

BIODATA-250

***BIOGRAPHICAL INFORMATION ABOUT OCCUPATIONALLY-DESCRIPTIVE
ATTITUDES, TRAITS, AND ABILITIES***

TECHNICAL MANUAL

DECEMBER, 2011

HR DECISIONS LTD.

Scientific Systems for Selection & Development

©2011 HR DECISIONS LTD.

TABLE OF CONTENTS

CONTENTS	PAGE
LIST OF TABLES AND FIGURE	ii
CHAPTER 1 - BACKGROUND, DEVELOPMENT, AND CURRENT STATUS	1
General Goals for the Inventory	1
Item Development	1
Analysis of Item Responses and Factor Analysis of the Inventory.....	2
Present Normative Reference Groups and Administration Modes	3
CHAPTER 2 - BIODATA-250 SCALES AND SCALE DESCRIPTIONS.....	5
The Broad Empirical Factor Scales	5
The Narrow Empirical Factor Scales.....	6
The Big Five Scales.....	6
Conceptually-Based Scales	7
Special Purpose Scales.....	7
CHAPTER 3 - FOCUSED RESEARCH STUDIES WITH THE BIODATA-250 INVENTORY AND SOME VALIDITY EVIDENCE FOR THE BIODATA-250 INVENTORY.....	9
The Counterproductive Behavior Study	9
The Safety Orientation Study.....	13
The Employment-related Motivational Distortion (EMD) Study	15
CHAPTER 4 - STATISTICAL DETAILS AND RESULTS WITH THE BIODATA-250 INVENTORY	19
Some Normative Data	19
Reliability Estimates for the BIODATA-250 Scales and Their Interpretation.....	20
BIODATA-250 Scale Intercorrelations	23
Evidence of Construct Validity: Correlations between 43 BIODATA-250 Scales and the 20 Main Dimensions of the California Psychological Inventory (CPI)	26
CHAPTER 5 - THE USE OF THE BIODATA-250 INVENTORY IN THE CONSTRUCTION OF ORGANIZATION-SPECIFIC DIMENSIONS—GATHERING EVIDENCE FOR VALIDITY BASED ON CONTENT.....	30
Some Background	30
Application of the Content Validity Approach with the BIODATA-250 Inventory.....	31
Flexibility of Measurement in Applied Settings	32
REFERENCES AND FURTHER READING.....	34

LIST OF TABLES AND FIGURE

TABLES/FIGURE	PAGE
Tables	
Table 1: Details Concerning Normative Groups available with the BIODATA-250 Inventory	3
Table 2: Means and Standard Deviations on the BIODATA-250 Constructiveness Scale for Several Samples	11
Table 3: Means and Standard Deviations on the 35-Item BIODATA-250 Integrity Scale for Several Samples	12
Table 4: Means and Standard Deviations on the BIODATA-250 Safety Orientation Scale for Several Samples	14
Table 5: Correlations between (a) the full BIODATA-250 Safety Orientation Scale, and the BIODATA-250 Absentmindedness and Risk-Taking Sub-Scales and (b) Selected Additional Variables (Subjects: Undergraduate University Students)	15
Table 6: Means and Standard Deviations of the Various Comparison Groups on the 25-Item BIODATA-250 Employment-related Motivational Distortion (EMD) Scale.....	17
Table 7: Means and Standard Deviations on the BIODATA-250 Scales for Four Examinee Groups .	19
Table 8: Reliability Estimates for the BIODATA-250 Scales in Several Samples of University Undergraduate Students	21
Table 9: BIODATA-250 Scale Intercorrelations for University Undergraduate Students (Correlations for Male Participants below the Diagonal; Those for Female Participants above the Diagonal)	24
Table 10: BIODATA-250 Scale Intercorrelations for Job Applicants (Correlations for 3,053 Non-Management Job Candidates below the Diagonal; Those for 1,826 Management Candidates above the Diagonal)	25
Table 11: BIODATA-250 Scales and CPI-Form 434 Scale Cross-Correlations for University Undergraduate Students (Mixed-Gender; $n = 192$)	27
Table 12: Internal-Consistency and Stability Reliability Estimates for six Content-Based BIODATA-250 Scales Developed for One Organization	32
Figure	
Figure 1: Constructiveness Scale Items Graphed According to Their Correlations with the Full Constructiveness Scale and Their Average Correlation with Other Substantive Scales	12

BIOGRAPHICAL INFORMATION ABOUT OCCUPATIONALLY-DESCRIPTIVE ATTITUDES, TRAITS, AND ABILITIES (BIODATA-250)

CHAPTER 1

BACKGROUND, DEVELOPMENT, AND CURRENT STATUS

General Goals for the Inventory

Unlike the objectives underlying the construction of most self-report inventories, those underlying the BIODATA inventory included—in addition to the usual goals of establishing effective measurement of empirically-derived traits—the establishment of an item-level database that would permit construction of new occupationally-descriptive and -relevant scales in response to the needs of various organizations that would use the instrument. This database would permit construction of such new scales—seen as more relevant to various job dimensions established by organizations than would be the more established (in academic research) scales—and would permit the assessment of reliability of these new scales and item-analytic work to improve their effectiveness.

Item Development

Work began on the BIODATA inventory in 1998. The initial phase involved scouring the literature, of the preceding five or more decades, on prediction of job performance by attitudes, personality traits, and other factors, and to isolate the dimensions found useful for this purpose. This phase of the project took about six months and led to the identification of about 30 or more dimensions that at least one researcher (and usually more than one) had found useful in employment-testing applications. These set of dimensions comprised attitudes, personality traits, and abilities and, as such, suggested the inventory's name: *Biographical Information about Occupationally-Descriptive Attitudes, Traits, and Abilities*, or **BIODATA** for short.

Following this work, items were written to measure these dimensions. At first, 500 items were constructed. The principles guiding the construction of items were: (a) writing items that clearly reflected the isolated dimensions, (b) couching the items, whenever possible, in work-related terms, (c) ensuring coverage of basic and well-established psychological personality traits (like the Big Five personality dimensions), (d) ensuring that the items were understandable with less than high-school reading ability, and (e) ensuring that the items were unlikely to provoke any challenges on privacy and human-rights grounds.

With respect to principle (d), all 500 items were submitted to a language-education expert at the University of British Columbia for examination and try-outs with students for whom English was not their first language. On the basis of this examination, a number of items were dropped, replaced with new items, or revised so as to be more readily understandable. At the end of this phase of the development, we had 500 items that we were confident were easily understood even by examinees for whom English was a second language or whose reading level was approximately Grade 8 or higher. With respect to principle (e), the 500 items that survived the above process were submitted to a lawyer specializing in privacy and human-rights law. This lawyer identified items that had: (a) a high probability of being problematic from a human-

rights perspective, (b) some, but not high, probability of being problematic, and (c) very little or no probability of being problematic. Items in Category (a) were all eliminated. Those in Category (b) were examined, and many were also dropped. Those not eliminated were retained—often in somewhat revised form—to mark particularly important dimensions. The result of this phase of the research (completed in late 1999) was 467 items that we regarded as usable and effective for the intended purposes.

Analysis of Item Responses and Factor Analysis of the Inventory

The next phase of the research saw administration of these 467 items to approximately 1,000 university students. The goal of this phase of the research was to (a) examine each item for its ability to discriminate among subjects and display sufficient variability to function effectively in the inventory and (b) perform a large-scale factor analysis to see how the items loaded factorially. On the basis of the factor analysis, 31 dimensions were clearly identified for the BIODATA items. Item analysis of the items marking each of the 31 items allowed a rank-ordering of these items in terms of their effectiveness in measuring that dimension.

A decision was then made to reduce the length of the battery to the best 250 items. This was accomplished by examining each of the 31 dimensions and retaining that subset of items for each that showed the highest reliability for measuring that dimension. This resulted in the 31 dimensions still being effectively measured, but now by the most effective 250 of the original 467 items. The length of 250 items was decided upon because this number (a) allowed a personality inventory that could be completed in well under one hour (approximately 40 – 45 minutes on average), (b) reliably captured the key dimensions measured by the inventory, and (c) was similar to the number of items on the best available personality inventories, the NEO PI-R (240 items), the revised California Psychological Inventory (260 items), and the PDI Global Personality Inventory (250 items). Personality inventories consisting of smaller numbers of items (many with 50 – 80) provide far too sparse sampling of item content to adequately measure important, predictive dimensions. Inventories consisting of more items have been found occasionally to present some logistical test-administration problems. Thus, this length seemed like the very best trade-off between information-yield and logistical constraints. The revised inventory was labeled **BIODATA-250**.

The next phase of the research—taking place in 2003-2005—called for further administration of the 250-item inventory to university undergraduates (students in their first or second year of their university programs). This allowed large-scale assessment of the reliability of the now-shortened BIODATA scales, decisions about the retention of some of the more-sparsely-defined factors, the addition of some logical scales suggested by the literature and by actual practical use of the inventory in applied personnel-selection contexts, and some final reliability analyses of a now-expanded list of 43 BIODATA-250 scales.

On the basis of large samples of both male and female university students, scale reliabilities were assessed by both internal-consistency and test-retest methods to assess both the relative homogeneity of the scales and their stability over time. These analyses were carried out separately by gender. Results showed that—partly because of the 4-point response grid for the items (as opposed to a true-false format used with many personality inventories)—scale reliabilities were high, in general higher than normally seen with self-report instruments. Since

the scale lengths varied widely (from a low of 5 items to a high of 104, with a mean of about 22 items), reliabilities also varied, but the mean scale reliability over 43 BIODATA scales was .751 for internal-consistency (coefficient α) and .845 for test-retest (stability) reliability. This means that aggregations of these scales can be expected to have reliabilities in the high .80 or .90 range. BIODATA-250 scale reliability estimates are presented and discussed in Chapter 4.

The original development design for the BIODATA inventory called for production of a database of item content that had relevance for employment applications. Thus, it was envisaged that not only would empirically-derived dimensions (like those established through factor analysis) be of use, but the item content would be sufficiently broad and relevant that special-purpose dimensions could be constructed around the competency dimensions identified by client organizations. This, it was believed, could be accomplished at the item level through aggregation of relevant items for the defined (by the organization) dimensions. As long as there was workplace normative information for the 250 items in the inventory, workplace-normed scales that corresponded to virtually any organization’s key dimensions could be derived.

Further, on the basis of the large university-student sample and the now-large samples from industry—both samples consisting of *item* data—any newly-developed scale can be assessed for its reliability. Internal-consistency reliability estimates can be obtained from both of these samples for any new aggregation of BIODATA items, and test-retest reliability estimates can be obtained from the student sample. This enables HR Decisions Ltd. to evaluate any new scales that have been constructed to meet organizational goals for their psychometric properties and, if necessary, revise and improve these scales until they reach acceptable levels of reliability before being used.

Present Normative Reference Groups and Administration Modes

At present, examinees’ raw scores can be referenced to three norm groups. Details about these groups follow:

Table 1

Details Concerning Normative Groups available with the BIODATA-250 Inventory

Norm Group	Size of Norm Group	Number of Organizations with Examinees in Norm Group
1. University Undergraduate Students ^a		
(a) Male University Undergraduate Students ^a	450	One Canadian university
(b) Female University Undergraduate Students ^a	1,250	One Canadian university
2. Job Applicants for Non-Management Positions Both Genders Combined ^b	3,053	3 Canadian organizations
3. Job Applicants for Management Positions Both Genders Combined ^b	1,826	5 Canadian organizations

^aAlmost entirely first- and second-year undergraduates. For two scales, Constructiveness and Safety Orientation, the norm-group *ns* are 171 male and 551 female subjects.

^bThese norm groups have the genders combined because job postings attract applicants from both genders, and these applicants are evaluated together as a group and not separately by gender.

The above norm groups represent the most recent (2011) updating of the norms.

In 2005-06, the inventory—which had existed in only pencil-and-paper format until that time—was put online, with the appropriate internet conventions. Large-sample comparisons of scores obtained via the pencil-and-paper version and the online version have revealed no significant differences in score levels. This means that this instrument can now be completed in either pencil-and-paper or online format, with scale scores strictly comparable across formats.

The plan going forward with the BIODATA inventory is to allow client organizations to present their corporate competencies or dimensions to HR Decisions Ltd., and for us to construct scales from the BIODATA item content that will measure these dimensions reliably and validly. In addition to this highly-tailored, organization-specific, customized self-report assessment, we can provide—from the 43 scales of the BIODATA inventory—other job-relevant information of interest in the employment-hiring and employee-development contexts.

At present, the BIODATA inventory is being widely used for employment assessment in connection with real-life personnel decisions (not research) in industry.

CHAPTER 2

BIODATA-250 SCALES AND SCALE DESCRIPTIONS

The current list of 43 BIODATA-250 scales is presented in this chapter. As noted, however, the BIODATA-250 item data base permits any number of client-oriented, special-purpose scales to be constructed. An example of this latter application is presented later in Chapter 5.

The Broad Empirical Factor Scales

Dominance: An eagerness to assume, and comfort with, leadership roles and decision-making; comfort with organizing, speaking to, and directing groups.

Orderliness: Having a methodical, structured approach to tasks and activities, along with a tendency to work in an organized and efficient fashion and to accomplish more than most others.

Sensitivity to Others: A tendency to listen to others, rather than to dominate conversations, and to a source of advice, nurturance, and understanding to others.

Optimism: An absence of worry and self-examination; a tendency to be happy, relaxed, easygoing, and generally in good spirits, and to be confident in one's decisions and future.

Energy: Drive and a preference for activity over rest and inactivity; a desire for a full slate of activities in one's life, and a busy, as opposed to relaxed life.

Assertiveness: A tendency to stand up for and effectively defend one's position on issues; being persuasive and credible when faced with opposing views.

Independence: A preference for working on one's own, rather than with others in groups; a preference to depend on oneself rather than on others.

Social Confidence: Ease in meeting and interacting with people; a natural tendency to express one's views and comfort in being the center of attention.

Perseverance: A tendency to stay with tasks and projects until they are completed; a tendency to work harder when adversity occurs, rather than giving up.

Service Orientation: A preference for work and activities that benefit others; a tendency to help and provide for others, particularly those in need.

Social Insight: Perceptiveness with respect to the feelings and motives of others, and the unspoken messages they send; an ability to "read" others.

Straightforwardness: Preference for completely honest, ingenuous relationships with others; a tendency to avoid any kind of deception, coercion, or manipulation of others.

Self-Esteem: A feeling of self-worth and a tendency to trust one's own opinions, wishes, and judgments when they are met with anger and opposition from others.

Resilience: A tendency to quickly and effectively recover emotionally from (or be unaffected by) assaults—such as criticism, insults, and anger—to one's ego from others.

The Narrow Empirical Factor Scales

Conformity: A tendency to view organizational and societal rules and values as necessary, and a desire to follow these rules and values.

Risk-Taking: A preference for high-risk activities and a high-risk approach to problems and work; a tendency to reject the safe course of action for a riskier one.

Attention to Detail: A preoccupation with the specific details and fine points of projects and ideas, rather than with the larger picture when the latter means overlooking these particulars.

Adaptability: A flexibility of behavior and thinking that enables one to adapt to changing circumstances without distress; a freedom from rigid thinking and behavior.

Impulse Control: A tendency to avoid spontaneous, impulsive, and possibly offensive, behaviors—particularly spoken utterances—when these could have problematic results.

Competitiveness: A preference for competing, rather than cooperating, with other people; a tendency to seek out and value competition and winning.

The Big Five Scales

Emotional Stability (Neuroticism Reflected): Freedom from neurotic tendencies and of fearfulness and anxiety; calmness, stability, and resiliency; self-confidence and a sense of well-being.

Extraversion: A tendency to be sociable, warm, and positive; a preference for excitement and activity and a tendency towards impulsiveness and expressiveness.

Openness: A preference for new and varied experiences and for the unconventional; tolerance for change, risk, and ambiguity; comfort with complexity and uncertainty.

Agreeableness: A tendency to trust others and to be sensitive to their feelings and well-being; a tendency to be open, trustworthy, modest and friendly with others.

Conscientiousness: A tendency to value hard work and achievement and to work towards these goals; perseverance, orderliness, self-discipline, and reliability.

Conceptually-Based Scales

Internal Locus of Control: A tendency to attribute success to one's own efforts—those attributed that are under a person's control—rather than to external, uncontrollable events.

Emotional Intelligence: A tendency to understand one's own emotions and motives and those of others, to express emotions appropriately, and to manage emotions in oneself and others.

Nurturance: A tendency to provide sympathy and comfort to others, to be empathic and sensitive, to provide help when needed, and to be interested in helping those less fortunate.

Self-Sufficiency: A preference for self-reliance and working independently and for standing up for one's beliefs and persuading others of their merits; confidence in and acceptance of one's self.

Results Orientation: An orientation towards the "bottom line"; the view that what is accomplished is what is truly important, whereas how it is accomplished is irrelevant.

Special-Purpose Scales

Leadership Potential: A tendency to assume leadership roles and to organize the activities of others; assertiveness and persuasiveness; interpersonal effectiveness; tolerance for stress.

Telemarketing Potential: Empirically-derived scale from a validity study with telemarketers in a large organization.

Customer Service Potential: Empirically-derived scale from a validity study with customer-service personnel in a large organization.

Constructiveness (Counterproductive Tendencies Reflected): A tendency to follow the rules and conform to accepted standards; honesty and integrity; dutifulness; self-control; avoidance of bad behavior and low standards. This scale resulted from an empirical study described in Chapter 3 of this document.

Safety Orientation: A tendency to engage in cautious, planful, and mindful behavior, to experience few accidents, and to be reliable and internally oriented. This scale resulted from an empirical study described in Chapter 3.

Social Ascendancy: A tendency towards being dominant, assertive, socially confident and self-assured, as well as having leadership potential and poise.

Note: This scale, along with the five following, is a large, broad-spectrum scale comprising several of the themes found in narrower BIODATA scales, and intended (and empirically validated) mainly for management-selection applications.

Disciplined Achievement: A tendency towards being conscientious, responsible, and having high achievement motivation and impulse control. *See note after Social Ascendancy.*

Calm Steadiness: A tendency towards being emotionally stable, resilient, self-sufficient, and having an internal locus of control. *See note after Social Ascendancy.*

Warmth and Concern for Others: A tendency towards being sensitive to others, patient, nurturing, emotionally intelligent, and having good social insight. *See note after Social Ascendancy.*

Openness to Change: A tendency towards being open, independent, adaptable, non-conforming, and willing to take risks. *See note after Social Ascendancy.*

Extraversion and Low Impulse Control: A tendency towards being gregarious, expressive, and have a lack of inhibition. *See note after Social Ascendancy.*

Integrity: A tendency towards honesty, trustworthiness, and reliability, developed as a narrower version of the much-longer Constructiveness scale described earlier. This scale resulted from empirical research described in Chapter 3.

Validity Scale – Employment-Related Motivational Distortion: A tendency towards distorting one's responses on a self-report inventory in such a way as to present an overly-favorable impression in order to improve one's chances of obtaining a desired job. This scale resulted from an empirical study described in Chapter 3 of this document. The meaning of scores on this scale should be clear from a reading of this material.

Chapter 3

FOCUSED RESEARCH STUDIES WITH AND SOME VALIDITY EVIDENCE FOR THE BIODATA-250 INVENTORY

Three large-scale, multi-year, focused research studies have been carried out with the BIODATA-250 inventory. These studies were carried out to determine the inventory's ability to measure certain behaviors effectively. The data that were collected in the course of these studies enabled us to enlarge the item database a little, but the purposes of the studies were focused on specific themes. All of these studies were presented at a total of six international conferences, four sponsored by the Canadian Psychological Association, and two by the Society of Industrial & Organizational Psychology. Each of these studies provided evidence of validity for special-purpose BIODATA scales. The three large-scale studies were focused on:

1. *Counterproductive Behavior*, with results presented at the Canadian Psychological Association Annual Meeting, 2003;
2. *Safety Orientation*, with results presented at the annual meetings of the Canadian Psychological Association in 2004 and 2006, and the Society of Industrial & Organizational Psychology in 2007;
3. *Employment-related Motivational Distortion*, with results presented at the annual meetings of the Canadian Psychological Association in 2008 and the Society of Industrial & Organizational Psychology in 2008.

The results of each of these studies provided us with three corresponding special-purpose BIODATA scales: (a) an 81-item counterproductivity scale, which we have reflected and renamed *Constructiveness* (including a number of organizational citizenship behaviors), (b) a 51-item *Safety Orientation* scale, and (c) a 25-item validity scale that we have labeled *Employment-Related Motivational Distortion*. Each of these special-purpose scales possesses adequate reliability for its purpose as well as empirically-established validity.

The Counterproductive Behavior Study

Background and Scale Development

The assessment of counterproductive tendencies has, over the past several decades, gone by the name of “integrity testing.” Both overt and personality-based measures have been developed for use in industrial applications. The goal with the BIODATA-250 inventory was to identify a subset of inventory items that would function like a personality-based integrity test. We saw an advantage in having this scale embedded within the much larger collection of themes represented by the items of the BIODATA-250 inventory, in that we could expect that the response set inculcated when taking a special-purpose integrity test would be largely absent.

In this study, which was begun in 2001 and spanned several years, university undergraduate students were administered the BIODATA-250 inventory along with a self-report questionnaire eliciting information about counterproductive behaviors in which the students had been involved earlier in their lives. Data on the latter were obtained via a

40-item Counterproductive Behaviors questionnaire, which covered such behaviors as: cheating, plagiarism, drug use, reckless recreational activities, laziness, tardiness and absence from work, stealing, drunkenness, general dishonesty, inappropriate aggressiveness, and reckless personal habits.

All items of the BIODATA-250 inventory were correlated with total scores on the Counterproductive Behaviors criterion instrument in a sample (collected over several years) of 722 undergraduate students at the University of British Columbia (171 male and 551 female students). The experimental design was such that the Counterproductive Behaviors criterion questionnaire was administered first in the testing session. This was followed by a filler period in which subjects were administered two speeded cognitive-ability tests in order to break up, as much as possible, any response sets that might have developed during the first part of the session. Since the BIODATA-250 inventory is a very broad-spectrum instrument (focusing on much more than just counterproductive behaviors), it was assumed that any counterproductive-behavior response sets that might have been formed would dissipate.

Those items for which the bivariate item-criterion correlation was statistically significant at the .005 level in both gender samples were selected for inclusion in the scale. This stringent significance level was used to reduce overall familywise Type 1 error in the study. With this criterion, we would expect just over one item to correlate at this level (and thus be included in the scale) by chance. Since 81 items met this statistical criterion, having one chosen erroneously was not considered important.

The final 81-item scale was found to correlate **.70** with the Counterproductive Behaviors criterion. This value can be viewed as a validity result, although concurrency of administration and method variance can be expected to have contributed considerably to this high value. We might expect this value to drop to the .30 to .40 range once the effects of capitalization on chance and the factors noted above had been adjusted for. This level would be more consistent with those found by Ones, Viswesvaran and Schmidt (1993) in their meta-analysis of integrity test validities, and would reflect what we would expect to find if we were to use the scale to predict on-the-job counterproductive behaviors.

To have a scale oriented in the same direction as most of the other BIODATA-250 scales, we have reflected the above-described 81-item scale to represent a positive trait, *constructiveness*, a term we see as denoting the opposite of counterproductive tendencies. Understanding the scale this way will allow us to represent it as the BIODATA-250 personality-based integrity test, with the validity potential demonstrated for other measures of this construct. We, therefore, refer to this scale as the BIODATA-250 **Constructiveness Scale**.

The 81-item scale that arose from this research was found to be highly reliable. The reliability estimates calculated for it appear later in this document in Chapter 4.

Some Results of Interest with the BIODATA-250 Constructiveness Scale

In the following table, mean scores on this scale, along with standard deviations, for

several groups are presented.

Table 2

*Means and Standard Deviations on the BIODATA-250 **Constructiveness** Scale for Several Samples*

Sample	Mean	Standard Deviation
University Undergraduate Students		
Male Students ($n = 171$)	198.11	22.40
Female Students ($n = 551$)	205.17	19.55
Genders Pooled (with equal numbers of each gender; $n = 342$)	201.64	21.34
Non-Management Job Applicants (both genders; $n = 3,053$)	231.07	16.34
Management Job Applicants (both genders; $n = 1,826$)	222.40	15.98

With respect to the means in Table 2 above, the gender difference with the university undergraduate students is statistically significant ($p < .0001$). This is not surprising; earlier studies of counterproductive behavior have found a gender difference, with male subjects exhibiting higher levels of this tendency. Also, not surprisingly, the means on this Constructiveness scale for the job-applicant samples are much higher than those for the students. These latter differences are on the order of standardized effect sizes of 1.0 to 2.0. The difference between management and non-management job applicants is highly significant ($p < .00001$), but is smaller in effect size than that just noted, although still substantial, with a standardized effect size of about .50. This difference is more difficult to explain than that between job applicants and students, and we will not attempt an explanation here.

*Development of a Shorter **Integrity Scale***

Rationale for this scale

The large 81-item Constructiveness scale detailed above was designed for use where this construct is of primary interest to the assessor. Its large number of items, however, means that it will correlate more highly than is desirable with other substantive scales of interest, since it will share items with many. For this reason, if assessment on many scales is needed in application of the BIODATA-250 inventory, a scale that captures the main themes of Constructiveness, but does so with far fewer items would be of benefit. This reasoning led to the development of a shorter scale capturing these key themes.

Scale development

The 81 items on the Constructiveness scale were examined in terms of (a) their correlation with the full Constructiveness scale and (b) their average correlation with a

number of widely-used BIODATA-250 scales. The 81 items were arrayed in a scatterplot-like figure as shown below in Figure 1:

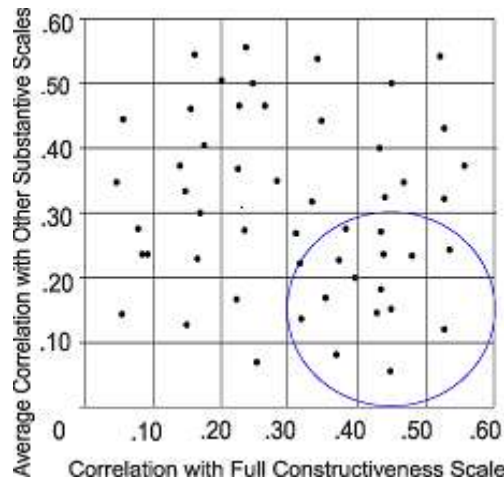


Figure 1. Constructiveness Scale Items Graphed According to Their Correlations with the Full Constructiveness Scale and Their Average Correlation with Other Substantive Scales.

Those items circled in Figure 1 are those we sought for the shorter Integrity scale; they are the items that tended to best measure the larger construct, but without overly-large correlations with the other important BIODATA-250 scales. One further consideration was that the main themes captured by the Constructiveness scale were also represented in the shorter Integrity scale. The resulting scale has 35 items.

Some results of interest with the 35-item BIODATA-250 Integrity Scale

Means and standard deviations on the BIODATA-250 Integrity scale, for several groups, appear below in Table 3.

Table 3

Means and Standard Deviations on the 35-Item BIODATA-250 Integrity Scale for Several Samples

Sample	Mean	Standard Deviation
University Undergraduate Students		
Male Students (<i>n</i> = 171)	87.64	11.13
Female Students (<i>n</i> = 551)	91.40	9.20
Genders Pooled (with equal numbers of each gender; <i>n</i> = 342)	89.72	10.25
Non-Management Job Applicants (both genders; <i>n</i> = 3,053)	107.25	9.44
Management Job Applicants (both genders; <i>n</i> = 1,826)	102.78	9.61

The Safety Orientation Study

Background and Scale Development

In this study, begun in 2002 and completed in 2005, 617 undergraduate students at the University of British Columbia (128 male and 489 female students) were administered the BIODATA-250 inventory, along with a 16-item Hazardous Behavior Scale (with a demonstrated internal-consistency reliability estimate of .77)—a questionnaire constructed to assess subjects' tendencies to engage in unsafe behaviors during the previous five years. Of these subjects, 227 responded to both the BIODATA-250 inventory and the Hazardous Behavior Scale during a single 1.5-hour session. The same filler procedure described above in connection with the Constructiveness Scale (with the interspersed speeded cognitive-ability tests) was employed with these study subjects. In addition, 390 subjects responded to the BIODATA-250 inventory and the Hazardous Behavior Scale in two separate sessions, administered two weeks apart. These latter subjects also completed a cognitive measure of hazardous behaviors, the Cognitive Failures Questionnaire (Broadbent, 1982) and several cognitive-ability measures. An additional sample of 140 undergraduate students were administered the BIODATA-250 inventory, the Hazardous Behavior Scale and series of measures of memory and attention.

In the scale-development phase of the study, BIODATA-250 items were individually correlated with scores on the Hazardous Behavior Scale, and those yielding a correlation coefficient that was significant at the .001 level were singled out for inclusion in the new BIODATA-250 scale. As was true in the development of the Constructiveness scale, this stringent significance level was used to reduce overall familywise Type 1 error in the study. With this criterion, we would expect far fewer than even one item to correlate at this level (and thus be included in the scale) by chance.

Next, double cross-validation of the proposed scale was performed. In this analysis, the sample of 617 subjects was randomly divided in half and items were chosen in each half using an alpha level of .005 (given the smaller, half-size samples). As a result of this examination of cross validity, 51 BIODATA-250 items demonstrated a significant correlation with the Hazardous Behavior Scale criterion in both subsamples. These items were retained for the BIODATA-250 *Safety Orientation Scale*. The computed cross-validity estimate for this scale and the Hazardous Behavior Scale criterion was **-.48**.

This value can thus be seen as a criterion-related validity coefficient for the scale, although it must be remembered that, because of concurrency of administration and method variance, this value would have to be seen as an overestimate of what we might expect for the scale in the prediction of job-related unsafe behaviors.

A factor analysis of the 51 items of the Safety Orientation scale, involving a properly-pooled (by gender) sample of 254 subjects (127 of each gender to balance the sample), revealed six sub-factors of this scale: (a) risk-taking, (b) absentmindedness, (c) assertiveness, (d) gregariousness, (e) planfulness/orderliness, and (f) counter-productivity, with all but (e) planfulness/orderliness reverse scored in the overall scale. Of these sub-factors, perhaps

only gregariousness is mildly surprising. Its negative relationship with safety orientation may suggest a lower attention to detail in those high on gregariousness.

The 51-item Safety Orientation scale that arose from this research was found to be highly reliable. The reliability estimates calculated for this scale appear in Chapter 4 of this document in a section devoted to BIODATA-250 scale reliability.

Some Results of Interest with the BIODATA-250 Safety Orientation Scale

Means and standard deviations on the 51-item scale for several groups are presented below in Table 4.

Table 4

Means and Standard Deviations on the BIODATA-250 Safety Orientation Scale for Several Samples

Sample	Mean	Standard Deviation
University Undergraduate Students		
Male Students ($n = 171$)	122.92	16.58
Female Students ($n = 551$)	128.43	14.14
Genders Pooled (with equal numbers of each gender; $n = 342$)	125.95	15.85
Non-Management Job Applicants (both genders; $n = 3,053$)	146.23	12.46
Management Job Applicants (both genders; $n = 1,826$)	139.57	12.20

With respect to the means in Table 4 above, the gender difference with the university undergraduate students is statistically significant ($p < .0001$). Given our results with the Constructiveness scale, along with other findings over the years, this is an expected result. As was true with the Constructiveness scale, the means for the two job-applicant samples on this Safety Orientation scale are considerably higher (with standardized effect sizes of 1.0 to 1.5, indicating very large differences) than for the student samples. Again, this is not surprising. The difference between management and non-management job applicants is also highly significant ($p < .00001$), but smaller in effect size than that just noted.

In the research conducted with the Safety Orientation scale, it, along with its sub-scales (identified earlier), was correlated with many other variables. One set of variables was constructed from items on the BIODATA-250 inventory. In another analysis, we developed a cognitive measure that had relevance for safe and unsafe behaviors. A total of 13 cognitive measures were subjected to a factor analysis. These measures tapped cognitive flexibility, inhibition of dominant response, selective and focused attention, alternating attention, visual search, sustained attention, shifts between activation and inhibition of quick responses, and cognitive impulsivity. Two large factors were obtained in this analysis, labeled: (a) Cognitive Errors and (b) Performance Speed. The first of these, Cognitive Errors, was of particular interest in our examination of safety orientation.

In Table 5 below, the correlations between (a) the overall Safety Orientation score and the Absentmindedness and Risk-Taking sub-scale scores and (b) these additional variables are presented.

Table 5

*Correlations between (a) the full BIODATA-250 **Safety Orientation** Scale, and Absentmindedness and Risk-Taking Sub-Scales and (b) Selected Additional Variables (Subjects: Undergraduate University Students)*

Other Variable	BIODATA-250 Scale		
	Full Safety-Orientation Scale	Absentmindedness Sub-Scale	Risk-Taking Sub-Scale
Scales Arising from the BIODATA-250 Inventory (<i>n</i> = 633, separate-gender data properly pooled) ^a			
Big Five Scales			
Emotional Stability	-.04	-.36*	.15*
Extraversion	-.54*	-.05	.45*
Openness	-.47*	.04	.51*
Agreeableness	.04	-.14*	.04
Conscientiousness	.45*	-.40*	-.29*
Impulse Control	.47*	-.36*	-.36*
Internal Locus of Control	.05	-.21*	.00
Cognitive Errors (<i>n</i> = 65, mixed-gender) ^b	-.30*	.38*	.12

**p* < .005, directional.

^aThese correlations reflect some common items, along with method variance, and thus are somewhat inflated.

^bThese correlations are between wholly-independent measures (one self-report, the other maximum-performance) and thus are free from any contamination due to common items or method variance.

The results in Table 5 suggest that the main Big Five correlates of Safety Orientation are Extraversion and Openness (both negatively correlated) and Conscientiousness (positively correlated). Emotional Stability does not appear, from the present results, to be related to the tendency towards safe behaviors, although it is significantly related to the Absentmindedness sub-scale. The correlations between the BIODATA-250 scales and the cognitive errors measure is interesting and encouraging in that it indicates that the self-report-based measure of safety orientation does have a cognitive-domain parallel. It is, in general, unusual for variables coming from these two different individual-differences domains (cognitive, affective) to correlate to any appreciable degree.

The Employment-related Motivational Distortion (EMD) Study

Background and Scale Development

In this study, begun in 2005 and completed in 2008, 633 undergraduate students at the University of British Columbia (129 male and 504 female students) were administered the

BIODATA-250 inventory with normal instructions—to respond to the items honestly. The median age of the respondents was 20, approximately 56% spoke English as a second language. This was the same sample as used by Hakstian and Ng (2005) in their study of employment-related motivational distortion. A second large sample of 1,003 undergraduate students (275 male and 728 female subjects) was administered the BIODATA-250 inventory with specific instructions to respond to each item as though they were being tested for a job that they wanted. This sample, having median age of 20, contained about 58% for whom English was a second language, and will be referred to as the "Cued-MD Response" group. The instructions read:

Imagine that you are applying for a job that you would like to have. This could be a career that you are pursuing or a job that you are currently applying for. You are asked to complete the inventory as part of the selection process. The responses you will provide will be valuable in making the hiring decision. Please respond to the items as you would if you were applying for this job.

The students were reminded, midway through their taking of the inventory, that they were applying for a very desirable job. After completion of the inventory, students were asked to disclose the job they were thinking of, to ensure that the resultant EMD (Employment-related Motivational Distortion) scale would not be limited to a narrow career path. Two to three weeks after the first administration, a smaller sample of 166 Cued-MD response subjects were re-administered the inventory under the same (Cued-MD) instructions in order to permit estimation of test-retest reliability of the scale developed. We felt it important to study the stability of this tendency (towards employment-related motivational distortion) in subjects in which the tendency was operating on both occasions, as this would give more accurate real-world information about the stability of this tendency.

Of the 250 BIODATA-250 items, 70 were found to manifest a significant difference between the normal control subjects and the Cued-MD subjects (with a p -value of at least .0001 to ensure a familywise Type 1 error rate of less than .05). Item-total correlations were then obtained, which permitted a rank-ordering of these items with respect to their contribution to scale reliability. In addition, we had data on a job-applicant sample of 1,494 subjects and calculated the correlations between each of the 70 items and a composite overall potential score for these applicants. The goal was to identify an optimal subset of the 70 items that had high reliability, but as low as possible a correlation with the overall job-potential score. The latter condition was to reduce as much as possible that portion of measured motivational distortion that might truly reflect job effectiveness. Scales of 20, 21, 22, 23, 24, and 25 items were tested against these two selection criteria (a shorter scale was also seen as desirable in not overlapping too much with the substantive scales of the BIODATA inventory). The final, optimal, scale—one balancing reliability on the one hand and lower overlap with the important substantive scales on the other—had 25 items. The calculated reliability estimates for this scale appear later in this document in a more general discussion of BIODATA-250 scale reliability.

Some Evidence of Scale Effectiveness

To determine the validity of this special scale, we calculated the differences between the scale means of the normal control group and of the cued-MD group. In addition, we

examined the difference between the scale means of the normal control group and of the job-applicant group. These scale means appear below.

Table 6

Means and Standard Deviations of the Various Comparison Groups on the 25-item BIODATA-250 Employment-related Motivational Distortion (EMD) Scale

Comparison Group	<i>n</i>	Mean	Standard Deviation
1. Normal-Response Students in Study (Honest Responding)			
Male Students	129	62.53	7.93
Female Students	504	63.77	6.44
Total Sample	633	63.52	6.78
2. Cued-MD Response Students in Study			
Male Students	275	76.45	9.67
Female Students	728	78.59	8.96
Total Sample	1,003	78.01	9.20
3. Job Applicants			
Male Job Applicants	978	75.09	6.94
Female Job Applicants	516	75.00	7.60
Total Sample	1,494	75.05	7.38

In the tabled results, the Cued-MD student mean (total sample) is significantly different from the normal-response student mean ($p < .00001$), this difference representing a standardized effect size of 2.14, an enormous difference. This corresponds to a point-biserial correlation (with group coded Normal = 0; Cued-MD = 1) of **.65**, the latter value representing a sort of validity coefficient. There were small (but statistically significant) gender differences in both student samples, but these are of little concern to us here.

In the job applicant sample, no gender differences appeared, and, again the means for that group were vastly different from those of the normal-response student group. Interestingly, the total-sample mean for the job-applicant sample was significantly lower than that for the Cued-MD student sample (although 1.70 standard deviations, a huge difference, above that of the normal-response students). Thus, it appears that the extent of motivational distortion can be expected to be higher in subjects all of whom have been instructed to respond this way than with regular job applicants, most, *but not all* of whom, will respond with this tendency. These results seem completely in line with what we might expect with this test-taking response set.

Potential Uses of the EMD Scale

Validity scales like the BIODATA-250 EMD scale, have a long history of use, with probably the best-known examples being those associated with the Minnesota Multiphasic Personality Inventory (MMPI). With that instrument, the tendencies of clinical patients either to “fake good” (analogous to our conceptualization of motivational distortion) to avoid treatment or to “fake bad” to prolong treatment are examined, with an eye to making sense of the substantive scale results.

With self-report assessment in applied industrial settings, we can either (a) examine the EMD score with an eye to warning recipients of the examinee’s report if there is evidence of excessive levels of this tendency (providing a general assessment of the trustworthiness of the reported scores) or (b) make numerical adjustments to the most relevant scale or competency scores as a function of the magnitude of the EMD score (as can be done with the MMPI). Assuming a near-zero correlation between these scores and actual work-performance assessments, EMD would operate as a suppressor variable in any optimal linear combination of predictive measures, and reasonably-precise numerical adjustments *could* be made using EMD scores.

At present, HR Decisions Ltd. uses the first approach noted above. As more criterion-related data are gathered, it may be possible to incorporate EMD score adjustments into the predictive scores reported to organizations.

Chapter 4
STATISTICAL DETAILS AND RESULTS WITH THE BIODATA-250 INVENTORY

Some Normative Data

Normative information for several examinee groups is presented below in Table 7. These values are, for the most part, based on large samples and can be expected to be quite stable. This normative information is what is available as of late 2011. As more examinees are administered the BIODATA-250 inventory in the future, the values in Table 7 (and, similarly, in our scoring programs) will be updated. In addition, the plan going forward is to develop some industry- and job-specific norms as the overall norm base size increases.

Table 7

Means and Standard Deviations on the BIODATA-250 Scales for Four Examinee Groups

BIODATA-250 Scale	No. of Items	University Undergraduate Students				Job-Applicants (Pooled Genders)			
		Male		Female		Non-Mgmt. Jobs ^c		Mgmt. Jobs ^d	
		Students ^a	Students ^b	Non-Mgmt. Jobs ^c	Mgmt. Jobs ^d	M	SD	M	SD
<i>Broad Empirical Factor Scales</i>									
Dominance	15	38.54	6.90	37.02	7.90	42.61	5.67	44.75	6.05
Orderliness	9	23.13	3.82	23.05	3.69	27.51	3.22	27.42	3.24
Sensitivity to Others	9	27.02	3.07	27.61	3.00	30.19	2.65	29.36	2.62
Optimism	8	21.29	3.54	20.33	3.54	24.85	2.71	24.64	2.88
Energy	10	20.04	3.51	20.17	3.18	24.12	3.09	24.00	3.46
Assertiveness	9	24.68	3.26	23.29	3.33	24.93	2.56	25.80	2.68
Independence	9	23.82	3.03	23.54	3.22	20.74	2.56	20.69	2.72
Social Confidence	16	40.30	7.41	39.59	7.39	48.10	6.19	48.04	6.91
Perseverance	14	38.29	4.66	38.26	4.69	41.78	3.94	41.90	3.84
Service Orientation	11	29.35	4.81	31.17	4.57	35.41	4.02	32.49	4.46
Social Insight	8	22.88	2.77	23.12	2.73	25.16	2.77	24.42	2.71
Straightforwardness	11	26.93	4.33	28.07	4.02	34.66	3.94	32.77	4.06
Self-Esteem	6	14.97	2.26	14.78	2.47	17.56	1.99	17.42	2.14
Resilience	9	19.77	4.01	17.78	3.75	24.05	3.28	22.74	3.45
<i>Narrow Empirical Factor Scales</i>									
Conformity	5	12.58	2.06	13.31	1.83	15.38	1.97	14.97	1.92
Risk-Taking	7	17.61	3.28	16.26	3.38	15.45	2.86	16.85	2.98
Attention to Detail	6	15.01	2.26	15.91	2.40	17.03	2.04	15.81	2.42
Adaptability	6	15.64	2.13	15.19	2.25	18.16	2.03	17.87	2.12
Impulse Control	6	14.49	2.44	15.33	2.39	17.22	2.07	16.39	2.07
Competitiveness	7	18.21	2.88	17.09	2.84	15.18	2.35	17.02	2.36
<i>The Big Five Scales</i>									
Emotional Stability (Refl. Neuroticism)	20	50.77	6.10	48.02	6.30	61.41	5.82	60.23	6.09
Extraversion	15	41.17	5.97	40.36	6.15	40.18	4.39	40.48	5.09
Openness	15	40.01	4.61	38.78	4.50	42.73	3.65	44.57	4.08
Agreeableness	14	37.84	4.34	39.02	4.11	45.63	3.86	43.14	3.97
Conscientiousness	39	96.33	10.21	97.56	9.67	112.51	8.95	112.22	9.06

Table Continues on Next Page

Table 7 (Continued)

BIODATA-250 Scale	No. of Items	University Undergraduate Students				Job-Applicants (Pooled Genders)			
		Male Students ^a		Female Students ^b		Non-Mgmt. Jobs ^c		Mgmt. Jobs ^d	
		M	SD	M	SD	M	SD	M	SD
Conceptually-Based Scales									
Internal Locus of Control	6	16.21	2.38	15.93	2.58	18.27	2.24	18.35	2.35
Emotional Intelligence	18	47.58	5.80	48.01	6.00	54.73	5.16	53.52	5.63
Nurturance	20	56.37	6.40	58.78	6.35	65.59	5.80	61.85	6.06
Self-Sufficiency	13	35.43	3.97	33.80	4.25	36.00	3.18	36.75	3.33
Results Orientation	7	16.93	2.99	15.81	2.51	14.91	2.15	16.31	2.16
Special-Purpose Scales									
Leadership Potential	25	64.20	10.50	62.24	11.40	72.10	8.42	74.45	9.30
Telemarketing Potential	104	258.48	19.54	258.99	18.19	301.02	21.24	298.56	21.72
Customer Service	78	205.35	17.59	203.59	17.82	238.47	17.56	234.38	18.55
Constructiveness (Refl. Counterprod.)	81	198.11	22.40	205.17	19.55	231.07	16.34	222.40	15.98
Safety Orientation	51	122.92	16.58	128.43	14.14	146.23	12.46	139.57	12.20
Social Ascendancy	41	105.79	14.32	102.16	15.78	118.56	11.73	121.02	13.20
Disciplined Achievement	43	106.01	10.68	107.83	10.10	124.33	9.65	123.55	9.71
Calm Steadiness	40	102.32	10.08	97.56	10.80	118.61	9.59	116.68	10.13
Warmth and Concern for Others	42	110.60	10.66	114.67	10.14	130.93	10.25	123.97	10.30
Adaptability (Openness to Change)	33	86.47	7.96	83.71	8.07	89.52	5.96	91.64	6.63
Extraversion and Low Impulse Control	19	51.49	6.66	50.09	6.86	52.96	4.68	54.09	5.45
Integrity	35	87.64	11.13	91.40	9.20	107.25	9.44	102.78	9.61
Validity Scale									
Employment-Related Motivational Distortion	25	See corresponding values in Table 6							

^a*n*_s: 170–300.

^b*n*_s: 500–800.

^c*n* = 3,053.

^d*n* = 1,826.

Reliability Estimates for the BIODATA-250 Scales and Their Interpretation

Estimates of reliability—both internal consistency and test-retest stability—follow on the next two pages in Table 8. These estimates have been obtained over a period of several years with different samples of university undergraduate students. As noted earlier, the mean value of coefficient α in Table 8 is **.751**, and the mean test-retest reliability estimate is **.845**. These values should be understood in the context of a 22-item scale, the mean scale length of those in Table 8. We note here that BIODATA-250 scale lengths range from a low of 5 items to a high of 104, with 30 of the 43 scales having 20 items or fewer. For these latter 30 scales, the mean coefficient α is **.707**, and the mean test-retest estimate is **.823**.

In comparison with reported reliability estimates for scales from other self-report inventories, these reliability values should be considered solid. The *median* reported coefficient α for the 20 scales of the considerably-longer California Psychological Inventory-Form 434 (averaging 35 items per scale) is **.76**; the *median* coefficient α for the 43 scales of the BIODATA-250 inventory (averaging 22 items per scale) is **.77**.

Table 8

Reliability Estimates for the BIODATA-250 Scales in Several Samples of University Undergraduate Students

Scale Name	Internal Consistency Reliability (α)			Test-Retest (Stability) Reliability ^a			
	No. of Items	Male Univ. Stud's (ns: 400-450)	Female Univ. Stud's (ns: 1,200-1,250)	Total (Averaged across Gender; ns: 1,600-1,700)	Male Univ. Stud's (ns: 70-80)	Female Univ. Stud's (ns: 299-330)	Total (Averaged across Gender; ns: 370-410)
Broad Empirical Factor Scales							
Dominance	15	.89	.92	.91	.91	.94	.93
Orderliness	9	.72	.70	.71	.89	.82	.84
Sensitivity to Others	9	.65	.65	.65	.75	.81	.80
Optimism	8	.74	.75	.75	.83	.81	.82
Energy	10	.68	.59	.61	.75	.77	.77
Assertiveness	9	.69	.73	.72	.83	.84	.84
Independence	9	.62	.69	.67	.70	.79	.77
Social Confidence	16	.87	.89	.89	.91	.94	.93
Perseverance	14	.75	.78	.77	.79	.79	.79
Service Orientation	11	.82	.82	.82	.84	.88	.87
Social Insight	8	.64	.67	.66	.84	.75	.77
Straightforwardness	11	.75	.74	.74	.88	.83	.84
Self-Esteem	6	.53	.62	.60	.69	.80	.78
Resilience	9	.77	.76	.76	.86	.81	.82
Narrow Empirical Factor Scales							
Conformity	5	.50	.44	.46	.76	.73	.74
Risk-Taking	7	.67	.71	.70	.89	.87	.87
Attention to Detail	6	.55	.65	.63	.72	.76	.75
Adaptability	6	.50	.56	.55	.68	.70	.69
Impulse Control	6	.58	.58	.58	.75	.82	.80
Competitiveness	7	.59	.62	.61	.79	.80	.80
The Big Five Scales							
Emotional Stability (Neuroticism Reflected)	20	.77	.79	.79	.89	.87	.87
Extraversion	15	.81	.83	.83	.90	.91	.91
Openness	15	.74	.74	.74	.85	.86	.85
Agreeableness	14	.69	.69	.69	.80	.79	.79
Conscientiousness	39	.83	.82	.82	.90	.87	.88

Table Continues

Table 8 (Continued)

SCALE NAME	No. of Items	Internal Consistency Reliability (α)			Test-Retest (Stability) Reliability ^a		
		Male Univ. Stud's (ns: 400-450)	Female Univ. Stud's (ns: 1,200-1,250)	Total (Averaged across Gender; ns: 1,600-1,700)	Male Univ. Stud's (ns: 70-80)	Female Univ. Stud's (ns: 299-330)	Total (Averaged across Gender; ns: 370-410)
Conceptually-Based Scales							
Internal Locus of Control	6	.52	.65	.62	.76	.78	.77
Emotional Intelligence	18	.77	.79	.79	.90	.88	.88
Nurturance	20	.80	.81	.81	.83	.87	.86
Self-Sufficiency	13	.68	.74	.73	.80	.85	.84
Results Orientation	7	.66	.58	.60	.81	.77	.78
Special-Purpose Scales							
Leadership Potential	25	.91	.92	.92	.93	.95	.94
Telemarketing Potential	104	.86	.84	.85	.91	.91	.91
Customer Service Potential	78	.88	.89	.89	.88	.91	.91
Constructiveness (Refl. Counterproduct'ity)^b	81	.90	.89	.89	.93	.92	.93
Safety Orientation^b	51	.89	.87	.88	.93	.93	.93
Social Ascendancy	41	.92	.93	.93	.93	.95	.94
Disciplined Achievement	43	.83	.81	.82	.90	.87	.88
Calm Steadiness	40	.82	.85	.84	.90	.87	.89
Warmth and Concern for Others	42	.84	.84	.84	.90	.87	.88
Adaptability (Openness to Change)	33	.77	.79	.79	.87	.89	.89
Extraversion and Low Impulse Control	19	.80	.82	.82	.89	.91	.91
Integrity^b	35	.85	.82	.83	.91	.87	.88
Validity Scale							
Employment-Related Motivational Distortion^b	25	.81 (n = 978 job applicants)	.77 (n = 516 job applicants)	.78 (n = 1,494 job applicants)	.80 (n = 40 cued ^c univ. students)	.79 (n = 126 cued ^c univ. students)	.80 (166 cued ^c univ. students)

^aTest-retest time interval: 2–3 weeks.

^bBolded scales discussed in more detail in the text.

^cThese students were in the *Cued-MD Response group* (see text) and were responding to the BIODATA inventory on both occasions with this motivational-distortion set.

Another widely-used non-clinical personality inventory is the 16 Personality Factor Questionnaire. This inventory has a long history, starting in 1949, and is now in its fifth edition. Coefficient α estimates for the 16 primary factors of the 16 PF-5th ed. average .746, very close to the mean coefficient α for the BIODATA-250 inventory of .751. Test-retest reliability estimates of the 16 primary factors—with a two-week inter-test interval—average .799. When the five higher-order factors are added to the primary 16, the average (2-week) test-retest reliability is .815. These latter two values can be compared to the .845 test-retest value computed for the 43 scales of the BIODATA-250 inventory.

A particularly good comparison is available with the scales of the NEO PI-R inventory, a widely-used measure of the Big Five personality factors. Since the BIODATA-250 inventory provides measurement of the same five factors, a direct comparison of reliabilities is possible. From the NEO PI-R manual, the five full-factor scales, each scale consisting of 48 items, provide an average coefficient α of .888. The corresponding scales of the BIODATA-250 provide an average coefficient α (seen in Table 8) of .774. When this latter value is adjusted for scale length, the associated mean Big Five scale coefficient α is .889—virtually identical to the NEO PI-R average. With the shorter Big Five inventory, the NEO-FFI, the mean coefficient α for the Big Five factor scales is .792, again very close to that of the mean BIODATA-250 value for these scales.

With respect to test-retest reliability, or stability, the picture is similar, although more favorable to the BIODATA-250 inventory. From an early version of the NEO PI, the average test-retest reliability (with an inter-test interval of three months) for the first three factors (Neuroticism, Extraversion, and Openness) is .880. Again, these scales each contain 48 items. For the corresponding three BIODATA-250 scales (Emotional Stability, Extraversion, and Openness), consisting of an average of only 16.67 items, the mean test-retest (with an inter-test interval of two-three weeks) reliability value is .877. The adjusted (for length) mean test-retest BIODATA-250 reliability is .954. From these analyses, then, it must be acknowledged that the BIODATA-250 inventory measures the Big Five factors about as reliably as does the NEO PI-R, if not more so.

Since a predictive composite will likely consist of several of the BIODATA-250 scales, composite reliability can be expected to exceed the values reported in the first paragraph in this section. For example, if we were to combine three average-length BIODATA-250 scales, having, on average, an inter-scale correlation of .40 (a very reasonable possibility), our composite coefficient α would be .86, and our composite test-retest stability estimate would be .91. If the inter-scale correlations averaged, instead, .50, the corresponding values would be .88 and .92. These levels of empirically-obtained reliability meet all reliability standards set down in measurement textbooks and testing associations for adequate scale reliability to support individual personnel decisions.

BIODATA-250 Scale Intercorrelations

The intercorrelations among the 43 BIODATA-250 scales discussed above were calculated in both (a) a university undergraduate student sample and (b) an occupational sample. The results for these two samples, broken down by gender in the case of the student sample and by the level of job applied for in the case of the occupational sample appear in Tables 9 and 10 following on the next two pages.

Table 9

BIODATA-250 Scale Intercorrelations for University Undergraduate Students
(Correlations for Male Participants below the Diagonal; Those for Female Participants above the Diagonal)

Table with 30 columns (do, ord, sens, opt, en, as, ind, scon, pv, sv, si, sf, se, rs, cnfo, rt, atd, adp, imp, comp, em, ex, op, ag, cnsc, ilo, ei, nu, ss, re, lp, tm, cu, co, so, sasc, dach, cst, wco, aoc, eic, int, emd) and 30 rows representing the same variables. It shows a lower triangular matrix of correlation coefficients for male students and an upper triangular matrix for female students.

Note: Correlations for Male participants (n = 481; n = 171) are presented above the diagonal, and correlations for Female participants (n = 1301; n = 551) are presented below the diagonal. do = Dominance; ord = Orderliness; sens = Sensitivity to Others; opt = Optimism; en = Energy; as = Assertiveness; ind = Independence; scon = Social Confidence; pv = Perseverance; sv = Service Orientation; si = Social Insight; sf = Straightforwardness; se = Self-Esteem; rs = Resilience; cnfo = Conformity; rt = Risk-Taking; atd = Attention to Detail; adp = Adaptiveness; imp = Impulse Control; comp = Competitiveness; em = Emotional Stability; ex = Extraversion; op = Openness; ag = Agreeableness; cnsc = Conscientiousness; ilo = Internal Locus of Control; ei = Emotional Intelligence; nu = Nurture; ss = Self-Sufficiency; re = Results Orientation; lp = Leadership Potential; tm = Telemarketing; cu = Customer Service; co = Constructiveness; so = Safety-Oriented; sasc = Social Ascendancy; dach = Disciplined Achievement; cst = Calm Steadiness; wco = Warmth and Concern for Others; aoc = Adaptability and Openness to Change; eic = Extraversion and low Impulse Control; int = Integrity; emd = Employment-Related Motivational Distortion. p < .05 for Male |rs| ≥ .09; p < .01 for Male |rs| ≥ .12; p < .05 for Male |rs| ≥ .16; p < .01 for Male |rs| ≥ .21; p < .05 for Female |rs| ≥ .06; p < .01 for Female |rs| ≥ .07; p < .05 for Female |rs| ≥ .09; p < .01 for Female |rs| ≥ .12.

It must be pointed out in connection with these tables, that there is considerable item overlap among the scales, with a total of 250 items being used to construct 43 scales. It should be remembered that the BIODATA-250 inventory was not developed to measure pure, unifactorial dimensions (as some self-report inventories are), but rather to assess occupationally-relevant themes (with many sharing certain sub-themes), as these have been revealed in the industrial psychology literature over several decades. Many pairs of scales get at somewhat similar themes and contain similar item content, although also differing in certain ways.

A good example of this would be the two BIODATA-250 scales Dominance and Leadership Potential. Dominance is a standard relatively narrow trait assessed by almost all personality inventories, and constitutes a basic factor in several theoretical models of personality. The BIODATA-250 Leadership Potential scale, on the other hand, is a much broader, workplace-oriented dimension that includes several themes. All 15 of the items on the Dominance scale also appear on the Leadership Potential scale, but an additional 10 items tapping such themes as social effectiveness, persuasiveness, verbal fluency, and self-confidence expand the meaning of this dimension well beyond merely a tendency to be dominant. However, with this item overlap, the two scales correlate .95 in all four samples used to compute the correlations in Tables 9 and 10. Using both of these scales in a predictive application would not make sense, of course; there may, however, be situations in which one of the two is to be preferred over the other. In a similar way many of the larger BIODATA-250 scales are very much like linear combinations of narrower scales.

The purpose, then, in presenting the scale Intercorrelations is simply to allow users to see how the scales *as they are constituted in the BIODATA-250 inventory* correlate with one another, rather than to make conceptual sense of the scales themselves from these correlations or to draw conclusions about possible higher-order factorial structures. These correlations do not provide any deeper theoretical information about the dimensions represented by the BIODATA-250 scales, but merely provide us with the correlations among these particular sets of items.

Evidence of Construct Validity: Correlations between 43 BIODATA-250 Scales and the 20 Main Dimensions of the California Psychological Inventory (CPI)

The cross-correlations between the 43 BIODATA-250 scales, on the one hand, and the 20 main scales of the CPI, Form 434, on the other, appear on the next page in Table 11. Thus, the item-overlap effects noted earlier with the BIODATA-250 scale intercorrelations are completely absent with these correlations. These correlations were calculated on a sample of 192 university undergraduate students. The two inventories were administered to these students on different days in order to reduce response-set and contamination effects as much as possible. Because of the relatively small sample sizes, scores from each gender sample were first mean-deviated and then both sets of mean-deviated scores were pooled in order to compute the correlations. This procedure is, of course, identical to pooling the separate-gender covariance matrices first and removes any between-groups (between-genders) correlation.

These correlations provide us with some evidence of construct validity for the BIODATA-250 scales. Such validity is evidenced from high correlations between pairs of scales that putatively measure much the same thing. In addition, the correlations give us some greater insights into the psychological underpinnings of the BIODATA-250 scales (and also the CPI scales).

Table 11

BIODATA-250 Scale and CPI-Form 434 Scale Cross-Correlations for University Undergraduate Students (Mixed-Gender; n = 192)

BIODATA Scale	CPI Scale																			
	Do	Cs	Sy	Sp	Sa	In	Em	Re	So	Sc	Gi	Cm	Wb	To	Ac	Ai	Ie	Py	Fx	F/M
do	.73	.48	.58	.45	.59	.54	.44	.18	.01	-.14	.07	.19	.25	.16	.30	.23	.41	.33	.01	-.23
ord	.27	.04	.14	.00	.12	.10	.07	.12	.11	.15	.19	.06	.15	.02	.29	.02	.06	.13	-.37	-.06
sens	.09	.06	.10	.02	.10	-.04	.09	.12	.11	.24	.31	.10	.20	.06	.22	.06	.05	.07	-.12	-.01
opt	.46	.30	.41	.39	.38	.39	.28	.06	.15	-.03	.15	.22	.36	.08	.20	.17	.34	.26	-.01	-.20
en	.11	.03	-.05	-.07	-.04	.07	-.03	.11	-.05	.15	.24	-.15	.01	.07	.08	.08	.04	.18	-.13	-.15
as	.53	.36	.42	.43	.49	.46	.32	.03	-.01	-.17	-.04	.19	.19	.08	.12	.14	.28	.27	.05	-.12
ind	.17	.02	.01	.07	.14	.18	-.03	-.21	-.09	-.21	-.24	.03	-.05	-.14	.00	-.07	.05	.06	-.17	-.17
scon	.61	.54	.65	.54	.59	.54	.46	.11	-.03	-.18	.04	.16	.27	.17	.16	.24	.37	.27	.23	-.24
pv	.26	.12	.18	-.01	.10	.09	.12	.31	.23	.21	.30	.19	.21	.10	.43	.08	.10	.18	-.29	.11
sv	.31	.24	.27	.13	.15	.12	.29	.49	.19	.25	.37	.29	.25	.27	.30	.22	.17	.24	.04	.10
si	.21	.17	.24	.16	.23	.13	.19	.11	.09	.06	.17	.18	.14	.12	.17	.06	.11	.14	-.07	.00
sf	.05	.03	.00	-.15	-.07	-.04	.02	.43	.15	.41	.46	.08	.17	.27	.25	.14	.02	.17	-.15	.07
se	.33	.29	.34	.41	.28	.41	.17	.04	.17	.01	.02	.12	.30	.13	.22	.28	.40	.29	.04	-.16
rs	.19	.18	.12	.18	.15	.27	.13	.11	-.04	.18	.32	.07	.24	.15	.04	.22	.19	.21	.13	-.25
cnfo	-.10	-.11	-.02	-.18	-.14	-.21	-.12	.13	.10	.07	.15	-.05	-.09	-.06	.06	-.15	-.16	-.13	-.25	.13
rt	.29	.35	.33	.45	.42	.30	.36	-.15	-.26	-.35	-.19	.04	.08	.07	-.10	.23	.25	.17	.34	-.37
atd	-.01	-.03	-.05	-.07	-.04	-.02	.02	.27	.25	.31	.20	.15	.16	.24	.29	.14	.09	.11	-.08	.19
adp	.26	.28	.28	.27	.23	.27	.28	.08	.02	.08	.20	.10	.28	.22	.10	.22	.26	.30	.27	-.16
imp	-.18	-.14	-.15	-.21	-.19	-.19	-.14	.13	.31	.41	.34	.04	.16	.09	.21	.00	-.02	.03	-.21	.17
comp	.01	-.05	.03	.06	.08	.00	-.12	-.21	-.14	-.39	-.38	-.10	-.18	-.29	-.12	-.14	-.09	-.14	-.12	-.05
em	.55	.43	.46	.49	.48	.57	.33	.15	.11	.10	.31	.25	.47	.19	.28	.34	.46	.44	.06	-.32
ex	.53	.53	.69	.59	.58	.43	.51	.18	.06	-.15	.08	.24	.33	.23	.20	.30	.38	.29	.26	-.15
op	.47	.47	.38	.46	.41	.41	.45	.18	.04	.00	.10	.29	.38	.36	.20	.46	.49	.46	.40	-.30
ag	.26	.28	.28	.12	.15	.11	.33	.48	.24	.40	.53	.27	.38	.38	.36	.31	.24	.30	.09	.05
cnsc	.30	.14	.16	-.01	.09	.14	.11	.33	.22	.31	.38	.12	.25	.18	.49	.16	.17	.31	-.34	.01
ilo	.32	.31	.32	.32	.32	.22	.27	.28	.15	.04	.10	.18	.29	.28	.28	.30	.27	.32	-.01	-.07
ei	.52	.45	.58	.45	.53	.39	.45	.18	.04	-.04	.18	.24	.31	.21	.24	.21	.33	.27	.13	-.16
nu	.28	.22	.25	.11	.16	.07	.27	.43	.20	.30	.43	.27	.28	.24	.34	.20	.16	.22	-.02	.07
ss	.58	.40	.45	.48	.53	.54	.34	.01	-.01	-.20	-.06	.19	.21	.08	.14	.19	.34	.32	.05	-.20
re	-.14	-.14	-.09	-.03	-.06	-.08	-.24	-.41	-.21	-.33	-.36	-.24	-.29	-.45	-.28	-.33	-.22	-.28	.16	-.08
lp	.72	.52	.64	.51	.61	.56	.47	.15	-.01	-.17	.05	.18	.27	.16	.25	.23	.40	.32	.08	-.23
tm	.48	.27	.36	.14	.27	.30	.21	.43	.26	.30	.46	.21	.32	.21	.52	.18	.24	.33	-.29	.00
cu	.70	.55	.62	.49	.56	.54	.50	.37	.17	.12	.36	.33	.47	.31	.44	.37	.46	.45	.04	-.18
co	-.14	-.26	-.31	-.47	-.37	-.24	-.25	.36	.31	.59	.51	.03	.07	.09	.34	-.06	-.16	.03	-.47	.28
so	-.24	-.33	-.42	-.51	-.46	-.26	-.37	.23	.28	.55	.40	-.04	.00	.01	.25	-.12	-.21	-.06	-.45	.29
sasc	.73	.55	.66	.58	.64	.61	.48	.13	.00	-.18	.04	.20	.30	.17	.23	.27	.44	.36	.12	-.26
dach	.26	.11	.13	-.03	.07	.10	.09	.34	.26	.36	.42	.13	.26	.20	.50	.16	.17	.31	-.34	.03
cst	.61	.47	.51	.54	.53	.58	.40	.18	.13	.07	.28	.28	.48	.25	.29	.37	.49	.47	.06	-.32
wco	.28	.25	.29	.13	.19	.10	.30	.46	.21	.34	.48	.27	.32	.32	.34	.23	.18	.26	-.01	.05
aoc	.49	.48	.42	.53	.49	.49	.49	.06	-.06	-.11	.03	.21	.33	.30	.13	.42	.47	.46	.42	-.37
eic	.52	.51	.64	.58	.57	.44	.49	.10	-.07	-.28	-.06	.19	.22	.16	.09	.26	.34	.24	.30	-.19
int	.07	-.02	-.03	-.22	-.13	-.06	.02	.53	.29	.49	.53	.15	.17	.28	.38	.11	.02	.17	-.25	.20
emd	.15	.03	.01	-.11	-.08	.03	.01	.50	.32	.52	.57	.15	.26	.25	.43	.14	.08	.21	-.31	.11

Note: All scale scores were mean-deviated by gender before pooling for the calculation of the correlations to eliminate between-groups correlation; n = 192.

BIODATA-250 scales: do = Dominance; ord = Orderliness; sens = Sensitivity to Others; opt = Optimism; en = Energy; as = Assertiveness; ind = Independence; scon = Social Confidence; pv = Perseverance; sv = Service Orientation; si = Social Insight; sf = Straightforwardness; se = Self-Esteem; rs = Resilience; cnfo = Conformity; rt = Risk-Taking; atd = Attention to Detail; adp = Adaptiveness; imp = Impulse Control; comp = Competitiveness; em = Emotional Stability; ex = Extraversion; op = Openness; ag = Agreeableness; cnsc = Conscientiousness; ilo = Internal Locus of Control; ei = Emotional Intelligence; nu = Nurturance; ss = Self-Sufficiency; lp = Leadership Potential; tm = Telemarketing; cu = Customer Service; co = Constructiveness; so = Safety-Oriented; sasc = Social Ascendancy; dach = Disciplined Achievement; cst = Calm Steadiness; wco = Warmth and Concern for Others; aoc = Adaptability and Openness to Change; eic = Extraversion and low Impulse Control; int = Integrity; emd = Employment-Related Motivational Distortion.

CPI scales (from Form 434): Do = Dominance; Cs = Capacity for Status; Sy = Sociability; Sp = Social Presence; Sa = Self-acceptance; In = independence; Em = Empathy; Re = responsibility; So = Socialization; Sc = Self-control; Gi = Good Impression; Cm = Communality; Wb = Well-being; To = Tolerance; Ac = Achievement via Conformance; Ai = Achievement via Independence; Ie = Intellectual Efficiency; Py = Psychological-mindedness; Fx = Flexibility; F/M = Femininity/Masculinity.

Correlations with $p < .01$ appear in boldface type, and those with $p < .05$ are italicized.

When viewing these cross-correlations with an eye to those larger ones for which correlation might be expected, it is important to keep in mind the attenuating effects of less-than-perfect reliability. In the present case, with the scales of the BIODATA-250 inventory cross-correlated with those of the CPI-Form 434 inventory, we are dealing with scale reliabilities averaging on the order of about .80 or slightly lower. If we were to use the two median coefficient α values—.76 for the CPI scales, .77 for the BIODATA-250 scales—our fully disattenuated cross-correlations would, on average, be about *1.31 larger* than the values in Table 11.

As an example, if we were to apply this to the cross-correlation between the two inventory Dominance scales (.73), we would have a value, for the disattenuated cross-correlation, of .95. This latter figure is an estimate of what the correlation between the construct measured by each of CPI-Dominance and BIODATA-250-Dominance would be if the two scales had been perfectly reliable. Clearly then, the two scales are assessing the same construct. Similarly, if we adjust the correlation between CPI-Sociability and BIODATA-250-Extraversion, our attenuated r of .69 becomes a disattenuated .90, again indicating that the two scales are measuring almost exactly the same construct.

There are few surprises in Table 11 for those who understand the conceptual nature of the two sets of scales. The large correlation between CPI-Do and BIODATA-250-Do has already been discussed. If we look at the CPI Class I scales as a group (dealing with interpersonal style and effectiveness, and comprising scales Do through Em in Table 11), we see correlations that make sense between these scales and many of the BIODATA-250 scales. Again, not surprisingly, CPI-Do correlates with BIODATA-250 Leadership Potential to almost the same extent as it does with BIODATA-250 Do, .72 and also with the broader BIODATA-250 scale Social Ascendancy, .73. These large correlations (having disattenuated values in the .90s) provide independent evidence that Dominance is a large component theme of these broader, applied BIODATA-250 scales.

In what follows, the disattenuated values will appear in parentheses to the right of the obtained unadjusted correlation coefficients. Still examining CPI Class 1 scales, we see high correlations between BIODATA-250 Social Confidence and CPI Sociability, .65 (.85) and Social Presence .54 (.71). BIODATA-250 Self Sufficiency correlates .53 (.69) with CPI Self Acceptance and .54 (.71) with CPI Independence. CPI-Independence can be seen to correlate .49 (.64) with BIODATA-250 Adaptability and Openness to Change and .54 (.71) with BIODATA-250 Self Sufficiency. Many other large correlations of interest appear in the first seven columns of Table 11 where the BIODATA-250 correlates of the CPI Class I scales appear. These correlations are conceptually meaningful and anticipated.

The CPI Class II scales (CPI Re through To in Table 11) tap into the understanding of and adherence to social norms. There are a number of BIODATA-250 scales assessing similar themes. The BIODATA-250 scale correlations with CPI Responsibility (Re) of Service Orientation, .49 (.64), Agreeableness, .48 (.63), and Integrity, .53 (.69) are all consistent with the thematic content of CPI-Re. The CPI Self Control (Sc) scale correlates predictably with BIODATA-250 scales Impulse Control, .41 (.54), Straightforwardness, .41 (.54), Constructiveness, .59 (.77), Safety Orientation, .55 (.72), and Integrity, .49 (.64)—all BIODATA-250 scales that have as one theme that of self-control and carefulness. The CPI scale Good Impression (Gi) has a very similar meaning conceptually as the BIODATA-250 scale Integrity, and, not surprisingly, the two scales correlate .57 (.75). The CPI Class II scale Well-Being (Wb) taps into themes of good

mental health and productiveness. This scale's correlations with the BIODATA-250 scales Optimism, .36 (.47), Emotional Stability, .47 (.62), Agreeableness, .38 (.50), and Calm Steadiness, .48 (.63), are, therefore, precisely what one would expect.

The CPI Class III scales (having a more cognitive/intellectual and achievement focus) are CPI-Ac (Achievement via Conformance), -Ai (Achievement via Independence), and -Ie (Intellectual Efficiency, since renamed "Conceptual Fluency"). With these CPI scales, we might expect to see significant correlations with the BIODATA-250 scales that capture the more cognitive and achievement-oriented themes. With respect to CPI-Ac, the correlations with BIODATA-250 Perseverance, .43 (.56), Conscientiousness, .49 (.64), and Disciplined Achievement, .50 (.66) make complete sense. CPI-Ai, stresses achievement through independent and creative channels. Its correlations with BIODATA-250 Openness, .46 (.60) and Adaptability and Openness to Change, .42 (.55) pick up on the openness, as opposed to rigid behavior, that is a central subtheme of CPI-Ai. Finally, CPI-Ie, perhaps better understood by its new name, Conceptual Fluency, also has a subtheme of openness in thinking, and, thus, its correlation of .49 (.64) with BIODATA-250 Openness is conceptually meaningful. Its correlations with BIODATA-250 Adaptability and Openness to Change, .47 (.62) and Calm Steadiness, .49 (.64) also support the fluency and clear-thinking nature of those high on the CPI-Ie scale.

Harrison Gough, author of the CPI, refers to the Class IV scales—Py (Psychological Mindedness, since renamed, fortunately, Insightfulness), Fx (Flexibility) and F/M (Femininity/Masculinity), as measuring stylistic aspects of thinking. The one aspect of CPI-defined Insightfulness that is picked up by the BIODATA-250 inventory is that of openness, represented in the BIODATA-250 by Openness and Adaptability and Openness to Change. Of the Big Five factors, Openness is the most cognitively-oriented. BIODATA-250 Openness correlates .46 (.60) with CPI-Py and BIODATA-250 Adaptability and Openness to Change also correlates .46 (.60). CPI Flexibility also correlates reasonably highly with BIODATA-250 Openness, .40 (.52) and Adaptability and Openness to Change, .42 (.55). Adaptability can certainly be seen as closely related to flexibility.

All things considered, these correlations between the BIODATA-250 scales and those of the CPI do establish the first evidence available of construct validity for the BIODATA-250 inventory scales. This evidence can provide confidence to users of the BIODATA-250 inventory, that the scales are actually measuring what they were intended to, and that users can use the scales to obtain reliable assessment of these important, basic psychological constructs. With more use, it will be possible to obtain cross-correlations between BIODATA-250 scales and those of other well-established self-report inventories, and to, thus, extend the accumulated evidence for construct validity for the inventory.

CHAPTER 5

THE USE OF THE BIODATA-250 INVENTORY IN THE CONSTRUCTION OF ORGANIZATION-SPECIFIC DIMENSIONS —GATHERING EVIDENCE FOR VALIDITY BASED ON CONTENT

Some Background

It has been noted above that a central goal in the development of the BIODATA-250 inventory has been to provide a content-rich database of items suitable for use in the construction of new, content-relevant dimensions suggested by organizational needs. Thus, although the empirically- and conceptually-based scales will be useful for personnel-selection purposes in many cases, there may well be a need for more organizationally-specific measurement.

Although personnel psychologists are inclined to regard criterion-related evidence of validity as the *sine qua non* of defensible validity, as this is understood in the personnel-selection context, such evidence is often either unavailable or logistically too difficult to obtain locally. This means that we must often search for other evidence of validity for the measures we use in our personnel assessments.

The Society for Industrial and Organizational Psychology provides a nice perspective on the establishment of validity evidence in their 2003 document, *Principles for the Validation and Use of Personnel Selection Procedures* (Fourth Edition). Generally speaking, three kinds of validity evidence are normally considered as “evidence of validity.” One of these is the criterion-related evidence noted in the preceding paragraph.

A second kind of validity evidence also involves the relationships between test variables and other variables and is usually referred to as *construct validity*. In establishing the existence of construct validity for a test or scale, we are demonstrating that the test or scale is actually measuring what it was designed to measure. We have provided some evidence of construct validity for a number of BIODATA-250 scales by correlating them with scales from the California Psychological Inventory (Form 434), as discussed in Chapter 4, with the numerical results appearing in Table 11. For example, the fact that BIODATA-250 scale Dominance correlates .73 with CPI 434 scale Dominance (as discussed in some detail in Chapter 4) provides strong evidence that we can count on the BIODATA-250 Dominance scale to capture the essentials of this construct.

Another kind of validity evidence, however, comes from an analysis of the *content* of the scales used in practice. Thus, we might attempt to replicate in the elements of a test or scale—to the extent possible—the content of a particular job or a job function. In such an applied setting, we might like to construct test elements (items in the present case) that reflect the important work behaviours, activities, and/or worker KSAOs (knowledge, skills, abilities, and other personal characteristics) necessary for successful work performance.

Linking such job elements to test items can be a challenging task when the test items have originally been developed to cover the whole range of human personality, motivational, attitudinal, and performance traits. Further, one can hardly include on even a test dedicated to selection for a particular job actual job elements when the persons taking the test have never

held that job or a similar one—the state of affairs that normally exists in the personnel-selection context. Further, if the work and worker requirements may change in the job or class of jobs, such close linkage is even more difficult.

Evidence of a thorough content-based validation study can come close to that of criterion-based validation study in justifying the use of a test or scale in the applied personnel-selection context. Ideally, such a content-based study—or a similar approach—would precede a full criterion-related validity study. Rather than being aimed at the prediction of a particular job or employment function, however, such a content-based study could be aimed at capturing evidence of relevant dimensions or job-related attributes in an organization. If, for example, a company value—one that permeated virtually all positions in the organization—were being a good team member, this could be addressed first via a content-based approach.

For this approach to produce high-quality measurement, a carefully thought-out strategy must be employed. First, the behaviours one would expect to see in an employee manifesting a high level of the attributes or job actions must be set down. To link these job-related qualities to the content of a self-report instrument like the BIODATA-250 inventory, some translation of these behaviours into personal traits that would be expected to be associated with them must take place. This translation is easier if the linkage is to job-related attributes than if the linkage is to actual behaviours. The thoroughness of the linking process will determine the extent to which the scales so developed will be useful in aiding in effective personnel selection.

Ideally, organizational personnel would be involved in establishing any linkages between attributes captured by inventory items and those understood by the organization to define the attributes in a purely job-related way. Organizational management persons could evaluate each of a sample of candidate items as to the extent to which it reflects important aspects of the job attribute or behaviour.

Application of the Content Validity Approach with the BIODATA-250 Inventory

To see this process illustrated, consider the following work done with the BIODATA-250 inventory in partnership with an organization that based work performance appraisal around six key organizational dimensions. We first obtained from the organization as many details as possible about the components of each organizational attribute or dimension *as it was understood in the organization*. From these details, we constructed a linkage grid in which we attempted to match BIODATA-250 item content (item by item) to the detailed components of each dimension. In some cases, the dimension name was identical to existing BIODATA-250 scale names, but the organizational definitions of these dimensions were unique to the organization and thus somewhat different from the content of the existing scales.

The organizational dimensions in this case were: (a) *Results Orientation*, (b) *Teamwork and Cooperation*, (c) *Service Orientation*, (d) *Organizational Commitment*, (e) *Flexibility*, and (f) *Listening, Understanding, and Responding*. Through a linkage process, we chose BIODATA-250 items that had as strong content overlap as possible with the details of each dimension. This process led to six new BIODATA-250 scales designed specifically for this one organization, dimensions that would aid it in selecting new employees that would best manifest the constituent qualities composing each dimension.

Given the relatively clear meaning of each dimension and the breadth of the BIODATA-250 item content, we were confident that we could find linkages among the BIODATA-250 items, but, given the heterogeneous nature of some of the dimensions, high levels of internal consistency reliability were not necessarily expected. We calculated these, and, since we had a very large item database, we were also able to evaluate the test-retest stability of the new scales. What follow are the results of this content-based study. The internal-consistency coefficients are based on a large sample of 2,111 job applicants. The stability coefficients are based on a sample of 383 mixed-gender (with mean-deviated scores to remove between-groups correlation) university undergraduate students.

Table 12

Internal-Consistency and Stability Reliability Estimates for Six Content-Based BIODATA-250 Scales Developed for One Organization

Content-Based BIODATA-250 Scale	No. of Items	Internal-Consistency Reliability (α) ($n = 2,111$)	Test-Retest (Stability) Reliability ^a ($n = 383$)
Results Orientation	24	.78	.84
Teamwork and Cooperation	21	.63	.82
Service Orientation	15	.78	.88
Organizational Commitment	27	.84	.89
Flexibility	24	.73	.88
Listening, Understanding, and Responding	19	.74	.77

^aTest-retest time interval: 2–3 weeks.

Interestingly, the mean length of these content-based scales is 21.67, very close to the average length (22) of the 43 BIODATA-250 scales described earlier. Further, the mean coefficient α of the six scales in Table 12 is .750. The mean coefficient α of all 43 scales is .751 (as noted earlier). Given the heterogeneity of some of the organizational dimensions (for example, *Teamwork and Cooperation*, and *Listening, Understanding, and Responding*), high estimates of internal consistency are just not realistic. As long as the scale scores demonstrate stability over time, the essential aspects of psychometric reliability in the applied context have been satisfied.

The mean test-retest reliability estimate in Table 12 is .847, which can be compared to the mean test-retest estimate given earlier for all 43 scales, .845. From these results, it can certainly be argued that these special-purpose, purely employment-related, scales, which were developed in partnership with the client organization using the BIODATA-250 inventory for personnel-selection and employing these six corporate dimensions to structure of their personnel-assessment, possess adequate reliability for general use. Since they will not be used alone for individual personnel decisions, linear combinations of two or three of them can be expected to produce composite reliabilities of approximately .84 (α) and .91 (stability), with the latter by far the more important in the personnel-selection context.

Flexibility of Measurement in Applied Settings

As noted at the beginning of this document, a central goal in the development of the

BIODATA-250 inventory was to assemble a comprehensive database of reliable items that would allow the construction of reliable and valid scales needed in particular personnel-selection applications. Although the measurement of a number of what might be considered “mainstream” traits has resulted from the developmental work with the BIODATA-250 inventory, we have seen, in the just-preceding example, that the item base is sufficiently rich to permit the construction of special-purpose scales that are more relevant than the mainstream scales in specific personnel-selection applications.

We consider this flexibility necessary if self-report information is to be useful in most job-related settings. In the past, self-report inventories that permit measurement of a fixed set of dimensions have been used in such settings under the belief that those dimensions were in some sense the most “real” and, further, those needed to explain occupational performance. A good example would be the use of the Big Five factors (Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness) in such applications. Although these factors may be the current set of most-researched personality factors in academic work, there is no reason to think that we can use self-report information that is restricted to only these five factors effectively in forecasting success in behaviours as complex as those related to job performance.

Application of a content-based approach to achieving validity in the job-applicant context requires comprehensive, detailed, and systematic analysis of both the observable aspects of the job behaviours or organizational dimensions and the items to be selected for a predictive scale. It is not enough simply to speculate about seemingly matching job or dimension elements and test item content. The job or dimension in question must first be broken down into the smallest-possible elements. As an example, one dimension noted above in our description of how we applied this approach with an organization, was *Organizational Commitment*. One observable aspect of this dimension in the organization is *completes projects that have been started*. This can be matched almost perfectly with Item 18 in the BIODATA-250 inventory: *I always complete projects that I start*. Thus, the extent to which the organizational job behaviours or dimensions have been analyzed into the most molecular units possible will determine the effectiveness of any linking of these units to test items. The linkage process will benefit greatly by having both organizational personnel, as well as the consulting test specialists, taking part, with some standards set for deciding on whether a match has been found or not.

The next logical step in the larger process of developing truly predictive scales involves empirical correlational work with incumbent employees, in which correlations are obtained between scores on the scale resulting from the above-described process and those on the corresponding job-performance measure arising from systematic performance evaluation. This is the criterion-related evidence of validity noted earlier that would provide confirmation about the success of the content-validity approach. This would be very desirable information to have, although more difficult to obtain than that accompanying the content-based process described in this chapter. Although such criterion-related validity evidence is more definitive in establishing the overall validity of a selection procedure than the other approaches noted, the content-based approach described here is a good first step.

REFERENCES AND FURTHER READING

- Cattell, R.B., Cattell, A.K., & Cattell, H.E.P. (1993). *16PF Fifth Edition Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Chuapetcharasopon, P., LeRoy, Z. P., Hakstian, A. R., Netchaeva, E., Desjardins, N., & Siu, L. W. (2010, June). *Does employment-related motivational distortion on the Big Five scales vary across Holland's RIASEC occupational types?* Symposium conducted at the 71st annual meeting of the Canadian Psychological Association, Winnipeg, MB.
- Costa, P.T., Jr., & McCrae, R.R. (1992). *NEO Personality Inventory-Revised professional manual*. Odessa, FL: Psychological Assessment Resources
- Gough, H. G. (1995). *California Psychological Inventory-434 professional manual*. Palo Alto, CA: Consulting Psychologists Press, Inc.
- See also: <http://www.inforefuge.com/california-personality-inventory>.
- Hakstian, A. R., & Ng, E. L. (2005). Employment-related motivational distortion: Its nature, measurement, and reduction. *Educational and Psychological Measurement*, 65, 405-441.
- LeRoy, Z. P., Florko, L., Hakstian, A. R., Gonsar, T., Sui, L., Desjardins, N., Chuapetcharasopon, P., & Cuttler, C. (2008, June). *The development and application of an employment-related motivational distortion scale in personality assessment*. Poster session presented at the annual meeting of the Canadian Psychological Association, Halifax, Canada.
- LeRoy, Z. P., Hakstian, A. R., Cuttler, C., Gonsar, T., Florko, L., & Siu, L. (2008, April). *Understanding employment-related motivational distortion through factor-analytic techniques*. Poster session presented at the annual conference of the Society for Industrial and Organizational Psychology, San Francisco, USA.
- LeRoy, Z. P., & Hakstian, A. R. (2007, April). *Unsafe behaviors: An empirically-developed scale and its personality correlates*. Poster session presented at the annual conference of the Society of Industrial and Organizational Psychology, New York, USA.
- LeRoy, Z. P., & Hakstian, A. R. (2006, June). *The development of a trait-based scale to assess safety-related behaviors*. Poster session presented at the annual convention of the Canadian Psychological Association, Calgary, Canada.
- LeRoy, Z. P., & Hakstian, A. R. (2004, June). *Examination of the predictive correlates of safety-oriented behaviors*. Poster session presented at the annual convention of the Canadian Psychological Association, St. John's, Canada.
- Ng, E. L., & Hakstian, A. R. (2001, June). *The prediction of hazardous behaviors by means of the California Psychological Inventory*. Poster session presented at the annual meeting of the Canadian Psychological Association, Ste.-Foy, Quebec, Canada.
- Ones, Viswesvaran, C., & Schmidt, F. L. (1993). Meta-analysis of integrity test validities. *Journal of Applied Psychology*, 78, 679-703.

Pirani, Z., & Hakstian, A. R. (2003, April). *The development of empirically-derived scales for the prediction of counterproductive and hazardous behaviors*. Undergraduate honours thesis paper presented at the 3rd Annual Undergraduate Psychology Conference, University of British Columbia, Vancouver, Canada.

Schmit, M. J., Kihm, J. A., & Robie, C. (2000). Development of a global measure of personality. *Personnel Psychology, 53*, 153-193. (This article describes the development of the Global Personality Inventory, published by Personnel Decisions International.)