

1 A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several Five-Factor models

Lewis R. Goldberg

University of Oregon and Oregon Research Institute, USA

This chapter is a plea for help in changing the way that we construct new measures of personality characteristics. Because I am going to propose a somewhat radical alternative to conventional practice, those who are satisfied with the pace of progress in the technology of personality assessment may not be pleased with my arguments.

In my view, however, the science of personality assessment has progressed at a dismally slow pace since the first personality inventories were developed over 75 years ago. What is usually taken to be the earliest personality instrument, Woodworth's Personal Data Sheet (PDS), was published in 1917, and since that time thousands of other instruments have been developed. Like the PDS, most of these have been of limited bandwidth, typically providing measures of one, two, or at most three traits. Virtually all of these *narrow-bandwidth* instruments are in the public domain--the items and their scoring keys having been published in scientific books, journal articles, or student theses or dissertations. The items are freely used by other scientists, either in their original form or quite commonly in some customized format. Examples of attributes measured by such narrow-bandwidth instruments include Achievement-Motivation, Adjustment, Conservatism, Coronary-Risk, Dogmatism, Empathy, Extraversion-Introversion, Guilt, Hostility, Locus of Control, Masculinity and/or-versus Femininity, Narcissism, Neuroticism, Openness to Experience, Optimism, Private and Public Self-Consciousness, Right-Wing Authoritarianism, Self-Disclosure, Self-Esteem, Self-Monitoring, Sensation-Seeking, Test-Anxiety, and Trust.

On the other hand, most *broad-bandwidth* personality inventories (like the MMPI, CPI, 16PF, and NEO-PI) are proprietary instruments, whose items are copyrighted by the test authors. As a consequence, the instruments cannot be used freely by other scientists, who thus cannot contribute to their further development and refinement. Indeed, broad-bandwidth inventories are rarely revised. At most, after many decades of commercial use, some of the most dated items might be changed and/or new norms established. For many inventories, nothing is ever done at all.

The manuals for some of these commercial inventories include tables of correlations between the scale scores and various criterion indices. But, such empirical findings are rarely used to actually influence scale development, much less to continually improve the quality of the scales. Even worse, virtually all of the findings from different inventories are incommensurate. Test authors are not encouraged to conduct *comparative* validity studies, pitting their instrument against one or more others as predictors of the same set of criterion indices. As a result, neither the science of personality assessment nor its applied practitioners have any information about the *comparative* performance of the different instruments available in the marketplace. There is no Consumers Union for testing our tests.

One basic problem is that scientific goals may become subjugated to commercial interests. I believe that it is time for a change: I envisage an international effort to develop and continually refine a broad-bandwidth personality inventory, whose items are in the public domain, and whose scales can be used for both scientific and commercial purposes. No one investigator alone has access to many diverse criterion settings; but the international scientific community has such access, and by pooling our findings we should be able to devise instruments over the next decade that make our present ones seem like ancient relics.

To get there, we need to start somewhere. To begin, we must agree on the solutions to at least three problems: (1) We need a *taxonomic framework* for organizing the nearly infinite variety of individual differences that might be measured. (2) We need a common *item format*, one that is amenable to faithful translation across diverse languages. And, (3) We need a *mode of communication*--an effective logistical procedure for investigators to easily obtain the items and the findings from previous studies, as well as the data for re-analyses; in addition, we need a way for investigators to add new items to the pool, along with findings about their properties. For the first time, the solutions to all three problems may now be at hand.

A FRAMEWORK FOR ORGANIZING PERSONALITY ATTRIBUTES

While the technology of personality assessment has remained stagnant over the last decades, the fundamental taxonomic problem in personality assessment may be close to a solution. In spite of strong denials by some vocal critics (e.g., Block, 1995), I think that most investigators would agree that the general framework for a comprehensive structure of phenotypic personality attributes seems finally to be visible (Digman, 1990; Goldberg, 1981, 1993b, 1995; John, 1990; Saucier & Goldberg, 1996b). In a variety of Indo-European and other languages, analyses of large samples of trait-descriptive adjectives have generally led to a structural representation--often referred to as the Big-Five factor structure--which seems to

incorporate most phenotypic personality attributes (Goldberg, 1990; Saucier & Goldberg, 1996a).

One way of viewing this model is as a hierarchical structure with the Big-Five factors at or near the top of the hierarchy, below which are located the various lower-level "facets" that are measured by particular narrow-bandwidth personality measures (Goldberg, 1993a). Although there is some agreement in the personality literature about the characteristics of the higher-level factors, there is no such agreement about an optimal set of lower-level facets. For example, there are 45 bipolar dimensions in the AB5C model of the Big Five proposed by Hofstee, De Raad, and Goldberg (1992); there are 30 bipolar dimensions in the Five-Factor model of Costa and McCrae as operationalized in their revised NEO inventory (NEO-PI-R); there are about 30 to 35 facets implied in the scales in Gough's California Psychological Inventory (CPI); and there are the well-known 16 primary factors in the hierarchical structure incorporated by Cattell in his Sixteen Personality Factors Questionnaire (16PF). *Because agreement has not yet been reached on the relative superiority of any one of these competing lower-level structures, it behooves us to incorporate them all in our preliminary inventory, so that they can be compared empirically.*

Although an inventory that includes a systematic set of lower-level facets can easily generate the higher-level Big-Five factors, the reverse is not true. Inventories that incorporate only five dimensions can not provide the specific variance associated with each of the lower-level facets. Because most of the variance in our instruments is specific to each particular trait, inventories that measure only the Big Five will necessarily be less useful than more comprehensive ones in most applied contexts. Indeed, the optimum number of variables to include in regression analyses of individual differences is limited only by considerations of statistical power, and thus of sample size. Recent empirical "demonstrations" of this psychometric principle by Mershon and Gorsuch (1988) and by Ashton, Jackson, Paunonen, Helmes, and Rothstein (1995) are hardly needed, unless one assumes that the only reliable variance in personality measures is that common variance associated with the Big-Five factors.

THE DEVELOPMENT OF A COMMON ITEM FORMAT

One major source of the Big-Five factor structure has been findings from analyses based on the "Lexical Hypothesis"--namely that the most important ways that individuals differ from each other will eventually come to be encoded as single attribute-descriptive terms (e.g., trait adjectives and type nouns) in the lexicons of the world's languages. Although the use of such single terms is necessary for the establishment of an indigenous structure in each new language

under study, these descriptors are not ideal for use as the items in multi-scale personality inventories. There are at least three interrelated problems with their use: (1) First of all, the same property that provides their major strength in fundamental taxonomic studies, namely their relatively finite number within any language, necessarily limits their utility as purveyors of the complex nuances of lower-level personality description; said another way, there are not enough of them--certainly not enough for redundant and thus reliable measurement in all regions of personality space. (2) In addition, trait adjectives and type nouns encode personality traits at an extremely high level of abstractness. Although research by Hampson, John, and Goldberg (1986) demonstrates substantial differences in breadth within the total set of English trait adjectives (e.g., Extraverted versus Talkative, or Reliable versus Punctual), even the most narrow of such terms (e.g., Talkative and Punctual) are still quite abstract. Most test authors prefer items that are more behaviorally and/or contextually specified. (3) Perhaps as a consequence of the abstractness of trait adjectives and type nouns, it is often not possible to find one-to-one translations for them in different languages, even languages as close linguistically as Dutch, German, and English (Hofstee, Kiers, De Raad, Goldberg, & Ostendorf, 1997). Given the desirability of international collaboration in the development of new assessment methods, this is a highly undesirable feature of their use as test stimuli.

Instead, I propose that we begin this project using an item format that is more contextualized and thus longer than trait adjectives, yet is more compact and thus shorter than the items in many modern personality inventories. The Groningen personality team of Hofstee, De Raad, and Hendriks have been the major proponents of this item format, and they have used it to develop an initial pool of 1,311 Dutch items which they hoped might cover many of the facets of the Big-Five factor structure; findings from analyses of 914 of these Dutch items can be found in Hendriks (1997). I worked with the Groningen team to translate most of these items into their English equivalents. From this initial English item pool, I selected about 750 in their original translations, and then added about 500 new English items that have as yet no Dutch translations¹.

The resulting pool of 1,252 English items--which I have dubbed the International Personality Item Pool (IPIP)--has now been administered in three parts to

¹ Of the initial 1,311 Dutch items, about 70% were written to measure the most highly filled facets of the AB5C structure based on Dutch trait adjectives, an additional 10% were derived from Dutch personality-related verbs, and a final 20% were written to represent various facets of the Intellect domain; the 1,311 items were reduced to 914 on the basis of ratings of their Difficulty, Observability, and Social Desirability, as well as their ease of translation into English (Hendriks, 1997). Most of the approximately 500 new English items were targeted at the weakest (least highly filled) facets of the AB5C structure based on an initial set of IPIP items.

participants in an adult community sample². Participants in this sample have also been administered an inventory of 360 trait-descriptive adjectives, which include 100 unipolar markers of the Big-Five factor structure (Goldberg, 1992), as well as an inventory of 525 of the most familiar person-descriptive adjectives in English. In addition, these participants have completed a variety of commercial personality inventories, including the NEO-PI-R, CPI, TCI, HPI, and the 16PF.

A MODE OF COMMUNICATION AMONG INVESTIGATORS

If scientists world-wide are to participate together to construct the next generation of personality inventories, they need an effective method for obtaining previous findings and data, and for adding their own new findings and data. With the rapid development of the World-Wide-Web (WWW) and the associated expansion of File Transfer Protocol (FTP) sites, it is now possible to access scientific data banks easily and economically through electronic means. Indeed, it is my prediction that over the next few years the phrase "public domain" will come to mean "accessible via the World-Wide-Web." To start this process, I have set up a WWW site (<http://ipip.ori.org/ipip/>), which includes the tables in this chapter, so as to provide easy access to this information from computer terminals throughout the world.

THE DEVELOPMENT OF SOME PRELIMINARY IPIP SCALES

Table 1 presents some of the characteristics of preliminary IPIP scales targeted at the 45 bipolar *AB5C* facets. For each of these new scales, Table 1 lists the number of items keyed positively, the number keyed negatively, and the total number of items in the scale; in addition, the mean item intercorrelation is provided, along with the Coefficient Alpha reliability estimate. My aspiration is to develop roughly 10-item scales, with Alphas that range from .70 to .90, and average .80. Of the 45 preliminary scales, 43 have Alphas of .70 or above, and 18 have reliabilities of .80 or above; the mean of all 45 scales is .78. The items included in each of the 45 preliminary *AB5C* scales are listed in the Appendix.

² The sample initially included approximately 500 men and 500 women, aged 18 to 85, recruited from the Eugene-Springfield [Oregon, USA] community, all of whom had agreed to work with the ORI personality team for at least five years; participants are periodically mailed questionnaires and inventories, which they are paid to complete. Over the four years since the sample was initially selected, approximately 20% of the participants have discontinued their participation; as a consequence, the number of participants for any single survey instrument (other than the first) varies from around 500 to around 800.

Community

Table 1. Characteristics of the 45 preliminary IPIP scales targeted at the AB5C facets

AB5C facet	Provisional label	No. of items		Mean item r	Coef. Alpha
		+	-		
Factor I					
I+/I+ vs. I-/I-	Gregariousness	4	6	.34	.83
I+/II+ vs. I-/II-	Friendliness	5	5	.37	.85
I+/III+ vs. I-/III-	Assertiveness	9	3	.20	.75
I+/IV+ vs. I-/IV-	Poise	5	5	.31	.82
I+/V+ vs. I-/V-	Leadership	5	5	.31	.82
I+/II- vs. I-/II+ ¹	Provocativeness	8	3	.19	.72
I+/III- vs. I-/III+	Self-Disclosure	8	2	.26	.78
I+/IV- vs. I-/IV+	Talkativeness	8	2	.35	.84
I+/V- vs. I-/V+ ¹	Sociability	3	7	.16	.66
Factor II					
II+/II+ vs. II-/II-	Understanding	5	5	.30	.81
II+/I+ vs. II-/I-	Warmth	9	2	.33	.84
II+/III+ vs. II-/III- ¹	Morality	5	7	.18	.73
II+/IV+ vs. II-/IV-	Pleasantness	6	6	.22	.76
II+/V+ vs. II-/V- ¹	Empathy	5	4	.20	.70
II+/I- vs. II-/I+ ¹	Cooperation	2	10	.18	.73
II+/III- vs. II-/III+ ¹	Sympathy	6	6	.20	.74
II+/IV- vs. II-/IV+ ¹	Tenderness	9	4	.18	.74
II+/V- vs. II-/V+	Nurturance	6	7	.16	.71
Factor III					
III+/III+ vs. III-/III-	Conscientiousness	6	7	.19	.75
III+/I+ vs. III-/I-	Efficiency	5	6	.30	.83
III+/II+ vs. III-/II- ¹	Dutifulness	6	7	.21	.78
III+/IV+ vs. III-/IV-	Purposefulness	5	7	.27	.81
III+/V+ vs. III-/V-	Organization	9	3	.23	.78
III+/I- vs. III-/I+ ¹	Cautiousness	5	7	.21	.77
III+/II- vs. III-/II+ ¹	Rationality	8	6	.13	.67
III+/IV- vs. III-/IV+	Perfectionism	7	2	.26	.76
III+/V- vs. III-/V+	Orderliness	7	3	.27	.78
Factor IV					
IV+/IV+ vs. IV-/IV-	Stability	5	5	.37	.86
IV+/I+ vs. IV-/I-	Happiness	5	5	.34	.84
IV+/II+ vs. IV-/II-	Calmness	4	6	.33	.83
IV+/III+ vs. IV-/III-	Moderation	4	6	.24	.76
IV+/V+ vs. IV-/V-	Toughness	4	8	.29	.84
IV+/I- vs. IV-/I+ ¹	Impulse Control	2	9	.24	.78
IV+/II- vs. IV-/II+ ¹	Imperturbability	2	7	.37	.84
IV+/III- vs. IV-/III+ ¹	Cool-headedness	0	10	.21	.73
IV+/V- vs. IV-/V+ ¹	Tranquillity	7	4	.22	.76
Factor V					
V+/V+ vs. V-/V-	Intellect	6	5	.27	.81
V+/I+ vs. V-/I-	Ingenuity	6	3	.37	.84
V+/II+ vs. V-/II- ¹	Reflection	8	2	.26	.75
V+/III+ vs. V-/III- ¹	Competence	8	0	.26	.74
V+/IV+ vs. V-/IV-	Quickness	7	3	.37	.84
V+/I- vs. V-/I+ ¹	Introspection	10	2	.18	.71
V+/II- vs. V-/II+ ¹	Creativity	5	5	.30	.81
V+/III- vs. V-/III+ ¹	Imagination	5	5	.27	.78
V+/IV- vs. V-/IV+ ¹	Depth	7	2	.27	.77
Mean				.26	.78

Note. All analyses are based on the responses of 501 adult subjects from the Eugene-Springfield Community Sample.

¹ These scales have been augmented with items from other AB5C facets.

Table 2 provides a comparison between some characteristics of the 30 facet scales from the *NEO-PI-R* (Costa & McCrae, 1992) and 30 similar constructs measured in the IPIP pool. Of the two sets of 30 constructs, only 13 are labeled identically; differences in scale labels are most pronounced in the Openness domain, where all six facets have different labels. On average, the IPIP scales include 10 items, with about half keyed in each direction. The average of the Coefficient Alpha values is a bit higher for the IPIP scales (.80) than for the NEO scales (.75). The average correlation between corresponding scales in the two sets is .73, which translates into a correlation of .94 when corrected for attenuation due to the unreliabilities of the two scales in each pair.

Table 3 presents a comparison between the scales on the *16PF* (Conn & Rieke, 1994) and 16 new IPIP scales constructed to measure the same constructs; the scales are ordered by their associations with the Big-Five factor structure. All but one of the IPIP scales include 10 items, about half keyed in each direction; of the 16PF scales, which are also balanced in their keying, nine include 10 items, five include 11 items, and one each includes 14 and 15 items. Cattell has preferred labeling his scales with letters (A through Q4), but the manual also lists the short verbal labels presented in Table 3; only two of these trait labels are the same as those of the corresponding IPIP scales. Because the average item intercorrelations are slightly higher for the IPIP than for the 16PF scales (.29 versus .21), the average IPIP Coefficient Alpha is also somewhat higher (.80 versus .74). The average correlation between corresponding scales in the two sets is .66, which translates into a corrected correlation of .86.

An analogous table focused on the lower-level constructs in Cloninger's Temperament and Character Inventory (TCI) is available from the author. Of the 31 TCI constructs, 30 are included in the corresponding IPIP scale set--11 with Coefficient Alpha reliability estimates of .80 or more, 15 others with reliabilities of .70 or more, and 4 others with reliabilities just slightly below .70. The average Alpha values of the two sets of scales are quite similar (.77 and .78); on average, the pairs of scales correlated .64, which corresponds to a correlation of .83 when corrected for the scale reliabilities.

Also available from the author is a similar table comparing 33 of the scales in the *CPI* (Gough, 1996) with the corresponding 33 preliminary IPIP scales targeted at those constructs. The original 33 CPI scales vary in length from 28 to 70 items, averaging 38; in contrast, the IPIP scales include about 10 items, usually with half keyed in each direction. The IPIP scales are far more homogeneous than the CPI scales, with mean item intercorrelations that average .26, as compared to a mere .08 for the CPI scales. As a consequence, although the IPIP scales are much shorter than the CPI scales, both sets of scales have similar Alpha coefficients; for the original CPI scales the Alphas vary from .53 to .88, averaging .74, whereas for the IPIP scales the values range from .62 to .87, averaging .76. The

Table 2. A comparison between the 30 facet scales in Costa and McCrae's NEO Personality Inventory (NEO-PI-R) and the corresponding 30 preliminary IPIP scales

Scale names		No. of items		Mean item r		Coef. Alpha		Correlation
IPIP	NEO	IPIP	NEO	IPIP	NEO	IPIP	NEO	IPIP vs. NEO
<i>Neuroticism</i>								
Anxiety	Anxiety (N1)	5	4	.32	.37	.83	.83	.75 [.90]
Anger	Angry Hostility (N2)	5	5	.42	.34	.88	.80	.76 [.91]
Depression	Depression (N3)	7	6	.43	.41	.88	.85	.80 [.92]
Self-consciousness	Self-Consciousness (N4)	6	5	.28	.26	.80	.74	.72 [.94]
Immoderation	Impulsiveness (N5)	5	4	.25	.24	.77	.72	.73 [.98]
Vulnerability	Vulnerability (N6)	5	3	.32	.35	.82	.79	.77 [.96]
<i>Extraversion</i>								
Friendliness	Warmth (E1)	5	6	.41	.33	.87	.80	.76 [.91]
Gregariousness	Gregariousness (E2)	5	4	.28	.34	.79	.80	.78 [.98]
Assertiveness	Assertiveness (E3)	5	4	.34	.33	.84	.80	.81 [.99]
Activity-level	Activity (E4)	5	5	.19	.25	.71	.72	.70 [.98]
Excitement-seeking	Excitement-Seeking (E5)	8	6	.28	.19	.78	.64	.67 [.95]
Cheerfulness	Positive Emotions (E6)	8	4	.30	.36	.81	.81	.77 [.95]
<i>Openness to Experience</i>								
Imagination	Fantasy (O1)	6	3	.32	.35	.83	.82	.74 [.90]
Artistic interests	Aesthetics (O2)	5	5	.36	.40	.84	.84	.80 [.95]
Emotionality	Feelings (O3)	5	5	.29	.28	.81	.75	.70 [.90]
Adventurousness	Actions (O4)	4	3	.24	.18	.77	.64	.71 [.99]
Intellect	Ideas (O5)	5	5	.38	.38	.86	.82	.80 [.95]
Liberalism	Values (O6)	3	3	.38	.30	.86	.78	.70 [.86]

Table 2 (cont.)

Scale names	Scale names	No. of items			Mean item r			Coef. Alpha			Correlation
		IPIP	NEO	+	IPIP	NEO	+	IPIP	NEO	+	
IPIP	NEO	+	-	-	+	-	+	-	-	+	IPIP vs. NEO
Agreeableness											
Trust	Trust (A1)	6	4	5	.32	.42	.82	.84	.79	.79	[.95]
Morality	Straightforwardness (A2)	2	8	3	.25	.26	.75	.74	.64	.64	[.86]
Altruism	Altruism (A3)	5	5	5	.25	.26	.77	.72	.67	.67	[.90]
Cooperation	Compliance (A4)	3	7	3	.22	.26	.73	.73	.71	.71	[.97]
Modesty	Modesty (A5)	4	6	4	.25	.27	.77	.75	.72	.72	[.95]
Sympathy	Tender-Mindedness (A6)	4	6	6	.23	.17	.75	.61	.61	.61	[.90]
Conscientiousness											
Self-efficacy	Competence (C1)	6	4	5	.27	.25	.78	.70	.66	.66	[.89]
Orderliness	Order (C2)	5	5	3	.33	.28	.82	.74	.77	.77	[.99]
Dutifulness	Dutifulness (C3)	5	5	6	.20	.23	.71	.67	.60	.60	[.87]
Achievement-striving	Achievement Striving (C4)	7	3	5	.27	.21	.78	.67	.70	.70	[.97]
Self-discipline	Self-Discipline (C5)	5	5	4	.36	.34	.85	.80	.76	.76	[.92]
Cautiousness	Deliberation (C6)	3	7	5	.24	.23	.76	.70	.69	.69	[.95]
Mean		5	5	4	.30	.29	.80	.75	.73	.73	[.94]

Note. All analyses are based on the responses of 501 adult subjects from the Eugene-Springfield Community Sample.
 [] Values in brackets are the correlations between the corresponding NEO and IPIP scales when corrected for attenuation due to scale unreliability.

Table 3. A comparison between the scales in Cattell's 16 Personality Factors Questionnaire (16PF) and the corresponding 16 preliminary IPIP scales

Scale names	Scale names	No. of items	Mean item r		Coef. Alpha		Correlation
			IPIP	16PF	IPIP	16PF	
IPIP	16PF						
		+					
<i>Factor I</i>							
Friendliness	Social Boldness (Factor H)	7 3 11	.30	.19	.80	.72	.64 [.84]
Openness	Privateness, reflected (N)	5 5 10	.39	.28	.86	.80	.74 [.89]
Assertiveness	Dominance (E)	6 4 10	.30	.18	.81	.68	.67 [.90]
Gregariousness	Liveliness (F)	6 4 10	.27	.18	.78	.69	.68 [.93]
Sociability	Self-reliance, reflected (Q2)	3 7 10	.22	.24	.73	.76	.64 [.86]
<i>Factor II</i>							
Warmth	Warmth (A)	7 3 11	.30	.19	.80	.72	.64 [.84]
Sensitivity	Sensitivity (I)	5 5 11	.22	.24	.73	.78	.68 [.93]
<i>Factor III</i>							
Orderliness	Perfectionism (Q3)	5 5 10	.31	.22	.81	.74	.77 [.99]
Dutifulness	Rule-Consciousness (G)	6 4 11	.35	.22	.84	.75	.69 [.87]
<i>Factor IV</i>							
Security	Emotional Stability (C)	5 5 10	.37	.25	.85	.76	.75 [.93]
Self-esteem	Apprehension, reflected (O)	3 7 10	.29	.24	.80	.76	.72 [.92]
Calmness	Tension, reflected (Q4)	3 7 10	.24	.24	.76	.76	.59 [.78]
Trust	Vigilance, reflected (L)	4 6 10	.29	.21	.80	.73	.57 [.75]
<i>Factor V</i>							
Complexity	Openness to Change (Q1)	5 5 14	.32	.17	.82	.74	.69 [.89]
Intellect	Reasoning (B)	8 5 15	.20	.15	.76	.73	.51 [.69]
Imagination	Abstractedness (M)	7 3 11	.28	.23	.80	.76	.65 [.83]
<i>Mean</i>		5 5 11	.29	.21	.80	.74	.66 [.86]

Note. All analyses are based on the responses of 501 adult subjects from the Eugene-Springfield Community Sample.

[] Values in brackets are the correlations between the corresponding 16PF and IPIP scales when corrected for scale unreliability.

average correlation between the corresponding CPI and IPIP scales is .62, which translates into a corrected correlation of .84.

Table 4 provides a summary of the reliability estimates for the five sets of preliminary IPIP scales, plus a comparison with the original 16PF, CPI, TCI, and NEO scales. Given the large size of this subject sample, one would not expect much attenuation of the IPIP reliabilities in new samples³. As a consequence, the most reasonable conclusion one can make from these findings is that the average reliability coefficients for the four sets of preliminary IPIP scales are quite similar, and all of them are at least as high as the values for the average NEO, CPI, TCI, and 16PF scales. Hopefully, with the help of other investigators, even more reliable IPIP scales will become available over the next decade.

Table 4. The mean reliabilities of the preliminary IPIP scales targeted at the AB5C, NEO, CPI, TCI, and 16PF facets, and a comparison with the original inventory scales

	Mean of the mean item correlations		Mean Coefficient Alpha		Mean correlation
	IPIP	Original	IPIP	Original	IPIP vs. original
AB5C	.26	—	.78	—	— —
NEO	.30	.29	.80	.75	.73 [.94]
CPI	.26	.08	.76	.74	.62 [.84]
TCI	.28	.34	.77	.78	.64 [.83]
16PF	.29	.21	.80	.74	.66 [.86]

Note. Most of the IPIP scales include 10 items; all of the NEO scales include 8 items, the average CPI scale includes 38 items, the average TCI scale includes 8 items, and the average 16PF scale includes 11 items. All analyses are based on the responses of 501 adult subjects from the Eugene-Springfield Community Sample.

IPIP = International Personality Item Pool.

[] Values in brackets are the correlations between the corresponding scales when corrected for attenuation due to scale unreliability.

³ Although generally one should expect lower reliabilities in new samples than in the initial derivation sample, this may not occur for the IPIP scales, given that the IPIP items were administered in three separate questionnaires over a period of nearly three years. One of the first tasks for future investigators must be to test the robustness of these scales in new samples.

A FIRST COMPARATIVE VALIDITY STUDY

Most of us would agree that the single most important question to be asked of new personality measures concerns their utility as predictors of diverse and important human outcomes. Given the great scientific interest of late in behavioral medicine and health psychology, I used as initial criterion variables some indices of health-related behaviors and practices. A factor analysis of the 39 items in a reasonably comprehensive Health Activities Questionnaire (HAQ: Vickers, Conway, & Hervig, 1990) yielded three orthogonal factors: (1) *Risk-Avoidance* (12 questions such as "I carefully obey traffic rules," "I do not drink," "I avoid high-crime areas" versus "I take chances crossing the street," "I speed while driving," "I drive after drinking"); (2) *Good Health Practices* (12 items such as "I exercise to stay healthy," "I eat a balanced diet," "I see a dentist for regular checkups," "I watch my weight," "I don't smoke"); and (3) *General Health Concerns* (15 items such as "I gather information on things that affect my health," "I avoid areas with high pollution," "I take health food supplements", "I watch for possible signs of major health problems").

The orthogonal factor scores on each of these three health factors were used as criterion variables, along with the factor scores on the first unrotated component from an analysis of all 39 HAQ items (*Total Health-Related Practices*). In stepwise regression analyses, the scores from each of the five IPIP preliminary scale sets and each of the four original 16PF, CPI, TCI, and NEO scale sets were included separately as the predictor variables. To control for any effects of subject sex, age, and educational level, all of the analyses were repeated using two additional procedures: (a) Using hierarchical regression procedures, the three demographic variables were entered first, followed by the personality scale scores, and (b) the residual factor scores from the HAQ were included as criterion variables after the effects of the three demographic variables were partialled out.

The validity findings are easy to summarize: At each step in the regression analyses, for each of the four criterion variables, and for each of the three types of regression procedures, the preliminary IPIP scale sets were typically more highly predictive than the original inventory scale sets. Of the four criteria, the most predictable were Risk Avoidance and Total Health-Related Practices, and for these two criteria the IPIP scales were always more predictive than the original inventory scales. For example, in the standard regression analyses predicting Risk Avoidance from the TCI and corresponding IPIP scale sets, the multiple correlations at each of the first four steps, for the TCI scales versus the IPIP scales were: (1) .42/.57; (2) .49/.63; (3) .52/.64; and (4) .54/.65. The corresponding values for predicting the same criterion using hierarchical regression analyses were (1) .49/.60; (2) .54/.65; (3) .55/.66; and (4) .57/.67.

In summary, then, the initial evidence regarding the reliability and predictive utility of the preliminary IPIP scales is quite favorable. With the help of other investigators throughout the world, we should be able to refine these scales, and develop new ones, so as to provide substantially improved measures of important personality attributes. As a reader of this chapter, will you join me in this scientific adventure?

Acknowledgements

This research program has been funded by Grant MH-49227 from the National Institute of Mental Health, U. S. Public Health Service. The author is enormously indebted to Willem K. B. Hofstee, Boele De Raad, and Jolijn Hendriks for their development of the initial Dutch item pool from which many of the IPIP items were translated.

References

- Ashton, M. C., Jackson, D. N., Paunonen, S. V., Helmes, E., & Rothstein, M. G. (1995). The criterion validity of broad factor scales versus specific facet scales. *Journal of Research in Personality*, 29, 432-442.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187-215.
- Conn, S. R., & Rieke, M. L. (1994). *The 16PF fifth edition technical manual*. Champaign, IL: Institute for Personality and Ability Testing.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology: Vol. 41* (pp. 417-440). Palo Alto, CA: Annual Reviews.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology: Vol. 2* (pp. 141-165). Beverly Hills, CA: Sage.
- Goldberg, L. R. (1990). An alternative "Description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Goldberg, L. R. (1993a). The structure of personality traits: Vertical and horizontal aspects. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 169-188). Washington, DC: American Psychological Association.

- Goldberg, L. R. (1993b). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26-34.
- Goldberg, L. R. (1995). What the hell took so long? Donald Fiske and the Big-Five factor structure. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research, methods, and theory: A Festschrift honoring Donald W. Fiske* (pp. 29-43). Hillsdale, NJ: Erlbaum.
- Gough, H. G. (1996). *CPI Manual: Third Edition*. Palo Alto, CA: Consulting Psychologists Press.
- Hampson, S. E., John, O. P., & Goldberg, L. R. (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology*, *51*, 37-54.
- Hendriks, A. A. J. (1997). *The construction of the Five-Factor Personality Inventory (FFPI)*. Groningen, The Netherlands: Rijksuniversiteit Groningen.
- Hofstee, W. K. B., De Raad, B., & Goldberg, L. R. (1992). Integration of the Big-Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*, 146-163.
- Hofstee, W. K. B., Kiers, H. A. L., De Raad, B., Goldberg, L. R., & Ostendorf, F. (1997). Comparison of Big-Five structures of personality traits in Dutch, English, and German. *European Journal of Personality*, *11*, 15-31.
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford.
- Mershon, B., & Gorsuch, R. L. (1988). Number of factors in the personality sphere: Does increase in factors increase predictability of real-life criteria? *Journal of Personality and Social Psychology*, *55*, 675-680.
- Saucier, G., & Goldberg, L. R. (1996a). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, *10*, 61-77.
- Saucier, G., & Goldberg, L. R. (1996b). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21-50). New York: Guilford.
- Vickers, R. R., Jr., Conway, T. L., & Hervig, L. K. (1990). Demonstration of replicable dimensions of health behaviors. *Preventive Medicine*, *19*, 377-401.

Scale item
Scoring key
prelim

Appendix. The 45 preliminary scales for the AB5C Facets

Facet	Label	Alpha	Facet	Label	Alpha
I+/I+ vs I-/I-	Gregariousness	(.83)	I+/IV+ vs I-/IV-	Poise	(.82)
Am the life of the party. Talk to a lot of different people at parties. Start conversations. Love large parties. Don't talk a lot. Keep in the background. Am quiet around strangers. Don't like to draw attention to myself. Bottle up my feelings. Keep my thoughts to myself.			Feel comfortable around people. Am comfortable in unfamiliar situations. Have a lot of fun. Am not embarrassed easily. Love life. Often feel uncomfortable around others. Find it difficult to approach others. Retreat from others. Give up easily. Only feel comfortable with friends.		
I+/II+ vs I-/II-	Friendliness	(.85)	I+/V+ vs I-/V-	Leadership	(.82)
Make friends easily. Am open about my feelings. Act comfortably with others. Radiate joy. Warm up quickly to others. Am hard to get to know. Am a very private person. Avoid contacts with others. Keep others at a distance. Reveal little about myself.			Take charge. Know how to captivate people. Express myself easily. Am the first to act. Never at a loss for words. Have little to say. Have difficulty expressing my feelings. Wait for others to lead the way. Am afraid to draw attention to myself. Let others make the decisions.		
I+/III+ vs I-/III-	Assertiveness	(.75)	I+/II- vs I-/II+	Provocativeness	(.72)
Automatically take charge. Can easily push myself forward. Try to lead others. Turn plans into actions. Stick up for myself. Am always busy. Come up with a solution right away. Do a lot in my spare time. Know what I want. Let myself be pushed around. Am not highly motivated to succeed. Need a lot of time to do things.			Dare to say anything. Am not afraid of providing criticism. Boast about my virtues. [Know no limits.] [Know how to get around the rules.] [Can take strong measures.] [Don't mind being the center of attention.] [Make demands on others.] Can't stand confrontations. Wait for my turn. [Hate to seem pushy.]		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
I+/III- vs I-/III+ Self-Disclosure		(.78)	II+/II+ vs II-/II- Understanding		(.81)
Act wild and crazy.			Sympathize with others' feelings.		
Am open about myself to others.			Respect others' feelings.		
Let myself go.			Take others' interests into account.		
Disclose my intimate thoughts.			Like to be of service to others.		
Laugh my way through life.			Appreciate the viewpoints of others.		
Express childlike joy.			Feel little concern for others.		
Joke around a lot.			Am not interested in other people's problems.		
Like to amuse others.			Am indifferent to the feelings of others.		
Seldom joke around.			Take no time for others.		
Prefer to deal with strangers in a formal manner.			Can't be bothered with others' needs.		
I+/IV- vs I-/IV+ Talkativeness		(.84)	II+/I+ vs II-/I- Warmth		(.84)
Do most of the talking.			Am interested in people.		
Talk too much.			Make people feel at ease.		
Speak loudly.			Know how to comfort others.		
Make myself the center of attention.			Inquire about others' well-being.		
Like to attract attention.			Take time out for others.		
Never stop talking.			Make people feel welcome.		
Make a lot of noise.			Show my gratitude.		
Demand to be the center of interest.			Make others feel good.		
Speak softly.			Feel others' emotions.		
Dislike talking about myself.			Am not really interested in others.		
I+/V- vs I-/V+ Sociability		(.66)	II+/III+ vs II-/III- Morality		(.73)
Can't do without the company of others.			Would never cheat on my taxes.		
[Enjoy being part of a loud crowd.]			Respect the privacy of others.		
[Enjoy being on the go.]			Like harmony in my life.		
Like to be alone.			[Try to follow the rules.]		
Seek quiet.			[Respect authority.]		
Enjoy silence.			Don't care about rules.		
[Don't like crowded events.]			Turn my back on others.		
[Dislike neighbors living too close.]			Only talk about my own interests.		
[Amuse myself easily.]			Overestimate my achievements.		
[Go my own way.]			Scheme against others.		
			[Act at the expense of others.]		
			[Break rules.]		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
II+/IV+ vs II-/IV-	Pleasantness	(.76)	II+/III- vs II-/III+	Sympathy	(.74)
Am easy to satisfy.			Am concerned about others.		
Trust what people say.			Am deeply moved by others' misfortunes.		
Have a good word for everyone.			Feel sympathy for those who are worse off than myself.		
Am on good terms with nearly everyone.			[Take an interest in other people's lives.]		
Trust others.			[Like to do things for others.]		
Respect others.			[Reassure others.]		
Am hard to satisfy.			Demand a lot from others.		
Am quick to judge others.			Don't fall for sob-stories.		
Insult people.			Listen to my brain rather than my heart.		
Find it hard to forgive others.			Tend to dislike soft-hearted people.		
Contradict others.			Try not to think about the needy.		
Criticize others' shortcomings.			Look down on any weakness.		
			Believe people should fend for themselves.		
II+/V+ vs II-/V-	Empathy	(.70)	II+/IV- vs II-/IV+	Tenderness	(.74)
Anticipate the needs of others.			Suffer from others' sorrows.		
Sense others' wishes.			Listen to my heart rather than my brain.		
[Love to reflect on things.]			Love children's movies.		
[Try to stay in touch with myself.]			Want to please others.		
[Work on improving myself.]			Remember my friends' birthdays.		
Pretend to be concerned for others.			Want to mean something to others.		
[Don't have a soft side.]			Cherish mementos.		
[Treat people as inferiors.]			[Believe crying helps me feel better.]		
[Am not in touch with my feelings.]			[Show my feelings.]		
			Don't understand people who get emotional.		
			[Don't get excited about things.]		
			[Don't call people just to talk.]		
			[Don't care about dressing nicely.]		
II+/I- vs II-/I+	Cooperation	(.73)	II+/V- vs II-/V+	Nurturance	(.71)
Value cooperation over competition.			Have a soft heart.		
Listen to my conscience.			Go out of my way for others.		
Impose my will on others.			Think of others first.		
Love a good fight.			Will do anything for others.		
Seek conflict.			Like to please others.		
Think too highly of myself.			Wouldn't harm a fly.		
Tell tall stories about myself.			Make enemies.		
Play tricks on others.			Oppose authority.		
Enjoy crude jokes.			Believe that I am better than others.		
[Comment loudly about others.]			Seek danger.		
[Enjoy being reckless.]			Put people under pressure.		
[Do dangerous things.]			Try to outdo others.		
			Believe only in myself.		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
III+/III+ vs III-/III- Conscientiousness (.75)			III+/IV+ vs III-/IV- Purposefulness (.81)		
<p>Accomplish my work on time. Do things according to a plan. Am careful to avoid making mistakes. Keep my checkbook balanced. Like to plan ahead. Return borrowed items.</p> <p>Often forget to put things back in their proper place. Neglect my duties. Take tasks too lightly. Leave my work undone. Do not plan ahead. Put off unpleasant tasks. Am often late to work.</p>			<p>Am always prepared. Carry out my plans. Get to work at once. Am not easily distracted. Handle tasks smoothly.</p> <p>Make a mess of things. Am easily distracted. Mess things up. Shirk my duties. Don't see things through. Do things at the last minute. Can't make up my mind.</p>		
III+/I+ vs III-/I- Efficiency (.83)			III+/V+ vs III-/V- Organization (.78)		
<p>Am exacting in my work. Make plans and stick to them. Get chores done right away. Follow through with my plans. Finish what I start.</p> <p>Waste my time. Find it difficult to get down to work. Postpone decisions. Have difficulty starting tasks. Need a push to get started. Frequently forget to do things.</p>			<p>Pay attention to details. Complete tasks successfully. Have an eye for detail. Demand quality. Set high standards for myself and others. Make well-considered decisions. Follow through on my commitments. Detect mistakes. Think ahead.</p> <p>Seldom notice details. Put little time and effort into my work. Don't pay attention.</p>		
III+/II+ vs III-/II- Dutifulness (.78)			III+/I- vs III-/I+ Cautiousness (.77)		
<p>Follow directions. Keep myself well-groomed. Check over my work. Behave properly. [Stick to the rules.] [Appreciate good manners.]</p> <p>Do improper things. Disregard rules. Do the opposite of what is asked. Pay no attention to my appearance. [Don't think laws apply to me.] [Make rash decisions.] [Say inappropriate things.]</p>			<p>Purchase only practical things. Tend to dislike impulsive people. Take precautions. [Never splurge.] [Never spend more than I can afford.]</p> <p>Do crazy things. Often make last-minute plans. Am easily talked into doing silly things. Laugh at the slightest provocation. Like to laugh out loud. [Like to act on a whim.] [Jump into things without thinking.]</p>		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
III+/II- vs III-/II+	Rationality	(.67)	IV+/IV+ vs IV-/IV-	Stability	(.86)
Do things in a logical order. Come straight to the point. Believe in a logical answer for everything. Get a head start on others. Dislike imperfect work. Believe in an eye for an eye. [Have no sympathy for criminals.] [Reason logically.] Sympathize with the homeless. Am not as strict as I should be. Let people pull my leg. [Do things in a half-way manner.] [Let my attention wander off.] [Believe that criminals should receive help rather than punishment.]			Seldom get mad. Am not easily bothered by things. Am not easily frustrated. Seldom take offense. Keep my cool. Get stressed out easily. Get upset easily. Am easily disturbed. Change my mood a lot. Get caught up in my problems.		
III+/IV- vs III-/IV+	Perfectionism	(.76)	IV+/I+ vs IV-/I-	Happiness	(.84)
Continue until everything is perfect. Want every detail taken care of. Want everything to be "just right." Want things to proceed according to plan. Demand perfection in others. Keep a sharp eye on others' work. Expect dedicated work from others. Am not bothered by messy people. Am not bothered by disorder.			Seldom feel blue. Feel comfortable with myself. Adapt easily to new situations. Look at the bright side of life. Am sure of my ground. Often feel blue. Worry about things. Feel threatened easily. Dislike myself. Am filled with doubts about things.		
III+/V- vs III-/V+	Orderliness	(.78)	IV+/II+ vs IV-/II-	Calmness	(.83)
Like order. Follow a schedule. Work according to a routine. Like to tidy up. Do things by the book. Take good care of my belongings. See that rules are observed. Leave my belongings around. Leave a mess in my room. Dislike routine.			Rarely get irritated. Am not easily annoyed. Take things as they come. Accept people as they are. Get angry easily. Am often in a bad mood. Get furious. Snap at people. Lose my temper. Have days when I'm mad at the world.		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
IV+/III+ vs IV-/III- Moderation (.76)			IV+/II- vs IV-/II+ Imperturbability (.84)		
Remain calm under pressure.			Seldom get emotional.		
Easily resist temptations.			Am not easily affected by my emotions.		
Rarely overindulge.			Get overwhelmed by emotions.		
Am able to control my cravings.			Cry easily.		
Am guided by my moods.			Burst into tears.		
Am not sure where my life is going.			Am easily moved to tears.		
Don't know why I do some of the things I do.			Cry during movies.		
Get out of control.			Wear my heart on my sleeve.		
Can't concentrate.			Have crying fits.		
Do things I later regret.					
IV+/V+ vs IV-/V- Toughness (.84)			IV+/III- vs IV-/III+ Cool-headedness (.73)		
Am calm even in tense situations.			(No positive items.)		
Don't lose my head.			Want everything to add up perfectly.		
Know how to cope.			Demand obedience.		
Can stand criticism.			Place demands on myself.		
Take offense easily.			Keep up an appearance.		
Panic easily.			[Love order and regularity.]		
Am easily hurt.			[Am attached to conventional ways.]		
Am easily offended.			[Want things done my way.]		
Feel crushed by setbacks.			[Am a creature of habit.]		
Become overwhelmed by events.			[Try to impress others.]		
Am easily frightened.			[Can't stand being contradicted.]		
Am easily confused.			[Want to be told I am right.]		
IV+/I- vs IV-/I+ Impulse Control (.78)			IV+/V- vs IV-/V+ Tranquility (.76)		
Keep my emotions under control.			Experience very few emotional highs and lows.		
Let others finish what they are saying.			Tend to feel the same every day.		
Demand attention.			Am always in the same mood.		
React intensely.			Rarely notice my emotional reactions.		
Talk even when I know I shouldn't.			[Am relaxed most of the time.]		
Often make a fuss.			[Am not easily stirred.]		
Shoot my mouth off.			[Am not disturbed by events.]		
Am easily excited.			Experience my emotions intensely.		
Blurt out whatever comes into my mind.			[Have frequent mood swings.]		
Barge in on conversations.			[Am swayed by my emotions.]		
Like to gossip.			[Can be stirred up easily.]		

Appendix (cont.)

Facet	Label	Alpha	Facet	Label	Alpha
V+/V+ vs V-/V-	Intellect	(.81)	V+/III+ vs V-/III-	Competence	(.74)
Have a rich vocabulary. Use difficult words. Make insightful remarks. Show a mastery of language. Enjoy thinking about things. Try to understand myself. Am not interested in abstract ideas. Will not probe deeply into a subject. Have a poor vocabulary. Dislike learning. Skip difficult words while reading.			Learn quickly. Use my brain. Excel in what I do. Look at the facts. Meet challenges. Seek explanations of things. Need things explained only once. [Know how to apply my knowledge.] (No negative items.)		
V+/I+ vs V-/I-	Ingenuity	(.84)	V+/IV+ vs V-/IV-	Quickness	(.84)
Am full of ideas. Have excellent ideas. Carry the conversation to a higher level. Come up with bold plans. Quickly think up new ideas. Am good at many things. Do not have a good imagination. Have difficulty imagining things. Can't come up with new ideas.			Can handle complex problems. Am quick to understand things. Catch on to things quickly. Love to read challenging material. Am able to find out things by myself. Can handle a lot of information. Quickly get the idea of things. Avoid difficult reading material. Try to avoid complex people. Don't understand things.		
V+/II+ vs V-/II-	Reflection	(.75)	V+/I- vs V-/I+	Introspection	(.71)
See beauty in things that others might not notice. Take time to reflect on things. Make beautiful things. Enjoy discussing movies and books with others. Enjoy the beauty of nature. [Like music.] [Love flowers.] [Love beautiful things.] Do not like concerts. [Do not enjoy watching dance performances.]			Spend time reflecting on things. Enjoy spending time by myself. Live in a world of my own. Enjoy my privacy. Don't mind eating alone. Do things at my own pace. [Enjoy contemplation.] [Prefer to be alone.] [Have a point of view all my own.] [Want to be left alone.] Can't stand being alone. Don't like to ponder over things.		

Appendix (cont.)

Facet	Label	Alpha
V+/II- vs V-/II+ Creativity		(.81)
Like to solve complex problems. Ask questions that nobody else does. Know the answers to many questions. Challenge others' points of view. Can easily link facts together. Have difficulty understanding abstract ideas. Avoid philosophical discussions. Am not interested in theoretical discussions. Consider myself an average person. Am not interested in speculating about things.		
V+/III- vs V-/III+ Imagination		(.78)
Have a vivid imagination. Prefer variety to routine. Believe in the importance of art. Need a creative outlet. Enjoy wild flights of fantasy. Do not like art. Do not enjoy going to art museums. Do not like poetry. Seldom get lost in thought. Seldom daydream.		
V+/IV- vs V-/IV+ Depth		(.77)
Look for hidden meanings in things. Like to get lost in thought. Think deeply about things. Need to understand my motives. [Tend to analyze things.] [Tend to think about something for hours.] [Enjoy examining myself and my life.] [Rarely look for a deeper meaning in things.] [Never challenge things.]		

Note. The IPIP items were administered with the following response options: (1) Very inaccurate, (2) Moderately inaccurate, (3) Neither inaccurate nor accurate, (4) Moderately accurate, and (5) Very accurate. Values in parentheses are Coefficient Alpha reliability estimates.

[] Items in brackets come from other AB5C facets.