10	-8	14
----	---	---	---	----	----	---	---	---	---	----	----	----	---	----	---	---	---	-----	----	----

E.C.: Training

25	-6	5	-1	11	3	11	4	-2	-1	29	0	8	4	-4	7	13	4	-5	2	27
----	----	---	----	----	---	----	---	----	----	----	---	---	---	----	---	----	---	----	---	----

Factor 12: Marketing/Sales

Freq: Customers/mktg/sales

7	-1	4	-4	-5	2	-2	-2	1	2	1	78	-3	-5	3	3	7	3	5	-2	-5
---	----	---	----	----	---	----	----	---	---	---	----	----	----	---	---	---	---	---	----	----

I.C.: Marketing/sales employees

-8	0	-6	-2	13	9	6	1	-1	2	2	61	-3	1	6	3	0	3	6	1	15
----	---	----	----	----	---	---	---	----	---	---	----	----	---	---	---	---	---	---	---	----

E.C.: Customers: sales-related

3	2	-1	-1	4	13	5	1	-1	1	40	53	-5	-1	5	1	14	0	-6	2	11
---	---	----	----	---	----	---	---	----	---	----	----	----	----	---	---	----	---	----	---	----

Factor 13: Meetings: Regulators/Gov't/Press

Attend Mtgs: Regulators/gov't/press

10	4	-2	-2	4	4	12	-2	2	3	6	-9	68	-1	4	0	20	2	-2	4	1
----	---	----	----	---	---	----	----	---	---	---	----	----	----	---	---	----	---	----	---	---

Chair Mtgs: Regul/gov't/press

4	1	-1	1	5	7	-13	22	0	4	-1	-6	56	-2	5	2	28	4	2	-1	0
---	---	----	---	---	---	-----	----	---	---	----	----	----	----	---	---	----	---	---	----	---

Attend Mtgs: Own/other exec.

7	17	0	5	14	17	27	12	6	-10	-2	11	37	2	-8	7	-19	1	15	-11	-4
---	----	---	---	----	----	----	----	---	-----	----	----	----	---	----	---	-----	---	----	-----	----

Factor 14: Technical/Scientific Equipment Use

Technical/scientific equipment

2	2	2	5	-1	-4	3	2	1	3	0	-1	1	79	9	0	0	-1	-1	6	2
---	---	---	---	----	----	---	---	---	---	---	----	---	----	---	---	---	----	----	---	---

Freq: Tech/scientific machs.

3	2	8	-2	-2	2	1	-1	0	0	0	-3	-1	78	-1	1	0	0	2	-4	-3
---	---	---	----	----	---	---	----	---	---	---	----	----	----	----	---	---	---	---	----	----

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

Factor 15: Office Machine Use

Freq: Office machines

0	4	-4	-4	-1	5	6	0	7	-6	2	8	1	2	63	1	-4	1	12	-4	-5
---	---	----	----	----	---	---	---	---	----	---	---	---	---	-----------	---	----	---	----	----	----

Office machines/equip.

0	3	2	5	1	0	7	2	5	-4	3	4	6	15	62	2	-4	-3	5	7	5
---	---	---	---	---	---	---	---	---	----	---	---	---	-----------	-----------	---	----	----	---	---	---

Factor 16: Strategic Planning: Products/Services

Freq: Planning: Products/services

2	11	1	2	0	6	5	4	5	2	0	3	-1	0	2	73	3	7	4	-1	-2
---	----	---	---	---	---	---	---	---	---	---	---	----	---	---	-----------	---	---	---	----	----

Strategic planning: Products/services

0	15	-1	-1	8	4	2	4	13	-2	2	3	3	1	0	70	3	8	-1	2	2
---	-----------	----	----	---	---	---	---	----	----	---	---	---	---	---	-----------	---	---	----	---	---

Factor 17: Meetings: Public/Customers

Chair Mtgs: Public/customers

5	-2	0	1	1	12	-5	20	0	-1	1	9	11	0	-3	5	67	-2	3	-1	2
---	----	---	---	---	----	----	-----------	---	----	---	---	----	---	----	---	-----------	----	---	----	---

Attend Mtgs: Public/customers

7	2	1	0	1	10	18	-5	1	-4	6	11	25	1	-5	3	59	-2	-1	1	2
---	---	---	---	---	----	-----------	----	---	----	---	----	-----------	---	----	---	-----------	----	----	---	---

Factor 18: Strategic Planning: Entire Organization

Strategic planning: Whole business

0	2	0	0	6	1	0	-1	5	-1	0	2	4	0	-1	7	0	76	-2	1	0
---	---	---	---	---	---	---	----	---	----	---	---	---	---	----	---	---	-----------	----	---	---

Freq: planning: Whole orgs.

2	2	1	0	-2	4	-1	1	-2	2	0	1	-1	-1	0	5	-2	74	2	0	-1
---	---	---	---	----	---	----	---	----	---	---	---	----	----	---	---	----	-----------	---	---	----

Factor 19: Internal Contacts: Exchanging Info

Freq: I.C.: Mgrs, prof/tech

6	6	2	-5	15	-2	9	-1	9	7	-2	11	1	1	25	1	5	0	51	-5	0
---	---	---	----	-----------	----	---	----	---	---	----	----	---	---	-----------	---	---	---	-----------	----	---

I.C.: Exchanging info

0	1	-7	-4	16	-6	18	4	0	18	8	20	1	3	28	1	1	-1	44	3	10
---	---	----	----	-----------	----	-----------	---	---	-----------	---	-----------	---	---	-----------	---	---	----	-----------	---	----

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

Factor 20: Vehicle Operation

Light highway vehicles

23 2 11 2 3 6 4 4 3 -2 5 -6 -7 3 13 2 7 -5 -5 43 -5

Water vehicles

3 2 -5 5 -2 0 1 0 3 -4 -5 0 9 5 -1 -2 -3 7 0 39 5

Hand-held tools

-11-3 30 31 -3 -4 2 2 3 7 10 2 -1 12 4 1 -1 -1 4 36 0

Heavy highway vehicles

-5 0 11 19 -2 -1 0 -2 -3 7 0 4 7 -8 -22 0 -7 0 3 31 0

Air vehicles

1 7 -3 -3 -5 6 0 3 -3 -3 -4 0 0 1 0 -1 -2 15 1 22 5

Factor 21: Entertaining/Serving

I.C.: Entertaining

3 5 5 0 0 1 -1 0 0 6 -2 6 -1 -3 3 -1 -1 0 5 0 51

E.C.: Entertaining

7 3 4 2 -8 9 -2 -2 3 -4 14 7 -1 -2 -2 -1 7 -1 2 1 51

I.C.: Providing treatment, therapy

10 -7 2 3 4 -1 6 -1 2 11 -10 3 4 0 -3 4 2 -1 -3 -3 19

Note: Values greater than 14.6 have been marked by bold-faced print.

Table 7. Intercorrelations of the 21 second-order factors.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21
F1	100	14	18	-9	15	9	12	7	3	6	26	8	13	8	5	4	20	3	-2	10	12
F2	14	100	0	-2	27	15	21	22	33	8	8	2	9	6	7	33	7	8	20	2	2
F3	18	0	100	25	-5	-1	3	-2	-3	22	1	1	-2	16	-5	0	2	1	-3	27	2
F4	-9	-2	25	100	-4	1	0	-1	1	17	3	-4	4	2	-12	-1	-3	1	-3	34	3
F5	15	27	-5	-4	100	25	19	21	16	11	21	9	16	1	8	19	5	6	22	-4	20
F6	9	15	-1	1	25	100	25	35	15	3	21	20	22	0	1	20	24	8	12	-1	15
F7	12	21	3	0	19	25	100	39	12	7	20	10	18	10	19	17	17	-1	30	3	7
F8	7	22	-2	-1	21	35	39	100	16	-1	18	3	29	4	5	21	23	1	21	0	3
F9	3	33	-3	1	16	15	12	16	100	6	15	7	10	2	11	27	2	10	18	-2	1
F10	6	8	22	17	11	3	7	-1	6	100	13	10	-1	4	-3	1	1	-1	8	11	10
F11	26	8	1	3	21	21	20	18	15	13	100	26	20	4	13	11	19	0	11	5	24
F12	8	2	1	-4	9	20	10	3	7	10	26	100	0	-5	14	7	14	5	18	-5	17
F13	13	9	-2	4	16	22	18	29	10	-1	20	0	100	-2	0	11	26	5	2	1	5
F14	8	6	16	2	1	0	10	4	2	4	4	-5	-2	100	18	6	1	-2	7	7	1
F15	5	7	-5	-12	8	1	19	5	11	-3	13	14	0	18	100	5	1	-3	28	-3	1
F16	4	33	0	-1	19	20	17	21	27	1	11	7	11	6	5	100	9	18	17	-1	8
F17	20	7	2	-3	5	24	17	23	2	1	19	14	26	1	1	9	100	-1	2	2	9
F18	3	8	1	1	6	8	-1	1	10	-1	0	5	5	-2	-3	18	-1	100	1	1	1
F19	-2	20	-3	-3	22	12	30	21	18	8	11	18	2	7	28	17	2	1	100	-2	-1
F20	10	2	27	34	-4	-1	3	0	-2	11	5	-5	1	7	-3	-1	2	1	-2	100	3
F21	12	2	2	3	20	15	7	3	1	10	24	17	5	1	1	8	9	1	-1	3	100

Note: Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 25.3 have been marked by bold-faced print.

Table 8. Multiple correlations, R²'s, and associated p-values for the four MBTI scales.

MBTI scale	Model Type					
	<u>6 Factor</u>			<u>21 Factor</u>		
	Multiple R	R ²	p	Multiple R	R ²	p
EI	.6046	.3656	.0001	.7321	.5360	.0001
SN	.5425	.2943	.0001	.6566	.4311	.0001
TF	.3494	.1221	.0336	.5550	.3082	.0233
JP	.3184	.1014	.0816	.5060	.2560	.1211

Table 9. Regression weights and associated p-values for the 21 predictor model for all four MBTI scales.

Extraversion/Introversion			
Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	51.770811	52.069	0.0001
FACTOR1	-0.215618	-0.189	0.8503
FACTOR2	0.430353	0.344	0.7320
FACTOR3	-1.553412	-1.400	0.1651
FACTOR4	-5.600840	-2.260	0.0263
FACTOR5	0.468998	0.194	0.8466
FACTOR6	2.300744	1.484	0.1415
FACTOR7	-1.455609	-1.078	0.2841
FACTOR8	3.805047	2.607	0.0107
FACTOR9	0.061027	0.051	0.9597
FACTOR10	-3.264674	-2.356	0.0207
FACTOR11	-1.570751	-1.001	0.3198
FACTOR12	4.564343	5.208	0.0001
FACTOR13	2.062340	1.350	0.1803
FACTOR14	-1.315067	-1.424	0.1580
FACTOR15	0.113640	0.096	0.9235
FACTOR16	0.903810	0.869	0.3872

FACTOR17	-1.070781	-0.865	0.3893
FACTOR18	2.434864	1.601	0.1130
FACTOR19	-0.719389	-0.577	0.5656
FACTOR20	2.673395	1.647	0.1032
FACTOR21	2.431499	1.110	0.2700

Sensing/Intuition

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	55.224519	26.921	0.0001
FACTOR1	3.817967	1.625	0.1078
FACTOR2	4.750080	1.838	0.0694
FACTOR3	2.809728	1.227	0.2231
FACTOR4	-6.262102	-1.225	0.2240
FACTOR5	-5.329582	-1.069	0.2881
FACTOR6	-0.734914	-0.230	0.8189
FACTOR7	-7.330481	-2.631	0.0101
FACTOR8	1.682522	0.559	0.5777
FACTOR9	-4.510958	-1.817	0.0726
FACTOR10	3.914977	1.369	0.1744
FACTOR11	-5.614251	-1.733	0.0865

FACTOR12	1.215354	0.672	0.5033
FACTOR13	1.199845	0.381	0.7043
FACTOR14	-2.299687	-1.207	0.2307
FACTOR15	0.244557	0.100	0.9202
FACTOR16	-4.126378	-1.923	0.0577
FACTOR17	-3.862327	-1.512	0.1340
FACTOR18	-2.558051	-0.815	0.4171
FACTOR19	-1.206371	-0.469	0.6404
FACTOR20	4.473842	1.336	0.1851
FACTOR21	5.468261	1.210	0.2295

Thinking/Feeling

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	49.921390	22.807	0.0001
FACTOR1	-4.745895	-1.893	0.0617
FACTOR2	5.349478	1.940	0.0556
FACTOR3	-1.585795	-0.649	0.5180
FACTOR4	3.422780	0.627	0.5321
FACTOR5	-4.273008	-0.803	0.4241
FACTOR6	4.701428	1.377	0.1720

FACTOR7	5.288530	1.779	0.0787
FACTOR8	-8.649566	-2.692	0.0085
FACTOR9	-2.351228	-0.888	0.3772
FACTOR10	3.357918	1.101	0.2741
FACTOR11	0.889642	0.257	0.7975
FACTOR12	-1.436093	-0.744	0.4587
FACTOR13	3.063697	0.911	0.3647
FACTOR14	1.350479	0.664	0.5083
FACTOR15	1.498582	0.577	0.5657
FACTOR16	2.466950	1.077	0.2842
FACTOR17	2.434885	0.894	0.3740
FACTOR18	4.497786	1.343	0.1826
FACTOR19	-2.499858	-0.910	0.3652
FACTOR20	4.259075	1.192	0.2366
FACTOR21	-4.840691	-1.004	0.3182

Judgment/Perception

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	61.168782	45.160	0.0001
FACTOR1	2.451531	1.580	0.1177

FACTOR2	4.124721	2.418	0.0177
FACTOR3	0.524581	0.347	0.7295
FACTOR4	-4.423222	-1.310	0.1936
FACTOR5	-2.650673	-0.805	0.4229
FACTOR6	2.778386	1.315	0.1919
FACTOR7	0.937791	0.510	0.6115
FACTOR8	-0.446313	-0.224	0.8229
FACTOR9	-0.722306	-0.441	0.6606
FACTOR10	2.098552	1.112	0.2694
FACTOR11	-1.908719	-0.893	0.3745
FACTOR12	-1.791014	-1.500	0.1372
FACTOR13	-0.192102	-0.092	0.9266
FACTOR14	-0.756215	-0.601	0.5494
FACTOR15	4.729675	2.941	0.0042
FACTOR16	-0.766217	-0.541	0.5900
FACTOR17	1.126470	0.668	0.5058
FACTOR18	-0.313255	-0.151	0.8802
FACTOR19	-3.610157	-2.124	0.0364
FACTOR20	2.218460	1.003	0.3185
FACTOR21	-0.601305	-0.202	0.8408

.

Table 10. Regression weights and associated p-values for the 6 predictor model for all four MBTI scales.

Extraversion/Introversion

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	52.085572	72.317	0.0001
F1 Meetings	0.705188	0.544	0.5875
F2 Management	2.364947	3.015	0.0032
F3 Tool Use	-3.303387	-2.774	0.0066
F4 Ext.Contcs	-1.182887	-1.075	0.2847
F5 Sales/Mktg	5.148621	4.322	0.0001
F6 Equip.Use	-2.409965	-3.101	0.0025

Sensing/Intuition

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	54.581655	38.560	0.0001
F1 Meetings	-10.271891	-4.034	0.0001
F2 Management	-1.690613	-1.097	0.2753

F3 Tool Use	3.355374	1.434	0.1546
F4 Ext.Contcs	3.890536	1.800	0.0749
F5 Sales/Mktg	0.306729	0.131	0.8960
F6 Equip.Use	-4.283861	-2.805	0.0060

Thinking/Feeling

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	49.663843	32.509	0.0001
F1 Meetings	-0.161767	-0.059	0.9532
F2 Management	3.928957	2.362	0.0201
F3 Tool Use	4.062510	1.609	0.1108
F4 Ext.Contcs	-4.992025	-2.140	0.0348
F5 Sales/Mktg	0.772770	0.306	0.7604
F6 Equip.Use	1.554655	0.943	0.3478

Judgment/Perception

Variable	Parameter Estimate	T for H0: Parameter=0	Prob > T
INTERCEPT	60.254847	65.338	0.0001
F1 Meetings	-1.774389	-1.069	0.2874
F2 Management	2.263938	2.254	0.0263
F3 Tool Use	-1.651194	-1.083	0.2813
F4 Ext.Contcs	2.372882	1.685	0.0951
F5 Sales/Mktg	-1.163371	-0.763	0.4474
F6 Equip.Use	0.932053	0.937	0.3511

Table 11. Regression results for interaction analyses for 6 and 21 predictor models.

6 Predictor Model

Model Type

Gender as a main effect

Gender as main effect and interaction

Change in R^2

MBTI scale	R^2	p	R^2	p	Fobt	sig?(Fcrit =2.21).
EI	.3870	.0001	.4757	.0001	2.31	Yes
SN	.3373	.0001	.4199	.0001	1.95	No
TF	.5242	.0001	.5865	.0001	2.06	No
JP	.0931	.2645	.1503	.3584	0.92	No

21 Predictor Model

Model Type

Gender as a main effect

Gender as main effect and interaction

Change in R^2

MBTI scale	R^2	p	R^2	p	Fobt	sig?(Fcrit =1.78).
EI	.5891	.0001	.7902	.0001	2.37	Yes
SN	.4603	.0005	.6167	.0112	1.01	No
TF	.6463	.0001	.7521	.0001	1.06	No
JP	.3196	.0818	.5017	.2475	0.90	No

113

Table 12. Regression results for analysis of gender within 6 and 21 predictor models.

6 Predictor Model			
<u>Results without Gender</u>		<u>Gender as main effect</u>	
		Change in R ²	
MBTI scale	R ²	R ²	Significant Change? .
EI	.3678	.3870	No
SN	.3298	.3373	No
TF	.1004	.5242	Yes
JP	.0931	.0931	No

21 Predictor Model			
<u>Results without Gender</u>		<u>Gender as main effect</u>	
		Change in R ²	
MBTI scale	R ²	R ²	Significant Change? .
EI	.5862	.5891	No
SN	.4557	.4603	No
TF	.4001	.6463	Yes
JP	.3168	.3196	No

Note: In order to have comparable models, only cases where gender domination data was obtained (n = 96) were used in the analyses without gender.

Table 13. Residual group means and ANOVA results for each scale.

Extraversion/Introversion					
Job Family Name		Mean of Absolute Value of Residual		S. D.	
Manager		4.02		3.17	
Supervisor		2.96		2.55	
Professional or Technical Specialist		4.16		3.44	
Marketing and Sales		3.60		2.34	
Administrative Support/General Service		3.96		3.53	
Precision Production Craft & Repair		2.53		1.59	
Operator, Fabricator, & Laborer		3.61		0.19	
Source	Df	Sum of Sq.	Mean Sq.	F-value	Pr > F
Between	6	24.997	4.166	0.40	.8787
Within	103	1077.67	10.463		
Total	109	1102.669			

Sensing/Intuition

<u>Job Family Name</u>	<u>Mean of Absolute Value of Residual</u>	<u>S. D.</u>
Manager	5.71	4.86
Supervisor	7.11	6.85
Professional or Technical Specialist	8.46	6.72
Marketing and Sales	4.62	1.51
Administrative Support/General Service	8.78	7.61
Precision Production Craft & Repair	6.86	6.44
Operator, Fabricator, & Laborer	8.10	1.07

<u>Source</u>	<u>Df</u>	<u>Sum of Sq.</u>	<u>Mean Sq.</u>	<u>F-value</u>	<u>Pr > F</u>
Between	6	131.377	21.896	0.50	0.811
Within	103	4555.60	44.229		
Total	109	4686.98			

Thinking/Feeling

<u>Job Family Name</u>	<u>Mean of Absolute Value of Residual</u>	<u>S. D.</u>
Manager	12.62	10.1
Supervisor	7.84	9.20
Professional or Technical Specialist	9.79	6.39
Marketing and Sales	5.03	5.20
Administrative Support/General Service	8.42	6.20
Precision Production Craft & Repair	6.33	6.15
Operator, Fabricator, & Laborer	2.99	2.32

<u>Source</u>	<u>Df</u>	<u>Sum of Sq.</u>	<u>Mean Sq.</u>	<u>F-value</u>	<u>Pr > F</u>
Between	6	356.459	59.41	1.36	.2360
Within	103	4484.939	43.54		
Total	109	4841.398			

Judgment/Perception

<u>Job Family Name</u>	<u>Mean of Absolute Value of Residual</u>	<u>S. D.</u>
Manager	6.19	7.26
Supervisor	4.32	2.90
Professional or Technical Specialist	5.50	4.13
Marketing and Sales	4.78	3.91
Administrative Support/General Service	5.11	4.14
Precision Production Craft & Repair	5.86	6.81
Operator, Fabricator, & Laborer	5.88	1.52

<u>Source</u>	<u>Df</u>	<u>Sum of Sq.</u>	<u>Mean Sq.</u>	<u>F-value</u>	<u>Pr > F</u>
Between	6	18.584	3.097	0.16	.9857
Within	103	1945.089	18.88		
Total	109	1963.673			

Table 14. Actual and predicted MBTI types for each job title.

Job Title	Modal Type	REGTYPE	DISCTYPE	CmpstType
Accountants	ISTJ	ESFJ	ISTJ	ISTJ
Admin: managers & sups misc	ESTJ	ENFJ	ENFP	ESTJ
Administrators: health	ENTJ	ESTJ	ENTJ	ESTJ
Auditors	ISTJ	ESTJ	ISTJ	ESTJ
Bank officers and fin. mgrs	ESTJ	ENTJ	ENFP	ESTJ
Bookkeepers	ISFJ	ENTJ	ISTJ	ISFJ
Carpenters	INFP	ISTJ	ISTJ	ISFP
Cashiers	ISFJ	ESFJ	ISFJ	ESFJ
Chain, rod & axe, surveying	ENFP	ISTJ	ESFJ	INFP
Child care workers	ESFJ	ESFJ	ISFJ	ESFJ
Cleaning services	ESTJ	ESFJ	ESTJ	ISTJ
Clerical supervisors: misc.	ISFJ	ESTJ	ESTJ	ESFJ
Clerical workers, misc.	ISFJ	ESFJ	ISFJ	ESFJ
Clinical & lab tech.	ISTJ	ISTJ	ISTJ	ISFJ
Computer & peripheral ops.	ENFP	ISFJ	ISTJ	ESFJ
Computer analysts, sprt rep	ESTJ	ENTJ	ESTJ	INTJ
Computer programmers	ISTJ	INTJ	ISTJ	INTJ

Job Title	Modal Type	REGTYPE	DISCTYPE	CmpstType
Computer specialists	ENTJ	ENTJ	ISTJ	ISTJ
Consultants: education	INFJ	ENFJ	INFJ	INFJ
Consultants: mangment anlst	ISTJ	ESFJ	ISTJ	ENTJ
Consultants: unspec.	ENTJ	ENTJ	ENTJ	ENTJ
Cooks	ESTJ	ESFJ	ISFJ	ISFJ
Coordinators: unspec.	ESTJ	ENFJ	ESTJ	ENFJ
Corrections & probation off	ISFJ	ENFJ	ENFP	ESTJ
Counselors: general	ENFP	ENFJ	ENFP	ENFJ
Counselors: rehab.	ENFP	ENFJ	ENFP	ENFP
Counselors: vocational & ed	ENFP	ENTJ	ENFP	ENFJ
Craft workers	ISTJ	ISFJ	ISTJ	ISTJ
Credit investigators & brok	ESTJ	ENTJ	ESTJ	ENTJ
Dental assistants	ESFJ	ESFJ	ESFJ	ESFJ
Dental hygienists	ISFJ	INFJ	ISFJ	ESFJ
Designers	ISTJ	INTJ	ISTJ	ENFJ
Dietitians, nutritionists	ISFJ	ISFJ	ISFJ	ISFJ
Editors & reporters	INFP	ENFJ	INFP	INFP
Electricians	ISTJ	ISTJ	ISTJ	ISTJ
Employment dev. specialists	ENTJ	ESFJ	ENTJ	ENTJ
Engineering & science tech.	ISTJ	ISFJ	ISTJ	ISTJ

<u>Job Title</u>	<u>Modal Type</u>	<u>REGTYPE</u>	<u>DISCTYPE</u>	<u>CmpstType</u>
Engineers	ISTJ	ISTJ	ISTJ	ISTJ
Engineers: electrical	ISTJ	INTJ	ISTJ	ISTJ
Engineers: mechanical	ISTJ	ISTJ	ISTJ	ISTJ
Factory & site supervisors	ESTJ	ISFJ	ISTJ	ESTJ
Food service workers, misc.	ISTJ	ESFJ	ESTJ	ESFJ
Guards & watch keepers	ISFJ	ESFJ	ISFJ	ISFJ
Health technicians	ISFJ	ENTJ	ENFP	ISFJ
Insurance agents & undrwtr	ESTJ	ESTJ	ESTJ	ESTJ
Lawyers	INTJ	ENFJ	ISFJ	INTJ
Librarians & curators	ISFJ	ISTJ	ISTJ	ISFJ
Lifeguards, rec attendants	ESTJ	ESFJ	ESTJ	ESFJ
Managers: restaurant & food	ESTJ	ESTJ	ESTJ	ESTJ
Managers: sales, unspec.	ESTJ	ESTJ	ESTJ	ESTJ
Marketing personnel	ENTJ	ENTJ	ENTJ	ENTJ
Mechanics & repairers unspc	ISTJ	ISTJ	INFP	ISTJ
Media specialists	ISTJ	INTJ	ISTJ	ISFJ
Medical assistants	ESTJ	ESFJ	ISFJ	ESFJ
Medical technologists	ISFJ	ISFJ	ISFJ	ISFJ
Mics. laborers	ESTJ	ISFJ	ISTJ	ESTP
Misc. ops. & factory worker	ISTJ	ISTJ	ISTJ	ISTJ

<u>Job Title</u>	<u>Modal Type</u>	<u>REGTYPE</u>	<u>DISCTYPE</u>	<u>CmpstType</u>
Nurses: aides & orderlies	ISFJ	ESFJ	ISFJ	ESFJ
Nurses: licensed practical	ISFJ	ISFJ	ISFJ	ISFJ
Nursing: administrators	ESTJ	ISTJ	ESTJ	ISTJ
Nursing: consultants	ESTJ	ESTJ	ISTJ	ESFJ
Nursing: registered, unspec	ISFJ	ISFJ	ISFJ	ISFJ
Office managers	ESFJ	ESFJ	ESFJ	ESFJ
Operations & systems anlst	ENTJ	INTJ	ENTJ	ENTJ
Personnel & labor workers	ISTJ	ISFJ	ISFJ	ESTJ
Pharmacists	ISTJ	ESFJ	ENTJ	ISTJ
Police & detectives	ESTJ	ISFJ	ISFJ	ISTJ
Police supervisors	ISTJ	ISTJ	ISTJ	ISTJ
Professional & tech workers	ESTJ	ISTJ	ISTJ	ESTJ
Psychologists	ENFP	ENFJ	ENFP	ENFP
Public relations workers	ENFP	ENFJ	ENFP	ENFJ
Public service & comm health	ISTJ	ESFJ	ISFJ	ESTJ
Radiologic tech.	ESFJ	ESFJ	ESFJ	ESFJ
Receptionists	ESFJ	ESFJ	ESFJ	ESFJ
Research assistants	ENFP	ISTJ	ESTJ	INFP
Research workers, unspec.	INTJ	ENFJ	INTJ	ENFJ
Resident housing assistants	ENFP	ESFJ	ENFP	ENFJ

<u>Job Title</u>	<u>Modal Type</u>	<u>REGTYPE</u>	<u>DISCTYPE</u>	<u>CmpstType</u>
Sales agents: retail trade	ESTJ	ESTJ	ESTJ	ESTJ
Sales clerks, retail trade	ESFJ	ESFJ	ESTJ	ESFJ
Sales reps, unspec.	ESTJ	ESTJ	ENTJ	ESTJ
Sales workers	ESTJ	ESFJ	ESTJ	ESFJ
Scientists: biological	ISTJ	INTJ	ISTJ	INTJ
Scientists: chemistry	ISTJ	ISTJ	ISTJ	INTJ
Scientists: life & physical	ISTJ	ISTJ	ISTJ	INTJ
Secretaries: exec & admin	ISTJ	ESFJ	ISTJ	ESFJ
Secretaries: legal	ISFJ	ESFJ	ISTJ	ISFJ
Secretaries: medical	ESFJ	INFJ	ESFJ	ESFJ
Secretaries: unspec.	ISFJ	ESFJ	ISFJ	ISFJ
Social scientists	ENFP	INTJ	ENFP	INFP
Social workers	ENFP	ESFJ	ISFJ	ENFJ
Speech pathologists	ENFP	ENFJ	ENFP	ENFJ
Stock clerks & storekeepers	ESTJ	ESFJ	ESTJ	ESTP
Teachers	ISTJ	ISFJ	ISFJ	ESFJ
Teachers' aides	ISFJ	ESFJ	ISFJ	ISFJ
Teachers: coaching	ESTJ	ESFJ	ISFJ	ESTJ
Teachers: elementary school	ISFJ	ISFJ	ISFJ	ESFJ
Teachers: health	ENFJ	ENFJ	ENFJ	ENFJ

<u>Job Title</u>	<u>Modal Type</u>	<u>REGTYPE</u>	<u>DISCTYPE</u>	<u>CmpstType</u>
Teachers: reading	ISFJ	ESFJ	INTJ	ESFJ
Teachers: trade & technical	ESTJ	ENTJ	ESTJ	ESTJ
Teachers: university	ISTJ	INFJ	ISTJ	INTJ
Teaching assistants	ISTJ	ISFJ	ISFJ	INFJ
Technicians: electrical	ISTJ	ISFJ	ISTJ	ISTJ
Technicians: misc.	ISTJ	ISTJ	ISTJ	ISTJ
Therapists: occupational	ENFP	ESFJ	ENFP	ENFJ
Therapists: physical	ISFJ	ISFJ	ISFJ	ISFJ
Therapists: respiratory	ISTJ	ISFJ	ISFJ	ISTJ
Transportation operatives	ISTJ	ESFJ	ESTJ	ISTJ
Typists	ISFJ	ESFJ	ISFJ	ESFJ
Waiters, waitresses	ENFP	ESFJ	ESTJ	ESFP
Writers & journalists	ENFP	ESFJ	ISFJ	ENFP

Table 15. Agreement rates between predicted and observed MBTI types.

<u>Agreement w/ modal type</u>	<u>Analytical Method</u>	
	<u>Regression on individual prefs</u>	<u>Disc. Analysis of type</u>
All four scales	29.10 %	65.45 %
At least three scales	72.73	81.81
At least two scales	93.64	90.91
At least one scale	98.18	99.10
No agreement	1.82	0.9

REAGAN D. BROWN

Current Address

750 Tall Oaks Dr. #11800D
Blacksburg, VA 24060
(540) 552-5587

Permanent Address

509 Ripple Creek
Houston, TX 77024
(713) 468-5612

EDUCATION

Master of Science Degree in Psychology, November 1995

Concentration: Industrial/Organizational Psychology

Virginia Polytechnic Institute and State University (Virginia Tech)

Thesis: An Examination of the Structure and Predictability of Myers-Briggs Type Indicator Preferences Using a Job Component Validity Strategy Based on the Common-Metric Questionnaire

Bachelor of Arts Degree in Psychology; Minor: Math, May 1993

The University of Texas at Austin

HONORS/AFFILIATIONS

Student Affiliate of the American Psychological Association

Student Affiliate of the Society for Industrial/Organizational Psychology

Dean's List, Spring 1992, University of Texas

President's Scholarship Award, University of Texas

Robert Dedmon Scholarship Recipient, 1990-1993

National Merit Scholarship Finalist, 1989

RESEARCH INTERESTS

- Selection: Validation and Fairness
- Personality Testing
- Psychopathy
- Job Analysis
- Validity Generalization
- Synthetic Validity

RELATED EXPERIENCE

Research

Thesis Research: Robert J. Harvey, Ph.D (chair), Virginia Tech

Designed and implemented a study to determine the feasibility of applying the job component validity technique to a personality inventory, the Myers-Briggs Type Indicator.

Other Research: Trained as a rater on the *Psychopathy Checklist*. Rated undergraduate students on a scale of Psychopathic Deviancy.

Teaching

Graduate Teaching Assistant, Virginia Tech, Blacksburg, Virginia

Laboratory for Psychometrics Laboratory - Expanded upon

material presented in the lecture portion of the course; instructed undergraduates on basic data entry, basic analytical procedures, and sophisticated analytical techniques on SAS system.

Laboratory Instructor for Introductory Psychology Course - Engaged students in discussion of psychology-related readings, assigned written homework, administered quizzes, and reviewed tests.

Reagan Brown