Separating Description and Evaluation in the Structure of Personality Attributes

Gerard Saucier

Personality attributes typically combine descriptive and evaluative aspects, as do most personality constructs and the factors in the Big Five taxonomy. However, an auxiliary representation, in which the descriptive and evaluative aspects are clearly separated, has technological utility in some areas of research and assessment. D. Peabody and H. G. Gough independently developed structural models of this type, each with 3 dimensions. Studies used a multiple-method, multiple-replication approach, with data sets involving a large, representative set of familiar personality adjectives, to uncover a structure with 5 orthogonal dimensions, 4 descriptive and 1 highly evaluative, easily measurable with adjective scales. These dimensions demonstrated a complex relation to previous Peabody factors and to the Big Five, and they may be useful in integrating some venerable constructs into the structure of personality attributes.

Just as there are different personality variables on which people might be compared, there are different ways of making the comparison. Over two decades ago, Peabody (1967) pointed out that personality comparisons normally represent a confound of evaluative and descriptive aspects. Consider the English adjective *conscientious*. If you tell me I am "conscientious," I feel complimented, because that is a desirable, positively evaluated attribute. But you are also providing specific descriptive information, because you used this specific positive attribute instead of *sympathetic* or *relaxed* or *creative*.

The typical personality adjective has a clear evaluative content, positive or negative, as well as a descriptive content (cf. Bochner & Van Zyl, 1985; Hampson, Goldberg, & John, 1987). Likewise, the typical personality scale involves evaluative as well as descriptive elements; one pole of each scale is typically more desirable than the other pole. Each of the five broadest domains from the language of personality (Goldberg, 1990, 1992; John, 1990; McCrae & Costa, 1985) confounds to some degree descriptive and evaluative elements. The consensual Big Five structure maximizes parsimony and simple structure and can integrate a broad range of personality constructs. The Big

Five structure performs these functions so well partly because it reflects the same confounding of description and evaluation found in most lay and expert judgments and in most personality scales.

However, Peabody's (1967) separation of descriptive and evaluative components made possible an alternative way of making personality comparisons. Separating description and evaluation in personality attributions proves useful in four specific ways.

First, as Peabody (1968, 1985) showed, the separation enables us to remove most ethnocentric biases from cross-cultural or cross-national comparisons of personality, because divergent groups often agree on the descriptive element with regard to one another, while disagreeing on the evaluative element.

Second, just as in comparisons of cultures, in certain contexts of assessment, personality comparisons are best left nonevaluative. For partners in egalitarian relationships—intimate, collegial, or work-related—it may be beneficial to learn to see one as "not better or worse, just different" from the other partner. In the United States, the production of such egalitarian assessments (although often with measures of questionable value) is a thriving industry (Moore, 1987).

Third, as illustrated by Gough's (1987) "structural scales" in the revised California Psychological Inventory (CPI), nonevaluative dimensions, when understood in conjunction with an evaluative dimension, permit a nonjudgmental assessment, wherein any personality style is conceptualized as a developmental continuum ranging from undesirable to neutral to desirable characteristics. Thus, both strengths and weaknesses of a style of personality can be revealed. When the person assessed, and not a court or an institution, is the only customer for the assessment results, this style of presentation is sometimes advantageous.

And fourth, nonevaluative dimensions may highlight important phenomena in human diversity. The interindividual variation defined by personality traits is likely to occur only where there are multiple solutions to adaptive challenges and thus where evolutionary selection pressure is fairly neutral (cf. Fisher, 1922; Mayr, 1988). Nonevaluative personality dimen-

sions could represent extremes of such neutrality, thus facilitating detection of effects known in population biology (E. O. Wilson & Bossert, 1971) as *frequency-dependent selection, diversifying* or *disruptive selection forces*, and *balanced polymorphisms*. Under these effects, highly divergent and even opposite traits are maintained within a population because each trait has its own adaptive utility. Just as black cats and spotted cats are each adaptive in divergent ways, some opposite human personality traits may be equally adaptive. Moreover, the diversity in adaptations may be useful to the population as a whole.

Some (e.g., Edwards, 1957), seeing social desirability as a contaminating and distorting element in responses to personality measures, have proposed that personality scales and constructs be reconceived with desirability removed, with the assumption that eliminating desirability eliminates the opportunity for faking. This argument has spawned a long and complex controversy (e.g., Dicken, 1963; Heilbrun, 1964; Paulhus, 1984; Taylor & Brown, 1988) that remains incompletely resolved. To this point, there seems to be inadequate support for separating description and evaluation in personality as a general rule. The confounding of description and evaluation is ubiquitous and generally useful. But given the ways in which separating description and evaluation seems useful, it is practical to propose that some personality scales and constructs should be unconfounded for some purposes.

Defining Basic Descriptive and Evaluative Dimensions

In recent personality research, one finds two distinct ways of ascribing the word *basic* to a trait or factor. One way, championed by Eysenck (1991) and others (Costa & McCrae, 1992), defines as *basic* a dimension on which many lines of research evidence, including biological evidence, converge. According to this view, one reviews evidence from scholarly and empirical studies to define a basic dimension. A second way of defining basic dimensions is provided by the *lexical hypothesis*, which states that the most important distinctions in personality are those encoded most prominently in the natural language. Following this hypothesis, rigorous study of the language of personality will reveal the basic dimensions of personality.

Fortunately, these two ways of defining basic dimensions frequently converge. The Big Five factor structure has been delineated by both the lexical approach (Goldberg, 1990; Peabody, 1987; see John, 1990, for a review of Dutch and German lexical studies) and the nonlexical approach (e.g., Costa & McCrae, 1988; McCrae, 1989; McCrae & Costa, 1985). Two approaches are doubtless better than one. In these studies I begin with the lexical approach, but then compare the results with relevant findings from prior nonlexical research.

Before investigating the content of descriptive dimensions, I must define criteria for the labels *descriptive* and *evaluative*. I assume that a descriptive dimension will show a near-zero relation to desirability values of variables; an evaluative dimension will show a near-unity relation to desirability values of variables and a near-zero relation to any descriptive dimensions.

Lessons From Previous Studies

Among previous studies, the largest contribution has been provided by Peabody (1967, 1970, 1984, 1985), who developed

one ingenious method for separating description and evaluation. Peabody created sets of four terms presenting both a descriptive and an evaluative contrast. For example, for the antonym pairs *cautious* versus *rash* and *bold* versus *timid* as two bipolar scales, evaluation and description are confounded. However, whereas *cautious* and *bold* present an evaluative contrast with *rash* and *timid*, *cautious* and *timid* present a descriptive contrast with *bold* and *rash*. Depending on whether one adds or subtracts the bipolar scales, one can obtain either a descriptive or an evaluative score.

Peabody (1967) originally constructed 15 sets-of-four from common personality adjectives; although this small selection of variables does not represent a rigorous application of representative sampling from a full pool of common personality adjectives (as in Peabody, 1987), it does represent an application of the lexical hypothesis. Using these sets of adjectives in studies of trait inferences, Peabody (1967, 1984) repeatedly found two descriptive factors in unrotated solutions or in deliberate rotations (De Boeck, 1978; Peabody, 1978, 1984).

Peabody's work provides the sole previous application of the lexical method to the task of separating description and evaluation. Nonetheless, other investigations using personality scales and items assembled by various other methods merit some review. Osgood and his collaborators (Osgood & Suci, 1955) found three "affective" dimensions—Evaluation, Potency, and Activity—used by people in describing diverse concepts. Because they are orthogonal to Evaluation, presumably Potency and Activity could be descriptive dimensions. But these orthogonal relations do not hold up well in descriptions of people (cf. Osgood, Suci, & Tannenbaum, 1957). Undeterred, a number of researches (Hallworth, 1965; Krieger, 1964; Kuusinen, 1969; Tzeng, 1975) sought to define structures of personality traits with Osgoodian dimensions as the first elements, but their studies did not converge on a common set of personality dimensions orthogonal to Evaluation.

Some frequently used personality inventories purport to separate description and evaluation. From the item pool of the CPI (Gough, 1987), Gough developed "structural" scales that measure two descriptive dimensions and a third evaluative dimension. Factor analyses of the CPI (see Levin & Karni, 1981) had repeatedly turned up three factors, one of which corresponded generally to CPI profile elevation and thus to evaluation, because almost all the CPI folk-concept scales are evaluation confounded and scored in the desirable direction. Through iterative item analyses, Gough refined three nearly orthogonal scales that are essentially slight rotations of these factors. The generalization of these factors to item sets other than the CPI remains to be determined.

Because the 20 scales of the Personality Research Form (PRF; Jackson, 1984) were developed with attention to minimizing social desirability responding, many of these scales may relate to broader nonevaluative, descriptive dimensions. However, the scales range from moderately evaluative to nonevaluative, as is indicated by the scales' mean item desirabilities. Likewise, the four scales of the Myers–Briggs Type Indicator (Myers & McCaulley, 1985), which have been shown to measure basic personality traits (McCrae & Costa, 1989), may relate to broad descriptive dimensions, although the consistent evaluative neutrality of these scales has not been established.

In summary, previous research has not taken advantage of the powerful rationale for discovering basic dimensions provided by the lexical hypothesis. Peabody's studies, which most approach the lexical rationale, were not based on a fully representative selection of variables and studied trait inferences rather than ratings of actual persons. Moreover, there is little evidence of convergence among studies by different researchers. Previous studies offer a hodgepodge of likely descriptive dimensions with no consensual organizing structure to link them together.

Drawing on studies involving large numbers of personality adjective variables and of subjects, the present studies are aimed at defining basic personality dimensions that clearly separate descriptive and evaluative elements, thereby organizing the present profusion of purportedly nonevaluative dimensions.

Study 1: Finding Basic Descriptive Factors

Method

Paulhus (1981) grouped methods used to control desirability into the categories of rational, covariate, and factor-analytic. Rational methods build control into the assessment instrument by using neutral items, subtle items, or a forced-choice format. Covariate methods statistically partial social-desirability responding from each subject's responses and thus from the resulting structural dimensions. Factor-analytic methods fix a factor identifiable with evaluation and thereby, through orthogonal rotation to simple structure, force subsequent factors into nonevaluative positions.

The strategy used in the present studies involves a combination of all three methods. In Study 1, rational and covariate methods are used to define descriptive dimensions that are replicable between methods and between samples. In Study 2 an evaluative dimension orthogonal to the descriptive dimensions is developed that provides, by way of an evaluative factor, a factor-analytic separation of description and evaluation.[1]

Selection of variables. Goldberg (1982) described the refinement of a set of 566 common English trait adjectives from the more unwieldy sets of 2,797 (Norman, 1967) and 1,710 adjectives culled in earlier dictionary-based lexical studies. The 566 include 46 amplifications (e.g., *overconfident* and *overfeminine*) included only for exploratory purposes. Removing the amplifications gives a set of 520 trait-descriptive adjectives (520-TDA), arguably a representative set of familiar trait terms, being the set of 540 terms studied by Hofstee, de Raad, and Goldberg (1992) minus 20.

In Study 1, I used two distinct subsets of the 520-TDA. One subset consisted of the 98 most neutral items in the 520. As a criterion for neutrality, social desirability scale values (SDSVs; Norman, 1967) for the 520 adjectives were examined; those 98 adjectives whose mean ratings of desirability by 100 (50 men and 50 women) judges fell between 4.00 and 6.00 on a 1–9 scale were retained. Table 1 lists the 98 adjectives (98-NA) with their respective SDSVs, followed by a very small sampling of the 422 nonneutral adjectives in the 520-TDA. Using only neutral items is a *rational* method for controlling desirability, restricting the evaluative range of variables.

The *covariate* method, in contrast, restricts the evaluative range of the targets of description. For an application of the covariate method, all terms, even the nonneutral ones, might be included; partialing item social desirability values from subject responses should remove evaluation from the factors. Because the set of 520 included some unfamiliar terms (e.g., *glib* and *circumspect*), a subset of 394 highly familiar trait-descriptive adjectives (the 394-TDA) studied previously (e.g., Hampson, Goldberg, & John, 1987) was chosen for the covariate analyses. A list of these adjectives is available from the author.

Samples. The two primary samples for this study were, first, a *self* sample including 320 college students who described themselves on 587 trait adjectives, including the entire 520-TDA, using the response format described in Studies 3 and 4 of Goldberg (1981; also Goldberg, 1990, Study 2); a variety of middle-response options described there were given a midscale value of 0 on a rating scale ranging from −3 to 3. A second, *peer* sample included 316 of the 320 subjects from the self sample who used the same inventory of 587 terms to describe someone of their sex and approximate age whom they knew well and liked.[2]

Data from a third sample were used in some analyses. The *pooled peers* sample included 205 students in law school or upper-division psychology courses, roughly one third of whom were randomly assigned to describe one of three types of peer targets: liked, neutral, or disliked (see Goldberg, 1990, Study 2; Goldberg, 1992, Study 4). These subjects used the set of 566 adjectives as a stimulus set, with an 8-step rating scale, and omitted responses assigned a midscale value (5) on a transformed 1–9 rating scale. The evaluative heterogeneity of targets in this sample led to decidedly evaluation-confounded factors in the analyses of the 98-NA. Partialing social desirability from subject responses, however, removed the differences in the evaluation of targets and resulted in more consistently nonevaluative factors. This sample provides an additional replication test for factors derived by the covariate method, but was not used for studies of the neutral terms.

Analyses. To eliminate individual differences in subjects' use of the rating scales, for the 98-NA analyses the responses of each subject were ipsatized on the basis of the full set of terms in the inventories. Factor analyses drew on the correlation matrix of the 98-NA in each data set.

Analyses of the 394-TDA were based on desirability-partialed responses. SDSVs (from Norman, 1967) were partialed from the responses of each subject across all terms in the inventory. Factor analyses were performed on the covariance matrix, rather than the correlation matrix, of social-desirability residuals in each data set. After partialing, intervariable correlations, unlike covariances, rescale the variables, magnifying the lesser variance left to be explained on the evaluatively extreme variables (e.g., *honest, unintelligent*) and minimizing the considerable variance remaining on more neutral variables (Cudeck, 1989; Rosenberg & Olshan, 1970).

For each selection of variables (98-NA or 394-TDA) in each sample, factor analysis, both principal factors and principal components,[3] was executed with normalized varimax rotation of four, five, and six factors.[4] The number of appropriate factors was determined by a multiple-replication approach (cf. Walkey & McCormick, 1985). I sought to define a maximum number of factors that would replicate between samples using each method (neutral-terms or covariate) and that also would converge between methods within each sample. Between-samples

---

[1] Paulhus (1981) outlined a method of factor-analytic control, assuming that the unrotated first factor can be identified with evaluation. This method is not used in the present studies; in practice (even in large or representative sets of trait adjectives), the unrotated first factor tends to gravitate toward the largest cluster of high-correlating variables, a cluster which almost never corresponds precisely to evaluation or social desirability.

[2] For the analysis of 394 variables, there are more variables than subjects. As explained in an earlier study (Saucier, 1992), the large sample sizes and huge variable-to-factor ratios make this characteristic unproblematic (cf. Guadagnoli & Velicer, 1988).

[3] Generally, few differences were found between principal-components and principal-factors solutions, as found previously for trait adjectives by Goldberg (1990, Study 1). As a result, only principal components, the less complex procedure, was used in the subsequent study.

[4] Seven- and eight-factor rotations were also examined for the 98-NA data sets, producing no evidence of convergent dimensions beyond that provided by rotations of fewer factors.

Table 1
List of 98 Neutral Personality Adjectives, With Illustrative Nonneutral Adjectives

SDSV	Adjective	SDSV	Adjective	SDSV	Adjective
			Neutral		
5.03	Acquiescent	4.91	Formal	5.93	Sentimental
5.77	Acquisitive	4.69	Glib	5.98	Sexy
5.35	Aggressive	5.91	Gregarious	4.35	Shrewd
5.96	Animated	5.24	Homespun	4.23	Shy
4.05	Argumentative	5.87	Imperturbable	4.20	Silent
4.33	Austere	4.53	Impetuous	5.94	Simple
4.16	Bashful	5.06	Impulsive	4.98	Skeptical
5.97	Bold	4.84	Indulgent	5.59	Soft
5.94	Circumspect	4.16	Malleable	4.87	Solicitous
4.62	Compulsive	4.12	Manipulative	4.81	Somber
4.69	Conservative	5.88	Meticulous	4.46	Stern
4.30	Conventional	5.80	Mischievous	4.36	Strict
4.65	Crafty	5.44	Nonchalant	5.39	Subjective
4.86	Critical	5.75	Nonconforming	5.14	Suggestible
5.16	Cunning	4.40	Opinionated	5.34	Talkative
4.07	Demanding	5.71	Opportunistic	4.40	Tempestuous
5.73	Demonstrative	5.29	Outspoken	5.63	Tenacious
4.29	Dependent	4.69	Patronizing	4.53	Terse
4.33	Detached	5.81	Perfectionistic	4.49	Theatrical
5.35	Devil-may-care	5.27	Placid	4.96	Tough
4.73	Docile	4.93	Predictable	4.83	Traditional
4.66	Dominant	4.50	Provincial	5.65	Unassuming
4.52	Dramatic	5.51	Quiet	5.75	Unconventional
5.90	Earthy	5.07	Rambunctious	4.09	Unexcitable
4.44	Eccentric	4.77	Rebellious	5.13	Unpredictable
5.69	Emotional	5.29	Reserved	4.31	Unrestrained
5.57	Exacting	5.21	Restrained	4.13	Untalkative
5.58	Excitable	4.45	Sarcastic	5.72	Urbane
5.79	Fastidious	4.18	Seclusive	4.05	Verbose
4.71	Flamboyant	5.04	Sedate	5.71	Vigilant
4.10	Flippant	4.09	Self-indulgent	4.61	Volatile
4.86	Flirtatious	4.92	Selfless	5.58	Wary
4.86	Folksy	5.44	Self-seeking		
			Illustrative Nonneutral		
1.64	Bigoted	2.63	Unimaginative	8.39	Intelligent
1.61	Deceitful	2.77	Unstable	8.48	Kind
2.14	Hypocritical	8.35	Considerate	7.67	Mature
3.71	Introverted	7.66	Efficient	8.06	Warm
2.78	Jealous	6.58	Extraverted	7.86	Wise

Note. SDSV = Social-desirability standard value, on a 1-to-9 scale, from Norman (1967; Hampson, Goldberg, & John, 1987).

(within-method) replication was indexed by correlating factor loadings across the 98 or the 394 items as applicable.[5] Between-methods (within-sample) replication was indexed by correlating factor scores across subjects for factors derived from the neutral-terms and the covariate approaches. For each factor, the degree of evaluative confounding was indexed by correlating factor loadings with Norman (1967) SDSVs across items.

Results

Table 2 presents the between-samples replication coefficients for rotations of six, five, and four factors.[6] The mean replication coefficients for the four-factor rotations are generally highest, the six-factor rotations generally being lowest.

Table 3 provides correlations between factor loadings and social desirability for each of the factors in the self and peer sam-

ples. None of the desirability-partialed factors exhibited a significant correlation with social desirability. Some of the neutral-terms factors did show significant correlations with desirability, although these correlations were taken across variables that are all neutral. The four-factor rotation of 98-NA variables in the self sample yielded four unconfounded factors. There were five unconfounded factors in the five-factor rotation in the self sam-

[5] In data sets with personality rating data, Saucier (1991) found a correlation of .99997 between (a) congruence coefficients determined by correlating factor loadings and (b) Tucker's coefficients of factor congruence (Mulaik, 1972, p. 335).

[6] Tables presenting factor loadings for all variables in all these analyses are available from the author.

Table 2
Replication Coefficients for Descriptive Factors

	Between-sample comparisons of factors			
	Neutral terms	Social desirability partialed		
Factor	Self–peer	Self–peer	Self-pooled	Peer-pooled
Six-factor rotation				
1	.93	.92	.88	.89
2	.85	.89	.82	.80
3	.76	.84	.77	.65
4	.71	.72	.64	.61
5	.73	.72	.70	.39
6	.29	.59	.33	.01
M	.69	.78	.69	.56
Five-factor rotation				
1	.93	.91	.87	.85
2	.52	.90	.81	.79
3	.72	.83	.78	.89
4	.70	.67	.51	.47
5	.17	.69	.67	.58
M	.61	.80	.73	.72
Four-factor rotation				
1	.91	.86	.86	.86
2	.69	.88	.60	.73
3	.88	.80	.80	.87
4	.78	.63	.53	.31
M	.81	.79	.70	.69

Note. Values are correlations between factor loadings across 98 items (neutral terms) or 394 items (social desirability partialed) for between-sample corresponding factors. Self sample $n = 320$; peer sample $n = 316$. Pooled sample ($n = 205$) neutral terms factors are not considered because of their obvious confound with desirability. Values based on principal-factors analyses.

ple, but this factor structure was inadequately replicated in the peer sample (see Table 2).

Thus, the replication coefficients pointed toward four descriptive factors and, as Table 3 shows, the four factors were usually unconfounded. Table 4 gives between-samples replication coefficients for the four-factor solution in greater detail.[7]

Using either the rational (neutral-terms) or the covariate method, four factors appear to be reasonably replicable across samples and not confounded with evaluation. However, this provides no evidence of between-methods convergence within either sample. Table 5 presents factor-score correlations between the four factors as derived by the two methods. These coefficients were generally lower than the between-samples, within-method coefficients but were high enough to suggest convergence.

Full interpretation of the four factors was deferred until the factors were further replicated, but provisional interpretations were ventured. The first factor highlighted terms such as *talkative* and *outspoken*, as opposed to *quiet* and *reserved*, clearly referring to Extraversion or interactional orientation. The second factor highlighted terms such as *emotional* and *sentimental*, as opposed to *firm* and *tough*, apparently referring to affective orientation. The third factor highlighted terms such as *traditional* and *religious*, as opposed to *nonconforming* and *unconventional*, apparently referring to norm orientation. The fourth factor highlighted terms such as *critical* and *sophisticated*, as

opposed to *casual* and *folksy*, apparently referring to degree of formality, or "form orientation."

Study 2: Replicating Factor Markers

Study 1 provided evidence that four is the maximum number of broad descriptive factors that can pass tests both of replication and of independence from social desirability. In Study 2, I sought to demarcate a set of markers for these four dimensions, develop markers for a highly evaluative dimension fully orthogonal to the four dimensions, subject these markers to a replication test in a new sample, and interpret the dimensions and their relation to previous personality factors.

Markers for the Four Dimensions and an Evaluative Dimension

Given findings from Study 1, between-samples, between-methods evidence for four descriptive factors, I sought to operationalize this four-factor structure. I examined correlations of

[7] An earlier study by Saucier (1991), representing an attempt to define descriptive factors by way of the desirability-partialing method alone, led to the same conclusion by a differing route. A fifth descriptive factor failed to replicate in an independent sample of 132 community college students.

Table 3
Correlations of Factor Loadings With Social Desirability

	Neutral terms		Social desirability partialed	
Factor	Self	Peer	Self	Peer
Six-factor rotation				
1	.14	.21*	.03	.04
2	.11	.07	−.01	.00
3	−.19	−.15	−.02	.06
4	.28**	.08	.03	.03
5	.26**	.37**	.05	.07
6	.08	.09	.00	.05
Five-factor rotation				
1	.16	.21*	.03	.04
2	−.01	−.01	−.01	.00
3	−.09	−.11	.02	−.03
4	.15	.04	.03	.05
5	−.13	−.36**	.06	.08
Four-factor rotation				
1	.14	.19	.03	.05
2	−.04	−.06	−.02	.00
3	−.12	−.09	.05	.05
4	−.14	−.36**	.01	−.05

Note. Self sample *n* = 320; peer sample *n* = 316. Neutral terms factors derived from correlation matrix of 98 adjectives with each subject's responses ipsatized; social desirability partialed factors derived from a covariance matrix of 394 adjectives, based on residuals after standard social desirability values were partialed from the responses of each subject. Values are based on principal-factors analyses. Factors corresponding most closely in content are placed in the same row.
* $p < .05$. ** $p < .01$.

the 587 adjectives in the peer and self samples with factor scores from the four-factor rotations in Study 1, seeking adjectives that correlated highly with one of the factors and that were unrelated to the remaining factors.

Table 4
Between-Sample Correlations of Factor Loadings for Descriptive Factors Derived by Two Methods

	Factors in self-ratings			
Factors in peer ratings	First	Second	Third	Fourth
Social desirability partialing method				
First	.86**	−.01	−.25**	−.13
Second	−.39**	.88**	−.02	−.21
Third	−.34**	−.02	.80**	−.41**
Fourth	−.54**	.12	.36**	.63**
Neutral-terms method				
First	.92**	.44**	−.37**	.40**
Second	−.08	.69**	.21*	−.56**
Third	−.36**	−.45**	.88**	−.21*
Fourth	.41**	.27**	−.15	.78**

Note. Self-rating sample *n* = 320; peer rating sample *n* = 316. Correlations between corresponding factors are listed in boldface type. Neutral terms factors are derived from a correlation matrix of 98 adjectives from responses ipsatized by subject, social desirability partialed factors from the covariance matrix of residuals from partialing social desirability values of adjectives from each subject's responses. Values are based on principal-factors analyses.
* $p < .05$. ** $p < .01$.

Through an iterative process of factor analyses with successive versions of a set of adjectives, a set was developed that could maintain its factor structure without desirability partialing, using subject responses that were merely ipsatized (as was the 98-NA in Study 1).

At all points in this iterative item-selection process, markers for each dimension were kept in evaluative balance, creating systematically nonevaluative scales. That is, the mean SDSV (from Norman, 1967) of the adjectives on one pole were kept within about 0.20 on a scale from 1.00 to 9.00 of the mean SDSV of the adjectives on the other pole.[8]

Delineation of an orthogonal fifth dimension identifiable with evaluation was deemed useful for three reasons: (a) An evaluative dimension would introduce factor-analytic control (cf. Paulhus, 1981) for evaluation in the other four dimensions when rotated with them; (b) such a dimension would facilitate comparison with structures like that of Peabody and Gough that contain such a dimension; and (c) such a dimension, un-

[8] The evaluative-balancing requirement had few consequences for any dimension except the fourth. The four-factor rotations in Study 1 yielded a fourth dimension with content centered on the adjectives *critical* and *fault finding* (vs. *uncritical, overtolerant, folksy,* and *earthy*). However, Machiavellian adjectives such as *sly, manipulative,* and *egotistical* were also associated with this factor, and inclusion of these traits, uniformly undesirable, in tentative scales made the scales impossible to balance evaluatively. To preserve the evaluative balance of the scale, this dimension was iteratively shifted to make some more desirable traits (e.g., *exacting, particular,* and *perfectionistic*), previously only peripheral facets of the factor, more central.

Table 5

Within-Sample Factor-Score Correlations Between Descriptive Factors Derived by Two Methods

Neutral-terms factors	Social desirability partialed factors			
	First	Second	Third	Fourth
Self-rating sample				
First	.83**	.39**	.05	-.23**
Second	-.21**	.55**	-.07	-.33**
Third	.08	.08	.54**	.33**
Fourth	.14*	-.07	-.44**	.49**
Peer rating sample				
First	.76**	.20**	.10	-.18**
Second	-.27**	.68**	-.01	-.10
Third	.04	-.03	.77**	-.07
Fourth	-.06	-.26**	-.05	.62**

Note. Self-rating sample *n* = 320; peer rating sample *n* = 316. Correlations between corresponding factors are listed in boldface type. Neutral-terms factors are derived from a correlation matrix of 98 neutral adjectives, social desirability partialed factors from the covariance matrix of residuals from partialing social desirability values of adjectives from each subject's responses. Values based on principal-factors analyses.
* *p* < .05. ** *p* < .01.

Table 6

Replication and Correlation With Social Desirability of Initial Set of 100 Adjectives Marking Four Descriptive Dimensions and One Evaluative Dimension

Factor	Correlation with social desirability		Replication coefficient
	Combined sample	Replication sample	Combined-replication samples
First	.05	.01	.95*
Second	.15	.19	.88*
Third	.09	.10	.86*
Fourth	.10	.01	.82*
Fifth	.88*	.86*	.90*

Note. Combined sample *n* = 636; Replication sample *n* = 250. Correlations of |.30| and above are listed in boldface type.
* *p* < .01.

derstood in conjunction with the descriptive dimensions, might permit one to understand each personality style as a developmental continuum, with a range from undesirable to neutral to desirable traits.

A set of markers for this evaluative dimension was derived in the following manner. First, 60 nonneutral adjectives from among the 587 showing very low multiple correlations with the four descriptive dimensions were selected. In each sample, the unrotated first factors from these 60 adjectives was extracted and used to derive an initial approximation of a marker scale. Eventually, through an iterative process of factor and item analyses, a marker scale was developed that interfered minimally with the already established factor axes of the four descriptive dimensions yet demonstrated considerable internal consistency. A set of 100 adjectives designed to measure the four descriptive and the single evaluative dimension was eventually demarcated.[9]

Because this initial set of adjectives was derived from item analyses in the same data sets in which the factor structure was derived, the robustness of the structure of four descriptive factors and one evaluative factor remained to be tested. Accordingly, a replication study was undertaken. This study had three goals: (a) Testing the replicability of the five factors from the 100 adjectives, (b) improving the convergent and divergent reliability of the scales by testing alternative items that might be added or substituted, and (c) defining the relation of these factors to the Peabody (1967, 1984, 1985) dimensions.

Method

Selection of variables. A total of 306 trait adjectives was administered in a separate alphabetized sequence of 102 on each of three pages, and the order of the three pages was counterbalanced across subjects. Subjects rated the self-descriptive applicability of each adjective using a 1–9 rating scale.

Added to the 100 adjectives described above were numerous candidate items for inclusion in the marker scales and 56 adjectives from Peabody's (1968, 1985) 14 "sets-of-four."[10]

Sample. Subjects were 250 university undergraduates (84 men and 166 women) enrolled in introductory psychology courses who agreed to participate in exchange for nominal course credit and the opportunity to receive nonevaluative information about their "personality style."

Analyses. Subject responses were ipsatized across the 306 adjectives, and the initial set of 100 adjectives was analyzed by principal components, with normalized varimax rotation of five factors; in Study 1 virtually no difference was found between principal-components and principal-factors solutions, so in this study the simpler procedure alone was used. As a replication index, correlations between loadings on these five factors and loadings on the corresponding five factors in the combined peer-self sample were calculated. Correlations of the loadings with social desirability (SDSVs from Norman, 1967) were calculated.

Next, correlations of the five sets of factor scores with the remaining terms in the 306-adjective inventory were calculated and used in refining the scales. Reliability statistics for these refined scales were then computed.

For a comparison with the Peabody factors, I computed sets of factors from 14 sets-of-four and from the 28 bipolar items. Because Peabody (1967, 1985) used original rather than ipsatized responses of subjects, original responses were used here also. Then factor scores on Peabody factors were compared with those from refined factor markers from the present study.

Results

Table 6 presents the replication coefficients and correlations with social desirability of the five factors from the initial set of

_____

[9] A list of these 100 adjectives is available from the author.

[10] Two of Peabody's adjectives, *cooperative* and *uncooperative*, were mistakenly left off the list, being replaced in analyses by two terms, *considerate* and *inconsiderate*, that performed very similarly to the two omitted terms in recent Big Five analyses (Hofstee, de Raad, & Goldberg, 1992). And because of ambiguity introduced by recent changes in word meaning, the adjective *gay* (Peabody, 1985, continued from Peabody, 1967) was replaced with the adjective *light-hearted*, at Peabody's suggestion (D. Peabody, personal communication, November 23, 1991). These substitutions had no apparent consequences on the results.

Table 7

Factor Loadings of Final Set of 100 Adjectives Marking Four Descriptive Factors and One Evaluative Factor

	Factor				
Adjective	First	Second	Third	Fourth	Fifth
Interactional orientation					
Talkative (+)	**.76***	.25	.04	.01	−.03
Vocal	**.70***	.00	.00	.03	.03
Chatty	**.63***	.35	.11	.08	−.05
Extraverted (+)	**.60***	.07	.04	−.09	.15
Outspoken (+)	**.57***	−.10	−.16	.04	.14
Boisterous (+)	**.53***	−.15	−.14	−.07	−.12
Rambunctious	**.49***	−.11	−.25	.02	−.07
Wordy (+)	**.49***	.19	.08	.22	−.11
Quiet (+)	**−.79***	.10	.15	.07	.03
Silent (+)	**−.70***	−.05	.10	−.02	.07
Soft-spoken	**−.67***	.21	.18	−.01	.09
Shy (+)	**−.66***	.31	.04	.10	−.16
Bashful	**−.64***	.19	.08	.07	−.11
Reserved (+)	**−.61***	−.02	.31	−.08	.07
Low-key	**−.59***	−.01	.20	−.12	−.04
Sedate (+)	**−.42***	−.11	.22	−.25	.01
Affective orientation					
Sensitive	−.05	**.60***	.00	.00	.08
Emotional (+)	.09	**.60***	−.03	.11	−.15
Tender-minded	.07	**.55***	.11	−.06	.02
Vulnerable	−.17	**.54***	.04	.02	−.16
Gullible (+)	−.01	**.54***	.08	−.02	**−.30**
Sentimental (+)	.08	**.53***	.13	.03	.08
Soft (+)	−.16	**.53***	.11	−.03	.12
Oversentimental (+)	−.01	**.46***	.08	.17	−.09
Rough (+)	.02	**−.67***	−.10	.04	−.11
Rugged	−.03	**−.63***	−.10	−.09	−.08
Tough (+)	.01	**−.63***	.00	.09	.04
Hardened	.04	**−.57***	−.20	.17	−.11
Iron-hearted	−.07	**−.57***	−.14	.05	−.12
Brave (+)	.21	**−.46***	−.11	−.15	.19
Firm (+)	.00	**−.43***	.09	.12	.12
Callous (+)	.02	**−.43***	−.19	.12	−.05
Norm orientation					
Conservative (+)	−.25	−.01	**.63***	.01	.06
Traditional (+)	−.23	.12	**.55***	.22	.02
Predictable (+)	−.21	.25	**.55***	−.01	−.11
Old-fashioned	−.23	−.09	**.45***	.03	.00
Religious (+)	.07	.13	**.44***	−.08	.03
Moralistic	−.07	.21	**.42***	−.12	.27
Conventional (+)	−.07	−.01	**.40***	.02	.14
Obedient	−.16	.18	**.40***	.03	.18
Reverent (+)	−.12	.04	**.39***	−.16	.04
Conforming	−.10	.26	**.38***	.02	−.14
Rebellious (+)	.24	−.26	**−.51***	−.12	−.16
Unpredictable (+)	.13	−.22	**−.47***	−.11	.00
Free-living	.16	−.16	**−.45***	**−.31**	.21
Erratic (+)	.18	−.14	**−.44***	.06	−.14
Nonreligious	−.04	−.06	**−.38***	−.03	−.08
Unmoralizing	−.13	−.20	**−.34***	.02	−.20
Liberal (+)	.12	.18	**−.32***	−.14	.15
Unorthodox	−.01	−.24	**−.32***	−.19	−.24
Nonconforming (+)	.04	−.10	**−.30***	−.09	.18
Unconventional (+)	.03	.09	**−.27***	−.11	−.16
Form orientation					
Fussy	.00	.07	−.05	**.56***	−.19
Perfectionistic (+)	.01	−.01	.07	**.53***	.19
Critical (+)	−.13	−.04	.02	**.52***	−.20
Particular (+)	−.07	.09	−.07	**.50***	.05
Faultfinding (+)	.01	.01	−.10	**.50***	−.18
Finicky (+)	−.05	.18	−.08	**.49***	−.20

Table 7 *(continued)*

	Factor				
Adjective	First	Second	Third	Fourth	Fifth
Form orientation					
(continued)					
Exacting (+)	.00	.03	−.01	**.47***	.18
Possessive (+)	.02	.00	−.02	**.40***	−.27
Materialistic	.04	−.07	.19	**.39***	−.29
Choosy	−.09	.00	.01	**.37***	.03
Sophisticated (+)	.02	−.02	.10	**.30**	**.41***
Dignified	.13	−.07	.29	.16	**.29***
Uncritical (+)	−.07	.13	.03	**−.49***	.00
Lenient (+)	−.11	.21	−.14	**−.41***	−.04
Overrelaxed	−.13	−.06	−.24	**−.40***	−.20
Easy-going (+)	.04	.08	−.27	**−.38***	**.31**
Laid-back	−.06	−.02	**−.35**	**−.38***	.05
Unsophisticated (+)	−.01	−.04	−.01	**−.37**	**−.43***
Leisurely	−.05	.06	−.22	**−.36***	.01
Overpatient	−.16	.17	.05	**−.34***	.05
Casual (+)	.01	.04	−.12	**−.33***	.06
Inarticulate (+)	−.15	−.03	.04	**−.32***	−.21
Overtolerant (+)	−.06	.25	−.13	**−.29***	.11
Overcasual (+)	−.02	−.08	−.21	**−.27***	−.14
General evaluation					
Self-controlled	−.16	−.04	.09	−.06	**.54***
Mature (+)	−.06	−.10	.20	.09	**.53***
Reasonable (+)	−.14	.12	−.02	−.18	**.51***
Broad-minded	.00	.02	**−.31**	−.10	**.46***
Wise	.00	−.19	−.08	−.03	**.45***
Alert	−.02	−.21	.05	.03	**.42***
Clear-headed	.04	−.14	.15	−.11	**.42***
Understanding (+)	.04	**.36**	.00	−.15	**.41***
Optimistic (+)	.22	.11	.08	−.20	**.41***
Purposeful (+)	.05	−.18	.19	−.03	**.36***
Considerate (+)	−.05	**.42***	.12	.01	**.33**
Generous	.07	.27	−.04	−.13	**.27***
Immature (+)	.04	.04	−.15	−.09	**−.62***
Closed-minded	.06	−.07	**.34**	.12	**−.51***
Greedy	.08	−.18	−.02	.24	**−.49***
Juvenile	.05	.01	−.18	−.22	**−.49***
Hypocritical (+)	.02	−.10	.03	.18	**−.48***
Inconsiderate	.00	**−.33**	−.15	−.14	**−.45***
Narrow-minded (+)	−.01	.00	**.41**	.13	**−.44***
Short-sighted	−.14	.04	−.02	−.10	**−.42***
Half-hearted	−.08	−.20	−.08	−.20	**−.42***
Unstable (+)	−.13	.13	−.22	.14	**−.40***
Bigoted (+)	.04	−.24	.16	.08	**−.40***
Deceitful (+)	.02	−.23	−.11	.08	**−.37***

Note. Principal-components factor loadings in replication sample (n = 250). Loadings |.30| and above are listed in boldface type. (+) = Adjective included in briefer 60-item inventory.
* Highest factor loading of each adjective.

100 adjectives. All replication correlations were .82 or higher. Whereas in the combined self and peer samples from Study 1 no less than 96 of the 100 variables loaded most highly on the intended factor, in the replication sample 86 of the 100 loaded most highly on the expected factor. Thus, the factors were robust, despite the imperfections of the initial marker set.

The desirability correlations showed the expected pattern in both samples; the four descriptive factors had nonsignificant correlations with desirability, and the fifth evaluative factor had a very high (.86 and .88) correlation with desirability.

From among the alternative candidate items included in the 306-item inventory, 35 substitutions were made in the process of refining the marker set. Table 7 presents the factor loadings of the final set of 100 marker adjectives for the five dimensions. The simple structure of the solution is manifest: 95 of the 100 items (95%) loaded at .30 and over on the expected factor, whereas only 15 of the 400 other factor-loading figures (4%) were .30 or more. Moreover, 96 of the 100 items loaded most highly on the expected factor.

Table 8 provides psychometric data on the reliability of the

Table 8
Reliability of the Sets of 60 and 100 Adjectives Derived in the Current Study in Comparison With 100 Big Five Markers

Dimension	No. of items	Coefficient alpha			Mean interitem correlation		
		Self	Peer	Replication	Self	Peer	Replication
		100 items					
First	16			.90			.36
Second	16			.86			.28
Third	20			.82			.19
Fourth	24			.80			.14
Fifth	24			.84			.18
M				.84			.23
		60 items					
First	10	.84	.84	.84	.34	.36	.35
Second	10	.72	.74	.77	.21	.22	.25
Third	12	.72	.76	.74	.19	.22	.20
Fourth	16	.66	.71	.72	.11	.14	.14
Fifth	12	.64	.77	.72	.14	.23	.18
M		.72	.76	.76	.20	.23	.22
		100 Big Five markers (Goldberg, 1992)					
I	20	.90	.90		.30	.31	
II	20	.84	.87		.22	.26	
III	20	.88	.90		.27	.31	
IV	20	.83	.82		.20	.18	
V	20	.82	.81		.19	.18	
M		.85	.86		.24	.25	

Note. Self sample $n = 320$; Peer sample $n = 316$; Replication sample $n = 250$. Adjectives from 100-item inventory are all found only in the replication data set; 100 adjectives from Big Five markers are all found in self and peer data sets but not in replication data set; briefer set of 60 adjectives is found in all three data sets. All statistics are based on ipsatized responses. Factor I is conventionally labeled Extraversion, Factor II Agreeableness, Factor III Conscientiousness, Factor IV Emotional Stability, and Factor V Intellect or Openness.

scales of the full 100-item inventory, as well as a 60-item subset including only adjectives from the data sets used in Study 1. The 60 items are marked with a (+) in Table 8; factor scores from the 60 items correlated from .92 to .97 with corresponding factor scores from the 100 items. Table 8 also depicts, for comparison, the reliabilities of the Big Five adjective scales from Goldberg (1992).

The alpha coefficients for the scales derived in the present studies were only slightly lower than those for the Big Five factors. But the fourth factor in the present studies proved more diffuse and more difficult to scale reliably than any Big Five factor. However, retaining only 16 items for some scales and up to 24 items for others (like the fourth) tended to equalize the alpha coefficients of scales with a broad range of mean interitem correlations. Alpha reliabilities of the briefer scales are typically about .10 lower than those for the full scales; to achieve brevity, some internal consistency is sacrificed.

Relation to Peabody factors. Peabody's (1967, 1984, 1985) Tight–Loose (T) and Assertive–Unassertive (A) factors were most accurately reproduced from the 14 sets of four adjectives. The first two unrotated factor loadings in this data set correlated .93 (T) and .97 (A) with the loadings published elsewhere (Peabody, 1985, Table A.3). The first three unrotated factors

from an analysis of the 28 scales corresponded fairly well to E (Evaluation), T, and A.

Table 9 presents the correlations of factor scores from Peabody factors with those from the present factors. The relations are clearly complex. Peabody descriptive factors from pairs demonstrated a near-zero correlation with the fifth (evaluative) factor, evidence that T and A are indeed descriptive and not evaluative. The A factor seems to be a blend of the present first (Interactional Orientation) and second (Affective Orientation) factors. The T factor represents a more equal blend of all four descriptive factors from the present studies; perhaps the present four factors can be seen as four orthogonal facets of Tight (e.g., *reserved, tough, traditional,* and *critical*) contrasted with four orthogonal facets of Loose (e.g., *talkative, emotional, nonconforming,* and *casual*). The correlations suggest that the Peabody Tight–Loose factor compresses variance from each of the four descriptive factors into a single broad megafactor.

Peabody (1967, 1984) originally recovered his dimensions in unrotated solutions. Indeed, initial unrotated factors from Study 1 data sets often resembled Peabody factors, particularly E and A. Unrotated solutions are typically regarded as less stable and interpretable than rotated solutions. In comparison with rotated solutions, which spread variance more equally

Table 9

Factor–Score Correlations Between Peabody Descriptive Factors and the Factors Derived in the Present Studies

	Peabody factors				
	From 14 pairs of scales		From 28 scales		
Present factors	T	A	E	T	A
First (Interactional Orientation)	**−.34**	**.49**	.04	**−.23**	**.55**
Second (Affective Orientation)	**−.27**	**−.59**	.26**	**−.32**	**−.50**
Third (Norm Orientation)	**.35**	**−.24**	.11	**.41**	**−.26**
Fourth (Form Orientation)	**.41**	.21**	**−.28**	**.37**	.14*
Fifth (General Evaluation)	−.09	−.07	**.67**	.24**	.17**
Multiple R	**.70**	**.83**	**.78**	**.72**	**.82**

Note. Replication sample N = 250. T = Tight–Loose; A = Assertive–Unassertive; E = Evaluation. Correlations of |.30| and above are listed in boldface type.
* p < .05.   ** p < .01.

among factors, unrotated solutions feature larger initial factors and smaller subsequent factors. Thus, the interpretable factors tend to be smaller in number but generally broader, as is the case with the set of Peabody factors.[11]

Alternatively, perhaps the present studies found four factors and Peabody only two because of differences in the selection of variables. Only 6 of the 100 adjectives in Table 7 are among the 106 adjectives in Peabody's (1987) representative selection. The omission in Peabody's selection of most neutral adjectives (e.g., *tough, emotional, traditional,* and *perfectionistic*) may prevent recovery of the present factors in that selection. Further studies of self- and peer ratings using Peabody's (1967, 1987) variable selections might clarify such issues.

The studies in this report provide evidence that four broad replicable descriptive dimensions can be recovered from personality-adjective ratings, once the evaluative element in these ratings is controlled. Moreover, a highly evaluative dimension orthogonal to the four descriptive dimensions is also replicable. These dimensions can be measured with personality adjectives in scales demonstrating internal consistencies approaching those of Big Five factor scales. And finally, these four dimensions have marked but complex relations to the two broader dimensions found in earlier investigations by Peabody (1967, 1984, 1985).

Interpreting the five dimensions. How do the present factors relate to the Big Five factors? Let us consider three interpretive hypotheses: (a) These five factors are in fact the Big Five personality factors, in which case each will have a very high correlation with one of the Big Five; (b) these factors are systematic rotations of the Big Five, in which case each will show high or moderate correlations with more than one of the Big Five, and thus a high multiple correlation with the Big Five; or (c) these dimensions represent something fundamentally different from the Big Five, as will be suggested by low or moderate multiple correlations. Table 10 provides evidence for weighing these hypotheses: factor–score correlations between Big Five factors (Goldberg, 1992) and the five factors derived in the present studies, as produced by the abbreviated 60-item inventory in samples from Study 1.

The first interpretive hypothesis holds only for the first factor, which is very highly correlated (r = .89) with Big Five Factor I (Extraversion) and no other. All other factors from the present study have substantial (over .30) correlations with more than one of the Big Five factors.

The second interpretive hypothesis holds, at least to a degree, for each of the remaining factors. The second factor represents a blend of Big Five factors II+ (Agreeableness) and IV- (lack of Emotional Stability; cf. Saucier, 1992, Figure 2). The third factor represents a blend of Big Five factors III+ (Conscientiousness) and V- (lack of Intellect or Openness), albeit with a lower multiple correlation. The fourth and fifth factors represent complex blends of more than two factors; the fourth is II-III+-IV-V+ and the fifth II+III+IV+ (or II+III+IV+V+).

The third interpretive hypothesis states that the present factors present something essentially different from the Big Five. It holds best for the third factor, which has no correlation greater than .34 with any Big Five factor and a multiple correlation of only .50 with all five Big Five factors. A considerable amount of variance on this third factor apparently lies outside the Big Five model.

The content of this third factor is, however, not new to psychology. Examination of the adjectives loading on this factor (Table 7) suggests an identification with well-developed constructs such as conformity (Asch, 1955; McDavid & Sistrunk, 1964), conservatism (Nisbet, 1986; G. D. Wilson, 1973), religiousness (Batson & Ventis, 1982), and a disposition to favor rather than question norms (Gough, 1987). The vital individual difference on which this factor touches involves different attitudes toward change, toward cultural traditions, and toward norms, thus the label *Norm Orientation* (No).

The term *orientation* suggests that some people look in one

---

[11] Post hoc analyses showed that factors resembling Peabody's T and A could be found in the desirability-partialed data sets (but not in the neutral-terms-only data sets) when two and only two factors are rotated. Rotation of more than two factors altered the first two factors in the direction of the positions exemplified in the present studies (e.g., Table 7).

Table 10

Factor-Score Correlations Between Big Five Factors and Factors Derived in the Present Studies

Big Five factors	Factors derived in present studies					
	First (Io)	Second (Ao)	Third (No)	Fourth (Fo)	Fifth (Ge)	R
I (Extraversion)	.89**	−.19**	−.09	.03	.08	.92
II (Agreeableness)	.08	.57**	.14**	−.22**	.50**	.81
III (Conscientiousness)	−.14**	−.19**	.32**	.42**	.41**	.71
IV (Emotional Stability)	−.10*	−.40**	.01	−.42**	.40**	.71
V (Intellect/Openness)	−.01	.04	−.34**	.34**	.18**	.51
Multiple R	.91	.75	.50	.72	.79	

Note. Combined self–peer sample n = 636, based on principal-components analyses. Correlations |.20| and over are in boldface type. Io = Interactional Orientation; Ao = Affective Orientation; No = Norm Orientation; Fo = Form Orientation; Ge = General Evaluation.
* $p < .01$. ** $p < .001$.

direction, some look in another direction, and neither view is clearly more desirable or correct. Human society might profit from this "polymorphism," this variety in orientation. Values of the past and values in new possibilities each have advocates.

The same notion of differences in "orientation" can apply to the second factor, which contrasts emotional, sensitive, and vulnerable traits with hardened, rough, and rugged traits. On one side of the contrast, emotions and sentiments are valued, whereas on the other side, such affects are suppressed, thus the label *Affective Orientation* (Ao).
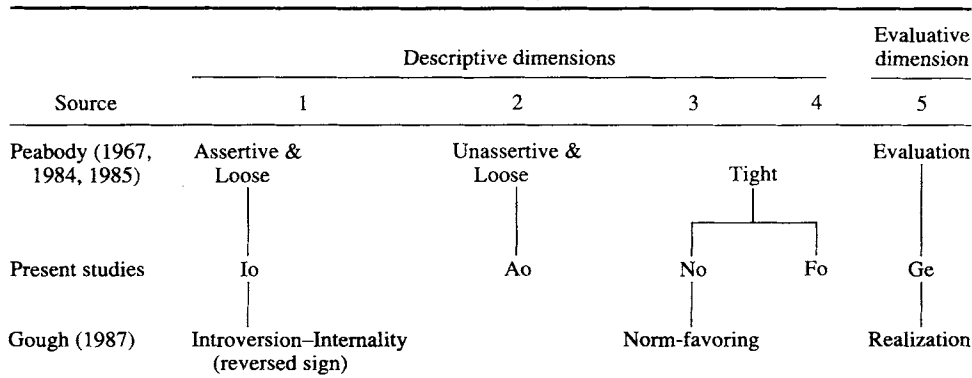
Factor scores on this Ao dimension correlate substantially (r = .58 in the self-description replication sample) with subject sex. A correlation of this magnitude suggests considerable gender differences, but also considerable overlap between the distributions of women and of men on the factor.

A comparison with gender-stereotype adjectives provides evidence for the relevance of the Ao dimension to the possible operation of gender-stereotype schemas. Among the 306 items on the Study 2 inventory were the 29 single-word adjectives from the Bem Sex-Role Inventory (BSRI; Bem, 1974). When raw

self-ratings on these 29 adjectives were rotated to a two-factor structure, factors clearly corresponding to agentic or instrumental Masculinity and communal or expressive Femininity were recovered. The Masculinity and Femininity factor scores correlated, respectively, −.59 and .58 with factor scores on the Ao dimension. An extreme score on Ao, then, may represent an extreme application of gender-stereotypic schemas in the description of self.

As for the first factor, its .89 correlation with Big Five Extraversion (Table 10) makes Extraversion an appropriate label. Extraversion is a construct measured on many inventories, including purportedly nonevaluative measures of the construct (e.g., Gough, 1987; Myers & McCaulley, 1985). An "orientation" interpretation for Extraversion deserves consideration. Chatty, boisterous, extraverted individuals have a highly interactional orientation; they may create and enjoy many social interactions. Conversely, quiet, sedate, reserved individuals tend not to participate in so many social interactions. Regardless of whether such attributes have significant biological and genetic determinants (cf. Eaves, Eysenck, & Martin, 1989), Extraver-

Table 11

Integration of Three Structural Models With Descriptive and Evaluative Axes

Source	Descriptive dimensions				Evaluative dimension
	1	2	3	4	5
Peabody (1967, 1984, 1985)	Assertive & Loose	Unassertive & Loose	Tight		Evaluation
Present studies	Io	Ao	No	Fo	Ge
Gough (1987)	Introversion–Internality (reversed sign)		Norm-favoring		Realization

Note. Tight and Loose are bipolar ends of a single dimension, as are Assertive and Unassertive. Io = Interactional Orientation; Ao = Affective Orientation; No = Norm Orientation; Fo = Form Orientation; Ge = General Evaluation.

sion can be conceived of as a difference in *Interactional Orientation* (Io).

The more diffuse fourth dimension refers to a less familiar individual difference. It might be interpreted as Culture, as Perfectionism, or as a contrast between work and leisure values. I prefer the label *Form Orientation* (Fo). This dimension seems to differentiate between those who tend to accept people or situations as they are naturally and those who are instead more particular as to what form is or is not acceptable.

Turning finally to the fifth, highly evaluative dimension, we encounter an individual difference not interpretable as a difference in "orientation." Traits such as *mature, wise,* and *understanding* are much more desirable than traits such as *immature, greedy,* and *hypocritical.* Factor scores on this dimension are highly predicted ($R = .76$; Table 10) by multiple regression from factor scores for Big Five Factors II through V. It might be interpreted as "Big Five profile elevation."

The broad range of extremely desirable or undesirable attributes is difficult to encompass within a single scale. The fifth factor approximates a central position within *General Evaluation* (Ge) that combines, at its positive pole, (a) *Likability* (e.g., *understanding, considerate,* and *generous*), (b) various desirable traits that have less obvious implications for one's likability (e.g., *wise, purposeful,* and *clear-headed*) but that may indicate a competent and prosocial use of personal judgment, and (c) perceived maturity. Ge may correspond to Ego Resiliency (Block, 1965) and to a convergence of both the "competence" and "ego-development" conceptions of maturity (Helson & Wink, 1987).

Self-ratings on Ge, unlike those on the descriptive dimensions, are likely to be quite susceptible to self-perception biases (cf. John & Robins, 1992), including both self-enhancement and denial effects (cf. Paulhus & Reid, 1991).

Conclusions and Implications

The present studies effected a rearrangement of the structure of personality attributes into an alternative representation in which description and evaluation are clearly separated. Although the relations with the Big Five structure are generally strong and systematic, this representation contrastingly highlights differences in values and in "orientation." The present representation can be considered auxiliary to the Big Five because it is likely to be less robust and inclusive and is only recovered by methodologies that restrict the evaluative range either of targets (cf. Peabody & Goldberg, 1989) or of variables.

Future investigations might further test the robustness of this representation of personality-attribute structure and the relation of its dimensions to constructs from other personality inventories, both nonevaluative (e.g., Norm-Favoring; Gough, 1987) and highly evaluative (e.g., Ego Resiliency; Block, 1965); Table 11 offers a hypothetical but testable integration of three structural models.

Finally, the present studies highlight a need to understand further the relation between individual differences in the valuing of attributes and individual differences in self-ascribed attributes. For this task the inventory derived in Study 2, containing evaluatively balanced scales, is well suited. Moreover, effects of denial and self-enhancement (Paulhus & Reid, 1991), with

regard to attributions on all five dimensions, merit close scrutiny. There is recent evidence that nonevaluative scales elicit the greatest agreement between descriptions by self and observer, whereas highly evaluative scales elicit the least agreement (John & Robins, 1992). Purely descriptive dimensions might prove unusually free of method variance in self-report, whereas a highly evaluative dimension will probably elicit such variance to an extreme degree. Accordingly, the separation of description and evaluation may afford us unprecedented control over troublesome method variance in the measurement of personality.

References

Asch, S. E. (1955, November). Opinions and social pressure. *Scientific American, 31*–35.

Batson, C. D., & Ventis, W. L. (1982). *The religious experience: A social psychological perspective.* New York: Oxford University Press.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155–162.

Block, J. (1965). *The challenge of response sets.* New York: Appleton-Century-Crofts.

Bochner, S., & Van Zyl, T. (1985). Desirability ratings of 110 personality-trait words. *Journal of Social Psychology, 125,* 459–465.

Costa, P. T., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology, 55,* 258–265.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653–665.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317–327.

De Boeck, P. (1978). On the evaluative factor in the trait scales of Peabody's study of trait inferences. *Journal of Personality and Social Psychology, 36,* 619–621.

Dicken, C. (1963). Good impression, social desirability, and acquiescence as suppressor variables. *Educational and Psychological Measurement, 23,* 699–720.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture, and personality: An empirical approach.* San Diego, CA: Academic Press.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York: Dryden.

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

Fisher, R. A. (1922). On the dominance ratio. *Proceedings of the Royal Society, Edinburgh, 42,* 321–341.

Goldberg, L. R. (1981). Unconfounding situational attributions from uncertain, neutral, and ambiguous ones: A psychometric analysis of descriptions of oneself and various types of others. *Journal of Personality and Social Psychology, 41,* 517–552.

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.

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

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103,* 265–275.

Hallworth, H. J. (1965). Dimensions of personality and meaning. *British Journal of Social and Clinical Psychology, 4*, 161–168.

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.

Heilbrun, A. B. (1964). Social learning theory, social desirability and the MMPI. *Psychological Bulletin, 61*, 377–387.

Helson, R., & Wink, P. (1987). Two conceptions of maturity examined in the findings of a longitudinal study. *Journal of Personality and Social Psychology, 53*, 531–541.

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.

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

John, O. P., & Robins, R. W. (1992). *Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self.* Unpublished manuscript, University of California, Berkeley.

Krieger, M. H. (1964). A control for social desirability in semantic differential. *British Journal of Social and Clinical Psychology, 3*, 94–103.

Kuusinen, J. (1969). Affective and denotative structures of personality ratings. *Journal of Personality and Social Psychology, 12*, 181–188.

Levin, J., & Karni, E. S. (1981). A note on the interpretation of the factor pattern of the California Psychological Inventory. *Journal of Personality Assessment, 45*, 430–432.

Mayr, E. (1988). *Toward a new philosophy of biology.* Cambridge, MA: Harvard University Press.

McCrae, R. R. (1989). Why I advocate the five-factor model: Joint factor analyses of the NEO-PI with other instruments. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 237–245). New York: Springer-Verlag.

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. (1989). Reinterpreting the Myers–Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality, 57*, 17–40.

McDavid, J. W., & Sistrunk, F. (1964). Personality correlates of two kinds of conformity behavior. *Journal of Personality, 32*, 421–435.

Moore, T. (1987, March 30). Personality tests are back. *Fortune,* pp. 76–81.

Mulaik, S. A. (1972). *The foundations of factor analysis.* New York: McGraw-Hill.

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.

Nisbet, R. A. (1986). *Conservatism: Dream and reality.* Minneapolis: University of Minnesota Press.

Norman, W. T. (1967). *2800 personality trait descriptors: Normative operative characteristics for a university population.* Unpublished manuscript, University of Michigan, Ann Arbor.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of Experimental Psychology, 50*, 325–338.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Paulhus, D. L. (1981). Control of social-desirability in personality inventories: Principal factor deletion. *Journal of Research in Personality, 15*, 383–388.

Paulhus, D. L. (1984). Two component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598–609.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*, 307–317.

Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monograph, 7* (4, Whole No. 644).

Peabody, D. (1968). Group judgments in the Philippines: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology, 10*, 290–300.

Peabody, D. (1970). Evaluative and descriptive aspects in personality perception: A reappraisal. *Journal of Personality and Social Psychology, 16*, 639–646.

Peabody, D. (1978). In search of an evaluative factor: Comments on De Boeck. *Journal of Personality and Social Psychology, 36*, 622–627.

Peabody, D. (1984). Personality dimensions through trait inferences. *Journal of Personality and Social Psychology, 46*, 384–403.

Peabody, D. (1985). *National characteristics.* Cambridge, England: Cambridge University Press.

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.

Rosenberg, S., & Olshan, K. (1970). Evaluative and descriptive aspects in personality perception. *Journal of Personality and Social Psychology, 16*, 619–626.

Saucier, G. (1991). *Evaluation-neutral personality description: A natural-language factor-replication study.* Unpublished doctoral dissertation, University of Oregon.

Saucier, G. (1992). Benchmarks: Integrating affective and interpersonal circles with the Big-Five personality factors. *Journal of Personality and Social Psychology, 62*, 1025–1035.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193–210.

Tzeng, O. C. S. (1975). Differentiation of affective and denotative meaning systems and their influence on personality ratings. *Journal of Personality and Social Psychology, 32*, 978–988.

Walkey, F. H., & McCormick, I. A. (1985). Multiple replication of factor structure: A logical solution for a number of factors problem. *Multivariate Behavioral Research, 20*, 57–67.

Wilson, E. O., & Bossert, W. H. (1971). *A primer of population biology.* Sunderland, MA: Sinauer.

Wilson, G. D. (Ed.). (1973). *The psychology of conservatism.* London: Academic Press.