

THE FIVE-FACTOR TAXONOMY: ROBUST DIMENSIONS OF PERSONALITY DESCRIPTION

Fritz OSTENDORF

University of Koblenz-Landau, Federal Republic of Germany

Alois ANGLEITNER

University of Bielefeld, Federal Republic of Germany

This paper summarizes the state of empirical research on the five-factor model (FFM). The basic assumptions of the lexical approach to personality psychology are presented, and the historical roots of the „Big Five” are traced. The results of several taxonomies indicate that the FFM provides a robust description of personality that – similar to Osgood’s dimensions of affective meaning – proves to be highly replicable. Relations between the five factors and the dimensions of alternative models of personality structure are considered, and the conceptual status of the Big Five is discussed.

Introductions to many recent major articles on personality psychology have welcomed the high level of consensus recognizing and confirming that the most important personality traits can be described almost completely and with sufficient precision by a model containing five orthogonal factors (Digman, 1990; Goldberg, 1993; John, 1990; McCrae & John, 1992). This factor model is considered to have a comprehensive and general structure that promises to finally bring order into the great number of competing constructs in personality psychology. Current findings suggest that these are robust factors that repeatedly appear when respondents are asked to make personality judgments on the basis of a large number of personality descriptors with the greatest possible heterogeneity in content, for example, personality-descriptive adjectives or questionnaire items (Ostendorf & Angleitner, 1992).

Table 1 presents the most frequently used labels for the five factors together with a selection of marker variables from various rating inventories. The use of several labels indicates that there is still no consensus on the most appropriate description of the factors.

Table 1. *Selection of Big-Five Marker Scales From Rating Inventories Developed by Norman (1963), McCrae and Costa (1987), and Peabody and Goldberg (1989)*

<i>Factors</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>
(I) Surgency or Extraversion					
talkative--silent	0.78	-0.05	0.02	-0.00	0.09
sociable--reclusive	0.79	0.09	0.02	0.18	0.02
fun-loving--sober	0.61	0.19	-0.24	0.06	0.18
spontaneous--inhibited	0.73	0.02	-0.09	0.26	0.14
frank--secretive	0.76	0.23	-0.03	0.03	0.14
(II) Agreeableness					
mild,gentle--headstrong	-0.03	0.69	-0.05	0.11	0.14
soft-hearted--ruthless	0.12	0.71	0.07	-0.13	0.15
forgiving--vengeful	0.25	0.64	0.10	0.17	0.07
trustful--distrustful	0.28	0.55	-0.02	0.07	-0.02
lenient--harsh	0.04	0.68	0.01	0.13	0.22
(III) Conscientiousness					
responsible--undependable	0.15	0.17	0.68	-0.04	0.25
scrupulous--unscrupulous	-0.15	-0.04	0.68	-0.03	-0.03
conscientious--negligent	0.13	0.07	0.80	0.07	0.07
hardworking--lazy	0.16	0.11	0.67	0.07	0.06
orderly--disorderly	0.04	0.07	0.76	0.18	-0.09
(IV) Emotional Stability					
poised--nervous,tense	0.11	0.24	0.04	0.75	0.05
calm--anxious	0.21	0.03	0.11	0.64	0.19
hardy--vulnerable	0.12	-0.07	0.15	0.64	-0.12
calm--worrying	0.11	0.12	0.04	0.64	0.06
relaxed--tense	0.08	0.24	-0.02	0.70	0.02
(V) Intellect / Openness					
intellectual--unreflective,narrow	-0.07	0.03	0.25	0.09	0.66
creative--uncreative	0.12	0.06	0.04	0.09	0.60
broad interests--narrow interests	0.18	0.09	0.18	0.21	0.58
intelligent-unintelligent	0.08	0.01	0.20	0.09	0.68
imaginative-unimaginative	0.27	0.10	0.00	0.04	0.63

Note. All loadings $\geq .30$ are listed in boldface. The factor loadings reported in the table are loadings on five varimax rotated principal components based on an analysis of 179 rating scales (see Ostendorf, 1990) published by Norman (1963), McCrae and Costa (1987), Peabody and Goldberg (1989), Goldberg (1983, 1992), and John, Goldberg, and Angleitner (1984). $N = 383$ peer ratings.

Naturally, the consensus did not last long, and Hans-Jürgen Eysenck and Lawrence Pervin, for example, have vehemently criticized some of the basic assumptions underlying the FFM and the empirical evidence in its support (Eysenck, 1991, 1994; Pervin, 1994). Whereas controversy is possible on how far the FFM meets the psychometric criteria demanded by, for example, Eysenck (1991, 1994), there should be agreement on the fact that the five factors have been obtained *inductively*, in other words: not postulated, but discovered. The model is not a theory; it only describes a potentially universal phenomenon.

We do not believe that this is a weakness. When developing a theory, one first has to determine the object of this theory. The object of research has to be analyzed and specimens have to be classified before one can construct a meaningful theory about its properties. Taxonomies precede the analysis of causal relationships. Although detailed and highly comprehensive taxonomies bring no guarantees, they nonetheless provide a good basis for the formulation of correct theories.

Goldberg and Saucier (1994) have noted that the current literature reveals at least two alternative five-factor models. On the surface, they appear to be very similar; but they have been developed for quite different purposes. In the tradition of the lexical approach, for example, the Big Five were viewed initially as only a reference framework for the taxonomy of phenotypical personality traits, whereas the authors of the NEO model (i.e., Costa & McCrae, 1992) presented their FFM explicitly as a model of genotypic traits. Whereas Costa and McCrae have tended to assume that the five NEO factors represent causal determinants of individual differences, supporters of the lexical approach usually have had far more modest goals: They have been looking for a parsimonious but, at the same time, comprehensive model to organize the thousands of terms that refer to interindividual differences in personality in the languages of a wide range of cultures. From such a modest perspective, it is meaningful to test the structural validity of the Big Five across several languages and to search for trans-language regularities.

HISTORICAL OVERVIEW

Even Cattell in 1943 formulated his theory of personality on the assumption that language contained clues on the most important individual differences.

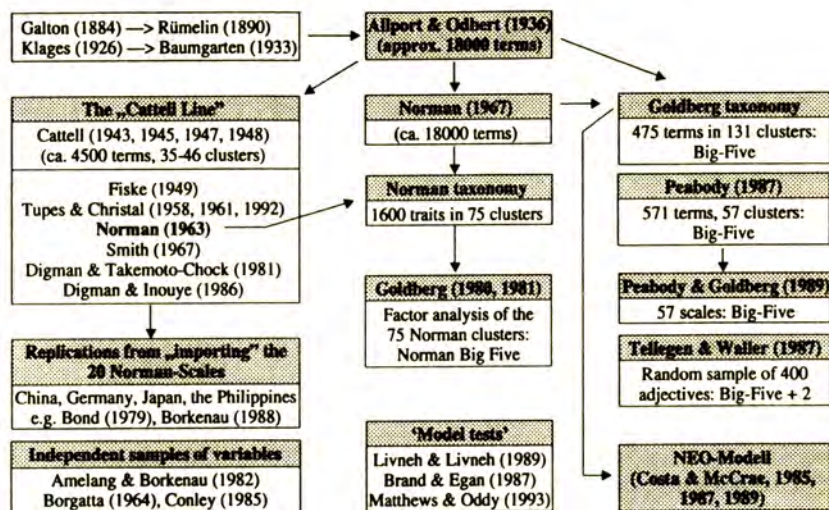


Figure 1. Historical roots of the Big Five.

He used several steps to reduce a list containing approximately 4,500 trait terms compiled by Allport and Odbert (1936) to only 35 rating scales and, finally, to 12 personality factors that formed the basis for his trait theory (see Figure 1). Because Cattell viewed factor analysis as a royal path for detecting underlying source traits, he did not hesitate to interpret his factors in the sense of causal determinants.

However, reanalyses of Cattell's variables in the domain of rating data consistently found that his set of variables were essentially based on the factors named in the FFM (see, e.g., Digman & Inouye, 1986; Digman & Takemoto-Chock, 1981; Fiske, 1949; Norman, 1963; Tupes & Christal, 1958, 1961, 1992).

The robustness of these findings (the most important studies are entered in the left-hand section of Figure 1 under the heading „Cattell-Line“) led many colleagues to assume that these five factors were universal factors of personality perception and description that provided an adequate description of the structure of the language of personality (e.g., Digman & Inouye, 1986). Indeed, the five-factor structure could be replicated on the basis of Cattell's variables in many countries including even China (see Bond, 1979, 1994; Borkenau, 1988).

However, the results of the „Cattell line“ studies do not provide any ideal support for the assumption that the Big Five could represent universal

factors of personality judgment.

First, they are all based on a selection of only approximately 35 rating variables. These variables were compiled by Cattell in a nonreplicable way and were supposed to cover the entire spectrum of the American personality language (see John, Angleitner, & Ostendorf, 1988). However, a few additional studies have found variants of the Big Five independent of Cattell's set of variables in samples of between 20 and 36 rating scales (e.g., Amelang & Borkenau, 1982; Borgatta, 1964; Conley, 1985). Nonetheless, if these are to be universal factors of personality description, it is necessary to perform a powerful test of the model with much more comprehensive samples of terms suitably representing the entire vocabulary of personality-descriptive terms in one language. Highly comprehensive samples of terms are also necessary for methodological reasons, because it is well-known that the selection of the sample of variables has by far the most significant impact on the outcomes of factor analysis (Goldberg & Digman, 1994; Gorsuch, 1983; Velicer & Fava, 1987). However, the „model tests” of Livneh and Livneh (1989), Brand and Egan (1989), and Matthews and Oddy (1993) presented in Figure 1 do not meet the criterion of representative item selection.

Second, it is generally not very difficult to import sets of homogeneous rating or questionnaire scales into one's own culture from another language and to show that the other culture's constructs can be replicated in one's own country. However, to confirm the universality of the Big Five, it is not sufficient to translate the NEO Inventory or the Norman Scales into the widest range of different languages and then to replicate the five-factor structure in each case. It is far more the case that the basic assumption of the lexical approach makes it necessary to study the dimensionality of person description in the vocabulary of each individual language.

The American Norman Big Five are based on a collection of approximately 1,600 trait terms taken from a total of more than 18,000 person-descriptive terms (Norman, 1967). Norman first classified the 1,600 terms to the 10 poles of the FFM according to semantic similarity, and then further differentiated the groups of terms into 75 clusters of synonyms. Goldberg (1980, 1981) factored these 75 clusters of synonyms and produced a clear replication of the Big Five. However, it could be claimed that this outcome was a direct result of the sorting procedure, because, in the first step, the Big Five were used as sorting criteria (John et al., 1988; Waller & Ben-Porath, 1987).

The studies by Goldberg and Peabody (Peabody, 1987; Peabody &

Goldberg, 1989) presented in the right-hand column of Figure 1 in which comprehensive lists of terms were also sorted into groups of synonyms and antonyms cannot be criticized in this way, because they were initially designed to test a completely different model of personality structure. These studies also clearly replicated the Big Five.

Nonetheless, when comprehensive samples of terms are presorted according to content criteria, the procedure is never free from the suspicion that implicit model ideas may have influenced the results.

Representative Variable Selections

A more convincing strategy would be simply to draw a random sample from the population of all personality descriptors.

Several studies have followed such a strategy by asking groups of judges to extract all personality-descriptive terms from the lexicon of a language that were suitable, in the sense of the operational definition of Allport and Odbert (1936, p. 24), to „distinguish the behavior of one human being from that of another.” One could also describe such a definition as a very *permeable, coarse filter* (see Figure 2).

To test the FFM, it would perhaps suffice to limit oneself to a random sample of pages from a dictionary. Tellegen and Waller (1987) extracted a random sample of 400 personality-descriptive adjectives from a smaller American dictionary presumably in an attempt to disconfirm the FFM (see Waller & Ben-Porath, 1987). The 400 adjectives were given to 585 respondents for self-ratings. Factor analyses of these self-ratings clearly confirmed the Big Five as well as two additional factors.

Such additional factors probably result when the selection of terms does not exclude person-descriptive terms referring to, for example, external appearances, physique, states, or attitudes. However, the breadth of validity of the FFM does not extend to the domain of all person-descriptive terms, but is limited, in a stricter sense, to the class of trait or personality terms. Therefore, most studies have used a further, *finer* filter to separate prototypical trait terms from other classes of terms describing emotions and attitudes, physical features, and so forth. However, there is still no unified concept on which filter, or, in other words, which definition of personality, should be used for the further selection of terms, and this may well explain the heterogeneity of findings.

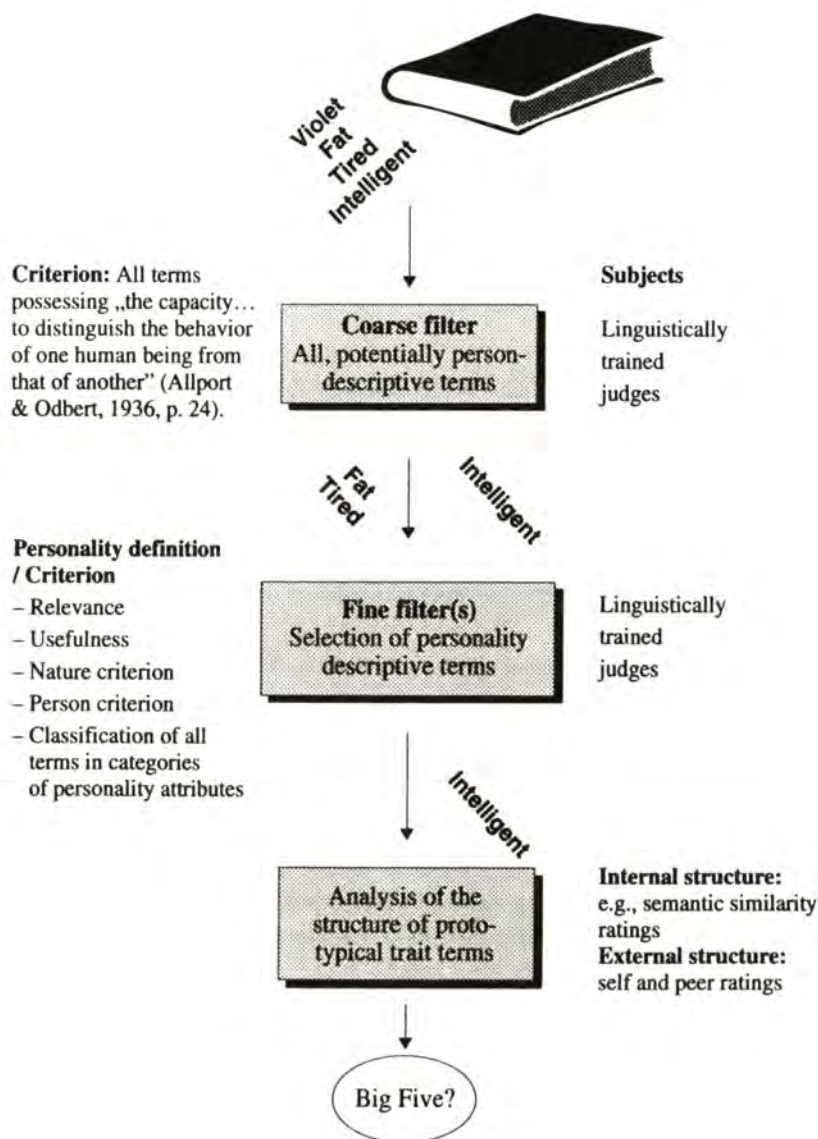


Figure 2. Typical strategies for the selection of personality-descriptive terms.

In the following, we shall only report on the general state of the findings in comprehensive taxonomies (see Figure 3).

Language	Dutch	Italian	German
Authors	De Raad, Hofstee, et al.	Caprara & Perugini	Angstlmer, Ostendorf, et al.
University	Groningen	Rome	Bielefeld
	Adjectives	Adjectives	Adjectives
Source	Dictionary (18,500 adjectives)	Dictionary	Dictionary (11,600 adjectives)
'Coarse filter'	8,690 adjectives; $\alpha = 0.78$	8,532 adjectives	5,160 adjectives; $\alpha = 0.73$
'Fine filter(s)'	Nature, person, fundamentality criteria; $\alpha = 0.84-0.94$	Utility criterion; $\alpha = 0.80$	Categories of personality attributes; $\alpha = 0.85$ (dispositions)
Variables	551 adjectives	492 adjectives	430 adjectives
Subjects	200 pairs of college students	274	400-500
Result	Big Five	Factors I,III,V; II+IV+, II-IV+	Big Five
	Dutch Nouns		
Source	Data base (80,838 nouns)		
'Coarse filter'	8,450 nouns		
'Fine filter(s)'	Usefulness; $\alpha = 0.65-0.76$		
Variables	571 nouns		
Subjects	200 students		
Result	Big Four (I, II, III, V)		
	Adjectives	Adjectives	Adjectives
Source	Dictionary (15,000 verbs)	Dictionary (12,110 verbs)	Dictionary
'Coarse filter'	1,557 verbs	3,914 adjectives	4,900/1,839 adjectives
'Fine filter(s)'	Personality criterion; $\alpha = 0.84$	Familiarity, relevance, trait-state-other categorization; $\alpha = 0.67-0.84$	German classification system; $\alpha = 0.89$ (dispositions)
Variables	507 verbs	624 adjectives	290 adjectives
Subjects	200 students	400	350 - 369 students
Result	Factors II, III, IV; various facets of Factor II	Big Four, I,II,III,IV, plus Integrity (facet of Factor II)	Big Five
		Hungarian	Polish
		Szirmák & De Raad	Szarota
		Groningen	Warsaw

Figure 3. The lexical approach to personality. Major taxonomies of personality descriptive terms.

For example, the Big Five in the Netherlands taxonomy of Hofstee, De Raad, and colleagues (Brokken, 1978; De Raad, 1992; De Raad, Hendriks, & Hofstee, 1992) are based on a representative selection of 551 terms taken from a total list of 8,690 potential person-descriptive adjectives collected, in turn, from the most comprehensive dictionary of the Dutch language available in the year 1978. This means that the 8,690 adjectives were initially extracted from the lexicon using the so-called coarse filter. The 551 adjectives given to 200 pairs for self- or peer-ratings were selected from the list of 8,690 adjectives with a finer filter (De Raad, 1992).

The results reported by Caprara and Perugini (1994) are based on a selection of 492 adjectives that were given to 274 respondents in the form of self-rating scales. These 492 terms were taken from a comprehensive list of 8,532 Italian adjectives. The results of the various projects are naturally not completely identical. For example, Caprara and Perugini's (1994) taxonomy clearly replicated Factors I, III, and V, whereas Factors II and IV

could not be discriminated from each other.

Findings that we consider to be in line with the other results come from initial studies in the Netherlands on personality-descriptive nouns and verbs (De Raad, 1992; De Raad & Hoskens, 1990; De Raad, Mulder, Kloosterman, & Hofstee, 1988). According to the Dutch results the dimensionality of personality descriptions using nouns and verbs seems to be less complex. The noun factors reported by De Raad (1992) show a high correspondence in content to the adjective factors I, II, III, and V. In the domain of verbs, no equivalent was found for the factors Extraversion and Emotional Stability. Nonetheless, it is significant that no new factors were found, and thus that all factors could be localized within the range of meaning covered by the adjective factors.

This is also the case with the findings from a Hungarian taxonomy reported by Szirmák and De Raad. Szirmák and De Raad (1994) clearly replicated the first four of the Big Five in a factor analysis of 624 adjectives, but there was no factor that could be named appropriately with the labels Intellect or Openness. However, the authors also reported that only 14 of their 624 adjectives complied directly with the range of meaning of the American fifth factor.

The Big Five could be replicated in two Slavic languages. Shmelyov and Pokhil'ko (1993) studied the Russian language of personality, and reported results „rather concordant with those of Goldberg (1990)” (p. 1). In Szarota's (1992) study independent judges extracted 4,900 potentially personality descriptive adjectives from a comprehensive dictionary of the Polish language. Szarota (1992) then selected a set of 290 prototypical disposition adjectives and gave these to a sample of approximately 350 subjects for self- and peer-ratings. Factor analyses of these data produced five factors that could be interpreted clearly in terms of the Big Five (personal communication, October 22, 1993).

Interpretation of the Factors

Most of the taxonomies (though unfortunately not all) performed empirical tests on factor interpretations through correlations with external criteria. We shall present examples for such a procedure from our own studies. Our results were based on a representative selection of 430 prototypical disposition adjectives that were given to between 400 and 500 respondents for self- or peer-ratings (see Figure 3; Angleitner & Ostendorf, 1989; Ostendorf, 1990).

One measure for testing the interpretation of the factors was to collect prototypicality indices of the adjectives for the FFM reported by Norman (1963) and for the constructs in several alternative models such as Osgood's (1962) three dimensions of affective meaning, the constructs of the circumplex model of Wiggins (1979), and the three factors in the P-E-N model of Eysenck and Eysenck (1985). We asked 16 independent judges to carry out prototypicality ratings. The mean prototypicality indices of the adjectives were then correlated with the loading patterns of the adjectives on self- and peer-rating factors in order to test the interpretation of these self- and peer-rating factors (Ostendorf, 1992, 1994).

Table 2 presents the results of a joint factor analysis of the prototypicality ratings and their correlations with the loading patterns of the Big Five in a sample of 394 peer-ratings.

In this procedure, the loading patterns of the first five factors in the peer-ratings were correlated with the prototypicality indices of the adjectives, and the resulting correlation matrix was then submitted to a principal component analysis. Five factors were extracted and varimax-rotated. The last five rows in Table 2 report the loadings of the peer-rating factors. We have called these representative factors, because they are based on a representative sample of German disposition adjectives. The other rows present the loadings of prototypicality ratings for Eysenck's and Osgood's dimensions, Wiggins' interpersonal constructs and, finally, the five factors of the Norman taxonomy. Table 2 shows that the prototypicality of the adjectives for the five Norman factors covaried highly with the size of their loadings on the factors of the peer-ratings. Thus, the first five factors in the peer-ratings could be interpreted clearly in the sense of the Big Five.

The loadings of the prototypicality factors for Eysenck's factors were lower, because they came from only one person. In fact, they were provided by Professor Eysenck in person.² The prototypicality ratings in the row Psychoticism, in contrast, came from the group of 16 raters at Bielefeld. As can be seen, there was a high level of agreement between Professor Eysenck's prototypicality ratings and the Big-Five factors Extraversion, Emotional Stability, and Agreeableness. Correspondingly, the Big-Five Factors III and IV seem to assess trait aspects that are not taken into account in Eysenck's P-E-N model.

A second measure for testing the factor interpretations was to correlate the German Big Five with relevant marker variables of the FFM. We collected marker variables by translating five published rating inventories (Goldberg, 1983, 1992; John, Goldberg, & Angleitner, 1984; McCrae &

Table 2. *Varimax Rotated Five-Factor Structure of Prototypicality Ratings and Peer-Rating Factors Based on a Representative Set of 430 German Trait Adjectives*

	Factors				
	(I)	(II)	(III)	(IV)	(V)
Prototypicality ratings					
Norman factors					
Surgency, Extraversion	0.96	-0.14	0.00	0.08	0.04
Agreeableness	0.07	0.95	0.02	0.08	0.13
Conscientiousness	-0.23	0.19	0.88	0.11	0.05
Emotional Stability	0.11	0.02	0.28	0.85	0.21
Culture	0.05	-0.03	0.01	-0.01	0.94
Wiggins circumplex					
Submissive-Dominant	-0.57	0.34	-0.41	-0.36	-0.15
Unassuming-Arrogant	-0.33	0.69	0.09	-0.01	0.01
Ingenuous-Calculating	-0.03	0.80	0.11	-0.14	-0.10
Warm-Cold	0.10	0.87	-0.00	-0.16	0.09
Agreeable-Quarrelsome	0.07	0.83	-0.02	0.23	0.14
Gregarious-Aloof	0.71	0.54	-0.05	0.11	0.04
Extraverted-Introverted	0.93	0.06	-0.00	0.08	0.03
Ambitious-Lazy	0.19	-0.09	0.81	0.17	0.30
Osgood dimensions					
Evaluation	0.20	0.55	0.38	0.12	0.65
Activity	0.68	-0.23	0.30	-0.03	0.49
Potency	0.37	0.12	0.47	0.30	0.66
Social desirability					
	0.21	0.55	0.41	0.14	0.63
Psychoticism					
	0.09	-0.92	-0.14	0.00	0.05
Prototypicality ratings provided by H.-J. Eysenck					
Psychoticism	0.10	-0.78	-0.05	0.05	0.02
Extraversion	0.78	-0.13	-0.28	0.10	-0.01
Neuroticism	-0.38	-0.10	-0.19	-0.71	0.03
Representative rating factors					
(I)	0.93	0.04	0.01	0.04	0.21
(II)	-0.04	0.94	0.03	0.09	0.17
(III)	-0.12	0.23	0.90	0.17	0.16
(IV)	-0.03	-0.07	0.04	0.90	0.03
(V)	0.04	0.06	0.15	0.07	0.96

Note. The prototypicality ratings are average ratings of 16 judges who rated the prototypicality of 430 German trait adjectives for the constructs listed in the table. H.-J. Eysenck rated the same representative set of 430 German trait adjectives for the factors Psychoticism, Extraversion, and Neuroticism. The representative rating factors are based on ratings of 394 subjects who judged one or two acquaintances on the set of 430 adjectives.

Costa, 1985, 1987; Norman, 1963; Peabody & Goldberg, 1989). We applied the NEO Personality Inventory of Costa and McCrae (1985) as a further validity criterion. The mean correlation between the Big Five (assessed on the basis of a representative sample of 430 adjectives) and these criteria was .74 for self-ratings and .73 for peer-ratings. Thus, our study produced extremely high convergent validity coefficients for the five factors (Ostendorf, 1990).

In all, results confirmed the structural validity of the FFM in the area of rating data. Although the five rating factors could not be replicated completely in, for example, the field of Dutch nouns and verbs, the dimensionality of person judgments in these data media seemed to be only less complex than in the field of personality-descriptive adjectives. In several languages, the five factors additionally proved to be robust across various samples of variables, various measurement instruments, different rater groups, and variations in the methods of factor analysis (e.g. Angleitner & Ostendorf, 1989; Caprara & Perugini, 1994; Goldberg, 1990; McCrae & Costa, 1987; Shmelyov & Pokhil' Ko, 1993).

DISCUSSION

The robustness of these findings contradicts the otherwise dominant skepticism toward factor-analytical procedures. It is commonly believed that the results of factor analyses depend to a large extent on sample selection, the method used to extract factors, communality estimations, and the rotation. However, the main findings of factor analyses of self- and peer-ratings seem to be highly robust across these methodological variants (Goldberg, 1990; Ostendorf, 1990). It seems highly probable that the FFM can always be replicated when a sufficiently large, representative set of personality-descriptive statements is studied. The decisive parameter therefore seems to be the size and representativeness of the sample of variables studied.

However, it has to be admitted that the FFM is in no way confirmed by all factor analytic studies. There are hundreds if not thousands of questionnaire studies that do not support the Big Five. However, the bulk of these studies examined samples of variables that had been selected systematically; that is, that have compiled questionnaire items in order to measure previously postulated constructs. Only a few studies have examined the structure of large pools of questionnaire items. But the statement from

Eysenck (1978, p. 478): „Traditional inventories have always been preoccupied with N and E variables, and have not paid much attention to Psychoticism,” also applies to comprehensive collections of questionnaire items – such as the 3,029 items that Browne and Howarth compiled from 17 personality inventories in 1977. We should like to add that traditional inventories have also not paid much attention to features of the factors Agreeableness, Conscientiousness, or Openness. In contrast, the FFM seems to be confirmed when (a) unselected item samples and (b) samples with the broadest possible content are factorized.

If the Big Five are in fact universal, this would indicate that we are dealing with a basic phenomenon that naturally requires explanation. Even critics who view the Big Five as an artefact of factor analysis and are more in favor of, perhaps, cluster analysis or certain multidimensional scaling procedures for methodological reasons, and would perhaps project somewhat different coordinate systems in the trait sphere, would nonetheless find a high degree of trans-cultural stability in the trait structure. And it is precisely these robust parts of the trait sphere that the Big Five describe. Or, in other words: Even with various coordinate systems, one would find that the pattern of stars in the night sky in Italy is hardly any different from that in Hungary, apart from a slight rotation.

However, to continue this analogy, it is well-known that factor analysis determines galaxies and not individual stars as reference points in a system of coordinates, and the crucial question is whether the galaxies that define the axes of the factors are of greater importance than the individual stars. In the personality lexicon, a galaxy is a cluster of synonymous terms that describe classes of similar behaviors. The axes of the five factors thus seem to be positioned at points at which differences in behavior are picked out by a very rich, redundant vocabulary. However, we cannot deduce the extent to which they are determined physiologically from the particularly extensive emphasis on certain individual differences. Thus, factor analysis delivers nothing other than the phenotypical structure of trait ratings, but not causal, for example, central biological mechanisms that might determine this phenotypical structure. Although it has been shown that the phenotypical variance of the Big Three or Big Five has a genetic component, this genetic variance is probably no larger than that of other traits such as dominance, aggressiveness, impulsiveness, sociability, and artistic or scientific interests, and it therefore does not mark the Big Three or Big Five as personality traits that are particularly based on biology (e.g., Loehlin, 1992).

In other words, there can be no doubt that there are genetic and biological

differences that lead to individual differences in behavior, but it is doubtful whether we can deduce the significance of very specific physiological differences from the dominance of the factors in personality judgments. From another perspective, they could be socially, linguistically determined constructions that are not for this reason fictitious, but may well reflect real covariations in behavior.

The question whether the structure of person judgments has anything to do with the covariation in observed behavior has been answered affirmatively, for example, in a study by Borkenau and Ostendorf (1987). This study clearly rejected Shweder and D'Andrade's (1980) systematic distortion hypothesis that trait ratings reflect fictitious and not real behavior covariations. In addition, trait ratings correlate with diverse external criteria, which also contradicts the hypothesis that they reflect only assumed covariations with behavior. Nonetheless, the central dimensions of person judgment are initially social, linguistically determined constructions, because language is the tool with which we not only express ourselves and communicate, but that we also use to think and structure reality. Thus, for example, language is not only involved in assigning trait terms to specific persons or in ascertaining the similarity between trait terms, but it is also linked inseparably to the coding and interpretation of elementary observations of behavior. If behavior observations and trait descriptions are represented in a common semantic space by the members of a language community (Borkenau, 1990, 1992), this could provide an initial explanation of why verbal categories or factors could be valid predictors. However, these assumptions provide no explanation for why the formation of categories should be universal and specifically follow the Big Five dimensions. The universality of the Big Five could be due to the fact that behaviors are generally evaluated with reference to universal social tasks or constraints.

If language has been the most important tool for structuring and conceiving reality since the beginnings of human development, then the five-factor structure could reflect the function of individual differences for the realization of important societal goals, or, in other words, that the structure of the language of personality has been strongly shaped during the course of its development by societal demands that have to be mastered universally in all societies (Hogan, 1983). One fact that supports this hypothesis is that the factors are strongly evaluative. Each judgment using the factors implies a value judgment; a comparison with a behavioral standard. Each judgment using the factors says something about the judged person's possible contribution to the effectiveness of his or her social group.

This means that the factors—or, in general terms, all trait descriptions—reflect thoroughly real behavior covariations, but are simultaneously a product of the instrument, namely, language, with which we structure reality. Other measures of individual differences, which are not based on self- or peer-ratings, may deliver a quite different personality structure. Thus, the factors obtained from rating and questionnaire data are initially nothing other than general dimensions of verbal person description.

REFERENCES

- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47, Whole No. 211.
- Amelang, M., & Borkenau, P. (1982). Über die faktorielle Struktur und externe Validität einiger Fragebogen-Skalen zur Erfassung von Dimensionen der Extraversion und emotionalen Labilität [On the factor structure and external validity of some questionnaire scales measuring dimensions of extraversion and emotional lability]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 119-146.
- Angleitner, A., & Ostendorf, F. (1989, July). *Personality factors via self- and peer-ratings based on a representative sample of German trait descriptive terms*. Paper presented at the First European Congress of Psychology, Amsterdam, The Netherlands, July 2-7, 1989.
- Baumgarten, F. (1933). Die Charaktereigenschaften [The character traits]. *Beiträge zur Charakter- und Persönlichkeitsforschung* (Whole No. 1). Bern: A. Francke.
- Bond, M. H. (1979). Dimensions used in perceiving peers: Cross-cultural comparisons of Hong Kong, Japanese, American and Filipino university students. *International Journal of Psychology*, 14, 47-56.
- Bond, M. H. (1994). Trait theory and cross-cultural studies on person perception. *Psychological Inquiry*, 5, 114-168.
- Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioral Science*, 9, 8-17.
- Borkenau, P. (1988). The multiple classification of acts and the big five factors of personality. *Journal of Research in Personality*, 22, 337-352.
- Borkenau, P. (1990). Traits as ideal-based and goal-derived social categories. *Journal of Personality and Social Psychology*, 58, 381-396.
- Borkenau, P. (1992). Implicit personality theory and the five-factor model. Special Issue: The five-factor model: Issues and applications. *Journal of Personality*, 60, 295-327.
- Borkenau, P., & Ostendorf, F. (1987). Retrospective estimates of act frequencies: How accurately do they reflect reality? *Journal of Personality and Social Psychology*, 52, 626-638.
- Brand, C. R., & Egan, V. (1989). The 'Big Five' dimensions of personality?

- Evidence from ipsative, adjectival self-attributions. *Personality and Individual Differences*, 10, 1165-1171.
- Brokken, F. B. (1978). *The language of personality*. Meppel, Holland: Krips.
- Browne, J. A., & Howarth, E. (1977). A comprehensive factor analysis of personality questionnaire items: A test of twenty putative factor hypotheses. *Multivariate Behavioral Research*, 12, 399-427.
- Caprara, G. V., & Perugini, M. (1994). Personality structure in Italian. *European Journal of Personality*, 8, in press.
- Cattell, R. B. (1943). The description of personality. 2. Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476-507.
- Cattell, R. B. (1945). The principle trait clusters for describing personality. *Psychological Bulletin*, 42, 129-161.
- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12, 197-220.
- Cattell, R. B. (1948). The primary personality factors in women compared with those in men. *British Journal of Psychology*, 1, 114-130.
- 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., & McCrae, R. R. (1985). *The NEO Personality Inventory. Manual. Form S and Form R*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory. Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- De Raad, B. (1992). The replicability of the big five personality dimensions in three word-classes of the Dutch language. *European Journal of Personality*, 6, 15-29.
- De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. B. (1992). Towards a refined structure of personality traits. *European Journal of Personality*, 6, 301-319.
- De Raad, B., & Hoskens, M. (1990). Personality-descriptive nouns. *European Journal of Personality*, 4, 131-146.
- De Raad, B., Mulder, E., Kloosterman, K., & Hofstee, W. K. B. (1988). Personality-descriptive verbs. Special Issue: Personality measurement. *European Journal of Personality*, 2, 81-96.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- 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., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, 16, 149-170.
- Eysenck, H. J. (1978). Superfactors P, E and N in a comprehensive factor space. *Multivariate Behavioral Research*, 13, 475-481.
- 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. (1994). The „Big Five” or „Giant Three”? Criteria for a paradigm. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 37-51). Hillsdale, NJ: Erlbaum.
- 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.
- Galton, S. F. (1884). Measurement of character. *Fortnightly Review*, 42, 179-185.
- Goldberg, L. R. (1980, May). *Some ruminations about the structure of individual differences: Developing a common lexicon for the major characteristics of human personality*. Symposium presentation at the meeting of the Western Psychological Association, Honolulu, Hawaii.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology* (Vol. 2, pp. 141-165). Beverly Hills, CA: Sage.
- Goldberg, L. R. (1983, June). *The magical number five, plus or minus two: Some considerations on the dimensionality of personality descriptors*. Paper presented at a Research Seminar, Gerontology Research Center, NIA/NIH, Baltimore, MD.
- 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., & Digman, J. M. (1994). Revealing structure in the data: Principles of exploratory factor analysis. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 216-242). New York: Springer.
- Goldberg, L. R., & Saucier, G. (1994). O.K., Contrarian — What do you propose we use instead? A reply to Block. Oregon Research Institute, Eugene, OR, USA.
- Gorsuch, R. L. (1983). *Factor analysis*. Philadelphia: W. B. Saunders.
- Hogan, R. T. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *1982 Nebraska Symposium on Motivation* (pp. 55-89). Lincoln: University of Nebraska Press.
- John, O. P. (1990). The ‘Big-Five’ factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality theory and research* (pp. 66-100). New York: Guilford.
- 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, 171-205.
- 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. C. J. Bonarius, G. L. M. Van Heck, & N. G. Smid (Eds.), *Personality psychology in Europe: Theoretical and empirical developments* (Vol. 1, pp. 83-100). Lisse: Swets & Zeitlinger.
- Klages, L. (1926). *Die Grundlagen der Charakterkunde [The science of character]*. Leipzig: Barth.
- Livneh, H., & Livneh, C. (1989). The five-factor model of personality: Is evidence of its cross-measure validity premature? *Personality and Individual Differences*, 10, 75-80.
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Newbury Park: Sage.
- Matthews, G., & Oddy, K. (1993). Recovery of major personality dimensions from trait adjective data. *Personality and Individual Differences*, 15, 419-431.
- McCrae, R. R., & Costa, P. T. (1985). 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. (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., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.
- 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). *2800 personality trait descriptors: Normative operating characteristics for a university population*. Department of Psychology, University of Michigan.
- Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, 17, 10-28.
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur. Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit. [Language and personality structure. On the structural validity of the Five Factor Model of personality]*. Regensburg: Roderer.
- Ostendorf, F. (1992, July). *Comparing the Big Five with Eysenck's Super-Three factor model. Paper presented at the 25th International Congress of Psychology*. Brussels, Belgium, July 19-24, 1992.
- Ostendorf, F. (1994). Zur Taxonomie deutscher Dispositionsbegriffe [Towards a taxonomy of German disposition terms]. In W. Hager & M. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen [Handbook of German word norms]* (pp. 382-441). Göttingen: Hogrefe.
- Ostendorf, F., & Angleitner, A. (1992). On the generality and comprehensiveness of the five-factor model of personality. Evidence for five robust factors in

- questionnaire data. In G. V. Caprara & G. van Heck (Eds.), *Modern Personality Psychology: Critical reviews and new directions* (pp. 73-109). New York: Harvester Wheatsheaf.
- Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, 52, 59-71.
- 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.
- Rümelin, G. (1890). Über die Temperamente [About temperaments]. *Deutsche Rundschau*, 64, 397-412.
- Shmelyov, A. G., & Pokhilko, V. I. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality*, 7, 1-17.
- Shweder, R. A., & D'Andrade, R. G. (1980). The systematic distortion hypothesis. In R. A. Shweder (Ed.), *Fallible judgment in behavioral research. New directions for the methodology of social and behavioral science* (Vol. 4, pp. 37-58). San Francisco: Jossey-Bass.
- Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement*, 27, 967-984.
- Szarota, P. (1992). *Classifying words into person descriptive categories*. Unpublished research report, University of Bielefeld, Germany.
- Szirmák, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality*, 8, 95-117.
- Tellegen, A., & Waller, N. G. (1987, August). *Re-examining basic dimensions of natural language trait descriptors*. Paper presented at the 95th annual convention of the American Psychological Association. New York, August 28 - September 1, 1987.
- Tupes, E. C., & Christal, R. E. (1958). *Stability of personality trait rating factors obtained under diverse conditions*. Research Report, USAF, Lackland Air Force Base, Texas.
- Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings*. Technical Report, USAF, Lackland Air Force Base, Texas.
- Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60, 225-251.
- Velicer, W. F., & Fava, J. L. (1987). An evaluation of the effects of variable sampling on Component, Image, and Factor Analysis. *Multivariate Behavioral Research*, 22, 193-209.
- Waller, N. G., & Ben-Porath, Y. S. (1987). Is it time for clinical psychology to embrace the five-factor model of personality. *American Psychologist*, 42, 887-889.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395-412.

AUTHOR NOTE

Translated from German into English by Jonathan Harrow at the University of Bielefeld. We would like to take the opportunity to thank Professor Eysenck for his helpful support.

Correspondence concerning this paper should be addressed to Fritz Ostendorf, University of Koblenz-Landau, Department of Psychology, D-76829 Landau or to Alois Angleitner, University of Bielefeld, Department of Psychology, D-33501 Bielefeld.