CHAPTER 4

Paradigm Shift to the Integrative Big Five Trait Taxonomy

History, Measurement, and Conceptual Issues

Oliver P. John
Laura P. Naumann
Christopher J. Soto

Since the first version of this chapter (John, 1990) was completed in the late 1980s, the field of personality trait research has changed dramatically. At that time, the Big Five personality dimensions, now seemingly ubiquitous, were hardly known. Researchers, as well as practitioners in the field of personality assessment, were faced with a bewildering array of personality scales from which to choose, with little guidance and no organizing theory or framework at hand. What made matters worse was that scales with the same name might measure concepts that were quite different, and scales with different names might measure concepts that were quite similar. Although diversity and scientific pluralism can be useful, systematic accumulation of findings and communication among researchers had become almost impossible amidst the cacophony of competing concepts and scales.

At the University of California, Berkeley, for example, researchers studied personality with as few as two, and as many as 20 concepts, including the two dimensions of ego-resilience and ego-control that Block and Block (1980) measured with their California Q-sort; the four scales on the Myers–Briggs Type Indicator (MBTI; Myers & McCaulley, 1985) that measure extraversion, feeling, judging, and intuition; and the 20 scales on the California Psychological Inventory (CPI; Gough, 1987) measuring folk concepts such as capacity for status, self-control, well-being, tolerance, and achievement via independence (see Table 4.1). At the time, many personality researchers were hoping to be the one who would discover the right structure that all others would then adopt, thus transforming the fragmented field into a community speaking a common language. However, we now know that such an integration was not to be achieved by any one researcher or by any one theoretical perspective. As Allport once put it, "each assessor has his own pet units and uses a pet battery of diagnostic devices" (1958, p. 258).

What personality psychology lacked was a descriptive model, or taxonomy, of its subject matter. One of the central goals of scientific taxonomies is the definition of overarching domains within which large numbers of specific instances can be understood in a simplified way. Thus, in person-

TABLE 4.1. Personality Dimensions in Questionnaires and in Models of Personality and Interpersonal Behavior: Classified by Big Five Domain

Theorist	Extraversion I	Agreeableness II	Conscientiousness III	Neuroticism IV	Openness/Intellect V
Bales (1970)	Dominant–Initiative	Social–Emotional Orientation	Task Orientation[a]		—
Block & Block (1980)	Undercontrol	Overcontrol		Ego-Resiliency[b] (R)	
A. H. Buss & Plomin (1975)	Activity Sociability	—	Impulsivity (R)	Emotionality	—
Cattell (1943)	Exvia (vs. Invia)	Pathemia (vs. Cortertia)	Superego Strength	Adjustment (R) (vs. Anxiety)	Independence
Clark & Watson (1999)	Positive Emotionality	Constraint (vs. Disinhibition)[c]		Negative Emotionality	—
Comrey (1970)	Extraversion, Activity	Femininity (vs. Masculinity)	Orderliness, Social Conformity	Emotional Stability (R)	Rebelliousness
Eysenck (1986)	Extraversion	Psychoticism[c] (R)		Neuroticism	—
Gough (1987)					
CPI Vectors	Externality	—	Norm-Favoring	Self-Realization[d] (R)	
CPI Scales	Sociability	Femininity	Achievement via Conformance	Well-Being (R)	Achievement via Independence
Guilford (1975)	Social Activity	Paranoid Disposition (R)	Thinking Introversion	Emotional Stability (R)	—
Hogan (1986)	Sociability	Likeability	Prudence (vs. Impulsivity)	Adjustment (R)	Intellectance
Jackson (1984)	Outgoing, Social Leadership	Self-Protective Orientation (R)	Work Orientation	Dependence (R)	Aesthetic–Intellectual
MMPI; Myers & McCauley (1985)	Extraversion (vs. Introversion)	Feeling (vs. Thinking)	Judging (vs. Perceiving)	—	Intuition (vs. Sensing)
MBTI; Tellegen (1982)	Positive Emotionality		Constraint	Negative Emotionality	Absorption
	Agentive	Communal			
Tellegen et al. (2003)	Histrionic	Paranoid (R)	Compulsive	Borderline	Schizotypal
Wiggins[e] (1979)	Power/Dominance	Nurturance	(Conscientiousness)	(Neuroticism)	(Openness)

Note. Based on John (1990) and McCrae and John (1992). (R) indicates that the dimension was reverse-scored in the direction *opposite* to that of the Big Five label listed above. MBTI, Myers–Briggs Type Indicator; MMPI-2, Minnesota Multiphasic Personality Inventory—2.

[a]This dimension contrasts a work-directed, emotionally neutral orientation with an erratic, emotionally expressive orientation (Bales, 1970), and thus seems to combine elements of both Conscientiousness and Neuroticism.

[b]Ego-resiliency seems to subsume aspects of both Openness and low Neuroticism, because an ego-resilient individual is considered both intellectually resourceful and effective in controlling anxiety (Block & Block, 1980). However, Robins, John, and Caspi (1994) found that in adolescents, ego-resiliency is related to all of the Big Five dimensions in the well-adjusted direction. The ego-control construct was related to Extraversion, Conscientiousness, and Agreeableness, with the undercontrolled pole most similar to Extraversion and overcontrolled pole most similar to Conscientiousness and Agreeableness.

[c]High scores on the GTS Constraint scale are correlated with both Agreeableness and Conscientious (L. A. Clark & Watson, 1999; Markon et al., 2005). Conversely, the EPQ Psychoticism scale is associated with low scores on both Agreeableness and Conscientiousness (Goldberg & Rosolack, 1994; McCrae & Costa, 1985a).

[d]The third vector scale on the CPI (Gough, 1987) measures levels of psychological integration and realization and should reflect aspects of both low Neuroticism (e.g., Well-Being) and high Openness (e.g., Achievement via Independence).

[e]Wiggins (1979) originally focused on Dominance and Nurturance, which define the interpersonal circumplex. The dimensions in parentheses indicate that Trapnell and Wiggins (1990) added adjective scales for Conscientiousness, Neuroticism, and Openness after the emergence of the Big Five (see also Wiggins, 1995).

ality psychology, a taxonomy would permit researchers to study specified domains of related personality characteristics, rather than examining separately the thousands of particular attributes that make human beings individual and unique. Moreover, a generally accepted taxonomy would facilitate the accumulation and communication of empirical findings by offering a standard vocabulary, or nomenclature.

After decades of research, the field has now achieved an initial consensus on a general taxonomy of personality traits, the "Big Five" personality dimensions. These dimensions do not represent a particular theoretical perspective but were derived from analyses of the natural-language terms people use to describe themselves and others. Rather than replacing all previous systems, the Big Five taxonomy serves an integrative function because it can represent the various and diverse systems of personality description in a common framework, as shown by the columns organizing Table 4.1.

OUTLINE AND GOALS OF THIS CHAPTER

The first version of this chapter (John, 1990) offered a comprehensive and detailed review of most of the available research. This is no longer possible as we are writing this chapter in 2007. What has happened? Figure 4.1 uses publication trends over the past 25 years to illustrate how fundamentally the field has changed. Specifically, we show the number of publications related to the Big Five personality traits for each 5-year interval, beginning in the early 1980s, obtained from keyword searches of the PsycINFO database. To provide a comparison, we also show the publication trend for the influential models developed earlier by Cattell and by Eysenck. Although both were then close to retirement age, their influence had continued during the 1980s. In fact, both Cattell (1990) and Eysenck (1990) had written chapters on personality traits for the first edition of this handbook.

What did we expect to find? Our intuitions suggested that publications on the Big

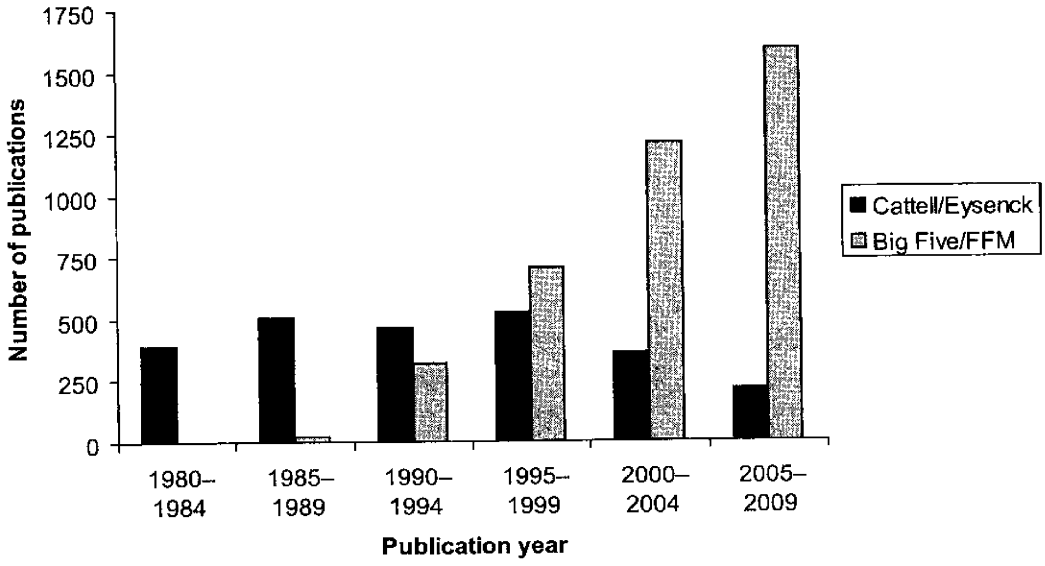


FIGURE 4.1. Number of publications related to either the Big Five personality traits or to the influential models developed earlier by Cattell and by Eysenck (in 5-year intervals), identified in keyword searches of the PsycINFO database. *Note*: The numbers identified in the figure as *Cattell/Eysenck* refer to the sum of all articles that used one of the measures developed by either Eysenck or Cattell as a keyword, such as "EPI," "EPQ," and "16PF"; those identified as *Big Five/FFM* are the sum of all articles that used as one of their keywords "Big Five," "Five Factor Model," "5 Factor Model," and "+ personality" (to rule out misidentifications of articles using these keywords in other literatures, such as the "big five game animals in Africa"). To estimate the projected publication trends for 2005–2009 (which were not yet available when this chapter was completed), we computed the sum of articles for 2005 and 2006 and multiplied that 2-year period by 2.5.

Five had increased substantially since the mid-1980s, with Cattell's and Eysenck's influence decreasing. But we were surprised by the data. First, the ascent of the Big Five happened much more gradually than we had expected, and Cattell's and Eysenck's influence held steady much longer. As Figure 4.1 shows, it took until the late 1990s for the number of Big Five publications to finally overtake the two older models. Second, whereas references to Cattell and Eysenck models have finally begun to decline in absolute numbers, their decline has been small compared to the amazing increase in research publications on the Big Five. By 2006, the last year for which we had figures available, the number of Big Five publications exceeded 300 per year, compared with less than 50 for the two older models.

In the 9 years since the previous version of this chapter (John & Srivastava, 1999) was completed, almost 2,000 new publications on the Big Five have appeared. As a result, we can now cover only a small fraction of all the relevant work in this chapter. Our main goal remains to provide a general overview and introduction to the field that focuses on the main issues and can serve as a useful reference resource. We therefore refer the reader to more specialized sources or reviews as needed. We begin our chapter with the history of the Big Five, including the discovery of the five dimensions, research replicating and extending the model, its convergence with research in the questionnaire tradition, and the development of several instruments to measure the Big Five. Then we compare three of the most frequently used instruments and discuss some new data regarding their reliability and validity. Finally, we address a number of conceptual issues, including how the Big Five taxonomy is structured hierarchically, how the five dimensions develop, whether they predict important life outcomes, and whether they are descriptive or explanatory concepts.

THE LEXICAL APPROACH AND DISCOVERY OF THE BIG FIVE

One starting place for a shared taxonomy is the natural language of personality description. Beginning with Klages (1932), Baumgarten (1933), and Allport and Odbert (1936), various psychologists have turned to the natural language as a source of attributes for a scientific taxonomy. This work, beginning with the extraction of all personality-relevant terms from the dictionary, has been guided by the lexical approach (see John et al., 1988; Saucier & Goldberg, 1996b). The *lexical hypothesis* posits that most of the socially relevant and salient personality characteristics have become encoded in the natural language (e.g., Allport, 1937). Thus, the personality vocabulary contained in the dictionaries of a natural language provides an extensive, yet finite, set of attributes that the people speaking that language have found important and useful in their daily interactions (Goldberg, 1981).

Allport and Odbert's Psycholexical Study: Traits, States, Activities, and Evaluations

Following Baumgarten's (1933) work in German, Allport and Odbert (1936) conducted a seminal lexical study of the personality terms in an unabridged English dictionary. They included all terms that could be used to "distinguish the behavior of one human being from that of another" (Allport & Odbert, 1936, p. 24) and identified almost 18,000 terms—"a semantic nightmare" (Allport, 1937, pp. 353–354) that would keep psychologists "at work for a life time" (Allport & Odbert, 1936, p. vi). Indeed, this task has preoccupied personality psychologists for more than 60 years (for details, see John et al., 1988; John, 1990).

What kinds of person descriptors are included in the dictionary? Allport and Odbert identified four major categories: (1) personality traits (e.g., *sociable*, *aggressive*, and *fearful*), defined as "generalized and personalized determining tendencies—consistent and stable modes of an individual's adjustment to his environment" (p. 26); (2) temporary states, moods, and activities, such as *afraid*, *rejoicing*, and *elated*; (3) highly evaluative judgments of personal conduct and reputation, such as *excellent*, *worthy*, *average*, and *irritating*—although these terms presuppose some traits within the individual, they do not indicate the specific attributes that gave rise to the individual's evaluation by others or by society in general; and (4) physical characteristics, capacities and talents, and other terms of doubtful relevance to personality. Norman

(1967) elaborated these classifications into seven content categories: Individuals can be described by their enduring *traits* (e.g., irrascible), by the *internal states* they typically experience (furious), by the *physical states* they endure (trembling), by the *activities* they engage in (screaming), by the *effects* they have on others (frightening), by the *roles* they play (murderer), and by social *evaluations* of their conduct (unacceptable, bad). Moreover, individuals differ in their anatomical and morphological characteristics (short) and in the personal and societal evaluations attached to these appearance characteristics (cute).

Both Allport and Odbert (1936) and Norman (1967) classified the terms culled from the dictionary into mutually exclusive categories. However, their categories clearly overlap and have fuzzy boundaries. Chaplin, John, and Goldberg (1988) proposed a prototype conception where each category is defined in terms of its clear cases rather than its boundaries. Chaplin and colleagues applied this prototype conception to traits, states, and activities. Prototypical *states* were seen as temporary, brief, and externally caused. Prototypical traits were seen as stable, long-lasting, internally caused, and needed to be observed more frequently and across a wider range of situations than states before they were attributed to an individual. These findings replicated the earlier classifications and confirmed that the lay conceptions of traits and states are widely shared and understood.

Identifying the Major Dimensions of Personality Description: Cattell's Early Efforts

Allport and Odbert's (1936) classifications provided some initial structure for the personality lexicon. However, to be of practical value, a taxonomy must provide a systematic framework for distinguishing, ordering, and naming individual differences in people's behavior and experience (John, 1989). Aiming for such a taxonomy, Cattell (1943) used the Allport and Odbert list as a starting point. Because the size of that list was too overwhelming for research purposes, Cattell (1943, 1945a, 1945b) began with the subset of 4,500 trait terms. Indeed, most taxonomic research has focused on the trait category, although the other categories are no less important; the emotional-state and social-evaluation categories have recently received more attention (Almagor, Tellegen, & Waller, 1995; Benet-Martínez & Waller, 1997).

Using both semantic and empirical clustering procedures as well as his own reviews of the literature available at the time (for reviews, see John, 1990; John et al., 1988), Cattell reduced the 4,500 trait terms to a mere 35 variables, eliminating more than 99% of the initial terms. This drastic reduction was dictated primarily by the data-analytic limitations of his time, which made factor analyses of large variable sets prohibitively costly and complex. Using this small set of 35 variables, Cattell conducted several oblique factor analyses (i.e., allowing for correlated factors) and concluded that he had identified 12 factors, which became part of his 16 Personality Factors (16PF) questionnaire (Cattell, Eber, & Tatsuoka, 1970).

Cattell also claimed that his factors showed excellent correspondence across methods, such as self-reports, ratings by others, and objective tests; however, these claims have not gone unquestioned (e.g., Becker, 1960; Nowakowska, 1973). Moreover, reanalyses of Cattell's own correlation matrices by others have not confirmed the number and nature of the factors he proposed (e.g., Tupes & Christal, 1961, 1992). Digman and Takemoto-Chock (1981) concluded that Cattell's "original model, based on the unfortunate clerical errors noted here, cannot have been correct" (p. 168), although the second-order factors of the 16PF show some correspondence between Cattell's system and the Big Five dimensions discovered later.

THE "BIG FIVE" FACTORS IN PERSONALITY TRAIT RATINGS

Initial Discovery of the Big Five in Cattell's Variable List

Cattell's pioneering work and the availability of a relatively short list of variables stimulated other researchers to examine the dimensional structure of trait ratings. Several investigators were involved in the initial discovery of the Big Five dimensions. First, Fiske (1949) constructed much simplified descriptions from 22 of Cattell's variables; the factor structures derived from self-ratings, ratings by peers, and ratings by psychologi-

cal staff members were highly similar and resembled what would later become known as the Big Five. To clarify these factors, Tupes and Christal (1961) reanalyzed correlation matrices from eight samples and found "five relatively strong and recurrent factors and nothing more of any consequence" (1961, p. 14). This five-factor structure has been replicated by Norman (1963), Borgatta (1964), and Digman and Takemoto-Chock (1981) in lists derived from Cattell's 35 variables. Following Norman (1963), the factors were initially labeled (I) Extraversion or Surgency (talkative, assertive, energetic); (II) Agreeableness (good-natured, cooperative, trustful); (III) Conscientiousness (orderly, responsible, dependable); (IV) Emotional Stability (calm, not neurotic, not easily upset); and (V) Culture (intellectual, polished, independent-minded).

These factors (see Table 4.2 for more recent labels, definitions, and examples) eventually became known as the "Big Five"—a name Goldberg (1981) chose not to reflect their intrinsic greatness but to emphasize that each of these factors is extremely broad. Thus, the Big Five structure does not imply that personality differences can be reduced to only five traits. Rather, these five dimensions represent personality at a very broad level of abstraction; each dimension summarizes a large number of distinct, more specific personality characteristics.

Testing the Big Five in a Comprehensive Set of English Trait Terms

After a period of dormancy during the 1970s and early 1980s, research on personality structure increased dramatically during the mid-1980s. Factor structures resembling the Big Five were identified in numerous sets of variables (e.g., Botwin & Buss, 1989; Conley, 1985; De Raad, Mulder, Kloosterman, & Hofstee, 1988; Digman & Inouye, 1986; Goldberg, 1981, 1990; John, 1990; McCrae & Costa, 1985a, 1987; Peabody & Goldberg, 1989; Saucier & Goldberg, 1996a). Because a number of these studies were influenced by Cattell's selection of variables (Block, 1995), it was important to test the generality of the Big Five in more comprehensive variable sets. To update the Allport and Odbert list and rectify the problems with Cattell's reduction steps, Norman (1967) compiled an exhaustive list of personality descriptive terms, which he sorted into 75 semantic categories. Goldberg (1990; see also 1981, 1982) used this list to clarify the composition of the Big Five factors and to test their generalizability across methodological variations and data sources. Goldberg (1990) constructed an inventory of 1,710 trait adjectives and had participants rate their own personality. He then scored Norman's semantic categories as scales and factor-analyzed their intercorrelations in the self-rating data. The first five factors represented the expected Big Five, replicated across a variety of different methods of factor extraction and rotation, and remained virtually invariant even when more than five factors were rotated.

To ensure independence from any a priori classification, Goldberg (1990) conducted two additional studies using abbreviated sets of more common terms. In one study, Goldberg obtained self- and peer ratings on 475 very common trait adjectives, which he had grouped into 131 sets of "tight synonym" clusters. The five-factor self- and peer-report structures were very similar to each other and to the structure obtained in the more comprehensive list of 1,710 terms. Most important were the null results from the search for replicable additional factors. Saucier and Goldberg (1996a) selected 435 highly familiar trait adjectives; a factor analysis of these adjectives closely replicated the Big Five. Another thorough search for factors beyond the Big Five showed that these five were the only consistently replicable factors (Saucier, 1997).

Assessing the Big Five with Trait Descriptive Adjectives: Simple and Circumplex Approaches

Goldberg (1990, 1992) distilled his extensive taxonomic findings into several adjective lists. A 50-item instrument using the so-called "transparent format" (Goldberg, 1992) is not used much for research but is excellent for instructional purposes (Pervin, Cervone, & John, 2005): 10 bipolar adjective scales (e.g., quiet–talkative) are grouped together under the factor name, thus making the constructs being measured transparent to the subject. The list used more commonly in research is the set of 100 unipolar trait descriptive adjectives (TDA). Goldberg (1992) conducted a series of factor-analytic studies

TABLE 4.2. The OCEAN[a] of Personality: Definition and Explication of the Big Five Domains

Factor initial (number)	Big Five domains				
	E (Factor I)	A (Factor II)	C (Factor III)	N (Factor IV)	O (Factor V)
Verbal labels	Extraversion Energy Enthusiasm	Agreeableness Altruism Affection	Conscientiousness Constraint Control of impulse	Neuroticism Negative Emotionality Nervousness	Openness Originality Open-Mindedness
Conceptual definition	Implies an *energetic approach* toward the social and material world and includes traits such as sociability, activity, assertiveness, and positive emotionality.	Contrasts a *prosocial and communal orientation* toward others with antagonism and includes traits such as altruism, tender-mindedness, trust, and modesty.	Describes *socially prescribed impulse control* that facilitates task- and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing, and prioritizing tasks.	Contrasts emotional stability and even-temperedness with *negative emotionality*, such as feeling anxious, nervous, sad, and tense.	Describes the breadth, depth, originality, and complexity of an individual's *mental and experiential life*.
Behavioral examples	Approach strangers at a party and introduce myself; Take the lead in organizing a project; Keep quiet when I disagree with others (R)	Emphasize the good qualities of other people when I talk about them; Lend things to people I know (e.g., class notes, books, milk); Console a friend who is upset	Arrive early or on time for appointments; Study hard in order to get the highest grade in class; Double-check a term paper for typing and spelling errors; Let dirty dishes stack up for more than one day (R)	Accept the good and the bad in my life without complaining or bragging (R); Get upset when somebody is angry with me; Take it easy and relax (R)	Take the time to learn something simply for the joy of learning; Watch documentaries or educational TV; Come up with novel set-ups for my living space; Look for stimulating activities that break up my routine
Examples of external criteria predicted	*High pole:* Social status in groups and leadership positions; selection as jury foreperson; positive emotion expression; number of friends and sex partners *Low pole:* Poorer relationships with parents; rejection by peers	*High pole:* Better performance in work groups *Low pole:* Risk for cardiovascular disease, juvenile delinquency, interpersonal problems	*High pole:* Higher academic grade-point averages; better job performance; adherence to their treatment regimens; longer lives *Low pole:* Smoking, substance abuse, and poor diet and exercise habits; attention-deficit/hyperactivity disorder (ADHD)	*High pole:* Poorer coping and reactions to illness; experience of burnout and job changes *Low pole:* Feeling committed to work organizations; greater relationship satisfaction	*High pole:* Years of education completed; better performance on creativity tests; success in artistic jobs; create distinctive-looking work and home environments *Low pole:* Conservative attitudes and political party preferences

Note. Conceptual definitions are based on John and Srivastava (1999). Behavioral examples are based on significant correlations between Big Five Inventory scales and self-reported act frequencies in an undergraduate sample ($N = 375$; John & Naumann, 2007). (R) denotes that the act was a reverse-keyed item (i.e., correlated negatively with the Big Five domain). For more examples of predictive validity criteria and relevant references, see the text. [a]The first letter of the Big Five dimensions form the anagram OCEAN.

to develop and refine the TDA, selecting only adjectives that uniquely defined each factor. These scales have high internal consistency, and their factor structure is easily replicated.

Another adjectival measure of the Big Five was developed by Wiggins (1995; Trapnell & Wiggins, 1990), who used trait adjectives to elaborate the two major dimensions of interpersonal behavior: dominance (or agency) and nurturance (or communion). As shown in Table 4.1, the first dimension resembles Extraversion in the Big Five, and the second resembles Agreeableness. Wiggins thus extended his circumplex scales by adding simple adjective measures for the other three Big Five factors (Trapnell & Wiggins, 1990).

The circumplex approach has also been extended to a perennial problem in lexical research on personality factors: namely, to describe more clearly those characteristics that fall in the fuzzy regions between the factors. Using 10 two-dimensional circumplexes, Hofstee, De Raad, and Goldberg (1992) devised a novel empirical approach, called the Abbreviated Big Five Circumplex (AB5C), to represent the two-dimensional space formed by each pair of factors and define eight facets that reflect various combinations of the two factors. The facets differ in whether they are more closely related to one or the other factor. For example, there are two facets that reflect high Agreeableness and high Conscientiousness, but they differ in which of the two factors is given more prominence. The *responsibility* facet represents agreeable Conscientiousness, whereas the *cooperation* facet represents conscientious Agreeableness (Hofstee et al., 1997).

Cross-Language and Cross-Cultural Studies

The results reviewed so far suggest that the Big Five structure provides a replicable representation of the major dimensions of trait description in English—the five dimensions generalize across different types of samples, raters, and methodological variations when comprehensive sets of variables are factored. Generalizability across languages and cultures is another important criterion for evaluating personality taxonomies (John, Goldberg, & Angleitner, 1984). The existence of cultural universals (Goldberg, 1981) would be consistent with an evolutionary perspective: If the tasks most central to human survival are universal, then the most important individual differences, and the terms people use to label these individual differences, should be universal as well (D. M. Buss, 1996; Hogan, 1983). Conversely, if cross-cultural research reveals a culturally specific dimension, variation on that dimension may be uniquely important within that culture's particular social context. Although central from the vantage point of the lexical approach, cross-language research is difficult and expensive to conduct, and until the 1990s it was quite rare. In the initial taxonomic studies, English was the language of choice, primarily because the researchers were American (see John et al., 1984; John, Angleitner, & Ostendorf, 1988).

Initial Studies in Dutch and German

The first two non-English taxonomy projects involved Dutch and German, both Germanic languages closely related to English. The Dutch projects, carried out by Hofstee, De Raad, and their colleagues at the University of Groningen (De Raad et al., 1988; Hofstee et al., 1997; see De Raad, Perugini, et al., 1998, for a review), yielded conclusions generally consistent with those from the American English research: Only five factors were replicable across different selections of trait adjectives and across different subject samples. Those five factors were similar to the English Big Five, although the fifth Dutch factor emphasized Unconventionality and Rebelliousness rather than Intellect and Imagination, as found in English.

The German taxonomy project, begun in Bielefeld, carried out a comprehensive "psycholexical" study of the German personality vocabulary (Angleitner, Ostendorf, & John, 1990). This study was explicitly based on the prototype conception and improved on earlier studies in several respects. In particular, 10 independent judges classified all the terms obtained from the dictionary as traits, states, social evaluations, etc., thus providing a continuous measure of the prototypicality of each term for each category and also a check on the reliability and validity of the judgments. The resulting German personality lexicon is more convenient to use than the unwieldy Allport and Odbert lists because prototypicality values are available for each term in 13

different content categories. Thus, it is easy to select subsets of prototypical traits, states, social evaluations, and so on, from the total pool for further studies. Angleitner and colleagues' (1990) research has served as a blueprint for subsequent taxonomic efforts in other languages.

To test the structure underlying the German trait terms, Ostendorf (1990) selected the most prototypical trait adjectives from the German taxonomy, and his factor analyses of about 450 traits yielded the clearest replication of the Big Five so far. In addition to the prototypical traits representing the distillation of the German trait lexicon, Ostendorf also included German translations of several English personality instruments—a combined emic–etic design that allows researchers to establish empirically the similarity of indigenous (emic) factors to the factors translated from other languages and cultures (etic). Using correlational analyses, Ostendorf conducted an a priori, quantitative evaluation of the fit between his emic German factors and the etic Big Five in the same sample of German subjects, and he found evidence for substantial cross-language convergence.

However, this combined emic–etic strategy is difficult to implement and, unfortunately, has not been used consistently in research. Thus, researchers often reach conclusions about factor similarity by "eyeballing" the item content of the factors in the indigenous language and comparing it to the typical factor definitions in English. That leaves much leeway to the investigators in finding (or not finding) a factor that another investigator might not have found. For example, a Hebrew factor defined primarily by traits such as sophisticated, sharp, knowledgeable, articulate, and impressive would lead some researchers to conclude that they had found a clear Intellect factor, whereas Almagor and colleagues (1995) interpreted it as Positive Valence.

Problems with Translations and Underestimating Cross-Language Congruence

Another methodological difficulty in cross-language research involves translations. Researchers working within their indigenous language have to translate their concepts into English to communicate their findings in scientific journals, and much too often considerable slippage occurs in the translation process. For example, because "temperamental" was listed as a definer of Extraversion in German, one might hypothesize an important cultural difference here until one realizes that the German trait was probably *temperamentvoll*, which has nothing to do with temper but means "full of life and energy," as in *vivacious*. Similarly, the Italian trait term *frizzante* (translated as *sparkling*) was not found related to intellect, as one might expect, but to extraversion and probably means something close to *bubbly*.

An initial study of German–English bilinguals, which provided support for cross-language generalizability (John et al., 1984), directly addressed the issue of translation equivalence. The unique advantage of the bilingual design is that sample differences can be controlled and translation checks can be made at the level of individual items because the same subject provides descriptions in both languages (see also Benet-Martínez & John, 1998). Using a careful back-translation procedure, translation equivalence between English and German trait adjectives was acceptable, with a mean correlation of .52 across a 2-week interval between administrations (John et al., 1984). However, a few translations proved to be inadequate, with item-translation correlations approaching zero. These findings, obtained for closely related languages, suggest that mistranslations are even more likely to occur in monolingual investigations of personality structure and lead to severe underestimations of cross-language generality.

These difficulties are illustrated in a study by Hofstee and colleagues (1997) who used 126 words they felt could be translated and matched across previous lexical studies in English, Dutch, and German to assess factor congruence coefficients among all pairs of factors in the three languages. Their findings seemed to show considerable congruence; with one exception (the Openness factors in Dutch and English), the pairwise congruence coefficients all exceeded .70. Strangely, the authors interpreted these levels of cross-language congruence as "disappointing" (Hofstee et al., 1997, p. 27). We are more optimistic about these findings. The observed levels of factor congruence can be taken as absolute estimates only if one as-

sumes that the translations are perfectly equivalent and that the factor structures in each language are perfectly stable. When the cross-language congruence coefficients were corrected for the imperfect reliabilities (replication) of the within-language factor structures, the corrected English–German congruence coefficients ranged from .84 to .93, impressive values given that they have not been corrected for the imperfect translations (John & Srivastava, 1999). Moreover, the correspondence for the fifth factor was .93, suggesting that the Intellect or Openness factor was defined almost identically in English and German. This reexamination suggests that translation-based comparisons across languages are heuristically useful but should not be interpreted in terms of absolute effect sizes. These results also suggest that the fifth factor in Dutch is defined differently than in the other two languages, and explanations for this finding need to be sought.

Rules for Including Trait Descriptors in Taxonomic Studies

In all likelihood, some of the differences observed among the factor structures in these three languages are due to the different inclusion rules adopted by the taxonomy teams. The selection criterion used by the Dutch researchers favored terms related to temperament, excluded terms related to intellect, talents, and capacities, and included a number of extremely negative evaluative terms, such as *perverse*, *sadistic*, and *criminal*. The German team explicitly included intellect and talent descriptors but omitted attitudes and evaluative terms, which were included as categories separate from traits. Finally, the American English taxonomy included attitudinal terms such as *liberal*, *progressive*, and *provincial*, along with a number of intellect terms. Given the diverse range of traits related to the fifth factor, it is less surprising that the German and English factors shared the intellect components, whereas the Dutch factor included some imagination-related traits (e.g., *inventive*, *original*, *imaginative*) but otherwise emphasized unconventionality and was thus interpreted initially as a "Rebelliousness" factor.

One Italian trait taxonomy (Caprara & Perugini, 1994) found a similar fifth factor interpreted as Unconventionality. Not surprisingly, these Italian researchers had followed the Dutch selection procedures rather than the German procedures, which likely would have represented more Intellect terms in the initial item pool. A second (and independent) team of Italian taxonomers (Di Blas & Forzi, 1998) failed to find the same factors as the first (see also De Raad, Di Blas, & Perugini, 1998). Given that both teams started with the same lexical material (Italian personality descriptors), this notable lack of convergence within the same language is disconcerting and serves to illustrate the inherent difficulties in standardizing taxonomic procedures and factor-analytic decisions across cultures and languages. How can we expect the Big Five to generalize across languages when two studies of the same language fail to show factor generalizability?

Thus, apparent failures to replicate the Big Five structure can be hard to interpret. For example, Szirmak and De Raad (1994) examined Hungarian personality descriptors and found strong support for the first four of the Big Five but failed to obtain a factor resembling the fifth of the Big Five. Instead, when they forced a five-factor solution, the Agreeableness factor split into two factors. Should this finding be counted as failure to replicate the Big Five, suggesting that Hungarians do not differ systematically on traits related to imagination, creativity, and intellect? Probably not: When six factors were rotated, an Intellect/Openness factor did emerge in the Hungarian data. Again, we suspect that this finding may be due to differences in the way the initial item pool was selected. In their review of this literature, Saucier, Hampson, and Goldberg (2000) conclude that "Given these and other differences among studies, is it any wonder that investigators might disagree about the evidential basis for a particular structural representation?" (p. 23).

Evidence in Non-Germanic Languages

Lexical research has now been extended to a growing range of non-Germanic languages, such as Chinese (Yang & Bond, 1990), Czech (Hrebickova & Ostendorf, 1995), Greek (Saucier, Georgiades, Tsouasis, & Goldberg, 2005), Hebrew (Almagor et al., 1995), Hungarian (Szirmak & De Raad, 1994), Italian (e.g., De Raad et al., 1998),

Polish (Szarota, 1995), Russian (Shmelyov & Pokhil'ko, 1993), Spanish (Benet-Martínez & Waller, 1997), Tagalog in the Philipines (e.g., Church & Katigbak, 1989; Church, Reyes, Katigbak, & Grimm, 1997), Turkish (Somer & Goldberg, 1999), and others. This literature has now grown far beyond the scope of this chapter, and we thus refer the reader to several in-depths reviews (e.g., Ashton et al., 2004; De Raad & Perugini, 2002; De Raad et al., 1998; Saucier & Goldberg, 2001; Saucier et al., 2000).

Most generally, our reading of this literature is that factors similar to the Big Five have been found in many other languages, but often more than five factors needed to be rotated, and sometimes two indigenous factors corresponded to one of the Big Five. The Big Five have been well-replicated in Germanic languages, but the evidence for non-Western languages and cultures tends to be more complex. Overall, the evidence is least compelling for the fifth factor, which appears in various guises, ranging from pure Intellect (in German) to Unconventionality and Rebelliousness (in Dutch and Italian). Moreover, when we only consider which factors are the *most* replicable, then structures with fewer factors (such as two or three) are often even more robust than the more differentiated Big Five (e.g., Saucier et al., 2005). Finally, a number of studies have suggested more than five factors. For example, the seven-factor solutions in Spanish and English (see Benet-Martínez & Waller, 1997) suggested additional separate positive and negative self-evaluation factors. The more recent six-factor solutions, obtained in reanalyses of data from several different languages (Ashton et al., 2004), suggest an additional honesty–humility factor.

While it is too early to decide whether these additional factors hold sufficient promise, two conclusions are apparent now. First, note that these factors are indeed "additional"; that is, they provide evidence for the generalizability of the Big Five *plus* one or two further factors. Second, when more factors than the Big Five have been identified, the additional factors rarely replicate across multiple studies conducted by independent investigators. Thus, we agree with De Raad and colleagues (1998), who concluded that the findings show "the general contours of the Big Five model as the best working hypothesis of an omnipresent trait structure" (p. 214). Although this cautious conclusion falls short of an unequivocal endorsement of the universality of the Big Five, it nonetheless offers a strong disconfirmation of the linguistic-relativism hypothesis that many of us had expected to hold true before the lexical data became available: There are no data to suggest that each culture and language has its own, unique set of personality dimensions; at least at the level of broad trait dimensions, cultures are more alike than we may have expected.

THE BIG FIVE IN PERSONALITY QUESTIONNAIRES

While researchers in the lexical tradition were accumulating evidence for the Big Five, the need for an integrative framework became more pressing among researchers who studied personality with questionnaire scales. Joint factor analyses of questionnaires developed by different investigators had shown that two broad dimensions, Extraversion and Neuroticism, appear in one form or another in most personality inventories. Beyond these "Big Two" (Wiggins, 1968), however, the various questionnaire-based models had shown few signs of convergence. For example, Eysenck (1991) observed that "Where we have literally hundreds of inventories incorporating thousands of traits, largely overlapping but also containing specific variance, each empirical finding is strictly speaking only relevant to a specific trait. ... This is not the way to build a unified scientific discipline" (p. 786).

Costa and McCrae's Research

The situation began to change in the early 1980s when Costa and McCrae were developing the NEO Personality Inventory (eventually published in 1985), labeled N-E-O because it was designed to measure the three dimensions of Neuroticism, Extraversion, and Openness to experience. Costa and McCrae (1976) had begun their work with cluster analyses of the 16PF (Cattell et al., 1970), which, as we described above, originated in Cattell's early lexical work. Their analyses again yielded the ubiquitous Extraversion and Neuroticism dimensions, but also convinced Costa and McCrae of the importance

of Openness, which originated from several of Cattell's primary factors (e.g., imaginative; experimenting).

In 1983 Costa and McCrae realized that their NEO system closely resembled three of the Big Five factors but did not encompass traits in the Agreeableness and Conscientiousness domains. They therefore extended their model with preliminary scales measuring Agreeableness and Conscientiousness. In several studies, McCrae and Costa (1985a, 1985b, 1987) demonstrated that their five questionnaire scales converged with adjective-based measures of the Big Five, although their conception of Openness was considerably broader than the Intellect or Imagination factor emerging from the lexical analyses (Saucier & Goldberg, 1996a). A series of influential papers in the late 1980s and early 1990s showed that these five factors could also be recovered in various other personality questionnaires, as well as in self-ratings on the California Adult Q-sort (see Costa & McCrae, 1992; McCrae & Costa, 2003).

The Revised NEO Personality Inventory

The initial NEO Personality Inventory (Costa & McCrae, 1985) included scales to measure six conceptually derived facets each for Neuroticism, Extraversion, and Openness but did not include facet scales for the newly added Agreeableness and Conscientiousness. In 1992, Costa and McCrae published the 240-item NEO Personality Inventory—Revised (NEO-PI-R; Costa & McCrae, 1992), which permits differentiated measurement of each Big Five dimension in terms of six more specific facets per factor (Costa & McCrae, 1995). Table 4.3 shows the six facets defining each of the factors. In contrast to most of the lexical studies, which relied on college student samples, the NEO-PI-R was developed in samples of middle-age and older adults, using both factor-analytic and multimethod validational procedures of test construction. The scales have shown substantial internal consistency, temporal stability, and convergent and discriminant validity against spouse and peer ratings (Costa & McCrae, 1992; McCrae & Costa, 2003).

For many research applications, the NEO-PI-R is rather lengthy. To provide a shorter measure, Costa and McCrae (1989, 1992) developed the 60-item NEO-FFI, an abbreviated version based on an item-level factor analysis of the 1985 version of the NEO PI (Costa & McCrae, 1985). The 12-item scales of the FFI consist of items that loaded highly on one of the five factors. The item content of the scales was adjusted somewhat to ensure adequate content coverage of the facets; however, these scales represent the core elements of each Big Five factor as defined on the NEO PI and therefore do not represent equally each of the six facets defining each factor. For example, the Agreeableness scale includes five items from the Altruism facet, three from Compliance, two from Trust, one from Tender-Mindedness, one from Straightforwardness, and none from Modesty. The reliabilities (Costa & McCrae, 1992) are adequate, with a mean of .78, and the NEO-FFI scales are substantially correlated with the NEO-PI-R scales.

DEFINING THE BIG FIVE ACROSS STUDIES: A PROTOTYPE APPROACH

So far, we have reviewed both Goldberg's (1990) lexically based research and Costa and McCrae's (1992) questionnaire-based research on the Big Five. Despite this extensive research, the Big Five structure was initially not widely welcomed in the field and, in fact, explicitly rejected by some senior researchers (e.g., Block, 1995; Eysenck, 1992, 1997; McAdams, 1992; Pervin, 1994). One problem, it seems, was the perception that there is no *single* Big Five, in the same way as there was just one 16PF, namely Cattell's, because he owned it. In contrast, the Big Five emerged in multiple labs and studies and is therefore not owned by any one person in the field, which makes it possible to ask questions such as "*which* Big Five?" or "*whose* Big Five?" (John, 1989). For example, the first factor has been labeled as surgency, confident self-expression, assertiveness, social extraversion, and power (see John, 1990, Table 3.1), and the second factor as social adaptability, likability, friendly compliance, agreeableness, and love. Of course, some variation across studies is to be expected with dimensions as broad and inclusive as the Big Five because researchers differ in the variables they include, thus representing different parts of the factor's total range of meaning. Moreover, researchers dif-

TABLE 4.3. Defining Facets for the Big Five Trait Domains: Three Approaches

Lexical facets (18) (Saucier & Ostendorf, 1999)	NEO-PI-R facets (30) (Costa & McCrae, 1992)	CPI-Big Five facets (16) (Soto & John, 2008)
Extraversion (E) facets		
E Sociability	E Gregariousness	E Gregariousness
E Assertiveness	E Assertiveness	E Assertiveness/Leadership
	E Activity	
E Activity/Adventurousness	E Excitement-Seeking	*[O Adventurousness]*
		E Social Confidence vs. Anxiety
	E Positive emotions	
E Unrestraint		
[A Warmth/Affection]	E Warmth	
Agreeableness (A) facets		
A Warmth/Affection	*[E Warmth]*	
A Modesty/Humility	A Modesty	A Modesty vs. Narcissism
	A Trust	A Trust vs. Suspicion
	A Tender-Mindedness	A Empathy/Sympathy
		A Altruism
A Generosity		
A Gentleness	A Compliance	
	A Straightforwardness	
Conscientiousness (C) facets		
C Orderliness	C Order	C Orderliness
C Industriousness	C Achievement Striving	
C Reliability	C Dutifulness	C Industriousness
C Decisiveness		
	C Self-Discipline	C Self-Discipline
[O Perceptiveness]	C Competence	
	C Deliberation	
Neuroticism (N) facets		
N Insecurity	N Anxiety	N Anxiety
N Emotionality		
N Irritability	N Angry Hostility	N Irritability
	N Depression	N Depression
		N Rumination–Compulsiveness
	N Self-Consciousness	*[E Social Confidence vs. Anxiety]*
	N Vulnerability	
	N Impulsiveness	
Openness (O) facets		
O Intellect	O Ideas	O Intellectualism
	O Aesthetics	O Idealism
O Imagination/Creativity	O Fantasy	
		O Adventurousness
	O Actions	
	O Feelings	
	O Values	
O Perceptiveness		

Note. The NEO-PI-R facets are listed in the middle column because that instrument makes the largest number of distinctions below the Big Five (30 facets), as compared with the 18 lexical facets and the 16 CPI facets. CPI-Big Five facets are matched with NEO-PI-R facets on the basis of both rational judgments by the authors and correlations between the two sets of facets in a sample of 520 adults (see Soto & John, 2008). Lexical facets are matched with NEO-PI-R facets on the basis of rational judgments by the authors. Some facets (e.g., CPI Adventurousness) are listed once under their primary Big Five domain (e.g., Openness) and again in brackets under another Big Five domain if their best-matching facet appears there (e.g., next to NEO Excitement-Seeking, which is an Extraversion facet on the NEO-PI-R but also has a substantial secondary correlation with Openness). Note that the Warmth facet belongs to the Extraversion domain in the NEO-PI-R, whereas a very similar Warmth/Affection facet belongs to the Agreeableness domain in the lexical approach.

fer in their preferences for factor labels even when the factor content is quite similar. The fact that the labels differ does not necessarily mean that the factors are different, too. Thus, there may be more commonality than meets the eye.

A prototype approach may help identify these commonalities across studies (John, 1990). Natural categories such as the Big Five typically have fuzzy and partially overlapping definitions (Rosch, 1978), but they can still be extremely useful when defined in terms of prototypical exemplars. Similarly, the Big Five may be defined with prototypical traits that occur consistently across studies. One way to integrate the various interpretations of the factors is to conceptually map the five dimensions into a common language. Human judges were used to abstract the common elements in these findings (John, 1989, 1990), and the 300 terms included in the Adjective Check List (ACL; Gough & Heilbrun, 1983) served as the standard language.

Conceptually Derived Prototype Descriptions of the Big Five and Their Validation in Observer Data

A set of 10 judges first developed a detailed understanding of the Big Five by reviewing the factor solutions and interpretations of all major Big Five articles published by 1988. The judges then independently sorted each of the 300 items in the ACL into one of the Big-Five domains or into a sixth "other" category, with substantial interjudge agreement. In all, 112 of the 300 ACL terms were assigned to one of the Big Five with 90% or better agreement. These terms form a relatively narrow, or "core," definition of the five factors because they include only those traits that appeared consistently across studies. As with any rationally constructed measure, the validity of these categorizations was tested empirically in a factor analysis of the 112 terms. Whereas most Big Five research has been based on college students' self- and peer ratings, this study used observer data: psychologists rated 140 men and 140 women who had participated in groups of 10–15 in one of the assessment weekends at the Institute of Personality and Social Research (IPSR, formerly IPAR) at Berkeley (John, 1990). Because each subject had been described on the ACL by 10 staff members, a factor analysis could be performed on more reliable, aggregated observer judgments. The varimax rotated factor loadings, shown in Table 4.4, provide a compelling confirmation of the initial prototypes. All but one item loaded on its hypothesized factor in the expected direction, and most of the loadings were substantial.

Note that the items defining each of the factors cover a broad range of content. For example, Factor I includes traits such as *active*, *adventurous*, *assertive*, *dominant*, *energetic*, *enthusiastic*, *outgoing*, *sociable*, and *show-off*. In light of this substantial bandwidth, the heterogeneity of the previous factor labels is understood more easily—different investigators have focused on different components, or facets, of the total range of meaning subsumed by each factor. In this study, the Extraversion factor includes at least five distinguishable components: activity level (active, energetic), dominance (assertive, forceful, bossy), sociability (outgoing, sociable, talkative), expressiveness (adventurous, outspoken, noisy, show-off), and positive emotionality (enthusiastic, spunky). Note that these five components are similar to five of the six facets Costa and McCrae (1992) included in their definition of the Extraversion domain: activity, assertiveness, gregariousness, excitement-seeking, and positive emotions (see Table 4.3)—and four of the five have been identified in a lexical study that empirically identified facets across two languages (Saucier & Ostendorf (1999), shown on the left-hand side of Table 4.3. Costa and McCrae's sixth Extraversion facet, warmth, is here considered a component of Factor II; all 10 judges interpreted past research to imply that warmth is part of Agreeableness, and the empirical loading of .82 confirmed this interpretation, just as the lexical facet of warmth/affection appears on Agreeableness. In addition to warmth (affectionate, gentle, warm), Factor II covers themes such as tender-mindedness (sensitive, kind, soft-hearted, sympathetic), altruism (generous, helping, praising), and trust (trusting, forgiving), as contrasted with hostility, criticality, and distrust; again, note the convergence with Costa and McCrae's facets. More generally, the adjectival definitions of the Big Five in Table 4.4 seem to capture the prototypical traits found in other studies and the facets shown in Table 4.3.

TABLE 4.4. Big Five Prototypes: Most Central Trait Adjectives Selected Consensually by Expert Judges and Their Factor Loadings in Personality Ratings by 10 Psychologists Serving as Observers

Extraversion		Agreeableness		Conscientiousness		Neuroticism		Openness	
Low	High	Low	High	Low	High	Low	High	Low	High
−.83 Quiet	.85 Talkative	−.52 Fault-finding	.87 Sympathetic	−.58 Careless	.80 Organized	−.39 Stable	.73 Tense	−.74 Commonplace	.76 Wide interests
−.80 Reserved	.83 Assertive	−.48 Cold	.85 Kind	−.53 Disorderly	.80 Thorough	−.35 Calm	.72 Anxious	−.73 Narrow interests	.76 Imaginative
−.75 Shy	.82 Active	−.45 Unfriendly	.85 Appreciative	−.50 Frivolous	.78 Planful	−.21 Contented	.72 Nervous	−.67 Simple	.72 Intelligent
−.71 Silent	.82 Energetic	−.45 Quarrelsome	.84 Affectionate	−.49 Irresponsible	.78 Efficient		.71 Moody	−.55 Shallow	.73 Original
−.67 Withdrawn	.82 Outgoing	−.45 Hard-hearted	.84 Soft-hearted	−.40 Slipshot	.73 Responsible		.71 Worrying	−.47 Unintelligent	.68 Insightful
−.66 Retiring	.80 Outspoken	−.38 Unkind	.82 Warm	−.39 Undependable	.72 Reliable		.68 Touchy		.64 Curious
	.79 Dominant	−.33 Cruel	.81 Generous	−.37 Forgetful	.70 Dependable		.64 Fearful		.59 Sophisticated
	.73 Forceful	−.31 Stern	.78 Trusting		.68 Conscientious		.63 High-strung		.59 Artistic
	.73 Enthusiastic	−.28 Thankless	.77 Helpful		.66 Precise		.63 Self-pitying		.59 Clever
	.68 Show-off	−.24 Stingy	.77 Forgiving		.66 Practical		.60 Temperamental		.58 Inventive
	.68 Sociable		.74 Pleasant		.65 Deliberate		.59 Unstable		.56 Sharp-witted
	.64 Spunky		.73 Good-natured		.46 Painstaking		.58 Self-punishing		.55 Ingenious
	.64 Adventurous		.73 Friendly		.26 Cautious		.54 Despondent		.45 Witty
	.62 Noisy		.72 Cooperative				.51 Emotional		.45 Resourceful
	.58 Bossy		.67 Gentle						.37 Wise
			.66 Unselfish						
			.56 Praising						
			.51 Sensitive						

Note. Based on John (1990). These items were assigned to one Big Five domain by at least 90% of the judges and thus capture the most prototypical (or central) content of each Big Five domain. The factor loadings, shown here only for the expected factor, were obtained in a sample of 140 males and 140 females, each of whom had been described by 10 psychologists serving as observers during an assessment weekend at the Institute of Personality and Social Research at the University of California at Berkeley (see also John, 1989).

The Prototypical Definition of Factor V: Culture, Intellect, or Openness?

The findings in Table 4.4 also address questions about the definition of the fifth factor. None of the items referring to aspects of "high" culture (e.g., civilized, polished, dignified, foresighted, logical) loaded substantially on Factor V (see John, 1990), and many loaded more highly on Factor III (Conscientiousness), thus discrediting an interpretation of Factor V as Culture. Apparently, the initial interpretation of Tupes and Christal's (1961) fifth factor as Culture was a historical accident (Peabody & Goldberg, 1989). The items that did load substantially on the fifth factor (see Table 4.4) include both the "open" characteristics (e.g., artistic, curious, original, wide interests) highlighted by McCrae and Costa (1985a, 1985b) and the "intellectual" characteristics (intelligent, insightful, sophisticated) emphasized by Digman and Inouye (1986), Peabody and Goldberg (1989), and Goldberg (1990).

These findings are also consistent with Goldberg's (1990) result that Factor V is defined as originality, wisdom, objectivity, knowledge, reflection, and art, thus involving facets of Openness related to ideas, fantasy, and aesthetics (Costa & McCrae, 1992). Similarly, Goldberg's analyses of the 133 synonym clusters showed intellectuality (intellectual, contemplative, meditative, philosophical, and introspective) and creativity (creative, imaginative, inventive, ingenious, innovative) with the highest loadings, only then followed by intelligence, versatility, wisdom, perceptiveness, art, logic, curiosity, and nonconformity (nonconforming, unconventional, rebellious), which loaded positively, and conventionality (traditional, conventional, unprogressive), which loaded negatively in all four samples. These and other lexical findings (see John & Srivastava, 1999) are inconsistent with both the Culture and a narrow Intellect interpretation and instead favor the broader Openness interpretation proposed by McCrae (1996); the inclusion of unconventionality and nonconformity also makes an important link to the definition of this lexical factor in Dutch and Italian (De Raad et al., 1998). Similarly, in a recent AB5C analysis designed to derive Big Five facets from the CPI item pool (summarized here in Table 4.3), we found three distinct Openness facets (being idealistic, adventurous, and intellectual) that were related both to adjective and NEO measures of the fifth factor (Soto & John, 2008). Moving away from a narrow Intellect interpretation, Saucier (1992, 1994) has suggested the label imagination, which is somewhat closer to Openness and emphasizes the emerging consensus that fantasy, ideas, and aesthetics, rather than intelligence, are most central to this factor. In this chapter, we therefore adopt the term *Openness*.

The Big Five Inventory: Measuring the Core Features of the Big Five with Short Phrases

To address the need for a short instrument measuring the prototypical components of the Big Five that are common across investigators, the Big Five Inventory (BFI) was constructed (John, Donahue, & Kentle, 1991; see also Benet-Martínez & John, 1998; John & Srivastava, 1999; Rammstedt & John, 2005, 2007). The 44-item BFI was developed to represent the Big Five prototype definitions described above (see Table 4.4)—a canonical representation of the factors intended to capture their core elements across the particulars of previous studies, samples, or instruments. The final items were selected on the basis of factor analyses in large samples of both junior college and public university students. Thus, Hampson and Goldberg (2006) were mistaken when they suggested that "John developed each of the five BFI scales to fall roughly between the lexical Big Five factors (Goldberg, 1992) and the five domain scores from the NEO PI-R" (p. 766), nor was this outcome either intended or entailed by the procedures used, as we will see below. The goal was to create a brief inventory that would allow efficient and flexible assessment of the five dimensions when there is no need for more differentiated measurement of individual facets. There is much to be said in favor of brevity: "Short scales not only save testing time, but also avoid subject boredom and fatigue. ... There are subjects ... from whom you won't get any response if the test looks too long" (Burisch, 1984, p. 219).

The BFI does not use single adjectives as items because such items are answered less consistently than when they are accompanied by definitions or elaborations (Goldberg & Kilkowski, 1985). Instead, the BFI uses short

phrases based on the trait adjectives known to be prototypical markers of the Big Five (John, 1989, 1990). One or two prototypical trait adjectives served as the item core to which elaborative, clarifying, or contextual information was added. For example, the Openness adjective *original* became the BFI item "Is original, comes up with new ideas" and the Conscientiousness adjective *persevering* served as the basis for the item "Perseveres until the task is finished." Thus the BFI items (which are reprinted here in Appendix 4.1) retain the advantages of adjectival items (brevity and simplicity) while avoiding some of their pitfalls (ambiguous or multiple meanings and salient desirability). Indeed, DeYoung (2006, p. 1140) hypothesized that with their more contextualized trait content, the BFI items should elicit higher interrater agreement than single-adjective items, and found that pairwise interrater agreement was indeed somewhat higher for the BFI.

Although the BFI scales include only eight to ten items, they do not sacrifice either content coverage or good psychometric properties. For example, the nine-item Agreeableness scale includes items related to at least five of the six facets postulated by Costa and McCrae (1992)—namely, trust (trusting, forgiving), altruism (helpful and unselfish), compliance (not quarrelsome), modesty (not faultfinding with others), and tender-mindedness (considerate and kind). In U.S. and Canadian samples, the alpha reliabilities of the BFI scales range from .75 to .90 and average above .80. Three-month test–retest reliabilities range from .80 to .90, with a mean of .85 (Rammstedt & John, 2005; 2007). In a middle-age sample, Hampson and Goldberg (2006) found a mean test–retest stability of .74, with stability correlations of .79 for Extraversion and Openness and about .70 for Agreeableness, Conscientiousness, and Neuroticism. Validity evidence includes substantial convergent and divergent relations with other Big Five instruments as well as with peer ratings (Rammstedt & John, 2005, 2007). DeYoung (2006) analyzed a large community data set with BFI self-reports and BFI ratings by three peers; however, he did not report validity correlations between self-reports and the aggregated peer ratings. We therefore reanalyzed these data and found validity correlations of .67 for Extraversion, .60 for Openness, .52 for Neuroticism, .48 for Agreeableness, and .47 for Conscientiousness, averaging .55. The sizes of these convergent correlations are even more impressive given that the (absolute) hetero-trait, hetero-method discriminant correlations averaged .09, and 19 of the 20 correlations were below .20, with only one reaching −.21 (indicating that individuals who described themselves as high in neuroticism were rated by their peers as slightly more disagreeable).

MEASUREMENT: COMPARING THREE BIG FIVE INSTRUMENTS

So far, we have discussed Goldberg's (1992) TDA, Costa and McCrae's (1992) NEO questionnaires, and the BFI. In addition, a variety of other measures are available to assess the Big Five in English and other languages (see De Raad & Perugini, 2002). Many of them were developed for specific research applications. Digman (e.g., 1989) constructed several different adjective sets to study teacher ratings of personality in children and adolescents, and Wiggins's (1995) scales were described above. Big Five scales have also been constructed using items from existing instruments. For example, scales were developed to measure the Big Five in adolescents using personality ratings on the California Child Q-sort obtained from their mothers (John et al., 1994). Measelle, John, Ablow, Cowan, and Cowan (2005) developed scales to measure the Big Five with a puppet interview in children ages 4–7. In behavior genetic research, Loehlin, McCrae, Costa, and John (1998) used Big Five scales constructed from the ACL (Gough & Heilbrun, 1983) and the CPI (Gough, 1987); for the latter, we (Soto & John, 2008) recently developed new Big Five domain and facet scales. As shown in Table 4.1, another broadband personality inventory that provides scores for the Big Five is the Hogan Personality Inventory (Hogan, 1986). The availability of so many different instruments to measure the Big Five makes clear that there is no single instrument that represents *the* gold standard.

Comparing the TDA, NEO-FFI, and BFI

In general, the NEO questionnaires represent the best-validated Big Five measures in the

questionnaire tradition. Goldberg's (1992) 100-item TDA and its abbreviated 40-item version (Saucier, 1994) are the most commonly used measures consisting of single adjectives. The BFI has been used frequently in research settings where subject time is at a premium, and its short-phrase item format provides more context than Goldberg's single-adjective items but less complexity than the sentence format used by the NEO questionnaires; the BFI items are also somewhat easier to understand (Benet-Martínez & John, 1998).

How well do these different Big Five measures converge? And are the five dimensions really independent? There has been concern that some of the Big Five dimensions are highly intercorrelated (Block, 1995; Eysenck, 1992). How high are these intercorrelations, and do they involve the same dimensions across instruments? A number of studies have reported on the psychometric characteristics of these instruments, and a few studies have compared two of them with each other (e.g., Benet-Martínez & John, 1998; DeYoung, 2006; Goldberg, 1992; McCrae & Costa, 1987). However, little is known about how all three compare to each other (see John & Srivastava, 1999, for an exception). To provide such a comparison, we summarize findings from a new large data set of self-reports on all three measures. The sample consisted of 829 undergraduates at the University of California, Berkeley (see John & Soto, 2007; Soto & John, 2008) who completed the BFI, Saucier's (1994) 40-item version of Goldberg's (1992) TDA, as well as Costa and McCrae's (1992) NEO-PI-R from which we scored both the 30 facets (see Table 4.3) and the NEO-FFI domain scores. The data thus represent a multitrait, multimethod (MTMM) design where the methods are the three Big Five self-report instruments (see John & Benet-Martínez, 2000).

Reliability of the Three Instruments

Overall, the coefficient alpha reliabilities, shown in Table 4.5, were impressive for these short scales, and relatively similar in size across instruments; the mean of the alphas was .84 for the TDA scales, .83 for the BFI, and .81 for the NEO-FFI, which had the longest scales (12 items compared to 8 for the TDA and about 9 for the BFI). Across instruments, Extraversion, Neuroticism, and Conscientiousness were measured most reliably (all clearly above .80 on all instruments), whereas Agreeableness and Openness tended to be somewhat less reliable. The scales with the lowest reliabilities were the NEO-FFI Openness and Agreeableness scale, similar to the values reported in the NEO-PI-R manuals and also replicating two other college samples (e.g., Benet-Martínez & John, 1998). Several NEO-FFI Openness items did not correlate well with the total scale, and these less reliable items came from particular openness facets, namely from openness to action (e.g., trying new and foreign foods) and from openness to values (e.g., looking to religious authorities for decisions on moral issues, reverse scored) (John & Srivastava, 1999).

Convergent Validity across the Three Instruments

Overall, we expected the convergent validities across the three instruments to be substantial. However, we already noted some potential differences in the way the three instruments define Extraversion and Openness. The NEO Extraversion domain had already been defined in terms of six facets before Costa and McCrae added domain scales for Agreeableness and Conscientiousness in 1985 and facet scales for these two factors in 1992. Thus, the warmth facet scale, included in the NEO-PI-R Extraversion domain (see Table 4.3), also correlates substantially with their Agreeableness domain scale (Costa & McCrae, 1992). In contrast, different lexical researchers (e.g., Goldberg, 1992; John, 1990; Saucier & Ostendorf, 1999) all found, in independent analyses, that trait adjectives related to warmth correlate more highly with Agreeableness than with Extraversion (see Tables 4.3 and 4.4). Thus, the NEO-FFI Extraversion scale showed much less convergence with either the TDA or the BFI than those two instruments with each other (John & Srivastava, 1999).

The other potential difference involves the fifth factor. As described above, Goldberg (1992) prefers to interpret this factor as Intellect or Imagination (Saucier, 1992), thus emphasizing openness to ideas and to fantasy over the other four facets. Similarly, the BFI Openness scale does not include items conceptually related to Costa and McCrae's

TABLE 4.5. Reliability and Convergent Validity Coefficients for Three Short Big Five Measures: Big Five Inventory, NEO Five-Factor Inventory, and Trait Descriptive Adjectives

Measures	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness	Mean
			Internal consistency			
BFI	.86	.79	.82	.87	.83	.83
NEO-FFI	.82	.75	.82	.87	.76	.81
TDA	.88	.84	.84	.83	.83	.84
Mean	.85	.80	.83	.85	.81	.83
			Uncorrected convergent validity correlations (across measures)			
BFI–TDA	.90	.75	.79	.70	.79	.80
BFI–NEO-FFI	.73	.76	.80	.81	.72	.77
TDA–NEO-FFI	.70	.66	.75	.64	.62	.68
Mean	.80	.73	.78	.73	.72	.75
			Corrected convergent validity correlations (across measures)			
BFI–TDA	.99	.93	.96	.82	.95	.95
BFI–NEO-FFI	.87	.99	.97	.94	.90	.95
TDA–NEO-FFI	.83	.83	.91	.76	.78	.83
Mean	.94	.95	.95	.86	.90	.93
			Standardized convergent validity coefficients from CFA (controlling for acquiescence factors)			
BFI–TDA	.99	.91	.91	.84	.97	.95
BFI–NEO-FFI	.83	.98	.95	.93	.90	.93
TDA–NEO-FFI	.76	.84	.87	.78	.74	.80
Mean	.92	.93	.92	.86	.91	.91

Note. $N = 829$ (see John & Soto, 2007). BFI, Big Five Inventory (John et al., 1991); TDA, Trait Descriptive Adjectives (Goldberg, 1992; 40-item mini-marker version, Saucier, 1994); NEO-FFI, NEO Five-Factor Inventory (Costa & McCrae, 1992); CFA, confirmatory factor analysis. Means are shown in **bold**. All means computed with Fisher *r*-to-*Z* transformations.

(1992) values and actions facets because preliminary BFI items, based on prototype items related to conventionality (i.e., relevant to the NEO-PI-R values facet) and behavioral flexibility (relevant to the action facet), failed to cohere with the other items on the BFI Openness scale (John et al., 1991). Thus, the NEO-FFI Openness scale showed less convergence with either the TDA or the BFI than those two instruments with each other (John & Srivastava, 1999).

As a first test of cross-instrument convergence, we examined the full 15 × 15 MTMM correlation matrix formed by the five factors crossed with the three instruments. The cross-instrument validity correlations, computed between pairs of instruments and shown in Table 4.5, were generally substantial in size. Across all five factors, the mean of the convergent validity correlations across instruments was .75, as compared with the much smaller discriminant correlations that averaged .19. As shown in Table 4.5, BFI and TDA showed the strongest overall convergence (mean $r = .80$), followed closely by BFI and NEO-FFI (.77), and finally TDA and NEO-FFI (mean $r = .68$).

To determine the extent to which the validity correlations simply reflect the imperfect reliability of the scales rather than substantive differences among the instruments, we corrected for attenuation using alpha. As shown in Table 4.5, the corrected validity correlations averaged .93. However, this excellent overall result masks some important differences. Across instruments, the first three of the Big Five (Extraversion, Agreeableness, and Conscientiousness) showed mean validities of about .95, suggesting very high equivalence of the reliable variance of the three instruments. However, Neuroticism (.86) and Openness (.90) were notably lower. Focusing on the pairwise comparisons between instruments, the patterns were more differentiated. BFI and TDA (corrected mean $r = .95$) shared virtually all of their reliable

variance, with the highest correlation for Extraversion; only the correlation for Neuroticism (.82) fell below .90. BFI and NEO-FFI showed the same substantial mean convergence (.95); here the highest correlation was for Agreeableness, and again only one correlation fell below .90 (for Extraversion). In contrast, TDA and NEO-FFI shared less in common (mean corrected r = .83); only one correlation (for Conscientiousness) exceeded .90, and those for Neuroticism and Openness did not even reach .80, suggesting that the conceptualization of four of the Big Five dimensions is not fully equivalent across these two instruments. On average, then, the BFI converged much better with both TDA and NEO-FFI than did TDA and NEO-FFI with each other. However, in contrast to Hampson and Goldberg's (2006) impression, the empirical findings show that the BFI does not simply occupy an intermediate position between the lexically derived TDA and the questionnaire-based NEO-FFI. Instead, the pattern of convergence correlations depends on the Big Five domain: The BFI achieved practical equivalence with the TDA for Extraversion (.99) but with the NEO-FFI for Agreeableness (.99), was much closer to the NEO-FFI than the TDA for Neuroticism (.94 vs. .82), a pattern that was reversed for Openness (.90 vs. .95), and finally converged equally well with both for Conscientiousness (.97 vs. .96).

Discriminant Correlations

Overall, discriminant correlations were low, with absolute values averaging .19 overall and .16 for the TDA and .20 for both the NEO-FFI and the BFI. Moreover, none of the discriminant correlations reached .35 on any of the instruments, and the largest correlations were .30 for the TDA, .34 for the NEO-FFI, and .31 for the BFI. Averaged across instruments, only four of the 10 discriminant correlations even exceeded .20: the mean correlation was .26 for Agreeableness and Conscientiousness, –.26 for Agreeableness and Neuroticism, –.26 for Conscientiousness and Neuroticism, and –.25 for Extraversion and Neuroticism. Thus, there was little support for Eysenck's (1992) contention that Agreeableness and Conscientiousness are highly correlated "primary" traits that should be combined into a broader dimension that contrasts Eysenck's Psychoticism with what might be called "good character." The size of these intercorrelations should also dampen some of the current enthusiasm (e.g., DeYoung, 2006; Markon, Krueger, & Watson, 2005) for higher-order factors above the Big Five (Digman, 1997). Yes, as has been noted repeatedly (e.g., John & Srivastava, 1999; Paulhus & John, 1998), the Big Five dimensions, as assessed by self and peer observers, are not strictly orthogonal, and scale intercorrelations of .26 are statistically significant. However, the size of these intercorrelations represents barely 10% shared variance—hardly enough, it would seem, to support the two substantively interpreted superordinate factors initially reported by Digman (1997). An alternative view showed that at least some of that covariance may be explained in terms of self-enhancing biases in self-reports (Paulhus & John, 1998).

Estimating Convergent and Discriminant Validity While Controlling Acquiescence

Finally, we tested whether individual differences in acquiescent response style (i.e., "yea-saying" vs. "nay-saying") might serve to influence cross-instrument validity estimates, such as inflating convergent validity correlations or depressing discriminant validity correlations (Soto, John, Gosling, & Potter, 2008). We found small but systematic acquiescence effects in the BFI; they appear as a small response-style factor in addition to the five substantive personality factors. We have therefore developed a new, content-balanced approach that controls acquiescence variance at the BFI item level and eliminates the response-style factor; this approach is described here in Appendix 4.2. For the present analyses, we used a bifactor approach to modeling an acquiescence factor, in addition to the five substantive personality factors, for each of the three instruments.

These structural equation models have a number of important properties. First, each instrument (BFI, TDA, NEO-FFI) has its own acquiescence factor, which was defined by setting the raw regression path from its acquiescence factor to each individual item equal to 1. By setting all of these loadings equally, we ensured that the acquiescence factor would represent *positive* covariance shared across all items on the instrument,

rather than letting it estimate something else (e.g., social desirability). By setting all of the loadings equal to 1 (rather than setting the factor variance equal to 1), we allow the variance of the acquiescence factor to be estimated and can thus compare the amount of acquiescence variance across instruments.

Second, the 15 substantive personality factors (Big Five times three instruments) were allowed to correlate with each other freely, both within and across each instrument, thus allowing us to estimate latent convergent and discriminant validity correlations across pairs of instruments. Also, the three acquiescence factors were allowed to correlate with each other across instruments, thus allowing us to test whether individual differences in acquiescent responding generalize across the Big Five instruments. Third, the substantive factors were not allowed to correlate with the acquiescence factors—further ensuring that the acquiescence factors did not contain any substantive personality variance.

We tested two general predictions about acquiescent responding. First, does the single-adjective item format of the TDA and its longer nine-step response scale elicit more acquiescence variance than the more contextualized item formats of the BFI and NEO-FFI, with their shorter five-step response scales? Indeed, the estimated variance of the acquiescence factor was .106 for the TDA, compared with .033 for the BFI and .013 for the FFI. Constraining the acquiescence factor variances to be equal across instruments significantly reduced fit; $\Delta\chi^2(2) = 375$, $p < .001$. Second, is acquiescent responding instrument-specific or a broad response disposition that generalizes across these instruments? Our findings suggested considerable generalizability; the correlations between the acquiescence factors were all significant: .62 for BFI–TDA, .61 for BFI–NEO-FFI, and .43 for TDA–FFI.

Third, what are the effects of including the acquiescence factors on the estimates of the convergent and discriminant correlations? The results for the convergent correlations with the acquiescence factors included (thus controlling the effects of acquiescence) are shown in Table 4.5. They were virtually identical to those without the acquiescence factors and, most important, the pattern was very similar to that for the corrected convergent validity correlations in Table 4.5, where acquiescence was not controlled. Including the acquiescence factors also did not change the mean estimated discriminant correlations, suggesting that at the scale level, the three Big Five instruments are not particularly susceptible to acquiescence effects. Together, the findings in this section show that the Big Five are fairly independent dimensions that can be measured by several instruments with impressive convergent and discriminant validity.

How Well Do BFI, TDA, and NEO-FFI Scales Represent the Six NEO-PI-R Facets?

The findings in Table 4.5 suggest that some pairs of scales define the intended Big Five domain in very similar ways (e.g., the BFI and TDA scales for Extraversion) whereas other pairs of scales do not (e.g., the TDA and NEO-FFI scales for Neuroticism). One way to explicate how the three Big Five instruments define each Big Five domain is to use the six facets included on the NEO-PI-R for each trait domain as a shared point of reference and correlate them, for each Big Five domain, with the scales on the three instruments. These correlations can then be graphed, yielding profiles that show how well each particular scale represents each of the six facets defined by the NEO-PI-R. The results are shown in Figure 4.2; we begin with Extraversion in the middle of the figure to illustrate how to read and interpret these profile graphs. In the Extraversion panel, it is immediately obvious that the profile curves for the BFI Extraversion scale and the TDA Extraversion scale are extremely similar. Both show correlations of about .55 with the four center facets (i.e., activity, gregariousness, warmth, and positive emotions), indicating that they weigh these facets all about equally. Both scales correlate most highly with assertiveness (about .70) and least highly with excitement-seeking (about .35), indicating that assertiveness is emphasized much more in their definition of Extraversion than is excitement-seeking, which is peripheral, at best. In contrast, the NEO-FFI Extraversion scale puts much more emphasis on positive emotions and warmth than do BFI and TDA, and less emphasis on assertiveness, as shown by the crossover pattern of their facet profiles. In other words, the NEO-FFI defines Extraversion as a somewhat different

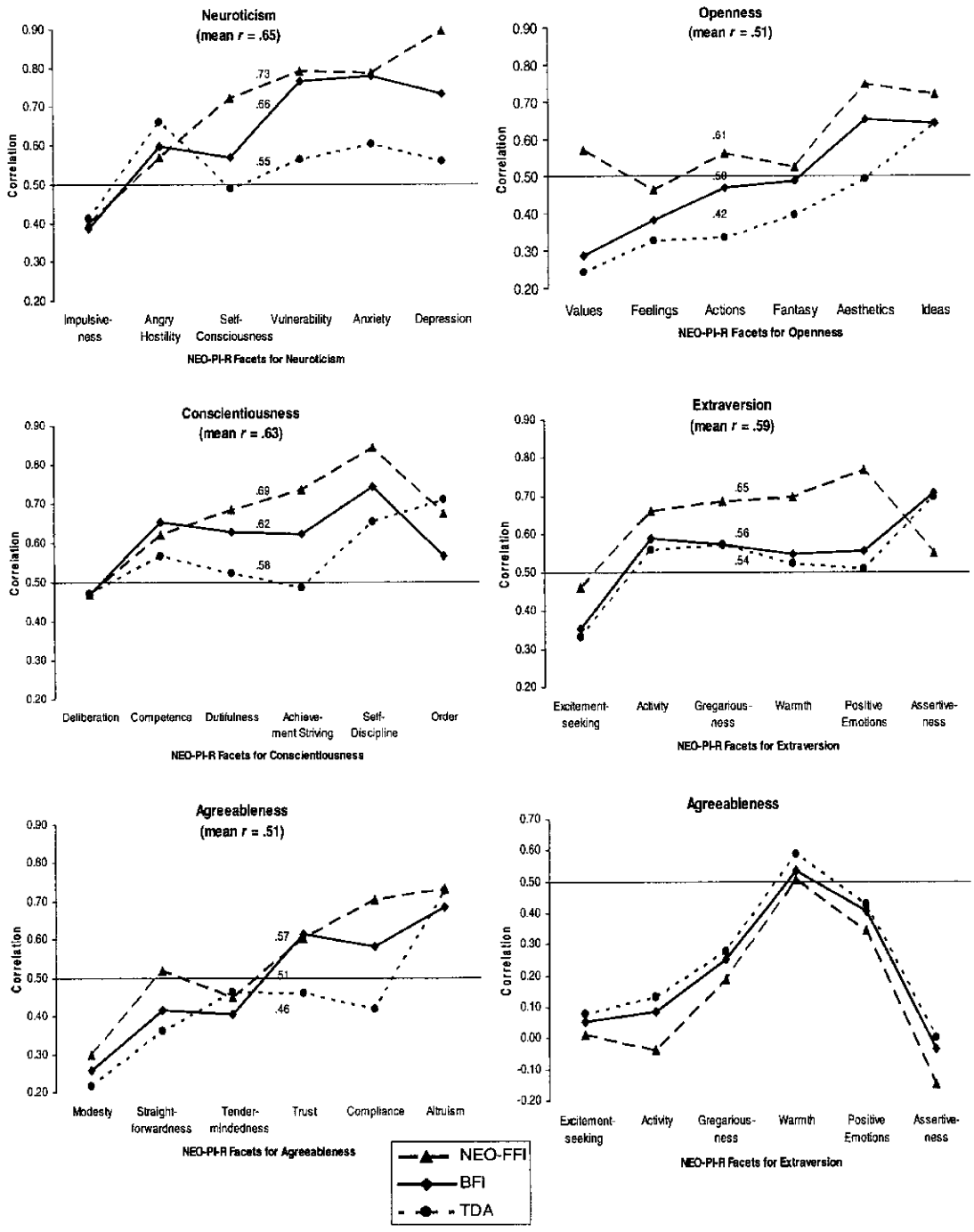


FIGURE 4.2. How do the three most commonly used, short Big Five instruments define each of the five broad trait domains? Profiles of the Big Five domains, as measured by the NEO Five-Factor Inventory (NEO-FFI), Big Five Inventory (BFI), and Trait Descriptive Adjectives (TDA) across the 30 facet traits defined by the NEO Personality Inventory—Revised (NEO-PI-R). *Note*: For each Big Five domain, the figure shows the convergent correlations of the scales of the three instruments, each shown as a profile curve, with the six NEO-PI-R facets for that domain. The average correlation between the domain scales from each Big Five instrument and the six NEO-PI-R facets is presented next to their respective profile curve. For the NEO-FFI (but not the other two instruments), these coefficients represent part–whole correlations; everything else being equal, the NEO-FFI profile curve should thus always be higher than the curves for the other two instruments (BFI and TDA). Therefore, the absolute elevation of the three profile curves is of less interest than their shape—that is, their relative similarity and differences. To illustrate that the interpersonal dimensions of Extraversion and Agreeableness show some overlap (see also Figure 4.3), the sixth panel (on the lower right) shows the discriminant validity correlations of the three Agreeableness domain scales with the Extraversion facets from the NEO-PI-R.

mixture of lower-level personality attributes than do the BFI and TDA.

We should comment on other features of these facet profiles. First, consider the overall elevation of these three Extraversion profiles, which is summarized by the mean correlation presented next to each profile line: .65 (highest overall) for the NEO-FFI, followed by .56 for the BFI and .54 for the TDA. The NEO-FFI's highest elevation here (and everywhere else in this figure) is due, in part, to partial item overlap and cannot be interpreted unambiguously—after all, the 60 items that make up the 12-item NEO-FFI scales are also included in the 30 facet scales that we used to generate the correlations for Figure 4.2. Thus, whereas the absolute elevation is informative for the BFI and TDA, it is less so for the NEO-FFI, and the shape of the profile is much more critical, like the crossover for assertiveness. Second, these graphs also tell us something about the facets that define the profiles. Note that all but one facet had correlations with all three Extraversion scales above the .50 line, suggesting that they are all substantially relevant to the way Extraversion is conceptualized on all three instruments. The finding that excitement-seeking had, by far, the lowest correlations with all three Extraversion scales suggests that it is a relatively peripheral facet within the Extraversion domain. Apparently, the NEO-PI-R facets are not all equal in their centrality to their Big Five domains.

This point is even more apparent in the Agreeableness panel on the lower left. The BFI, TDA, and NEO-FFI Agreeableness scales all correlate about .70 with the altruism facet, showing remarkable Agreement about the centrality of this facet to this domain. In contrast, all three correlated less than .30 with modesty, suggesting that this facet is peripheral to the Agreeableness domain. In terms of profile similarity, BFI and NEO-FFI are much closer to each other than either is to the TDA, especially for trust and compliance, which seem relatively underrepresented on the TDA. Finally, consider the sixth panel (on the lower right), which is the one panel that shows discriminant validity correlations, relating the three Agreeableness scales with the facets from the NEO-PI-R Extraversion domain. Here the Agreeableness scales show impressive profile similarity, indicating that all three correlate above .50 with warmth and about .40 with positive emotions.

This fuzzy boundary between the Extraversion and Agreeableness domain is also illustrated in Figure 4.3 which shows the Big Five trait version of the interpersonal circumplex (Wiggins, 1979). This figure shows factor loadings for the Extraversion and Agreeableness scales from the NEO-FFI, BFI, and TDA, which were factored along with the six NEO-PI-R facets defining Extraversion, the six defining Agreeableness, and the angry hostility facet from the Neuroticism domain, because it is also highly negatively related to Agreeableness (as indicated by its substantial negative loading on that factor). As we noted earlier, warmth and positive emotions are the Extraversion facets with the largest (positive) loadings on Agreeableness. The Agreeableness facets such as trust and altruism have positive loadings on Extraversion, whereas modesty and compliance have negative loadings. Finally, the locations of the TDA, BFI, and NEO-FFI scales are also compatible with our earlier observations. BFI and TDA Extraversion are almost in exactly the same spot in this two-dimensional space, and certainly closer to assertiveness than is NEO-FFI Extraversion, which is rotated to the right toward Agreeableness and thus closer to positive emotions and warmth. Both TDA and BFI Agreeableness are located very close to the altruism facet, which turned out to be so central to all three Agreeableness scales in Figure 4.2.

Although two-dimensional plots of factor loadings, such as those in Figure 4.3, have long been used in trait taxonomic and especially circumplex work, they seem to offer less specific information than the facet profiles in Figure 4.2. For Openness, for example, the biggest difference involves the value facet, which seems barely represented on either TDA or BFI. In contrast, for Neuroticism it is the underweighting of depression, anxiety, and vulnerability and the relative overweighting of hostility that make the TDA scale so different from both the BFI and NEO-FFI Neuroticism scales.

Big Five Measurement: Conclusions and Limitations

One of the limitations of the findings presented here is that we did not examine external (or predictive) validity. However, both the NEO questionnaires and the BFI have

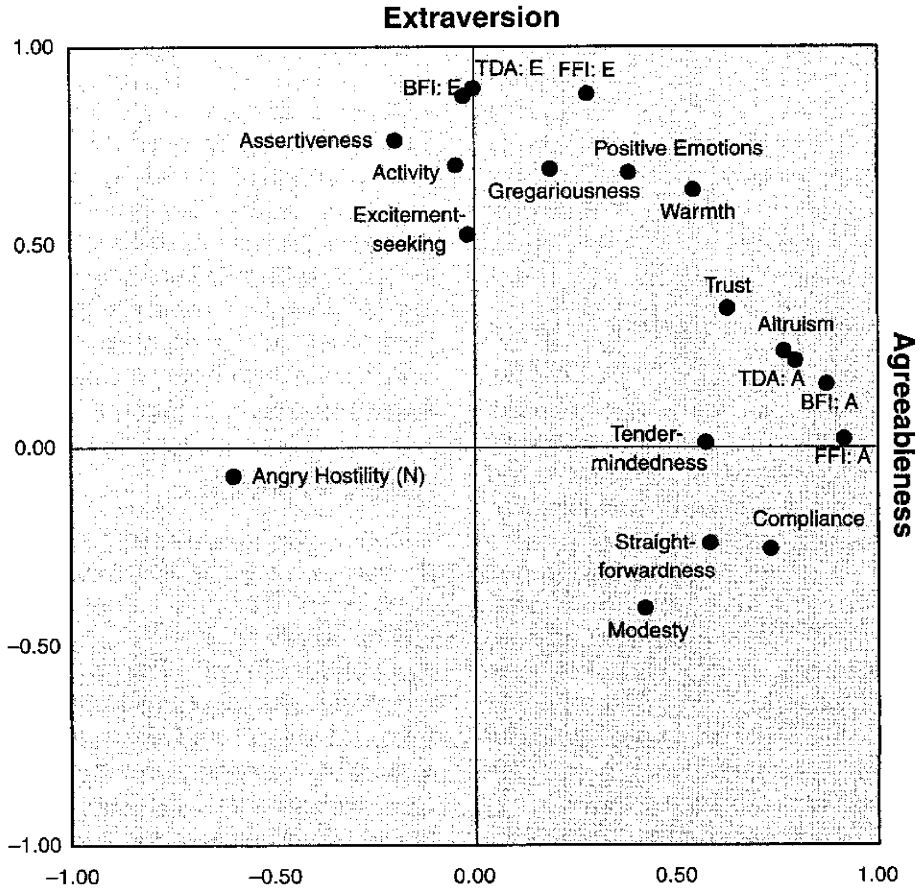


FIGURE 4.3. The interpersonal circumplex formed by the Big Five domains of Extraversion and Agreeableness. The Extraversion and Agreeableness domain scales from three Big Five instruments (NEO-FFI, BFI, and TDA) were factored along with the six NEO-PI-R facets defining Extraversion, the six defining Agreeableness, and the angry hostility facet from the Neuroticism domain (because it is also highly related to Agreeableness). FFI, NEO Five-Factor Inventory; BFI, Big Five Inventory; TDA, Trait Descriptive Adjectives; E, Extraversion; A, Agreeableness; N, Neuroticism.

been shown to predict peer ratings; and initial evidence is now available for the TDA scales (DeYoung, 2006). Future research needs to study the comparative validity of all three instruments using peer ratings and other external criteria. One of the advantages of the BFI is its efficiency, taking only about 5 minutes of administration time, compared with about 15 minutes for the NEO-FFI and the 100-item TDA. Moreover, the BFI items are shorter and easier to understand than the NEO-FFI items (Benet-Martínez & John, 1998; Soto et al., 2008). The 100 (or 40) adjectives on the TDA are even shorter; however, single-trait adjectives can be ambiguous in their meanings.

When should researchers use each of these instruments? When participant time is not at a premium, participants are well educated and test-savvy, and the research question calls for the assessment of multiple facets for the Big Five, then the full 240-item NEO-PI-R would be most useful. Otherwise, the 44-item BFI would seem to offer a measure of the core attributes of the Big Five that is at least as efficient and easily understood as the 60-item NEO-FFI and the 100-item TDA. At this point, we cannot recommend the use of even shorter instruments, with as few as 10 items (e.g., Gosling, Rentfrow, & Swann, 2003; Rammstedt & John, 2007), unless a researcher encounters truly exceptional circumstances, such as the need to measure the Big Five as part of a national phone survey. From our perspective, the gains in time achieved by moving from a measure such as

the BFI (i.e., 5 minutes of subject time) to an even shorter measure can rarely compensate for the potential losses in reliability and validity one has to risk with such minimalist measurement.

FACTOR NAMES, NUMBERS, OR INITIALS: WHICH SHALL WE USE?

Problems with the English Factor Labels

Now that we have considered both the history of the Big Five and their measurement, it is time to revisit the names or labels assigned to the factors (see Table 4.2). Although the constructs that will eventually replace the current Big Five may be different from what we know now, labels are important because they imply particular interpretations and thus influence the directions that theorizing might take. Norman's (1963) factor labels have been frequently used, but Norman offered little theoretical rationale for their selection. Norman's labels differ vastly in their breadth or inclusiveness (Hampson, Goldberg, & John, 1987); in particular, Conscientiousness and Culture are much too narrow to capture the enormous breadth of these two dimensions. Moreover, researchers have abandoned Culture as a label for Factor V, in favor of Intellect or Imagination (Saucier & Goldberg, 1996a) or Openness to Experience (McCrae & Costa, 1985b). Neither label is truly satisfactory, however, because Intellect is too narrow and Openness, while broad enough, is somewhat vague.

Agreeableness is another problematic label. For one, it refers to the behavioral tendency to *agree* with others, thus incorrectly implying submissiveness, which is more closely related to the introverted pole of Factor I. Agreeableness is also too detached, too neutral a label for a factor supposed to capture intensely affective characteristics, such as love, compassion, and sympathy. Freud viewed love and work as central; following this lead, we could call Factor II simply *Love* (Peabody & Goldberg, 1989).

However, *Work* is too narrow a label for Factor III. Even Conscientiousness is too narrow because it omits a central component that Peabody and Goldberg (1989) called "favorable impulse control." Thus, Responsibility or even Degree of Socialization (see Gough, 1987) might be labels more appropriate for Factor III than Conscientiousness.

More could be said about the many shortcomings of the traditional labels, but better labels are hard to come by. The unsurpassed advantage of the traditional labels is that they are commonly known and used, thus preventing Babel from taking over the literature on the Big Five. Moreover, before any new names are devised, the definition of the factors in terms of facets or components needs to be elaborated and sharpened. At this point, it seems premature to settle the scope and theoretical interpretation of the factors by devising new names.

Preliminary Definitions

Because the traditional labels are so easily misunderstood, we provide short definitions of the five dimensions in Table 4.2 (cf., Costa & McCrae, 1992; John, 1990; Tellegen, 1985). Briefly, Extraversion implies an *energetic approach* toward the social and material world and includes traits such as sociability, activity, assertiveness, and positive emotionality. Agreeableness contrasts a *prosocial and communal orientation* toward others with antagonism and includes traits such as altruism, tender-mindedness, trust, and modesty. Conscientiousness describes *socially prescribed impulse control* that facilitates task- and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing, and prioritizing tasks. Neuroticism contrasts emotional stability and even-temperedness with *negative emotionality*, such as feeling anxious, nervous, sad, and tense. Finally, Openness to Experience (vs. closed-mindedness) describes the breadth, depth, originality, and complexity of an individual's *mental and experiential life*.

The numbering convention from I to V, favored by Saucier and Goldberg (1996b) and Hofstee and colleagues (1997), is useful because it reflects the relative size of the factors in lexical studies. Factors I and II, which primarily summarize traits of an interpersonal nature, tend to account for the largest percentage of variance in personality ratings, followed by Factor III, whereas the last two factors are the smallest, by far, in lexical studies (De Raad et al., 1998). However, the Roman numerals are hard to remember, and the order of the factors is not invariant across studies. Thus, we favor the mnemonic

convention suggested by the initials given in Table 4.2. These initials evoke multiple associations that represent more fully than a single word the broad range of meaning captured by each of the factors: *E* stands for Extraversion, Energy, or Enthusiasm; *A* for Agreeableness, Altruism, or Affection; C for Conscientiousness, Control, or Constraint; *N* for Neuroticism, Negative Affectivity, or Nervousness; and O for Openness, Originality, or Open-Mindedness. The reader intrigued by anagrams may have noticed that these letters form the OCEAN of personality dimensions.

CONVERGENCE BETWEEN THE BIG FIVE AND OTHER STRUCTURAL MODELS

McCrae and Costa's (1985a, 1985b, 1985c, 1987) findings, like the cross-instrument convergence described above, showed that the factor-analytic results from the lexical tradition converged surprisingly well with those from the questionnaire tradition. This convergence eventually led to a dramatic change in the acceptance of the five factors in the field. With regard to their empirical status, the findings accumulated since the mid-1980s show that the five factors replicate across different types of subjects, raters, and data sources, in both dictionary-based and questionnaire-based studies. Indeed, even more skeptical reviewers were led to conclude that "Agreement among these descriptive studies with respect to *what* are the appropriate dimensions is impressive" (Revelle, 1987, p. 437; see also McAdams, 1992). The finding that it doesn't matter whether Conscientiousness is measured with trait adjectives, short phrases, or questionnaire items suggests that the Big Five dimensions have the same conceptual status as other personality constructs. For example, Loehlin and colleagues (1998) found that all five factors show substantial and about equal heritabilities, regardless of whether they are measured with questionnaires or with adjective scales derived from the lexical approach.

One of the great strengths of the Big Five taxonomy is that it can capture, at a broad level of abstraction, the commonalities among most of the existing systems of personality traits, thus providing an integrative descriptive model for research. Table 4.1 summarizes the personality dimensions proposed by a broad range of personality theorists and researchers. These dimensions, although by no means a complete tabulation, emphasize the diversity of current conceptions of personality. However, they also point to some important convergences. First, almost every one of the theorists includes a dimension akin to Extraversion. Although the labels and exact definitions vary, nobody seems to doubt the fundamental importance of this dimension. The second almost universally accepted personality dimension is Emotional Stability, as contrasted with Neuroticism, Negative Emotionality, and Proneness to Anxiety. Interestingly, however, not all the researchers listed in Table 4.1 include a separate measure for this dimension. This is particularly true of the interpersonal approaches, such as Wiggins's and Bales's, as well as the questionnaires primarily aimed at the assessment of basically healthy, well-functioning adults, such as Gough's CPI, the MBTI, and even Jackson's Personality Research Form (PRF; 1984). In contrast, all of the temperament-based models include Neuroticism. There is less agreement on the third dimension, which appears in various guises, such as Control, Constraint, Super-Ego Strength, or Work Orientation, as contrasted with Impulsivity, Psychoticism, or Play Orientation. The theme underlying most of these concepts involves the control, or moderation, of impulses in a normatively and socially appropriate way (cf. Block & Block, 1980). However, Table 4.1 also points to the importance of Agreeableness and Openness, which are neglected by temperament-oriented theorists such as A. H. Buss and Plomin, Eysenck, and Zuckerman. In a comprehensive taxonomy, even at the broadest level, we need a "place" for an interpersonal dimension related to Communion, Feeling Orientation, Altruism, Nurturance, Love Styles, and Social Closeness, as contrasted with Hostility, Anger Proneness, and Narcissism. The existence of these questionnaire scales, and the cross-cultural work on the interpersonal origin and consequences of personality, stress the need for a broad domain akin to Agreeableness, Warmth, or Love.

Similar arguments apply to the fifth and last factor included in the Big Five. For one, there are the concepts of Creativity, Originality, and Cognitive Complexity, which are measured by numerous questionnaire scales (Gough, 1979; Helson, 1967, 1985). Al-

though these concepts are cognitive, or, more appropriately, *mental* in nature, they are clearly different from IQ. Second, limited-domain scales measuring concepts such as Absorption, Fantasy Proneness, Need for Cognition, Private Self-Consciousness, Independence, and Autonomy would be difficult to subsume under Extraversion, Neuroticism, or Conscientiousness. Indeed, the fifth factor is necessary because individual differences in intellectual and creative functioning underlie artistic interests and performances, inventions and innovation, and even humor. Individual differences in these domains of human behavior and experience cannot be neglected by personality psychologists.

Finally, the matches between the Big Five and other constructs noted in Table 4.1 should be considered with a healthy dose of skepticism. Some of these correspondences are indeed based on solid research findings, but others are conceptually derived and await empirical confirmation. These matches reflect broad similarities, ignoring some important, implicative, and useful differences among the concepts proposed by different investigators. Nonetheless, at this stage in the field, we are more impressed by the newly apparent similarities than by the continuing differences among the various models. Indeed, the Big Five are useful primarily because of their integrative and heuristic value, a value that becomes apparent in Table 4.1. The availability of a taxonomy, even one that is as broad and incomplete as the Big Five, permits the comparison and potential integration of dimensions that, by their names alone, would seem entirely disparate.

CRITICAL ISSUES AND THEORETICAL PERSPECTIVES

Like any scientific model, the Big Five taxonomy has limitations. Critics have argued that the Big Five does not provide a complete theory of personality (e.g., Block, 1995; Eysenck, 1997; McAdams, 1992; Pervin, 1994), and we agree. In contrast to McCrae and Costa's (1996; see also McCrae & Costa, Chapter 5, this volume) five-factor theory, the Big Five taxonomy was never intended as a comprehensive personality theory; it was developed to account for the structural relations among personality traits (Goldberg, 1993). Thus, like most structural models it provides an account of personality that is primarily descriptive rather than explanatory, emphasizes regularities in behavior rather than inferred dynamic and developmental processes, and focuses on variables rather than on individuals or types of individuals (cf. John & Robins, 1993, 1998). Nonetheless, the Big Five taxonomy of trait terms provides a conceptual foundation that helps us examine these theoretical issues. In this section, we begin with the hierarchical structure defined by the Big Five and then review how the Big Five predict important life outcomes, how they develop, how they combine into personality types, and how different researchers view their conceptual status.

Hierarchy, Levels of Abstraction, and the Big Five

A frequent objection to the Big Five is that five dimensions cannot possibly capture all of the variation in human personality (e.g., Block, 1995; McAdams, 1992; Mershon & Gorsuch, 1988), and that they are much too broad. However, the objection that five dimensions are too few overlooks the fact that personality can be conceptualized at different levels of abstraction or breadth. Indeed, many trait domains are hierarchically structured (Hampson, John, & Goldberg, 1986).

The advantage of categories as broad as the Big Five is their enormous bandwidth. Their disadvantage, of course, is their low fidelity. In any hierarchical representation, one always loses information as one moves up the hierarchical levels. For example, categorizing something as a guppy is more informative than categorizing it as a fish, which in turn is more informative than categorizing it as a vertebrate. Or, in psychometric terms, one necessarily loses item information as one aggregates items into scales, and one loses scale information as one aggregates scales into factors (John, Hampson, & Goldberg, 1991).

The Big Five dimensions represent a rather broad level in the hierarchy of personality descriptors. In that sense, they are to personality what the categories "plant" and "animal" are to the world of biological objects—extremely useful for some initial rough distinctions but of less value for predicting specific behaviors of a particular individual. The hierarchical level a researcher selects depends on the descriptive and pre-

dictive tasks to be addressed (Hampson et al., 1986). In principle, the number of specific distinctions one can make in the description of an individual is infinite, limited only by one's objectives.

Norman, Goldberg, McCrae and Costa, and Hogan all recognized that there was a need in personality, just as in biology, "to have a system in which different levels of generality or inclusion are recognized" (Simpson, 1961, p. 12). A complete trait taxonomy must include middle-level categories, such as assertiveness, orderliness, and creativity, and even narrower descriptors, such as talkative, punctual, and musical (John et al., 1991). At this point, Costa and McCrae's (1992) 30 facets, shown in Table 4.4, represent the most widely used and empirically validated model. Soto and John (2008) developed 16 facets (see Table 4.4) from the CPI item pool that tend to be broader than Costa and McCrae's facets. Hofstee and colleagues' (1992) circumplex-based AB5C approach defines 45 facets as unique, pairwise combinations or "blends" of the Big Five factors (e.g., Poise as a blend of high Extraversion and low Neuroticism); measures of these facets have now developed as part of Goldberg's collaborative and Web-based International Personality Item Pool (IPIP) project. Saucier and Ostendorf (1999) provide a thoughtful discussion of the fundamental issues in developing empirically based facets and present 18 facets (see Table 4.4) that show initial cross-language generalizability. Although Table 4.4 shows some promising convergences, the three approaches differ substantially in the number and nature of the facets they propose, indicating that further conceptual and empirical work is needed to achieve a consensual specification of the Big Five factors at lower level of abstraction.

Person–Environment Interactions: Do the Big Five Predict Important Life Outcomes?

Given that the Big Five dimensions were derived initially from analyses of the personality lexicon, one might wonder whether they merely represent linguistic artifacts. Do the Big Five actually predict important behavioral and life outcomes in people's lives? Initially, external validity and predictive utility did not receive much attention from researchers working in the Big Five tradition. Indeed, Eysenck (1991) challenged the field, arguing that "Little is known about the social relevance and importance of openness, agreeableness, and conscientiousness. ... What is lacking is a series of large-scale studies which would flesh out such possibilities" (p. 785).

Over the past two decades, however, researchers have taken up the task of identifying the particular Big Five dimensions that predict particular life outcomes in such fundamental domains as physical and mental health, work, and relationships (see Table 4.2 for examples). This research is based on the assumption that personal factors (such as the individual's traits) and environmental factors (such as aspects of a job or a relationship partner) interact to jointly produce behavioral and experiential outcomes that accumulate over the individual's lifespan (e.g., Caspi & Bem, 1990; Scarr & McCartney, 1984). In other words, personality traits are important because they influence the way individuals interact with particular environments. As we review below, traits influence how individuals construe and interpret the personal meaning a particular environment or situation has for them (e.g., how they interpret a potential health risk), and to which aspects of the environment they attend (e.g., a doctor's prescribed treatment regimen). In addition to these cognitive processes of perceiving and attending to the environment, traits also influence the way individuals select both social and nonsocial environments (e.g., college classes, jobs, places to live, relationship partners, even music) and how they then modify those environments (e.g., their bedrooms). It is through their systematic interaction with environmental affordances and risks that traits are hypothesized to influence the behavioral, emotional, social, and material life outcomes of the individual.

Links to Health, Health Behaviors, and Longevity

Although research on personality and health has a long tradition in the field, the emergence of the Big Five taxonomy has greatly helped clarify and organize the links between personality, health behaviors, illness, and mortality across the lifespan. Multiple studies have provided converging evidence that Conscientiousness, for example, predicts good health habits, health outcomes, and longevity (for a review, see Hampson

& Friedman, Chapter 31, this volume). For example, low Conscientiousness predicts the likelihood of engaging in risky behaviors such as smoking, substance abuse, and poor diet and exercise habits (Bogg & Roberts, 2004; Hampson, Andrews, Barckley, Lichtenstein, & Lee, 2000; Trull & Sher, 1994). Moreover, highly conscientious individuals, when diagnosed with an illness, are more likely to adhere to their treatment regimens (Kenford et al., 2002) and have been shown to live longer lives (Danner, Snowden, & Friesen, 2001; Friedman, Hawley, & Tucker, 1994; Weiss & Costa, 2005). Other Big Five dimensions are also related to health-related risk factors. Low Agreeableness (especially hostility) predicts cardiovascular disease (Miller, Smith, Turner, Guijarro, & Hallet, 1996). High Neuroticism predicts less successful coping and poorer reactions to illness, in part because highly neurotic individuals are more likely to ruminate about their situation (David & Suls, 1999; Scheier & Carver, 1993). Individuals high in Extraversion, on the other hand, have available more social support and close relationships important for coping with illness (Berkman, Glass, Brissette, & Seeman, 2000).

Links to Psychopathology, Personality Disorders, and Adjustment Problems

The availability of the Big Five taxonomy has also renewed interest in the links between personality and psychopathology, especially personality disorders (e.g., Costa & Widiger, 2002; Wiggins & Pincus, 1989); findings from this burgeoning literature have been reviewed by Krueger and Tackett (2006) and Widiger and Smith (Chapter 30, this volume). From a developmental perspective, the Big Five dimensions may serve as risks or buffers for subsequent adjustment problems. In adolescents, for example, both low Agreeableness and low Conscientiousness predict delinquency and externalizing problems, whereas high Neuroticism and low Conscientiousness predict internalizing problems such as depression and anxiety (John et al., 1994; Measelle et al., 2005; Robins et al., 1994; 1996). Low Conscientiousness is also the personality trait most strongly linked to ADHD, at least when diagnosed in adulthood (Nigg et al., 2002); more specifically, low Conscientiousness predicts attentional and organizational problems that can lead to broader adjustment problems in school settings and even relationships. Ultimately, findings linking Big Five profiles to developmental or adjustment problems may help us identify children at risk and design appropriate interventions, such as teaching children relevant behaviors and skills (e.g., strategies for delaying gratification).

Links to Academic and Work Outcomes

Industrial and organizational researchers have also rediscovered the importance of personality traits, and a growing body of research has linked the Big Five to academic and work achievement. Early studies showed that Conscientiousness as well as Openness predict school performance as measured with objective tests in early adolescence (John et al., 1994; Robins et al., 1994). In college, Conscientiousness predicts higher academic grade-point averages (Noftle & Robins, 2007; Paunonen, 2003), whereas Openness predicts the total years of education completed by middle adulthood (Goldberg, Sweeney, Merenda, & Hughes, 1998).

Beyond primary and secondary schooling, Conscientiousness has emerged also as a general predictor of job performance across a wide range of jobs (for reviews, see Barrick & Mount, 1991; Mount, Barrick, & Stewart, 1998). The other Big Five dimensions relate to more specific aspects of job performance, such as better performance or satisfaction in specific job types or positions. For example, Agreeableness and Neuroticism predict performance in jobs where employees work in groups, Extraversion predicts success in sales and management positions, Openness predicts success in artistic jobs, and Conscientiousness predicts success in conventional jobs (Barrick, Mount, & Gupta, 2003; Larson, Rottinghaus, & Borgen, 2002). Neuroticism is an important predictor of job satisfaction. Highly neurotic individuals are more likely to experience burnout and to change jobs, whereas more emotionally stable individuals feel satisfied and committed to their organizations (Thoresen, Kaplan, Barsky, Warren, & de Chermont, 2003). These trait-by-job interactions help researchers develop a more fine-grained understanding of how different traits are instrumental to performance and satisfaction in various job environments.

Links to Social Outcomes in Relationship and Group Contexts

The Big Five dimensions are also relevant for social behaviors and experiences, such as relationship maintenance and satisfaction, both in dyadic relationships and in groups. In terms of family relationships, adolescents high in Neuroticism, low in Conscientiousness, and low in Extraversion tend to have poorer relationships with parents (Belsky, Jaffee, Caspi, Moffitt, & Silva, 2003). Individuals low in Agreeableness and Extraversion are more likely to experience peer rejection (Newcomb, Bukowski, & Pattee, 1993). Extraversion, Conscientiousness, and low Neuroticism predict greater relationship satisfaction and less conflict, abuse, or dissolution (Karney & Bradbury, 1995; Robins, Caspi, & Moffitt, 2002; Watson, Hubbard, & Wiese, 2000). Because most social groups develop formal or informal status hierarchies, one important social goal is to attain status (respect, influence, and prominence) in one's social groups. Across several types of groups (Anderson, John, Keltner, & Kring, 2001), Extraversion substantially predicted higher status attainment for both sexes; high Neuroticism, incompatible with male gender norms, predicted lower status only in men. Consistent with these findings, extraverted individuals were more likely to be chosen as the jury foreperson (J. Clark, Boccaccini, Caillouet, & Chaplin, 2007) and more likely to have a firmer "power" handshake (Chaplin, Phillips, Brown, Clanton, & Stein, 2000).

Whereas status and leadership positions are not chosen by the individual but awarded by peers or coworkers, individuals do select or modify their environments. Individuals differ in their selection and modification efforts and successes, and these individual differences are related to personality traits. For example, several Big Five dimensions predict how and where people spend their time. In a study monitoring students' daily life (Mehl, Gosling, & Pennebaker, 2006), students high in Conscientiousness spent more time in the classroom or on campus; students high in Openness spent more time in coffee houses and restaurants; and students high in Extraversion engaged in more conversations and spent less time alone. In a series of studies relating independent assessments of bedrooms and offices to their inhabitants' personality traits (Gosling, Ko, Mannarelli, & Morris, 2002), the Big Five dimensions predicted how individuals shape and modify their physical environments: Highly open individuals tend to create distinctive-looking work and living spaces (with a large variety of different kinds of books), whereas highly conscientious individuals tend to keep their rooms well-organized and clutter-free. High Openness also predicts the expression of a wider variety of interests (e.g., complex music genres) and preferences in personal and social networking websites (Rentfrow & Gosling, 2006; Vazire & Gosling, 2004).

This brief review of Big Five research on person–environment interactions illustrates that the nomological network emerging for each of the Big Five domains now includes an ever-broadening range of life outcome variables. These findings have been summarized in greater detail in several reviews (e.g., Graziano & Eisenberg, 1997; Hogan & Ones, 1997; McCrae, 1996; Ozer & Benet-Martínez, 2006; Watson & Clark, 1997). In interpreting these findings, however, it is important to realize that although individual differences in personality traits are relatively stable over time (e.g., Roberts & DelVecchio, 2000), they are not fixed or "set like plaster" (e.g., Srivastava, John, Gosling, & Potter, 2003). Many people have the capacity to change their patterns of behavior, thought, and feeling, for example, as a result of therapy or intervention programs. Thus, the links between the Big Five and the life outcomes reviewed above are neither fixed nor inevitable for the individual. Instead, they point to critical domains of behavior and emotion that the individual may target for personal development and change. In the health domain, for example, people can improve how conscientiously they adhere to a diet, exercise regimen, or medical treatment plan (Friedman et al., 1994), thus greatly influencing their ultimate health outcomes and longevity.

The Big Five and Personality Development

Historically, personality psychology has concerned itself with a range of developmental issues that are relevant to the Big Five: the antecedents of adult personality traits, how traits develop, the timelines for the emer-

gence and peak expression of traits, their stability or change throughout the lifespan, and the effects of traits on other aspects of personal development. Some critics have suggested that Big Five researchers have not paid enough attention to issues of personality development in childhood and adolescence (Pervin, 1994). This criticism has some merit: Although the Big Five taxonomy has influenced research on adult development and aging (e.g., Helson, Kwan, John, & Jones, 2002; McCrae & Costa, 2003; Roberts, Walton, & Viechtbauer, 2006; see also Roberts, Wood, & Caspi, Chapter 14, this volume), there has been much less research on personality structure in childhood. Developmental and temperament psychologists have studied a number of important traits (e.g., sociability, fearful distress, shyness, impulsivity), but many studies examine one trait at a time, in isolation from the others, and the available research has not been integrated in a coherent taxonomic framework. Until this work is done, however, research on personality development across the lifespan is likely to remain fragmented (Halverson, Kohnstamm, & Martin, 1994).

The adult personality taxonomy defined by the Big Five can offer some promising leads. In our view, the Big Five should be examined in developmental research for two reasons (John et al., 1994). Theoretically, it may be necessary to examine the developmental origins of the Big Five: Given that the Big Five emerge as basic dimensions of personality in adulthood, researchers need to explain how they develop. Practically, the Big Five have proven useful as a framework for organizing findings on adult personality in areas as diverse as behavior genetics and industrial psychology. Thus, extension of the Big Five into childhood and adolescence would facilitate comparisons across developmental periods.

Work on these issues has now begun, and researchers are drawing on existing models of infant and child temperament to make connections to the Big Five dimensions in adulthood (for reviews, see Caspi, Roberts, & Shiner, 2005; De Clerq, De Fruyt, Van Leeuwen, & Mervielde, 2006; Halverson et al., 1994; Shiner & Caspi, 2003). Some research suggests that the Big Five may provide a good approximation of personality structure in childhood and adolescence (Digman, 1989; Graziano & Ward, 1992). Extending Digman's (1989) earlier work on Hawaiian children, Digman and Shmelyov (1996) examined both temperament dimensions and personality dimensions in a sample of Russian children. Based on analyses of teachers' ratings, they concluded that the Big Five offer a useful model for describing the structure of temperament. Studies using free-response techniques found that the Big Five can account for a substantial portion of children's descriptions of their own and others' personalities (Donahue, 1994), as well as teachers' and parents' descriptions of children's personality (Kohnstamm, Halverson, Mervielde, & Havill, 1998). In a large Internet sample providing self-reports on the BFI, we analyzed data from children and adolescents, ages 10–20, and found substantial changes in the coherence and differentiation of the Big Five domains, even though a variant of the adult Big Five factor structure was apparent as early as age 10 when individual differences in response acquiescence were controlled (see Appendix 4.2; Soto et al., 2008). In an even younger sample, Measelle and colleagues (2005) found that using an age-appropriate puppet interview task, children as young as ages 5–7 were able to self-report on their personality; by age 6, their self-reports were beginning to show evidence of coherence, longitudinal stability, and external validity (using ratings by adults) for the Big Five domains of Extraversion, Agreeableness, and Conscientiousness.

Two large-scale studies suggest that the picture may be more complicated. One study tested whether the adult Big Five structure would replicate in a large and ethnically diverse sample of adolescent boys (John et al., 1994). The California Child Q-sort provided a comprehensive item pool for the description of children and adolescents that was not derived from the adult Big Five and does not represent any particular theoretical orientation. Factor analyses identified five dimensions that corresponded closely with a priori scales representing the adult Big Five. However, two additional dimensions emerged. *Irritability* involved negative affect expressed in age-inappropriate behaviors, such as whining, crying, tantrums, and being overly sensitive to teasing. *Activity* was defined by items involving physical activity, energy, and high tempo, such as running, playing, and moving

and reacting quickly. In several Dutch samples of boys and girls ages 3–16 years, van Lieshout and Haselager (1994) also found the Big Five plus two additional factors—one similar to Activity factor observed earlier in the United States (John et al., 1994), and a Dependency dimension defined primarily by eagerness to please and reliance on others. Although these findings need to be replicated and extended, they leave open the possibility that the structure of personality traits may be more differentiated in childhood than in adulthood. Specifically, the additional dimensions may originate in temperamental features of childhood personality (e.g., activity level) that become integrated into adult personality structure over the course of adolescence (John et al., 1994).

These studies illustrate how the Big Five can help stimulate research that connects and integrates findings across long-separate research traditions. They also provide initial insights about the way in which personality structure may develop toward its adult form. Yet a great deal of work still lies ahead. Studies need to examine the antecedents of the Big Five and their relations to other aspects of personality functioning in childhood and adolescence. In this way, the Big Five can help connect research on adult personality with the vast field of social development.

Theoretical Perspectives on the Big Five: Description and Explanation

Over the years, researchers have articulated a number of different perspectives on the conceptual status of the Big Five dimensions. Because of their lexical origin, the factors were initially interpreted as dimensions of trait description or attribution (John et al., 1988). Subsequent research, however, has shown that the lexical factors converge with dimensions derived in other personality research traditions, that they have external or predictive validity (as reviewed above), and that all five of them show about equal amounts of heritability (Loehlin et al., 1998). We thus need to ask how these differences should be conceptualized (e.g., Wiggins, 1996); below we briefly summarize some of the major theoretical perspectives.

Researchers in the lexical tradition tend to take an agnostic stance regarding the conceptual status of traits. For example, Saucier and Goldberg (1996b) argued that their studies of personality description do not address issues of causality or the mechanisms underlying behavior. Their interest is primarily in the language of personality. This level of self-restraint may seem dissatisfactory to psychologists who are more interested in personality itself. However, the findings from the lexical approach are informative because the lexical hypothesis is essentially a functionalist argument about the trait concepts in the natural language. These concepts are of interest because language encodes the characteristics that are central, for cultural, social, or biological reasons, to human life and experience. Thus, Saucier and Goldberg argue that lexical studies define an agenda for personality psychologists because they highlight the important and meaningful psychological phenomena (i.e., phenotypic characteristics) that personality psychologists should study and explain. Thus, issues such as the accuracy of self and peer descriptions and the causal origin of traits (i.e., genotypes) are left as open questions that need to be answered empirically. However, important characteristics may exist that people may not be able to observe and describe verbally; if so, the agenda specified by the lexical approach may be incomplete and would need to be supplemented by more theoretically driven approaches (Block, 1995; Tellegen, 1993).

Several theories conceptualize the Big Five as relational constructs. Interpersonal theory (Wiggins & Trapnell, 1996), emphasizes the individual in relationships. The Big Five are taken to describe "the relatively enduring pattern of recurrent interpersonal situations that characterize a human life" (Sullivan, 1953, pp. 110–111), thus conceptualizing the Big Five as descriptive concepts. Wiggins and Trapnell emphasize the interpersonal motives of agency and communion and interpret all of the Big Five dimensions in terms of their interpersonal implications. Because Extraversion and Agreeableness are the most clearly interpersonal dimensions in the Big Five, they receive conceptual priority in this model.

Socioanalytic theory (Hogan, 1996) focuses on the social functions of self- and other perceptions. According to Hogan, trait concepts serve as the "linguistic tools of observers" (1996, p. 172) used to encode and communicate reputations. This view implies

that traits are socially constructed to serve interpersonal functions. Because trait terms fundamentally reflect reputation, individuals who self-report their traits engage in a symbolic–interactionist process of introspection (i.e., the individual considers how others view him or her). Hogan emphasizes that individuals may distort their self-reports with self-presentational strategies; another source of distortion are self-deceptive biases (cf. Paulhus & John, 1998), which do not reflect deliberate impression management but honestly held, though biased, beliefs about the self.

The evolutionary perspective on the Big Five holds that humans have evolved "difference-detecting mechanisms" to perceive individual differences that are relevant to survival and reproduction (D. M. Buss, 1996, p. 185; see also Botwin, Buss, & Shackelford, 1997). Buss views personality as an "adaptive landscape" where the Big Five traits represent the most salient and important dimensions of the individual's survival needs. The evolutionary perspective equally emphasizes person perception and individual differences: Because people vary systematically along certain trait dimensions, and because knowledge of others' traits has adaptive value, humans have evolved a capacity to perceive those individual differences that are central to adaptation to the social landscape. The Big Five summarize these centrally important individual differences.

McCrae and Costa (1996; see also Chapter 5, this volume) view the Big Five as causal personality dispositions. Their five-factor theory (FFT) is a general trait theory that provides an explanatory interpretation of the empirically derived Big Five taxonomy. One central tenet of the FFT is based on the finding that all of the Big Five dimensions have a substantial genetic basis (e.g., Loehlin et al., 1998; Plomin, DeFries, Craig, & McGuffin, 2003; see also Krueger & Johnson, Chapter 10, this volume) and must therefore derive, in part, from biological structures and processes, such as specific gene loci, brain regions (e.g., the amygdala), neurotransmitters (e.g., dopamine), hormones (e.g., testosterone), and so on (e.g., Canli, 2006; see also Canli, Chapter 11, this volume); it is in this sense that traits are assumed to have causal status. McCrae and Costa distinguish between "basic tendencies" and "characteristic adaptations." Personality traits are basic tendencies that refer to the abstract underlying potentials of the individual, whereas attitudes, roles, relationships, and goals are characteristic adaptations that reflect the interactions between basic tendencies and environmental demands accumulated over time. According to McCrae and Costa, basic tendencies remain stable across the life course, whereas characteristic adaptations can undergo considerable change. From this perspective, then, a statement such as "Paul likes to go to parties because he is extraverted" is not circular, as it would be if *extraverted* were merely a description of typical behavior (Wiggins, 1997). Instead, the concept *extraverted* stands in for biological structures and processes that remain to be discovered. This view is similar to Allport's (1937) account of traits as neuropsychic structures and Eysenck's view of traits as biological mechanisms (Eysenck & Eysenck, 1985).

The idea that personality traits have a biological basis is also fundamental to Gosling's proposal that personality psychology must be broadened to include a comparative approach to study individual differences in both human and nonhuman animals (Gosling, 2001; Gosling, Kwan, & John, 2003). Although scientists are understandably reluctant to ascribe personality traits, emotions, and cognitions to animals, evolutionary theory predicts cross-species continuities not only for physical but also for behavioral traits; for example, Darwin (1872/1998) argued that emotions exist in both human and nonhuman animals. A review of 19 studies of personality factors in 12 nonhuman species showed substantial cross-species continuity (Gosling & John, 1999). Chimpanzees and other primates, dogs, cats, donkeys, pigs, guppies, and octopuses all showed reliable individual differences in Extraversion and Neuroticism, and all but guppies and octopuses varied in Agreeableness as well, suggesting that these three Big Five factors may capture fundamental dimensions of individual differences. Further evidence suggests that elements of Openness (such as curiosity and playfulness) are present in at least some nonhuman animals. In contrast, only humans and our closest relatives, chimpanzees, appear to show systematic individual differences in Conscientiousness. Given the relatively

complex social-cognitive functions involved in this dimension (i.e., following norms and rules, thinking before acting, and controlling impulses), it makes sense that Conscientiousness may have appeared rather recently in our evolutionary history. The careful application of ethological and experimental methodology and the high interobserver reliability in these studies make it unlikely that these findings merely reflect anthropomorphic projections (c.f., Gosling et al., 2003; Kwan, Gosling, & John, 2008). Rather, these surprising cross-species commonalities suggest that personality traits reflect, at least in part, biological mechanisms that are shared by many mammalian species.

In summary, researchers subscribe to a diversity of perspectives on the conceptual status of the Big Five, ranging from purely descriptive concepts to biologically based causal concepts. This diversity might be taken to imply that researchers cannot agree about the definition of the trait concept and that the field is in disarray. It is important to recognize, however, that these perspectives are not mutually exclusive. For example, although Saucier and Goldberg (1996b) caution against drawing inferences about genotypes from lexical studies, the lexical hypothesis does not preclude the possibility that the Big Five are embodied in biological structures and processes. In our view, "what is a trait?" is fundamentally an empirical question. Research in diverse areas such as behavior and molecular genetics, personality stability and change, and accuracy and bias in self-reports and interpersonal perception will be instrumental in building and refining a comprehensive theoretical account of the Big Five.

CONCLUSIONS AND IMPLICATIONS

At the beginning of this chapter, we argued that a personality taxonomy should provide a systematic framework for distinguishing, ordering, and naming the behavioral, emotional, and experiential characteristics of individuals. Ideally, that taxonomy would be built around principles that are causal and dynamic, exist at multiple levels of abstraction or hierarchy, and offer a standard nomenclature for scientists working in the field of personality. The Big Five taxonomy does not yet meet this high standard. It provides descriptive concepts that still need to be explicated theoretically, and a nomenclature that is still rooted in the "vernacular" English.

The Big Five structure has the advantage that everybody can understand the words that define the factors. Moreover, the natural language is not biased in favor of any existing scientific conceptions; although the atheoretical nature of the Big Five dimensions makes them less appealing to some psychologists, it also makes them more palatable to researchers who reject dimensions cast in a theoretical mold different from their own. Whatever the inadequacies of the natural language for scientific systematics, broad dimensions inferred from folk usage are *not* a bad place to start a taxonomy. Even in the biological taxonomy of animals, "the technical system evolved from the vernacular" (Simpson, 1961, pp. 12–13).

Obviously, a system that initially derives from the natural language does not need to reify such terms indefinitely. Indeed, several of the dimensions included among the Big Five, most notably Extraversion and Neuroticism, have been the target of various physiological and mechanistic explanations (e.g., Canli et al., 2001; see also L. A. Clark, 2005). In research on emotion regulatory processes, the links between the Big Five, the chronic use of particular regulatory strategies, and their emotional and social consequences are being articulated (John & Gross, 2007). Similarly, the conceptual explication of Extraversion and Neuroticism as persistent dispositions toward thinking and behaving in ways that foster, respectively, positive and negative affective experiences (e.g., Tellegen, 1985; see also Clark & Watson, Chapter 9, this volume) promises to connect the Big Five with individual differences in affective functioning, which, in turn, may be studied in more tightly controlled laboratory settings (see Gross, Chapter 28, this volume). At this point, the Big Five differentiate domains of individual differences that have similar surface manifestations—just like the early animal taxonomy that was transformed by better accounts of evolutionary processes and by the advent of new tools, such as molecular genetics. Likewise, the structures and processes underlying these personality trait domains are now beginning to be explicated. Explanatory and

mechanistic terms will likely change the definition and assessment of the Big Five dimensions as we know them today.

As Gould (1981) observed, even in the biological and physical sciences, "taxonomy is always a contentious issue because the world does not come to us in neat little packages" (p. 158). Since the early 1980s when the Big Five barely registered a blip in the published personality literature (see Figure 4.1), we have come a long way in understanding the "messy packages" that are personality traits. Researchers have made enormous progress on the Big Five trait taxonomy, producing an initial consensus that we can differentiate five replicable domains of personality as summarized by the broad concepts of *Extraversion*, *Agreeableness*, *Conscientiousness*, *Neuroticism*, and *Openness to experience*. Viewed from a historical vantage point, the emergence of the Big Five structure, and the fact that multiple groups of researchers worked on it jointly, brought about a major change in the field of personality that is akin to a paradigm shift. Personality trait research has moved from a stage of early individualistic pioneers to a more mature stage of normal scientific inquiry: Researchers interested in studying the effects of personality traits on important theoretical or applied phenomena, such as emotion, social behavior and relationships, work and achievement, or physical and mental health, now use a commonly understood framework to conceptualize their research and choose from several well-validated instruments to operationalize these personality domains. Literature reviews and meta-analyses are commonly done to organize all the available empirical findings on a phenomenon, such as whether and how much personality traits change during particular periods of adulthood (e.g., Helson et al., 2002; Roberts et al., 2006), into one coherent set of hypotheses and findings. This is indeed a paradigm shift (or a seismic shift here in California) in a field dominated, until recently, by seemingly incompatible systems that caused fragmentation and competition, rather than fostering commonalities and convergences. As illustrated in Table 4.1, the Big Five structure captures, at a broad level of abstraction, the commonalities among the existing systems of personality description and thus provides an integrative descriptive taxonomy for personality research.

ACKNOWLEDGMENTS

This chapter summarizes and updates previous reviews by John (1990) and John and Srivastava (1999). The preparation of this chapter was supported in part by research grants from the Retirement Research Foundation and the Metanexus Foundation, and by National Science Foundation Predoctoral Fellowships to Laura Naumann and Christopher Soto. The support and resources provided by the Institute of Personality and Social Research are also gratefully acknowledged.

REFERENCES

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Allport, G. W. (1958). What units shall we employ? In G. Lindzey (Ed.), *Assessment of human motives* (pp. 238–260). New York: Rinehart.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47*, No. 211.

Almagor, M., Tellegen, A., & Waller, N. G. (1995). The Big Seven model: A cross-cultural replication and further exploration of the basic dimensions of natural language trait descriptors. *Journal of Personality and Social Psychology, 69*, 300–307.

Anderson, C., John, O. P., Keltner, D., & Kring, A. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology, 81*, 116–132.

Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality, 4*, 89–118.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., et al. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology, 86*, 356–366.

Bales, R. F. (1970). *Personality and interpersonal behavior.* New York: Holt, Rinehart & Winston.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26.

Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-analysis of the relationship between the five factor model of personality and Holland's occupational types. *Personnel Psychology, 56*, 45–74.

Baumgarten, F. (1933). *Die Charktereigenschaften* [The character traits]. In Beitraege zur Charak-

ter- und Persoenlichkeitsforschung (Whole No. 1). Bern, Switzerland: A. Francke.

Becker, W. C. (1960). The matching of behavior rating and questionnaire personality factors. *Psychological Bulletin, 57*, 201–212.

Belsky, J., Jaffee, S. R., Caspi, A., Moffitt, T., & Silva, P. A. (2003). Intergenerational relationships in young adulthood and their life course, mental health, and personality correlates. *Journal of Family Psychology, 17*, 460–471.

Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait–multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75*, 729–750.

Benet-Martínez, V., & Waller, N. G. (1997). Further evidence for the cross-cultural generality of the Big Seven Factor model: Indigenous and imported Spanish personality constructs. *Journal of Personality, 65*, 567–598.

Berkman, L. F., Glass, T., Brissette, I., & Seeman, T. E. (2000). From social integration to health. *Social Science and Medicine, 51*, 843–857.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187–215.

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 13, pp. 39–101). Hillsdale, NJ: Erlbaum.

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health behaviors: A meta-analysis. *Psychological Bulletin, 130*, 887–919.

Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioral Science, 9*, 8–17.

Botwin, M. D., & Buss, D. M. (1989). Structure of act–report data: Is the five-factor model of personality recaptured? *Journal of Personality and Social Psychology, 56*, 988–1001.

Botwin, M. D., Buss, D. M., & Shackelford, T. K. (1997). Personality and mate preferences: Five factors in mate selection and marital satisfaction. *Journal of Personality, 65*, 107–136.

Burisch, M. (1984). Approaches to personality inventory construction. *American Psychologist, 39*, 214–227.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Buss, D. M. (1996). Social adaptation and five major factors of personality. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180–207). New York: Guilford Press.

Canli, T. (Ed.). (2006). *Biology of personality and individual differences*. New York: Guilford Press.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience, 115*, 33–42.

Caprara, G. V., & Perugini, M. (1994). Personality described by adjectives: The generalizability of the Big Five to the Italian lexical context. *European Journal of Personality, 8*, 357–369.

Caspi, A., & Bem, D. J. (1990). Personality continuity and change across the life course. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 549–575). New York: Guilford Press.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453–484.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38*, 476–506.

Cattell, R. B. (1945a). The description of personality: Principles and findings in a factor analysis. *American Journal of Psychology, 58*, 69–90.

Cattell, R. B. (1945b). The principal trait clusters for describing personality. *Psychological Bulletin, 42*, 129–161.

Cattell, R. B. (1990). Advances in Cattellian personality theory. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 101–110). New York: Guilford Press.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, IL: IPAT.

Chaplin, W. F., John, O. P., & Goldberg, L. R. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality and Social Psychology, 54*, 541–557.

Chaplin, W. F., Phillips, J. B., Brown, J. D., Clanton, N. R., & Stein, J. L. (2000). Handshaking, gender, personality, and first impressions. *Journal of Personality and Social Psychology, 79*, 110–117.

Church, A. T., & Katigbak, M. S. (1989). Internal, external, and self-report structure of personality in a non-Western culture: An investigation of cross-language and cross-cultural generalizability. *Journal of Personality and Social Psychology, 57*, 857–872.

Church, A. T., Reyes, J. A. S., Katigbak, M. S., & Grimm, S. D. (1997). Filipino personality structure and the Big Five model: A lexical approach. *Journal of Personality, 65*, 477–528.

Clark, J., Boccaccini, M. T., Caillouet, B., & Chaplin, W. F. (2007). Five factor model personality traits, jury selection, and case outcomes in criminal and civil cases. *Criminal Justice and Behavior, 34*, 641–660.

Clark, L. A. (2005). Temperament as a unifying basis for personality and psychopathol-

ogy. *Journal of Abnormal Psychology, 114,* 505–521.

Clark, L. A., & Watson, D. (1999). Temperament: A new paradigm for trait psychology. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality* (2nd ed., pp. 399–423). New York: Guilford Press.

Comrey, A. L. (1970). *Manual for the Comrey Personality Scales.* San Diego, CA: Educational and Industrial Testing Service.

Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait–multimethod–multioccasion analysis. *Journal of Personality and Social Psychology, 49,* 1266–1282.

Costa, P. T., Jr., & McCrae, R. R. (1976). Age differences in personality structure: A cluster analytic approach. *Journal of Gerontology, 31,* 564–570.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1989). *NEO PI/FFI manual supplement.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64,* 21–50.

Costa, P. T., Jr., & Widiger, T. A. (Eds.). (2002). *Personality disorders and the five factor model of personality* (pp. 203–214). Washington, DC: American Psychological Association.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology, 80,* 804–813.

Darwin, C. (1998). *The expression of emotions in man and animals* (3rd ed.). New York: Oxford University Press. (Original work published 1872)

David, J., & Suls, J. (1999). Coping efforts in daily life: Role of Big Five traits and problem appraisal. *Journal of Personality, 67,* 119–140.

De Clercq, B., De Fruyt, F., Van Leeuwen, K., & Mervielde, I. (2006). The structure of maladaptive personality traits in childhood: A step toward an integrative developmental perspective for DSM-V. *Journal of Abnormal Psychology, 115,* 639–657.

De Raad, B., di Blas, L., & Perugini, M. (1998). Two independently constructed Italian trait taxonomies: Comparisons among Italian and between Italian and Germanic languages. *European Journal of Personality, 12,* 19–41.

De Raad, B., Mulder, E., Kloosterman, K., & Hofstee, W. K. (1988). Personality-descriptive verbs. *European Journal of Personality, 2,* 81–96.

De Raad, B., & Perugini, M. (Eds.). (2002). *Big Five assessment.* Kirkland, WA: Hogrefe & Huber.

De Raad, B., Perugini, M., Hrebickova, M., & Szarota, P. (1998). *Lingua franca* of personality: Taxonomies and structures based on the psycholexical approach. *Journal of Cross-Cultural Psychology, 29,* 212–232.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91,* 1138–1151.

Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality-descriptive adjectives in the Italian language. *European Journal of Personality, 12,* 75–101.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57,* 195–214.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73,* 1246–1256.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50,* 116–123.

Digman, J. M., & Shmelyov, A. G. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology, 71,* 341–351.

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis and comparison of six major studies. *Multivariate Behavioral Research, 16,* 149–170.

Donahue, E. M. (1994). Do children use the Big Five, too?: Content and structural form in personality description. *Journal of Personality, 62,* 45–66.

Eysenck, H. J. (1986). Models and paradigms in personality research. In A. Angleitner, A. Furnham, & G. Van Heck (Eds.), *Personality psychology in Europe (Vol. 2): Current trends and controversies* (pp. 213–223). Lisse, The Netherlands: Swets & Zeitlinger.

Eysenck, H. J. (1990). Biological dimensions of personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 244–276). New York: Guilford Press.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3?—criteria for a taxonomic paradigm. *Personality and Individual Differences, 12,* 773–790.

Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences, 13,* 667–673.

Eysenck, H. J. (1997). Personality and experi-

mental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology, 73*, 1224–1237.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum Press.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44*, 329–344.

Friedman, H. S., Hawley, P. H., & Tucker, J. S. (1994). Personality, health, and longevity. *Current Directions in Psychological Science, 3*, 37–41.

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: Sage.

Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. 1, pp. 203–234). Hillsdale, NJ: Erlbaum.

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. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26–34.

Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology, 48*, 82–98.

Goldberg, L. R., & Rosolack, T. K. (1994). The Big Five factor structure as an integrative framework: An empirical comparison with Eysenck's P-E-N model. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 7–35). Hillsdale, NJ: Erlbaum.

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J.E. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences, 24*, 393–403.

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, 127, 45–86.

Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science, 8*, 69–75.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*, 379–398.

Gosling, S. D., Kwan, V. S. Y., & John, O. P. (2003). A dog's got personality: A cross-species comparative approach to evaluating personality judgments. *Journal of Personality and Social Psychology, 85*, 1161–1169.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality, 37*, 504–528.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology, 37*, 1938–1405.

Gough, H. G. (1987). *The California Psychological Inventory administrator's guide*. Palo Alto, CA: Consulting Psychologists Press.

Gough, H. G., & Heilbrun, A. B., Jr. (1983). *The Adjective Check List manual*. Palo Alto, CA: Consulting Psychologists Press.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–824). San Diego, CA: Academic Press.

Graziano, W. G., & Ward, D. (1992). Probing the Big Five in adolescence: Personality and adjustment during a developmental transition. *Journal of Personality, 60*, 425–439.

Guilford, J. P. (1975). Factors and factors of personality. *Psychological Bulletin, 82*, 802–814.

Halverson, C. F., Kohnstamm, G. A., & Martin R. P. (1994). *The developing structure of temperament and personality from infancy to adulthood*. Hillsdale, NJ: Erlbaum.

Hampson, S. E., Andrews, J. A., Barckley, M., Lichtenstein, E., & Lee, M. E. (2000). Conscientiousness, perceived risk, and risk-reduction behaviors: A preliminary study. *Health Psychology, 19*, 247–252.

Hampson, S. E., & Goldberg, L. R. (2006). A first large cohort study of personality trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology, 91*, 763–779.

Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category-breadth and social-desirability values for 573 personality terms. *European Journal of Personality, 1*, 241–258.

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.

Helson, R. (1967). Personality characteristics and developmental history of creative college women. *Genetic Psychology Monographs, 76*, 205–256.

Helson, R. (1985). Which of those young women with creative potential became productive? Personality in college and characteristics of parents. In R. Hogan & W. H. Jones (Eds.), *Perspectives in personality theory, measurement, and interpersonal dynamics* (Vol. 1, pp. 49–80). Greenwich, CT: JAI Press.

Helson, R., Kwan, V. S. Y., John, O. P., & Jones, C. (2002). The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of Research in Personality, 36*, 287–306.

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., De Raad, B., & Goldberg, L. R. (1997). A comparison of Big-Five structures of personality traits in Dutch, English, and German. *European Journal of Personality, 11*, 15–31.

Hogan, R. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *Nebraska Symposium on Motivation, 1982: Personality—current theory and research* (Vol. 29, pp. 55–89). Lincoln: University of Nebraska Press.

Hogan, R. (1986). *Hogan Personality Inventory manual*. Minneapolis, MN: National Computer Systems.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180–207). New York: Guilford Press.

Hogan, R., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 849–870). San Diego, CA: Academic Press.

Hrebickova, M., & Ostendorf, F. (1995). *Lexikalni pristup k osobnosti. V: Klasifikace pridavnych jmen do kategorii osobnostni deskripce* [Lexical approach to personality: V. Classification of adjectives into categories of personality description]. *Ceskoslovenska Psychologie, 39*, 265–276.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd ed.). Port Huron, MI: Research Psychologists Press.

John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 261–271). New York: Springer-Verlag.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*, 171–203.

John, O. P., & Benet-Martínez, V. (2000). Measurement, scale construction, and reliability. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York: Cambridge University Press.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "Little Five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160–178.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California at Berkeley, Institute of Personality and Social Research.

John, O. P., Goldberg, L. R., & Angleitner, A. (1984). Better than the alphabet: Taxonomies of personality-descriptive terms in English, Dutch, and German. In H. Bonarius, G. van Heck, & N. Smid (Eds.), *Personality psychology in Europe: Theoretical and empirical developments* (pp. 83–100). Berwyn, PA: Swets North America.

John, O. P., & Gross, J. J. (2007). Individual differences in emotion regulation. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 351–372). New York: Guilford Press.

John, O. P., Hampson, S. E., & Goldberg, L. R. (1991). Is there a basic level of personality description? *Journal of Personality and Social Psychology, 60*, 348–361.

John, O. P., & Naumann, L. P. (2007). *Correlations of BFI scales and self-reported act frequencies in an undergraduate sample*. Unpublished data, Institute of Personality and Social Research, University of California at Berkeley.

John, O. P., & Robins, R. W. (1993). Gordon Allport: Father and critic of the five-factor model. In K. H. Craik, R. T. Hogan, & R. N. Wolfe (Eds.), *Fifty years of personality research* (pp. 215–236). New York: Plenum Press.

John, O. P., & Robins, R. W. (1998). Recent trends in Big Five research: Development, predictive validity, and personality types. In J. Bermudez et al. (Eds.), *Personality psychology in Europe*

(Vol. 6, pp. 6–16). Tilburg, The Netherlands: Tilburg University Press.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). New York: Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford Press.

Karney, B. R., & Bradbury, T. N. (1995). Assessing longitudinal change in marriage: An introduction to the analysis of growth curves. *Journal of Marriage and the Family, 57*, 1091–1108.

Kenford, S. L., Smith, S. S., Wetter, D. W., Jorenby, D. E., Fiore, M. C., & Baker, T. B. (2002). Predicting relapse back to smoking: Contrasting affective and physical models of dependence. *Journal of Consulting and Clinical Psychology, 70*, 216–227.

Klages, L. (1932). *The science of character*. London: Allen & Unwin.

Kohnstamm, G. A., Halverson, C. F., Mervielde, I., & Havill, V. L. (Eds.). (1998). *Parental descriptions of child personality: Developmental antecedents of the Big Five?* Mahwah, NJ: Erlbaum.

Krueger, R. F., & Tackett, J. L. (Eds.). (2006). *Personality and psychopathology*. New York: Guilford Press.

Kwan, V. S. Y., Gosling, S. D., & John, O. P. (2008). Anthropomorphism as a special case of social perception: A cross-species comparative approach and a new empirical paradigm. *Social Cognition, 26*, 129–142.

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six interests and Big Five personality factors. *Journal of Vocational Behavior, 61*, 217–239.

Loehlin, J. C., McCrae, R. R., Costa, P. T., Jr., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality, 32*, 431–453.

Markon, K. E, Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*, 139–157.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality, 60*, 329–361.

McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin, 120*, 323–337.

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

McCrae, R. R., & Costa, P. T., Jr. (1985b). Openness to experience. In R. Hogan & W. H. Jones (Eds.), *Perspectives in personality* (Vol. 1, pp. 145–172). Greenwich, CT: JAI Press.

McCrae, R. R., & Costa, P. T., Jr. (1985c). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences, 6*, 587–597.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81–90.

McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175–215.

Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology, 89*, 90–106.

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*, 862–877.

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.

Miller, T. Q., Smith, T. W., Turner, C. W., Guijarro, M. L., & Hallet, A. J. (1996). A meta-analytic review of research on hostility and physical health. *Psychological Bulletin, 119*, 322–348.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145–165.

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

Newcomb, A. F., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: A meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychological Bulletin, 113*, 99–128.

Nigg, J. T., John, O. P., Blaskey, L. G., Huang-Pollock, C. L., Willcutt, E. G., Hinshaw, S. P., et al. (2002). Big Five dimensions and ADHD symptoms: Link between personality traits and clinical syndromes. *Journal of Personality and Social Psychology, 83*, 451–469.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of Academic Outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*, 116–130.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574–583.

Norman, W. T. (1967). *2,800 personality trait descriptors: Normative operating characteristics for a university population.* Unpublished paper, Department of Psychology, University of Michigan.

Nowakowska, M. (1973). The limitations of the factor-analytic approach to psychology with special application to Cattell's research strategy. *Theory and Decision, 4*, 109–139.

Ostendorf, F. (1990). *Sprache und Persoenlichkeitsstruktur: Zur Validitaet des Fuenf-Faktoren-Modells der Persoenlichkeit* [Language and personality structure: On the validity of the five factor model of personality]. Regensburg, Germany: S. Roderer Verlag.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401–421.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66*, 1025–1060.

Paunonen, S. V. (2003). Big Five factors of personality and replicated prediction of behavior. *Journal of Personality and Social Psychology, 84*, 411–424.

Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57*, 552–567.

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry, 5*, 103–113.

Pervin, L. A., Cervone, D., & John, O. P. (2005). *Personality: Theory and research* (9th ed.). Hoboken, NJ: Wiley.

Plomin, R., DeFries, J. C., Craig, I. W., & McGuffin, P. (Eds.). (2003). *Behavioral genetics in the postgenomic era.* Washington, DC: American Psychological Association.

Rammstedt, B., & John, O. P. (2005). Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality. *Diagnostica, 51*, 195–206.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203–212.

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science, 17*, 236–242.

Revelle, W. (1987). Personality and motivation: Sources of inefficiency in cognitive performance. *Journal of Research in Personality, 21*, 436–452.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3–25.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25.

Robins, R. W., Caspi, A., & Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70*, 925–964.

Robins, R. W., John, O. P., & Caspi, A. (1994). Major dimensions of personality in early adolescence: The Big Five and beyond. In C. F. Halverson, J. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 267–291). Hillsdale, NJ: Erlbaum.

Robins, R. W., John, O. P., Caspi, A., Moffitt, T. E., & Stouthamer-Loeber, M. (1996). Resilient, overcontrolled, and undercontrolled boys: Three replicable personality types? *Journal of Personality and Social Psychology, 70*, 157–171.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science, 17*, 236–242.

Saucier, G. (1992). Openness versus Intellect: Much ado about nothing. *European Journal of Personality, 6*, 386–391.

Saucier, G. (1994). Mini-markers: A brief version

of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506–516.
Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology, 73*, 1296–1312.
Saucier, G., Georgiades, S., Tsouasis, I., & Goldberg, L. R. (2005). Factor structure of Greek personality-descriptive adjectives. *Journal of Personality and Social Psychology, 88*, 856–875.
Saucier, G., & Goldberg, L. R. (199a). 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 Press.
Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality, 69*, 847–879.
Saucier, G., Hampson, S. E., & Goldberg, L. R. (2000). Cross-language studies of lexical personality factors. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 1–36). London: Routledge.
Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*, 613–627.
Scarr, S., & McCartney, K. (1984). How people make their own environments: A theory of genotype–environment effects. *Annual Progress in Child Psychiatry and Child Development*, 98–118.
Scheier, M. F., & Carver, C. S. (1993). On the power of positive thinking: The benefits of being optimistic. *Current Directions in Psychological Science, 2*, 26–30.
Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*, 2–32.
Shmelyov, A. G., & Pokhil'ko, V. I. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality, 7*, 1–17.
Simpson, G. G. (1961). *Principles of animal taxonomy*. New York: Columbia University Press.
Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of Personality and Social Psychology, 76*, 431–450.
Soto, C. J., & John, O. P. (2008). *Measuring Big Five domains and 16 facets using the California Psychological Inventory*. Manuscript submitted for publication.
Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94*, 718–737.
Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84*, 1041–1053.
Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.
Szarota, P. (1995). *Polska Lista Przymiotnikowa (PLP): Narzedzie do diagnozy Pieciu Wielkich czynnikow osobowosci* [Polish Adjective List: An instrument to assess the five-factor model of personality]. *Studia Psychologiczne, 33*, 227–256.
Szirmak, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality, 8*, 95–117.
Tellegen, A. (1982). *Brief manual for the Differential Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.
Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–716). Hillsdale, NJ: Erlbaum.
Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4*, 122–130.
Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *MMPI-2 Restructured Clinical (RC) Scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.
Thoresen, C. J., Kaplan, S. A., Barsky, A. P., Warren, C. R., & de Chermont, K. (2003). The affective underpinnings of job perceptions and attitudes: A meta-analytic review and integration. *Psychological Bulletin, 129*, 914–945.
Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology, 59*, 781–790.
Trull, T. J., & Sher, K. J. (1994). Relationship between the five-factor model of personality and Axis I disorders in a nonclinical sample. *Journal of Abnormal Psychology, 103*, 350–360.
Tupes, E. C., & Christal, R. C. (1961). *Recur-*

rent personality factors based on trait ratings. Technical Report, U.S. Air Force, Lackland Air Force Base, TX.

Tupes, E. C., & Christal, R. C. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60,* 225–251.

van Lieshout, C. F. M., & Haselager, G. J. T. (1994). The Big Five personality factors in Q-sort descriptions of children and adolescents. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 293–318). Hillsdale, NJ: Erlbaum.

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology, 87,* 123–132.

Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 767–793). San Diego, CA: Academic Press.

Watson, D., Hubbard, B., & Wiese, D. (2000). General traits of personality and affectivity as predictors of satisfaction in intimate relationships: Evidence from self- and partner-ratings. *Journal of Personality, 68,* 413–449.

Weiss, A., & Costa, P. T., Jr. (2005). Domain and facet personality predictors of all-cause mortality among Medicare patients aged 65–100. *Psychosomatic Medicine, 67,* 724–733.

Wiggins, J. S. (1968). Personality structure. In P. R. Farnsworth (Ed.), *Annual review of psychology* (Vol. 19, pp. 320–322). Palo Alto, CA: Annual Reviews.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412.

Wiggins, J. S. (1995). *Interpersonal Adjective Scales: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Wiggins, J. S. (Ed.). (1996). *The five-factor model of personality: Theoretical perspectives.* New York: Guilford Press.

Wiggins, J. S. (1997). In defense of traits. In R. Hogan, J. A. Johnson, & S. R. Briggs (Ed.), *Handbook of personality psychology* (pp. 649–679). San Diego, CA: Academic Press.

Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment, 1,* 305–316.

Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic–interactional perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180–207). New York: Guilford Press.

Yang, K. S., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of Personality and Social Psychology, 58,* 1087–1095.

APPENDIX 4.1. BIG FIVE INVENTORY RESPONSE FORM AND INSTRUCTIONS TO PARTICIPANTS

Instructions: Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who *likes to spend time with others*? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

1 Disagree strongly	2 Disagree a little	3 Neither agree nor disagree	4 Agree a little	5 Agree strongly

I see myself as someone who ..

1. ____ Is talkative
2. ____ Tends to find fault with others
3. ____ Does a thorough job
4. ____ Is depressed, blue
5. ____ Is original, comes up with new ideas
6. ____ Is reserved
7. ____ Is helpful and unselfish with others
8. ____ Can be somewhat careless
9. ____ Is relaxed, handles stress well
10. ____ Is curious about many different things
11. ____ Is full of energy
12. ____ Starts quarrels with others
13. ____ Is a reliable worker
14. ____ Can be tense
15. ____ Is ingenious, a deep thinker
16. ____ Generates a lot of enthusiasm
17. ____ Has a forgiving nature
18. ____ Tends to be disorganized
19. ____ Worries a lot
20. ____ Has an active imagination
21. ____ Tends to be quiet
22. ____ Is generally trusting
23. ____ Tends to be lazy
24. ____ Is emotionally stable, not easily upset
25. ____ Is inventive
26. ____ Has an assertive personality
27. ____ Can be cold and aloof
28. ____ Perseveres until the task is finished
29. ____ Can be moody
30. ____ Values artistic, aesthetic experiences
31. ____ Is sometimes shy, inhibited
32. ____ Is considerate and kind to almost everyone
33. ____ Does things efficiently
34. ____ Remains calm in tense situations
35. ____ Prefers work that is routine
36. ____ Is outgoing, sociable
37. ____ Is sometimes rude to others
38. ____ Makes plans and follows through with them
39. ____ Gets nervous easily
40. ____ Likes to reflect, play with ideas
41. ____ Has few artistic interests
42. ____ Likes to cooperate with others
43. ____ Is easily distracted
44. ____ Is sophisticated in art, music, or literature

Please check: Did you write a number in front of each statement?

Note. From John, Donahue, and Kentle (1991). Copyright 1991 by Oliver P. John. Reprinted by permission.

APPENDIX 4.2. SCORING THE BFI SCALES AND ACQUIESCENCE INDEX, AND IPSATIZING THE BFI ITEMS

Computing Simple BFI Scale Scores

BFI scale scoring: Reverse score the items labeled "R" and compute scale scores as the mean of the following items:

Extraversion (8 items): 1, 6R, 11, 16, 21R, 26, 31R, 36
Agreeableness (9 items): 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42
Conscientiousness (9 items): 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R
Neuroticism (8 items): 4, 9R, 14, 19, 24R, 29, 34R, 39
Openness (10 items): 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

Computing the Content-Balanced Acquiescence Index and Ipsatizing the BFI Items

To index individuals' acquiescent response style (where high scores mean "yea-saying" and low scores mean "nay-saying"), we compute their acquiescence score as their mean response across 32 BFI items that form 16 pairs of items with opposite implications for personality (e.g., item 1 "Is talkative" and item 21 "Tends to be quiet"). As described in Soto, John, Gosling, and Potter (2008), we devised these 16 pairs of opposite items on the basis of item content and the size of their (negative) interitem correlations. Using the standard BFI item numbers (see Appendix A; see also Benet-Martínez & John, 1998; John & Srivastava, 1999), these item pairs are 1 and 21, 6 and 16, 31 and 36, 2 and 17, 7 and 12, 27 and 42, 32 and 37, 3 and 43, 8 and 13, 18 and 33, 23 and 28, 9 and 19, 24 and 29, 34 and 39, 5 and 35, 30 and 41.

The Statistical Package for the Social Sciences (SPSS) syntax given below first computes each person's acquiescence score ("bfiave"; the average of his or her 16 × 2 = 32 item responses) and response extremeness ("bfistd"; the standard deviation of that person's 32 item responses). Both of these individual-difference scores may be retained and used in research on individual or group differences in response-scale use (e.g., comparing Asian Americans and European Americans). The syntax below shows how these two scores can be used to ipsatize the full set of 44 BFI items, by removing from each item score the individual's acquiescence score (i.e., content-balanced response mean) and then adjusting the resulting deviation scores by dividing them by the individual's standard deviation, resulting in person-centered standard (or Z) scores. The syntax below assumes that the variables for the 44 BFI items are named bfi1 to bfi44 in standard order.

SPSS Syntax to Ipsatize the 44 BFI Items before Scoring the Scales

```
* Compute within-person response means (bfiave) and standard deviations
(bfistd).

COMPUTE bfiave = mean(bfi1, bfi6, bfi16, bfi21, bfi31, bfi36, bfi2, bfi7, bfi12,
bfi17, bfi27, bfi32, bfi37, bfi42, bfi3, bfi8, bfi13, bfi18, bfi23, bfi28, bfi33,
bfi43, bfi9, bfi19, bfi24, bfi29, bfi34, bfi39, bfi5, bfi30, bfi35, bfi41).

COMPUTE bfistd = sd(bfi1, bfi6, bfi16, bfi21, bfi31, bfi36, bfi2, bfi7, bfi12, bfi17,
bfi27, bfi32, bfi37, bfi42, bfi3, bfi8, bfi13, bfi18, bfi23, bfi28, bfi33, bfi43, bfi9,
bfi19, bfi24, bfi29, bfi34, bfi39, bfi5, bfi30, bfi35, bfi41).
EXECUTE.

* Compute ipsatized BFI items (zbfi).

COMPUTE zbfi1 = (bfi1 - bfiave)/bfistd.
COMPUTE zbfi2 = (bfi2 - bfiave)/bfistd.
COMPUTE zbfi3 = (bfi3 - bfiave)/bfistd.
                .
                .
                .
COMPUTE zbfi44 = (bfi44 - bfiave)/bfistd.
EXECUTE.
```

Then use the ipsatized item scores to compute scale scores as mean item responses, as described above.