

Predictive Index® Technical Overview

Todd C. Harris, Ph.D.

**Director of Research: PI Worldwide
Visiting Professor: Department of Psychology
University of Massachusetts at Amherst**

Allison J. Tracy, Ph.D.

**Methodologist and Senior Research Scientist:
Wellesley Centers for Women
Wellesley College**

Gwen G. Fisher, Ph.D.

**Assistant Research Scientist:
Institute for Social Research
University of Michigan**

PI Worldwide 

16 Laurel Avenue, Wellesley Hills
Massachusetts 02481-7532
Telephone: +1 781 235 8872
Toll-Free in USA: 800 832 8884
Fax: +1 781 235 0959
www.PIworldwide.com

Introduction

If you have ever heard of the phrase “the people make the place”, then you’ll understand why the use of personality assessments in business and industry continues to grow rapidly. Since approximately the late 1980’s, the academic study of personality and the application of personality theory toward the solution of key organizational challenges has undergone a marked renaissance. Interest in personality has also expanded past traditional domains such as personnel selection and hiring (Roberts & Hogan, 2001) to touch upon diverse areas such as the influence of personality on team performance, leadership, organizational culture and climate, entrepreneurship and innovation, stress and well-being, work motivation, job satisfaction, and a host of others. Personality assessments tap into each employee’s unique “behavioral DNA,” yielding key insights into people’s individual drives, temperaments and motivations (e.g. why I might enjoy analyzing financial statements for hours on end, while you’d prefer to be out of the office developing personal relationships with customers.) Hundreds of empirical research studies, conducted in a wide variety of settings, have conclusively demonstrated the quantitative connection between personality and job performance (Barrick & Mount, 1991; Barrick, Mount & Judge, 2001). Quite simply, scores on well-developed personality measures are stable across time and situations, and are useful predictors of behavior and job performance.

Personality Defined

Personality has been defined in many different ways by a variety of theorists. Harvard’s Gordon W. Allport, in his seminal textbook on the field (Allport, 1937), defined personality thusly:

“Personality is the dynamic organization within the individual of those psychophysical systems that determine his unique adjustment to his environment.”

Another pioneer in the field of personality, Raymond B. Cattell of the University of Illinois viewed personality as (Cattell, 1945):

“That which permits a prediction of what a person will do in a given situation. The goal of psychological research is to establish laws about what different people will do in all kinds of social and general environmental situations.”

MacKinnon (1944) suggested that personality has two conceptually distinct definitions. Used one way, personality refers to the distinctive and unique impression that one makes on others. This perspective refers to personality from the viewpoint of the *observer*, and is functionally equivalent to a person’s *reputation*. Used in a second way, personality refers to the structures inside of a person that are useful in explaining why a person creates a particular impression on others.

This is personality from the perspective of the *actor*, concerned with how a person perceives him or herself, and is functionally equivalent to a person's *identity*. This "dual definition" of personality is also espoused by more contemporary personality researchers and theorists (e.g., Hogan, 2004).

The Five-Factor Model (FFM) of personality normally attributed to Tupes and Christal (1961) suggests that many existing personality constructs can be organized in terms of five broad traits: (1) Extroversion, (2) Agreeableness, (3) Conscientiousness, (4) Emotional Stability, and (5) Openness.

The Growth of Personality Assessment

What factors have led to the increased usage of personality assessments around the world? First, the U.S. population is increasingly diverse, as are the populations of many other countries. Personnel selection systems that rely solely or primarily on measures of cognitive ability significantly adversely affect most protected groups, especially African Americans, Native Americans and Hispanics. White people are often hired at a disproportionately high rate when typical cognitive ability tests are the primary selection and screening tools. These adverse impacts created substantial pressure on companies to find equally valid but less discriminatory selection techniques. Research on personality variables indicates that they have much less, and often no, adverse impact on members of protected groups (Hough, 1998), a tremendous advantage when dealing with increasingly heterogeneous customer, supplier and employee bases.

Second, the research findings from the U.S. Army's Selection and Classification Project (known as Project A), a multi-million dollar, seven-year research effort clearly indicated that the Army could improve its forecasts of overall job performance via the addition of personality assessments to its battery of cognitive tests.

Third, the appearance of the Five-Factor Model (FFM) in the early 1960's spurred a large amount of academic research in the area of personality, and provided a reasonably comprehensive yet parsimonious taxonomy to organize the measurement of personality. The amount of academic research on personality continues to increase, as evidenced by the 40 different personality-based papers presented at the 2010 Annual Society for Industrial and Organizational Psychology Conference, and the appearance of personality-themed articles in leading journals such as the Journal of Applied Psychology.

Fourth, influential meta-analytic reviews such as those by Barrick and Mount (1991) and Tett, Jackson, and Rothstein (1991), in which the relationships between a variety of measures of personality and a variety of job performance metrics were examined provided further empirical evidence for the predictive validity of personality measures.

Fifth, “real-world” views held by non-psychologists continue to hold that the qualities and characteristics within individuals have a crucial impact on their performance at work. For example, an examination of most job postings indicates that employers frequently seek personality-driven factors such as social skills, initiative, creativity, flexibility, etc. at least as often, if not more so, than specific technical skills, experience or intellectual abilities. Thus, there appears to be a widely held view that personality really matters.

The last factor that may be driving the increased importance and usage of personality assessments in business and industry is the nature of work and the economy itself. The world of work has changed more rapidly in the past 15 years than in the prior 100 combined. Eight fundamental forces are shaping the 21st Century workplace:

1. Increased technology, especially information technology.
2. Increased diversity and globalization.
3. Increased prevalence of collaborative, team-based work structures.
4. Increased concern for person-organization “fit.”
5. The importance of work/non-work balance.
6. The frequency and pace of change.
7. An increased emphasis on learning.
8. New expectations and definitions of leadership.

These factors do not operate independently, but rather accentuate and multiply one another. For example, leaders who operate in today’s global environment have to possess fundamentally different skill sets than their predecessors of prior generations did. As a group, these conditions make the explicit consideration of applicant and employee personality factors by organizations more critical than ever before.

The Predictive Index®

The Predictive Index® (PI®) is a theory-based, self-report measurement of normal, adult, work-related personality, and has been developed and validated exclusively for use within occupational and organizational populations. The PI employs a free-choice (as opposed to forced-choice) response format, in which individuals are presented with two lists of descriptive adjectives, both containing 86 items, and are asked to endorse those which they feel describe them (the “SELF” domain), and then those which they feel coincide with how they feel others expect them to behave (the “SELF-CONCEPT” domain). Summing across these two domains yields a third implied domain (the “SYNTHESIS”), which can be interpreted as reflecting an employee’s observable behavior in the workplace. The assessment is un-timed, generally takes approximately five to ten minutes to complete, and is available in paper-and-pencil, desktop and internet formats. The PI® has been in wide-spread commercial use since 1955, with minor revisions to the assessment occurring in 1958, 1963, 1988 and 1992. These minor revisions were undertaken to both improve the PI’s psychometric properties and to insure that each of the individual items on the assessment conformed to appropriate and contemporary language norms.

The PI[®] is currently used by over 7,800 organizations across a wide variety of industries and company sizes, including 51 companies listed in the 2009 Fortune 500, and 82 companies listed in the 2009 Fortune Global 500. Organizations that use the PI[®] are located in 143 different countries, with approximately 30% of PI[®] utilization occurring outside of North America. In 2010, over one million people around the world completed the Predictive Index[®] assessment. The PI[®] is used for a variety of talent management purposes, such as personnel selection, executive on-boarding, leadership development, succession planning, performance coaching, team-building, and organizational culture change among others.

PI[®] Construct Definitions

The PI[®] measures four primary and fundamental personality constructs:

1. **DOMINANCE:** The degree to which an individual seeks to control his or her environment. Individuals who score high on this dimension are independent, assertive and self-confident. Individuals who score low on this dimension are agreeable, cooperative and accommodating.
2. **EXTROVERSION:** The degree to which an individual seeks social interaction with other people. Individuals who score high on this dimension are outgoing, persuasive and socially-poised. Individuals who score low on this dimension are serious, introspective and task-oriented.
3. **PATIENCE:** The degree to which an individual seeks consistency and stability in his or her environment. Individuals who score high on this dimension are patient, consistent and deliberate. Individuals who score low on this dimension are fast-paced, urgent and intense.
4. **FORMALITY:** The degree to which an individual seeks to conform to formal rules and structure. Individuals who score high on this dimension are organized, precise and self-disciplined. Individuals who score low on this dimension are informal, casual and uninhibited.

The PI[®] also measures two “secondary” personality constructs, which are derived from a combination of each of the four “primary personality” constructs described previously:

1. **DECISION-MAKING:** Measures how an individual processes information and makes decisions. Individuals who score high on this dimension are objective, logical and are primarily influenced by facts and data. Individuals who score low on this dimension are subjective, intuitive and are primarily influenced by feelings and emotions.
2. **RESPONSE-LEVEL:** Measures an individual’s overall responsiveness to the environment, which is reflected in his or her energy, activity level and stamina. Individuals who score high on this dimension have an enhanced capacity to sustain activity and tolerate stress over longer periods of time. Individuals who score low on this dimension have less of this capacity.

The scoring of the Predictive Index® checklist produces a behavioral pattern with three elements, known as the Self, the Self-Concept and the Synthesis. The Self measures a person's natural, basic and enduring personality. The Self Concept measures the ways in which a person is trying to modify his or her behavior to satisfy perceived environmental demands. Lastly, the Synthesis, which is a combination of the Self and Self-Concept, measures the ways in which a person typically behaves in his or her current environment.

Users of the Predictive Index® should be aware of the following: (1) as demonstrated by our nearly 500 criterion-related validity studies, dating back to September of 1976, the PI's results are job-related and significantly predict job performance across a wide range of jobs and performance criteria; (2) PI's job analysis tool, known as the PRO, can be used to independently establish the behavioral requirements and benchmarks for a job; and (3) at any time, as part of the service contract, a client can request us to perform, free of charge, a local criterion-related validity study that uses their own employees and performance measures, quantifiably demonstrating the connection between the PI® and job performance in their workplace.

Please note that all research studies on the PI® conducted by PI Worldwide are designed and executed in a manner that is consistent with the U.S. Federal Regulations of the Equal Employment Opportunity Commission's (EEOC) Uniform Guidelines for The Development and Use of Personnel Selection Procedures (1978). These guidelines are designed to provide a framework for determining the proper use of tests and other selection procedures, as well as preventing discriminatory employment practices. Research on the PI® is also conducted in accordance with the guidelines for test development established by professional organizations such as the American Psychological Association (APA), and the Society for Industrial and Organizational Psychology (SIOP). This research indicates that: (a) traits measured by the Predictive Index® add incremental validity to the prediction of job performance, and (b) use of the Predictive Index® within a personnel selection system does not produce adverse impact.

At present, PI Worldwide conducts and supports scientific research in the following areas:

The Reliability of PI®

Reliability refers to the consistency or stability of a measure (Nunnally & Bernstein, 1994). If the concept being measured is assumed to be consistent, such as a personality trait, then the measure should yield similar results if the same person responds to it a number of times. If the concept being measured is assumed to be inconsistent, such as mood, then the measure should yield dissimilar results if the same person responds to it a number of times. One way to estimate reliability is by computing the measure's "test-retest reliability". Test-retest reliability is perhaps the easiest assessment of a measuring device's reliability to conceptualize and understand. Using the same group of people, a construct is measured at two separate points in time and then

the two sets of scores are compared. This technique yields a correlation often known as the *coefficient of stability*, because it reflects the stability of test scores over time. If the measure under study is reliable, people will have scores that are similar across trials. Note that the shorter the time interval between administrations of the test (e.g. two weeks versus three months), the higher will be the test-retest coefficient.

The test-retest reliability of the PI[®] was first examined in a 1983 study by Perry & Lavori, who reported the following coefficient of stability values:

Table 1: PI[®] Test-retest Reliability Results from Perry-Lavori

Sample Size	N = 85	N = 87	N = 86
Time Interval	Three Months – Two Years	Two Years – Four Years	Four Years – Eight Years
Self A (Dominance)	.71	.53	.49
Self B (Extroversion)	.67	.57	.56
Self C (Patience)	.54	.58	.52
Self D (Formality)	.70	.61	.56
Self-Concept A (Dominance)	.56	.45	.47
Self-Concept B (Extroversion)	.48	.47	.50
Self-Concept C (Patience)	.60	.51	.44
Self-Concept D (Formality)	.66	.46	.47
Synthesis A (Dominance)	.67	.55	.52
Synthesis B (Extroversion)	.60	.55	.56
Synthesis C (Patience)	.61	.60	.49
Synthesis D (Formality)	.69	.59	.54
Synthesis M (Response Level)	.60	.57	.51

A 1999 study by Everton reported the following coefficient of stability values:

Table 2: PI[®] Test-retest Reliability Results from Everton

Sample Size	N = 77	N = 58
Time Interval	Two Weeks	Six Months
Self A (Dominance)	.80	.75
Self B (Extroversion)	.71	.80
Self C (Patience)	.76	.71
Self D (Formality)	.80	.57
Self-Concept A (Dominance)	.68	.65
Self-Concept B (Extroversion)	.61	.73
Self-Concept C (Patience)	.71	.48
Self-Concept D (Formality)	.56	.57
Synthesis A (Dominance)	.84	.76
Synthesis B (Extroversion)	.69	.82
Synthesis C (Patience)	.83	.62
Synthesis D (Formality)	.74	.60
Synthesis E (Decision-Making)	.71	.58
Synthesis M (Response-Level)	.81	.71

2009 and 2011 studies by Harris, utilizing samples of 61 and 44 undergraduate students respectively, reported the following coefficient of stability values over a two-week interval:

Table 3: PI[®] Test-retest Reliability Results from Harris

PI[®] Factor	Stability Coefficient: 2009 Study	Stability Coefficient: 2011 Study
Self A (Dominance)	.84	.79
Self B (Extroversion)	.77	.81
Self C (Patience)	.75	.70
Self D (Formality)	.83	.87
Self-Concept A (Dominance)	.72	.70
Self-Concept B (Extroversion)	.69	.63
Self-Concept C (Patience)	.65	.73
Self-Concept D (Formality)	.77	.74
Synthesis A (Dominance)	.81	.78
Synthesis B (Extroversion)	.82	.84
Synthesis C (Patience)	.78	.78
Synthesis D (Formality)	.75	.68
Synthesis E (Decision-Making)	.78	.80
Synthesis M (Response-Level)	.70	.73

Taken as a whole, these results indicate that the PI[®] demonstrates acceptable levels of test-retest reliability.

A second way to estimate reliability is by computing the measure's "internal consistency reliability." This is accomplished by determining whether the individual items on the assessment intended to measure the same construct (such as Dominance) are mathematically related. Internal consistency methods

estimate the reliability of a test based solely on the number of items within the test and the average inter-correlation among those items. The internal consistency reliability of the PI[®] has been examined in three different studies. The average internal consistency reliability of PI[®] Factors across these studies is .85, with a range of .82 to .87. Although estimates vary, the lower boundary for the acceptability of internal consistency reliability is often taken as .70. There may also be an upper boundary of acceptability as well, perhaps .90 or above, as that may signal measurement redundancy across some of the items.

Note that the last exploration of the internal consistency of PI[®] constructs was undertaken in 2009, using a randomly selected sample of 3,991 applicants to and employees from a variety of US businesses in a variety of industries: Information Technology, Financial Institutions, Surface Transportation, Air Transportation, Manufacturers, Wholesalers, Retail, Hospitality, and Restaurants. The sample contained individuals from all levels of their respective organizations, and included a variety of jobs, such as warehouse workers, tellers, accountants, pilots, inside and outside sales, and vice presidents.

The following alpha coefficients (Cronbach, 1951) were obtained in this study: Self Concept (Factor A=0.78, Factor B=0.87, Factor C=0.77 & Factor D=0.84) and Self domains (Factor A=0.80, Factor B=0.88, Factor C=0.82 & Factor D=0.86). Note that in 1996 study based on data from a sample of 1,411 individuals, internal consistency reliability estimates for Self Factors ranged from 0.79 to 0.87 (Wolman & Pratt, 1996).

PI[®] Measurement and Scaling Structure

The PI[®] assumes that a single personality dimension is being tapped by each PI[®] factor scale. For example, Factor A is expected to assess the broad personality trait of dominance. Nuances of this trait are captured in the individual adjectives and summed to broadly define dominance. This assumption of unidimensionality within each factor is examined empirically in this evaluation.

According to classical test theory, the variability in an item set can be separated into shared, or “true score” variance and unique, or “measurement error” variance (see Allen & Yen, 2002). In other words, differences between individuals’ responses to an adjective can be attributed to actual differences in the personality dimension it is designed to measure (“true score”) and to unrelated influences (“measurement error”). For instance, the word “popular” can evoke images of specific attributes of an individual's past schoolmates in addition to more abstract notions of popularity associated with extroversion. The degree to which an item is influenced by factors other than that intended is referred to as measurement error. All psychological assessment instruments contain some proportion of measurement error.

In this evaluation, individual items are evaluated in an effort to discover the extent to which they contribute true score variance to the factor score estimates.

Note that the adjectives reflecting each PI[®] factor are words of degree - not true opposites. For example, a peaceful person can be somewhat belligerent. Item Response Theory (IRT) scaling and other measurement models assume that items are ordered according to their ability to detect differences in degree at different points along the underlying personality continuum. In other words, one item may be endorsed at very low levels of a personality dimension and, as the degree of the dimension increases, another item may also be endorsed.

Because psychological dimensions cannot be directly observed but must be inferred from measurement characteristics of the items in the assessment instrument, changes in the measurement model (item changes, shifts in the proportion of measurement error, etc.) imply changes in the interpretation of the factor. Meaningful comparison of factors requires the assumption that factor scores have the same meaning in both the Self Concept and Self domains. In other words, scaling must be identical across domains if one is to say, for instance, that an individual perceives himself to be less formal than he feels others expect him to be. Since factor scores are composites of item sets, this assumption implies that each item reflects a personality dimension in the same way in the two domains and has a similar proportion of measurement error.

The measurement properties of the items underlying the construction of factor scores allow us to explicitly test the assumption of equivalence of scale across domains.

A key feature of the PI[®] is that factor scores can be used to construct an individual's personality profile, or factor pattern. Because the relative intensity levels of the four personality dimensions measured by Factors A through D are essential to the interpretation of personality profiles, there is an implied assumption of comparability of the metric of factor scores. In other words, personality tendencies representing a similar degree of intensity should be plotted at similar points on the PI[®] factor plotting template. Only by making this assumption regarding the scaling of the factors can one say, for example, that an individual has a higher tendency toward dominance than extroversion (High A/Low B).

Principal component analysis, a special case of exploratory factor analysis, can be used to evaluate the assumption of unidimensionality in the sets of items comprising each PI[®] Factor. Utilizing the same sample of 3,991 applicants and job incumbents referenced above, eigenvalue decomposition was used to create a scree plot for each item set, which shed light on the optimal number of components (factors) that can be extracted. Scree plots obtained from models for each PI[®] factor showed evidence of unidimensionality. As can be seen in the Appendix of this document, Figures 1 (Self-Concept) and 2 (Self), there is a decidedly sharp drop-off after the first eigenvalue for each of the Factors A through D, followed by a relatively flat slope.

This implies that the most parsimonious solution describing the item set is one with a single underlying dimension. Additional dimensions lend only negligible

explanatory value to the solution. This examination supports the assumption that a single personality construct is described by adjectives in each item set.

Comparative Fit Index (CFI) values, which is a standardized scale assessing overall model fit, for Self Factors A, B, C and D were CFI = .92, .96, .95 and .90 respectively. The Root Mean Square Error of Approximation (RMSEA), which summarizes the degree of local fit, or the consistency of fit throughout all areas of the model (Rigdon, 1996) = 0.06.

Individual item loadings in Factor A ranged from 0.49 to 0.76 in the Self-Concept domain (mean=0.64) and from 0.52 to 0.78 in the Self domain (mean=0.64). In Factor B, the loadings in the Self-Concept domain ranged from 0.51 to 0.75 (mean=0.64) and in the Self domain, loadings ranged from 0.56 to 0.74 (mean=0.67). Two items had loadings lower than 0.40: "Amiable" (Factor C, Self Concept and Self domains) and "Passive" (Factor D, Self-Concept domain). The remaining loadings in Factor C ranged from 0.60 to 0.78 in the Self-Concept domain (mean=0.66) and from 0.57 to 0.72 in the Self domain (mean of 0.66). Remaining loadings for Factor D ranged from 0.44 to 0.71 (mean=0.57) in the Self-Concept domain and from 0.45 to 0.70 (mean=0.61) in the Self domain.

In 2003, an examination of the factor structure of the Swedish version of the PI[®] was undertaken by Dr. Lennart Sjoberg of the Stockholm School of Economics (Sjoberg, 2003), starting with two factors and going up to ten. The results of this analysis indicated only small drops in eigenvalue size after the fourth factor, giving support to the notion of a maximum of four factors.

The Validity of the PI[®]

While reliability refers to the consistency of a measure, validity refers to the accuracy of a measure. A measure is valid if it actually measures what it purports to measure. The validity of a measure can be assessed in a number of ways.

Construct validity is demonstrated when a measure is statistically compared with another measure of similar and/or different concepts. To be successful, the comparison measure must have been soundly constructed and be generally accepted. Such research on the PI[®] has been conducted twice. Both of these studies compared the PI[®] with Raymond Cattell's 16PF (The 16PF is a well-respected and well-researched personality assessment). A construct validity study involves looking at patterns of correlations. Correlations are mathematical measures that can identify the presence and strength of the relationship between two variables.

A pattern should emerge that meets the following expectations: Factors that are defined in a similar way by both the PI[®] and the 16PF should prove to be very similar statistically (e.g., PI's extroversion factor and the 16PF extroversion

factor), and factors that are defined in a dissimilar way on both the PI[®] and the 16PF should prove to be mathematically unrelated (e.g., PI's extroversion factor and the 16PF emotional stability factor).

In both of these studies, the PI[®] successfully demonstrated construct validity: the relationships you would intuitively expect should be related were mathematically related, and the relationships you would intuitively not expect should be related were mathematically unrelated. For example, the correlation between PI's Dominance factor and the 16PF's Independence factor was .47 ($p < .01$), between PI's Extroversion factor and 16PF's Extroversion factor was .34 ($p < .01$) and between PI's Patience factor and 16PF's Tension Factor was .35 ($p < .01$), indicating that respondents who had relatively low Patience scores as measured by the PI[®] tended to also be tense, impatient and driven as measured by the 16PF.

A 2010 pilot study (N = 186) comparing the relationship between the Predictive Index[®] and analogous constructs measured by the NEO PI-R, discovered the following correlations, all in the hypothesized direction and all significant at the $p < .01$ level:

- PI[®] Factor A (DOMINANCE) and NEO PI-R "AGREEABLENESS": $r = -.61$
- PI[®] Factor B (EXTROVERSION) and NEO PI-R "EXTROVERSION": $r = .63$
- PI[®] Factor D (FORMALITY) and NEO PI-R "CONSCIENTIOUSNESS": $r = .61$

Criterion-related validity is demonstrated when a measure is statistically compared with behaviors it claims to predict. We say that the PI[®] is related to, and can predict, behaviors in the workplace. Criterion-related validity studies objectively show whether these relationships exist, and if so, they show the nature of these relationships.

The PI[®] has been investigated in nearly 500 concurrent (in which data for the predictor and criteria are collected at the same time) and predictive (in which there is some time-lapse between when the data for the predictor and criteria are collected) criterion-related validity studies since September of 1976, for a variety of jobs, in a variety of industries, in a variety of countries and utilizing a wide range of job performance metrics, such as tenure, turnover, sales, and customer satisfaction. This body of evidence supports the fact that the PI[®] is indeed consistently related to important workplace outcomes, with studies typically yielding uncorrected correlations between PI[®] Factors and individual job performance criteria in the .20 to .40 range. These correlations indicate that the PI[®] can be an effective predictor of workplace performance.

Meta-Analytic Results for the Predictive Index[®]

Meta-analysis is a statistical procedure designed to combine the results of many individual, independently conducted empirical studies into a single result or

outcome (Hunter & Schmidt, 1990; Raju et al., 1991). The logic underlying meta-analysis is that researchers can arrive at a more accurate or “truer” conclusion about a particular topic via combining or aggregating the findings of a wide cross-section of many studies that address the topic, instead of relying on the results obtained in a single study. In 2008, a PI[®] meta-analysis examining the most recent 57 predictive and concurrent criterion-related validity studies was conducted. These 57 studies were conducted over a time period ranging from April, 2003 to February, 2008. Cumulative sample size was 5,765 (Mean N = 101.1; Median N = 74.0; Maximum N = 431; Minimum N = 32). The studies comprising the meta-analysis were conducted on behalf of 51 different PI[®] clients (including 21 Fortune 500 clients), drawn from 20 different industry classifications (e.g. “financial services”) and 15 different occupational classifications (e.g. “customer service”). 14 of the 57 studies were conducted with managerial-level jobs, with the remainder focused on individual-contributor roles such as bank tellers, customer service associates, truck drivers, machine operators, etc.

Meta-analytic results were as follows:

Table 4: PI[®]- Job Performance Correlations:

PI [®] Factor	Mean	Standard Deviation	High	Low
Self A (Dominance)	.17	.09	.49	.02
Self B (Extroversion)	.18	.10	.50	.00
Self C (Patience)	.18	.10	.47	.00
Self D (Formality)	.19	.10	.48	.03
Synthesis E (Decision-Making)	.18	.10	.50	.00
Synthesis M (Response-Level)	.18	.12	.51	.01

Table 5: PI[®] - Job Tenure Correlations:

PI [®] Factor	Mean	Standard Deviation	High	Low
Self A (Dominance)	.13	.08	.35	.01
Self B (Extroversion)	.15	.11	.47	.01
Self C (Patience)	.14	.08	.41	.02
Self D (Formality)	.14	.10	.39	.01
Synthesis E (Decision-Making)	.11	.08	.33	.01
Synthesis M (Response-Level)	.14	.09	.37	.02

A group of notes about the meta-analytic methodology employed in this meta-analytic project warrants mention:

- All of the above results were weighted by the respective sample sizes of the original studies.

- The correlations reported above have *not* been corrected for unreliability, a fact that in all likelihood somewhat reduced the PI[®] – job performance and PI[®] – tenure correlations reported here.
- Both statistically significant (i.e. $p < .05$) and non-significant results were included in the meta-analysis.
- The correlations reported above did not differ significantly by either industry or occupational classification. Additionally, the strength of the relationships between the PI[®] and job performance and job tenure was the same for managerial versus non-managerial jobs.
- A wide variety of job performance criteria were included in the meta-analysis, including objective production data (e.g. average units sold per month), personnel data (e.g. number of unexcused absences over an eight-month period) and judgmental data (e.g. supervisor ratings of work quality). This may at least partially explain the variability of PI[®] – job performance correlations reported in the individual criterion-related validity studies.

The Impact of Demographic Variables on PI[®] Scores

We have performed research to determine whether the Predictive Index[®] discriminates against protected classes. In a report written by Dr. Richard Wolman of Harvard University, the PI[®] was analyzed to determine whether men and women tended to score differently on the PI[®], and whether African-Americans, Hispanics, and Caucasians tended to score differently on the PI[®]. His analyses showed that neither gender nor race was significantly related to PI[®] scores.

In a more recent study by Dr. Todd Harris, the PI[®] was analyzed to determine whether the PI[®] produces adverse impact based on age. The study showed that for all PI[®] factors, people over age 40 (the protected class) had PI[®] patterns that were no different than people under age 40, confirming similar findings initially obtained by Dr. Wendi Everton in 1998.

In a 2008 investigation of the impact of gender and race on PI[®] scores in a sample of 347 employees in the banking industry, covering a variety of jobs (e.g. teller, branch manager, loan officer, etc.), gender and race accounted for less than 2% and 3% of the variability, on average, in PI[®] Factor scores. The results of this body of research indicate that the PI[®] is age-, gender- and race-neutral, and we believe that the inclusion of a well-validated personality assessment such as the PI[®] in a company's personnel selection system may lead to a more demographically diverse workforce.

Note that there is no evidence to indicate that the inclusion of the PI[®] in a company's personnel selection system, either in a compensatory or "multiple-hurdle" selection model, results in adverse impact against any protected class. When examining adverse impact via the four-fifths rule, Adverse Impact (AI) Ratios for large-scale studies of the PI[®] typically range from a low of .86 to a high of 1.25. When examining the issue of adverse impact via the "two Standard

Deviations” rule which we believe can be a more defensible and appropriate approach in situations in which there are low sample sizes in either the applicant or selected pool, we again see no evidence of the PI[®] producing adverse impact. For example, in a recent multi-year study of bank tellers for a Fortune 1000 financial services company, a Z-score comparison of the selection ratios for blacks and whites was $Z = 1.51$. In a second study of retail store managers, a Z-score comparison of the selection ratios for males and females was 1.20. When requested by the client, the PI Worldwide Research Department will partner with PI[®] clients to examine adverse impact and other selection-related statistics, incorporating both PI[®] and non-PI[®] predictors, and make any appropriate recommendations. Custom research studies, including pass-rate and adverse impact analyses, can also be conducted for PI[®] clients, designed in full accordance with Uniform Guidelines and other professional standards.

Cross-Cultural Research on the PI[®]

The PI[®] has now been translated into 63 different languages, has been used globally since 1958, and is seeing strong growth in Asia and India. We continue to take steps to insure the PI’s validity in non-U.S. markets. For example, over the past five years, we have conducted criterion-related validity studies involving employees from China, Canada, India, Britain, Spain, Portugal, Germany, Hungary, Russia, Australia and the Netherlands. The results of these international job validity efforts demonstrate the same quantitative connections between the PI[®] and job performance that our U.S.-based studies have shown.

Recent studies have also documented the PI’s internal consistency reliability, freedom from bias, and construct validity in non-U.S. markets. For example, in 2000 and 2003 in studies of the Swedish version of the PI[®] checklist, Dr. Lennart Sjöberg of the Stockholm School of Economics demonstrated that the PI[®] had acceptable reliability, validity and internal factor structure. The report “The Construct Validity of the Predictive Index[®] in a South African Sample” (1996) demonstrated PI’s construct validity as compared to the well-regarded 16PF personality assessment.

The 2006 report entitled “The Impact of Race and Gender on PI[®] Scores in a Large South African Sample” indicated that PI[®] scores were not significantly impacted by race or gender in 1,620 employees from South Africa. Three additional reports, published in 2007, demonstrated that the PI[®] is neither impacted by age or nor gender in samples of employees drawn from the U.K., India or Romania.

A March 2011 study of 197 Predictive Index[®] profiles drawn from Poland (128 males; 61 females; 8 with sex data not available), covering a time period from November, 2004 to February, 2011, and a range of occupational classifications, found no statistically significant ($p < .05$) relationships between sex and any personality construct measured by the Predictive Index[®].

The Impact of Differing Modes of Administration on PI[®] Scores

The results of our research indicate that an individual's PI[®] scores are not significantly impacted by mode of administration (i.e. paper-and-pencil vs. desktop computer-administered vs. internet). For example, a 2007 project used a stratified random sample of 15,000 PI's selected from our client database and covering a wide range of industries, companies, occupations and jobs. One third of this sample had taken the PI[®] via paper-and-pencil (N=5,000), one third of the sample had taken the PI[®] via an electronic kiosk/desktop mode (N=5,000) and one third had taken the PI[®] via the internet (N=5,000). For each individual PI[®] Factor, mode of administration accounted for much less than 1% of the variability in individual scores on that Factor.

Additionally, the shapes of the distributions for all PI[®] Factors were consistent across each of the modes of administration, being normally distributed (Skew and Kurtosis values <1.0) with equivalent means and medians for all PI[®] Factors at the Self, Self Concept and Synthesis levels of analysis. Additionally, the results of factor-analytic work indicated that PI[®] factor structure was invariant across these three modes of administration.

The Relationship between the PI[®] and General Cognitive Ability

When personality variables that are correlated with a job performance criterion are added to a battery of other personnel selection instruments measuring constructs such as general cognitive ability, psychomotor abilities, technical skills, mechanical aptitude and vocational interests, predictive validity is typically increased because personality variables are essentially uncorrelated with these other potential predictors of job performance (Judge, Higgins, Thoresen & Barrick, 1999). For example, in samples ranging in size from 7,188 to 8,547, Hough, Kamp and Ashworth (1993) reported median correlations between personality variables and measures of cognitively ability of .04. On a conceptual level, personality and intelligence are typically viewed as separate constructs. Intelligence has traditionally reflected the "can do" aspect of an individual; in effect, the employee can do the job because he or she possesses an adequate level of intelligence to do so. Personality has traditionally reflected the "will do" aspect of the individual; in effect, the employee will do the job because he or she has the motivation or temperament to do so. Personality and intelligence can both be predictive of job performance, each in their own way. For example, employees comparably more dependent on personality factors might achieve job success via being dependable, conscientious, helpful, team-oriented and so on. Employees comparably more dependent on intelligence might achieve the same level of job success by being excellent problem-solvers, synthesizers of information, and analyzers of data.

Our research has indicated that the PI[®] is unrelated to general cognitive ability. In four recent studies in which this relationship was examined (aggregate N=915), each of which utilized samples from different clients, industries and jobs, the median absolute correlation between individual PI[®] Factors and a variety of measures of general cognitive ability was .06 (Mean = .07; SD = 0.6).

Table 6: PI[®] – Cognitive Ability Correlations:

PI[®] Factor	Study One 2002 N = 82	Study Two 2004 N = 172	Study Three 2005 N = 182	Study Four 2008 N = 178	Study Five 2010 N = 301
Self A (Dominance)	.14	.19	.05	-.02	.09
Self B (Extroversion)	-.19	.05	-.02	-.06	-.07
Self C (Patience)	.05	-.18	.01	.09	.12
Self D (Formality)	.06	-.17	-.03	.02	.15
Synthesis E (Decision-Making)	.03	-.07	.04	.00	.02
Synthesis M (Response-Level)	-.01	.08	.01	.04	.06

PI Worldwide Academic Collaborations

The Research Department at PI Worldwide welcomes the opportunity to collaborate with academic researchers across disciplines. Here is a sampling of some of the research projects that utilized the Predictive Index[®] in 2010:

Institution	Department	Research Topic
Indiana University	Management	<i>The Role of Personality in Volunteerism Choices and</i>

		<i>Satisfaction</i>
Eastern Kentucky University	Psychology	<i>Administrators' Personality Traits and Nursing Home Performance</i>
Johnson and Wales University	Hospitality	<i>The Relationship between Personality Factors and Student Attitudes</i>
University of Sao Paulo (Brazil)	Management	<i>Six Sigma and the Project Portfolio Management Process (PPMP): Predictors of Success</i>
Curtin University (Australia)	Center for Research on Aging	<i>The Impact of Personality on Staff Retention in the Community Care Environment</i>
Capella University	Education	<i>At a Distance: A Comparative Study of Distance Delivery Modalities for Ph.D. Nursing Students</i>
University of Zaragoza (Spain)	Psychology	<i>Role Evolution Among Managers</i>

References

- Allen, M.J., & Yen, W. M. (2002). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.
- Allport, G.W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Barrick, M.R., & Mount, M.K. (1991). The Big Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology*, *44*, 1-26.
- Barrick, M.R., Mount, M.K., & Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9-30.
- Cattell, R.B. (1945). The principal trait clusters for describing personality. *Psychological Bulletin*, *42*(3), 129-161.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297 – 334.
- Everton, W. (1999). *A further investigation of the construct validity of the Predictive Index®*. (Validity Report No. 1999.11.01). Wellesley Hills, MA: Praendex, Inc.
- Hogan, R. (2004). Personality psychology for organizational researchers. In B. Schneider & D.B. Smith (Eds.), *Personality and organizations* (pp. 1-23). Mahwah, N.J.: LEA.
- Hough, L.M., Kamp, J. & Ashworth, S.D. (1993). *Development of Project A temperament inventory: Assessment of background and life experiences (ABLE)* (Institute Report No. 259). Minneapolis, MN: Personnel Decisions Research Institutes.
- Hough, L.M., Ones, D.S., & Viswesvaran, C. (1998, April). Personality correlates of managerial performance constructs. In R.C. Page (Chair), *Personality determinants of managerial potential performance, progression and ascendancy*. Symposium conducted at the 13th annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Hunter, J.E., & Schmidt, F.L. (1990). *Method of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Judge, T.A., Higgins, C.A., Thoresen, C.J. & Barrick, M.R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*, 621-652.
- MacKinnon, D.W. (1944). The structure of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 4-43). New York: Ronald.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.) New York: McGraw-Hill.

Perry, J.C., & Lavori, P.W. (1983). *The Predictive Index[®]: A report on reliability and construct validity*. (Validity Report No. 1983.09.01). Wellesley Hills, MA: Praendex, Inc.

Raju, N.S., Burke, M.J., Normand, J. & Langlois, G.M. (1991). A new meta-analytic approach. *Journal of Applied Psychology*, 76, 432-446.

Rigdon, E.E (1996). CFI versus RMSEA: a comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling*, 3, 369–379.

Roberts, B.W. & Hogan, R. (Eds.) (2001). *Personality psychology and the workplace*. Washington, D.C. : American Psychological Association.

Sjoberg, L. (2003). *Psychometric properties of the Swedish version of the Predictive Index[®] Checklist*. Stockholm, Sweden. Department of Economic Psychology, Stockholm School of Economics.

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

Tupes, E.C., & Christal, R.E. (1961). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60, 225-251.

Wolman, R.N., & Pratt, J.H. (1996). *A Normative Reliability Investigation of the Predictive Index[®] Organization Survey Checklist*. Wellesley Hills, MA: Praendex Incorporated.

Appendix

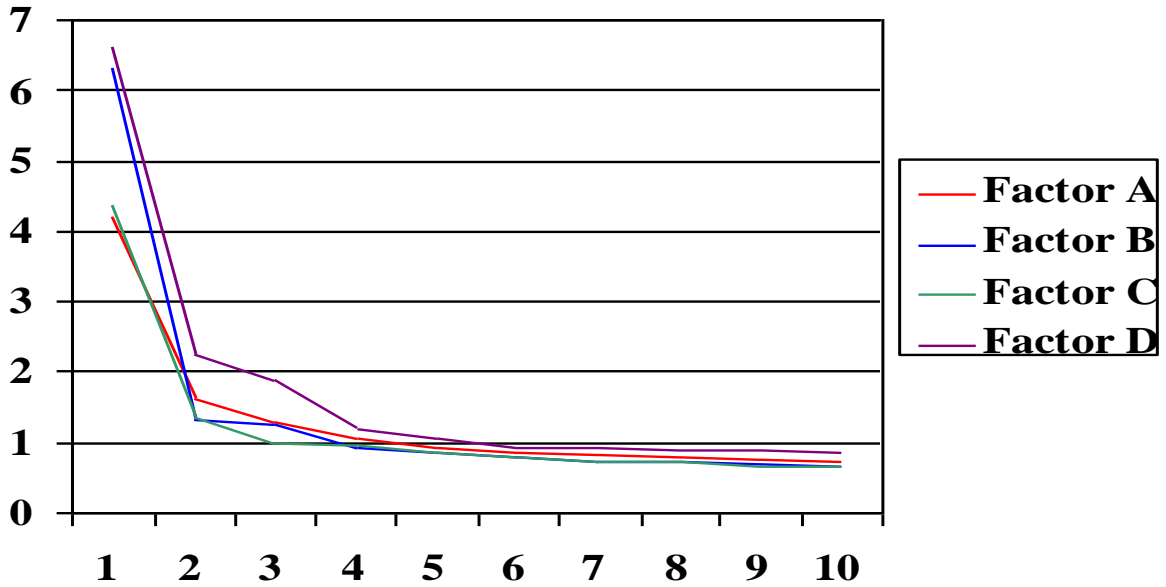


Figure 1. Scree plots of the eigenvalues for the Self Concept domain (Form IV Evaluation Study, subsample #1: n=3,991).

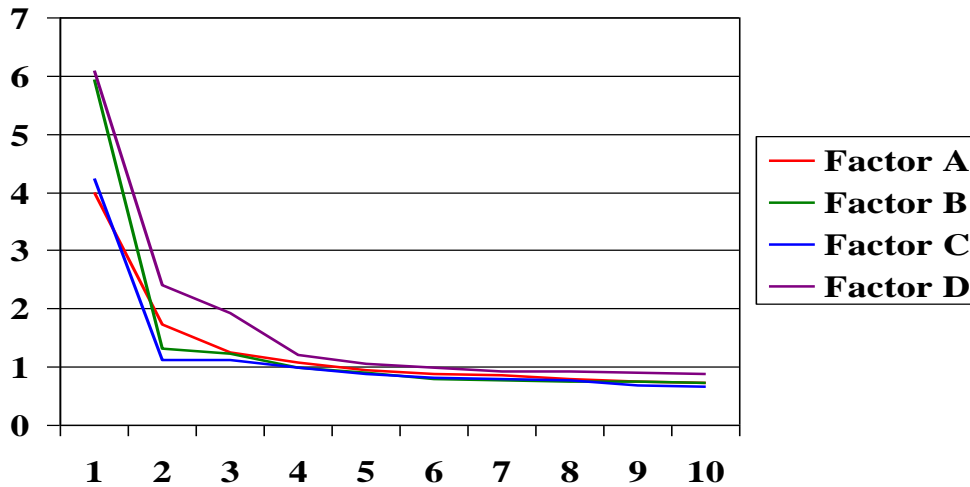


Figure 2. Scree plots of the eigenvalues for the Self domain (Form IV Evaluation Study, subsample #1: n=3,991).