

Validation of the Five-Factor Model of Personality Across Instruments and Observers

Robert R. McCrae and Paul T. Costa, Jr.
Gerontology Research Center,
National Institute on Aging, National Institutes of Health

Two data sources—self-reports and peer ratings—and two instruments—adjective factors and questionnaire scales—were used to assess the five-factor model of personality. As in a previous study of self-reports (McCrae & Costa, 1985b), adjective factors of neuroticism, extraversion, openness to experience, agreeableness-antagonism, and conscientiousness-undirectedness were identified in an analysis of 738 peer ratings of 275 adult subjects. Intraclass correlations among raters, ranging from .30 to .65, and correlations between mean peer ratings and self-reports, from .25 to .62, showed substantial cross-observer agreement on all five adjective factors. Similar results were seen in analyses of scales from the NEO Personality Inventory. Items from the adjective factors were used as guides in a discussion of the nature of the five factors. These data reinforce recent appeals for the adoption of the five-factor model in personality research and assessment.

Perhaps in response to critiques of trait models (Mischel, 1968) and to rebuttals that have called attention to common inadequacies in personality research (Block, 1977), personologists in recent years have devoted much of their attention to methodological issues. Lively discussions have centered on the merits and limitations of idiographic versus nomothetic approaches (Kenrick & Stringfield, 1980; Lamiell, 1981), aggregation and its effects on reliability (Epstein, 1979; Rushton, Brainerd, & Pressley, 1983), and alternative methods of scale construction (Burisch, 1984; Wrobel & Lachar, 1982). The veridicality of traits (beyond the realm of cognitive categories) has been tested by examining the correspondence between traits and behaviors (Mischel & Peake, 1982; Small, Zeldin, & Savin-Williams, 1983) and between self-reports and ratings (Edwards & Klockars, 1981; Funder, 1980; McCrae, 1982). As a body, these studies have simultaneously increased the level of methodological sophistication in personality research and restored confidence in the intelligent use of individual difference models of personality.

In contrast, there has been relatively little interest in the substance of personality—the systematic description of traits. The variables chosen as vehicles for tests of methodological hypotheses often appear arbitrary. Bem and Allen (1974) gave no rationale for the use of conscientiousness and friendliness in their classic paper on moderators of validity. McGowan and Gormly's (1976) decision to examine activity and Small et al.'s (1983) choice of prosocial and dominance behavior appear to have been made to facilitate their research designs. Indeed,

Kenrick and Dantchik (1983) complained that “catalogs of convenience” have replaced meaningful taxonomies of personality traits among “most of the current generation of social/personality researchers” (p. 299).

This disregard of substance is unfortunate because substance and method are ultimately interdependent. Unless methodological studies are conducted on well-defined and meaningful traits their conclusions are dubious; unless the traits are selected from a comprehensive taxonomy, it is impossible to know how far or in what ways they can be generalized.

Fortunately, a few researchers have been concerned with the problem of structure and have recognized the need for a consensus on at least the general outlines of a trait taxonomy (H. J. Eysenck & Eysenck, 1984; Kline & Barrett, 1983; Wiggins, 1979). One particularly promising candidate has emerged. The five-factor model—comprising extraversion or surgency, agreeableness, conscientiousness, emotional stability versus neuroticism, and culture—of Tupes and Christal (1961) was replicated by Norman in 1963 and heralded by him as the basis for “an adequate taxonomy of personality.” Although it was largely neglected for several years, variations on this model have recently begun to reemerge (Amelang & Borkenau, 1982; Bond, Nakazato, & Shiraishi, 1975; Conley, 1985; Digman & Takemoto-Chock, 1981; Goldberg, 1981, 1982; Hogan, 1983; Lorr & Manning, 1978; McCrae & Costa, 1985b).

Some researchers (Goldberg, 1982; Peabody, 1984) have chiefly been concerned with the representativeness and comprehensiveness of this model with respect to the natural language of traits; others have sought to provide a theoretical basis for the taxonomy (Hogan, 1983). Our major concern has been the convergent and discriminant validity of the dimensions of the model across instruments and observers. If the five-factor model is a reasonable representation of human personality, it should be recoverable from questionnaires as well as from adjectives and from observer ratings as well as from self-reports. This line of research addresses substantive questions from the methodological perspective developed in the past few years.

We wish to thank Petra Netter of the University of Giessen for her assistance in translating Amelang and Borkenau (1982), and Lewis R. Goldberg and Auke Tellegen for many stimulating comments.

Correspondence concerning this article should be addressed to Robert R. McCrae, Section on Personality, Stress, and Coping, Gerontology Research Center, Francis Scott Key Medical Center, Baltimore, Maryland 21224.

Five Factors in Self-Reports and Ratings

One of the strongest arguments in favor of the five-factor model has been its appearance in both self-reports and ratings. Norman (1963) reported the structure in peer ratings. Goldberg (1980) showed parallel structures in both ratings and self-reports. As early as the 1960s, convergence across observers was also demonstrated (Borgatta, 1964; Norman & Goldberg, 1966). However, with a few exceptions (e.g., Norman, 1969), these studies used only adjective-rating scales, and few attempts were made to compare adjective factors with standardized questionnaires that are more widely used in personality research.

In a recent publication (McCrae & Costa, 1985b), we examined the correspondence between adjective and questionnaire formats to see if the same substantive dimensions of personality would be obtained in each. Our adjective-rating instrument was an extension of one devised by Goldberg (1983); our questionnaire was the NEO Inventory (McCrae & Costa, 1983a), which measures three broad dimensions identified in analyses of standard personality measures. Self-reports on five adjective factors were compared with both self-reports and spouse ratings on the inventory dimensions of neuroticism, extraversion, and openness to experience. In brief, the study showed that a version of the five-factor model could be recovered from the adjectives, that there were clear correspondences for neuroticism and extraversion dimensions across the two instruments, that Norman's culture factor was better interpreted as openness to experience, and that validity coefficients above .50 could be obtained with both self-reports and spouse ratings.

Three major questions were left unanswered by that study. As Kammann, Smith, Martin, and McQueen (1984) pointed out, research using spouses as raters differs in some respects from more traditional peer-rating studies. Spouses may "more often disclose their feelings to each other through verbal self-statements" (p. 117), and spouses may be more willing to adopt and support the self-concept thus communicated than would peers. Further, the design of our earlier study allowed comparison only between an observer and a self-report; no comparisons were possible between different external observers. A first question, then, concerned the generalizability of our findings to agreement among peer ratings and between peer ratings and self-reports.

A second question involved the particular five-factor structure observed in our set of 80 adjectives. In most studies the fifth and smallest factor has been labeled *culture* and has been thought to include intelligence, sophistication, and intellectual curiosity. The latter element, in particular, suggested correspondence with the questionnaire factor of openness to experience (McCrae & Costa, 1985a). In Goldberg's 40-item instrument (1983) the terms *curious* and *creative* fell on a factor defined primarily by self-rated intelligence. By adding 40 additional items, including some intended to measure such aspects of openness as preference for variety and imaginativeness, we tested the hypothesis that the fifth factor might better be construed as openness rather than as culture. Results confirmed this expectation by showing a factor with only small loadings from *intelligent* and *cultured* but large loadings from *original*, *imaginative*, and *creative*, and including other forms of openness (*independent*, *liberal*, *daring*) that were clearly distinct

from intelligence. This factor correlated .57 with the NEO Inventory Openness scale. Because of the conceptual importance of this reformulation of the Norman model, it was essential to replicate the adjective-factor structure among peer ratings—the data source on which Norman (1963) had originally relied.

Finally, the NEO Inventory included no measures of two of the Norman factors: agreeableness and conscientiousness. These two dimensions have occurred less frequently in questionnaire measures, and they have been thought by some to represent merely the respondent's evaluation of the target. Consensual validation across observers is therefore particularly important for these two dimensions. For that purpose we developed questionnaire measures of agreeableness–antagonism and conscientiousness–undirectedness, and we examined agreement for both dimensions across instruments and observers in the present study.

Method

Subjects

Individuals who provided self-reports and who were targets for peer ratings were members of the Augmented Baltimore Longitudinal Study of Aging. The Baltimore Longitudinal Study of Aging (BLSA) sample is composed of a community-dwelling, generally healthy group of volunteers who have agreed to return for medical and psychological testing at regular intervals (Shock et al., 1984). The sample has been recruited continuously since 1958, with most new subjects referred by friends or relatives already in the study. Among the men, 93% are high school graduates and 71% are college graduates; nearly one fourth have doctorate-level degrees. The Augmented BLSA sample consists of 423 men and 129 women who participate in the BLSA, and it includes 183 wives and 16 husbands who are not themselves BLSA participants but who have agreed to complete questionnaires at home. Some participants chose not to participate in this study, and some provided incomplete data. Results are based on the 156 men and 118 women for whom complete data were available, except for one subject who scored more than five standard deviations below the mean on the conscientiousness adjective factor and whose adjective data were thrown out. Comparison of individuals who chose to participate in the peer-rating study with others showed no significant differences in age or sex. Somewhat surprisingly, there were also no differences in self-reported personality as measured by the five adjective factors. Participating subjects were slightly more open to experience than were others, $F(1, 634) = 4.15, p < .05$, when NEO Inventory scores were examined. At the time of the peer ratings, ages ranged from 29 to 93 ($M = 59.9$ years) for men and from 28 to 85 ($M = 53.8$ years) for women.

Peer Raters

Subjects were asked to nominate

three or four individuals who know you very well *as you are now*. They can be friends, neighbors, or co-workers, but they should *not* be relatives. These should be people who have known you for at least one year and have seen you in a variety of situations.

A few subjects nominated more than four raters, and names and addresses for 1,075 raters were obtained, a few of whom were dropped because they were themselves members of the BLSA or had already been nominated by another subject (peers rated only one subject). Of those contacted, 747 (69%) provided rating data. Raters were assured of the confidentiality of their responses and were specifically instructed not to discuss the ratings with the subject.

For purposes of item factor analyses, all ratings were pooled. For intra-class and peer/self-report correlations, data were analyzed in four subsamples defined by the number of ratings available for each subject: 49 subjects had exactly one rater, 71 had two raters, 90 had three, and 63 had four or more raters. Too few subjects had five or more raters to allow analyses of these data, and only the first four raters' data were examined in these cases.

A background sheet completed by raters was used to characterize the peers and their relationships to the subjects. As a group, the raters resembled the subjects. They ranged in age from 19 to 87 ($M = 54.2$ years), and the correlation of rater's age with subject's age was .72. Like BLSA participants, the raters were well educated: 78% were high school graduates, 57% were college graduates, and 41% had some graduate or professional education. Most of the raters (91% of the men and 75% of the women) were of the same sex as the subjects they rated. However, when asked if they thought they were similar in "personality, attitudes, temperament, and feelings" to the subject, only 8% of the raters considered themselves very similar, and 51% considered themselves similar; 34% considered themselves different, and 7% considered themselves very different from the subject.

The raters appeared to be well acquainted with the subjects. They reported knowing the subjects for an average of 18.3 years (range = 1 to 74 years). Currently, 57% reported seeing the subject weekly, 27% monthly, and 14% once or twice a year. In addition, 9% of the subjects volunteered the information that they had seen the subject more frequently at some time in the past. Furthermore, raters appeared to have some depth of acquaintance: 61% reported that the subject sometimes shared confidences or personal feelings with them, and 35% said he or she often did. In addition, raters said that the subject sometimes (67%) or often (15%) came to them for advice and support. When asked for their own assessment, 56% said they knew the subject pretty well, 40% said they knew the subject very well.

Most (75%) of the raters described their relationship with the subject as a close personal friend, 29% as a family friend, 28% as a neighbor, and 34% as a coworker. Only 8% listed themselves as an acquaintance. Most raters had seen the subjects in a variety of settings: at parties or social events (89%), with same-sex friends (67%), with family (85%), at work (51%), during subject's personal crisis (46%), on vacation (39%), or at religious services (31%).

As a group, these raters seemed particularly well qualified to give personality ratings. They had known the subjects for many years, seen them frequently in a variety of settings, and shared their confidences. The raters themselves believed they could give accurate ratings: 86% believed they were good at understanding others, and 89% thought the subject they rated was straightforward and easy to understand.

Measures and Procedure

Data from two kinds of instruments—adjective-rating scale factors and questionnaire scales—were obtained by mail administration over a period of 4 years. The schedule of administration of the personality measures is given in Table 1.

Adjective factors. On the basis of a series of analyses of English-language trait names, Goldberg (1983) developed a 40-item bipolar¹ adjective-rating scale instrument to measure five major dimensions of personality. In subsequent work (McCrae & Costa, 1985b) we supplemented his list with an additional 40 items. Subjects rated themselves on this 80-item instrument with the use of Goldberg's 9-point scale. Five factors were derived from these self-reports and identified as neuroticism, extraversion, openness, agreeableness, and conscientiousness. Similar factors were found for men and women. Neuroticism, extraversion, and openness factors were validated against NEO Inventory measures of the same constructs from both self-reports and spouse ratings, with convergent correlations ranging from .52 to .65 (McCrae & Costa,

Table 1
Schedule of Administration of Personality Measures

Instrument	Date
Self-reports	
NEO inventory	February 1980
Adjective-rating scales, agreeableness and conscientiousness items (preliminary)	March 1983
Peer ratings	
Adjective-rating scales	
NEO Personality Inventory (Form R)	July 1983

1985b). To examine the factor structure of this instrument in peer ratings is one of the aims of this article.

NEO Personality Inventory. A questionnaire measure of the five-factor model is provided by the NEO Personality Inventory (Costa & McCrae, 1985), which comprises the NEO Inventory (Costa & McCrae, 1980; McCrae & Costa, 1983a) along with newly developed scales to measure agreeableness and conscientiousness. The original NEO Inventory is a 144-item questionnaire developed through factor analysis to fit a three-dimensional model of personality. Eight-item scales are used for each of six facets or specific traits within each of three broad trait dimensions, and overall scores are obtained by summing the scores of the six facets of neuroticism, extraversion, and openness. Item scoring is balanced to control for acquiescence, and socially desirable responding does not appear to bias scores (McCrae & Costa, 1983b). Internal consistency and 6-month retest reliability for the three global scores range from .85 to .93 (McCrae & Costa, 1983a). A third-person form of the NEO Inventory has been developed for use by raters and has shown comparable reliability and validity when spouse ratings are obtained (McCrae, 1982).

Questionnaire scales to measure agreeableness and conscientiousness were developed as part of the present research. Two 24-item rational scales were created, and joint factor analysis of these items with NEO Inventory items led to the identification of the hypothesized five factors. Ten items loading on the agreeableness and conscientiousness factors were tentatively adopted as measures of those factors. These items also correlated more highly with the appropriate adjective factor than with any of the other adjective factors in self-reports.

When the NEO Inventory Rating Form was administered to peers, 60 items intended to measure agreeableness and conscientiousness were interspersed. These included the best items from the pilot study undertaken on self-reports along with new items written subsequently. Two final 18-item scales measuring agreeableness and conscientiousness² were derived from analyses of these data. Two criteria were used for item selection. First, joint factor analysis with the NEO items in peer ratings again showed five factors that could be identified as neuroticism, extraversion, openness, agreeableness, and conscientiousness. Items selected for the final scales were required to have their highest loadings on their hypothesized factor. Second, scales were created by using the 10-item preliminary scales developed on self-reports. The 60 proposed items were correlated with these two scales as well as with the three domain

¹ It is sometimes questionable whether the adjectives in bipolar scales are true psychological opposites. If they are not, one pole may load on one factor, the opposite pole on another. In consequence, the scale may appear factorially complex.

² Items, along with normative data, are available as Document No. 04440 from ASIS/NAPS, Microfiche Publications, P.O. Box 3513, Grand Central Station, New York, New York 10163. Remit in advance \$4 for a microfiche copy or \$7.75 for a photocopy.

scores from the NEO Inventory Rating Form. To be included in the final selection, items were required to show higher correlations with the Agreeableness or Conscientiousness scale than with any of the other four scales. Because the final selection included some items written after the self-report data had been collected, the final self-report scales consisted of only 10 agreeableness and 14 conscientiousness items.

Coefficient alpha for the Conscientiousness scale was .91 within peer ratings and .84 within self-reports; for the Agreeableness scale it was .89 within peer ratings but only .56 within self-reports. In part, this low internal consistency was due to the inclusion of only 10 items in the self-report scale; in part, it was due to lower average interitem correlations. Correlations between questionnaire and adjective measures of agreeableness and conscientiousness in self-reports were .48 and .65, respectively; neither scale correlated over .20 with any of the other adjective factors.

It is essential to note that although the development of the Agreeableness and Conscientiousness scales was conducted in parallel on peers and self-reports, correlations across these two methods did not influence item selection in any way. Thus, the correlation between self-reports and ratings was not inflated by the capitalization on chance inherent in some types of item selection.

Results

The results will be considered in three sections. First, we will examine the factor structure of the 80-item adjective-rating scales in peer ratings and validate the factors by correlation with peer ratings on the NEO Personality Inventory. Second, we will consider agreement among peers on the personality characteristics of the targets they have rated by examining intraclass correlations among raters for both adjective factor and questionnaire measures of the five-factor model. Finally, we will present correlations between self-reports and peer ratings.

Adjective Factors in Peer Ratings

Everett (1983) has recently suggested that the number of factors to be retained and interpreted should be determined by comparing rotated solutions in different samples or subsamples and adopting the solution that can be replicated. Coefficients of factor comparability should be used as the measure of similarity, and Everett suggested that coefficients above .90 be required to consider two factors to be a match. When peer ratings on the 80 adjective scales were submitted to principal components analysis, a scree test suggested that approximately five factors would be needed. Everett's procedure was then used to compare solutions in self-reports and peer ratings for the third through eighth factors. Varimax-rotated three-factor solutions were obtained independently in self-report data from 503 subjects (McCrae & Costa, 1985b) and in ratings from 738 peers. Comparability coefficients were calculated by applying the scoring weights derived from both analyses to the data from peers and by correlating the resulting factor scores. This process was repeated for four-, five-, six-, seven-, and eight-factor solutions (results are shown in Table 2). Only the five-factor solution showed replication of all factors, and comparabilities were very high in this case, ranging from .95 to .98. This is clear evidence that the five-factor solution, and only the five-factor solution, was invariant across observers.

Table 3 shows factor loadings for the five-factor solution in peers. The factors in Table 3 have been reordered, and variables

Table 2
Comparabilities for Varimax-Rotated Principal Components in 738 Peers Using Factor-Scoring Matrices from Ratings and Self-Reports

Components rotated	Factor comparabilities after varimax rotation							
	1st	2nd	3rd	4th	5th	6th	7th	8th
8	.94	.87	.85	.73	.72	.70	.68	.14
7	.95	.95	.86	.85	.58	.53	.08	
6	.96	.93	.91	.89	.82	.61		
5	.98	.98	.97	.96	.95			
4	.93	.86	.81	.76				
3	.87	.84	.74					

are arranged using the structure observed in self-reports (McCrae & Costa, 1985b) for comparison. The match between factors in the two data sets is clear; the great majority of items loaded on the same factor in peer ratings as they did in self-reports. The most notable difference appeared to be in the peer agreeableness-antagonism factor, which included, as definers of the antagonistic pole, aspects of dominance (e.g., *dominant*, *bold*) from the extraversion factor and hostility (e.g., *temperamental*, *jealous*) from the neuroticism factor.

The similarity of structure was particularly important in the case of the openness factor. In peer ratings, as in self-reports, *broad interests*, *prefer variety*, *independent*, and *liberal* were among the definers of this factor; *intelligent* and *cultured* showed small loadings. From this set of 80 items in both data sources, a factor emerged in which concern with rich and varied experience was more central than cognitive ability.

The interpretation of the peer-rating factors were confirmed by correlating factor scores with scale scores from the NEO Personality Inventory ratings. Convergent correlations between adjective factors and corresponding NEO scales were .73 for neuroticism, .70 for extraversion, .70 for openness, .80 for agreeableness, and .76 for conscientiousness ($N = 722$, $p < .001$). The largest divergent correlation was .33. These findings demonstrate that raters were highly consistent across instruments in the ways in which they described their targets on each of the five dimensions. The findings do not, however, speak to the accuracy of the ratings, as judged against external criteria.

Consensual Validation Across Peer Raters

The extent to which different peers agreed on the attribution of traits to the same individual was calculated by examining the intraclass correlations between factor scores for raters. Intraclass correlations were equivalent to the Pearsonian correlation between all possible pairs of raters (Haggard, 1958). The top half of Table 4 gives intraclass correlations for groups of subjects with two, three, or four raters. All were statistically significant, and values ranged from .30 to .65, with a median of .38. Levels of cross-peer agreement were approximately equal for all five factors. These data provide evidence of consensual validation for all five dimensions in three independent subsamples.

Agreement across observers on questionnaire measures is seen in significant intraclass correlations given in the bottom

Table 3
Varimax-Rotated Factor Loadings for 80 Adjective Items From Peer Ratings

Adjectives	Factor					Adjectives	Factor				
	N	E	O	A	C		N	E	O	A	C
Neuroticism (N)						Agreeableness vs. antagonism (A)					
Calm-worrying	79	05	-01	-20	05	Irritable-good natured	17	34	09	61	16
At ease-nervous	77	-08	-06	-21	-05	Ruthless-soft hearted	12	27	-01	70	11
Relaxed-high-strung	66	04	01	-34	-02	Rude-courteous	03	18	09	55	36
Unemotional-emotional	44	40	14	03	-03	Selfish-selfless	-07	-02	04	65	22
Even-tempered-temperamental	41	01	01	-56	-21	Uncooperative-helpful	01	23	14	44	45
Secure-insecure	63	-16	-08	-07	-39	Callous-sympathetic	04	29	11	67	20
Self-satisfied-self-pitying	53	-17	-07	03	-17	Suspicious-trusting	-14	19	15	62	08
Patient-impatient	41	02	-03	-57	02	Stingy-generous	02	24	17	55	22
Not envious-envious/jealous	29	01	-10	-46	-19	Antagonistic-acquiescent	-02	-06	-09	66	-02
Comfortable-self-conscious	57	-30	-17	-16	-16	Critical-lenient	-13	09	00	65	-14
Not impulse ridden-impulse ridden	20	26	22	-16	-38	Vengeful-forgiving	-15	11	07	70	16
Hardy-vulnerable	50	-14	-13	23	-26	Narrow-minded-open-minded	-14	15	48	54	16
Objective-subjective	17	10	-31	-20	-36	Disagreeable-agreeable	14	24	06	59	26
Extraversion (E)						Stubborn-flexible	-18	08	12	61	00
Retiring-sociable	-14	71	08	08	08	Serious-cheerful	-10	58	08	26	02
Sober-fun loving	-08	59	12	14	-15	Cynical-gullible	14	14	-17	40	16
Reserved-affectionate	-01	65	12	25	-15	Manipulative-straightforward	-15	06	-02	47	31
Alloof-friendly	-16	58	02	45	06	Proud-humble	01	-18	-09	45	13
Inhibited-spontaneous	-21	52	49	01	-02	Conscientiousness vs. undirectedness (C)					
Quiet-talkative	01	64	06	-19	00	Negligent-conscientious	-01	02	08	18	68
Passive-active	-26	42	28	-23	37	Careless-careful	-08	-07	-01	11	72
Loner-joiner	-14	53	-08	14	12	Undependable-reliable	-07	04	05	23	68
Unfeeling-passionate	14	43	28	31	09	Lazy-hardworking	-07	17	14	03	66
Cold-warm	-05	57	09	54	06	Disorganized-well organized	14	-02	05	-05	68
Lonely-not lonely	-49	30	-01	10	11	Lax-scrupulous	05	03	03	10	53
Task oriented-person oriented	-04	36	09	35	-29	Weak willed-self-disciplined	-26	-01	23	-03	62
Submissive-dominant	-16	20	20	-57	27	Sloppy-neat	-01	00	-04	12	59
Timid-bold	-21	33	31	-44	10	Late-punctual	-05	-09	-05	05	60
Openness (O)						Impractical-practical	-24	01	-04	05	54
Conventional-original	-06	12	67	08	-04	Thoughtless-deliberate	-03	-08	05	14	45
Down to earth-imaginative	16	03	54	-10	-12	Aimless-ambitious	-09	12	21	-08	52
Uncreative-creative	-08	09	56	11	25	Unstable-emotionally stable	-57	09	07	27	45
Narrow interests-broad interests	-15	20	52	18	27	Helpless-self-reliant	-29	19	21	-01	53
Simple-complex	16	-13	49	-20	08	Playful-businesslike	00	-26	02	-09	49
Uncurious-curious	00	12	41	00	24	Unenergetic-energetic	-14	34	27	-06	46
Unadventurous-daring	-18	31	55	-06	08	Ignorant-knowledgeable	-12	-03	53	13	43
Prefer routine-prefer variety	-11	30	43	14	-21	Quitting-persevering	-09	13	27	00	62
Conforming-independent	-22	09	49	-14	21	Stupid-intelligent	-04	03	41	17	44
Unanalytical-analytical	-15	-13	43	-13	30	Unfair-fair	-14	04	19	59	33
Conservative-liberal	04	08	46	15	-13	Imperceptive-perceptive	-16	07	46	24	39
Traditional-untraditional	02	-01	45	-05	-36	Uncultured-cultured	01	00	36	15	33
Unartistic-artistic	10	15	36	21	18						

Note. These are varimax-rotated principal component loadings for 738 raters. The loadings above .40 given in boldface. Decimal points are omitted.

half of Table 4 for three subsamples. These correlations closely resemble those for adjective factors, and they suggest that peers agreed as well on questionnaire as on adjective checklist descriptions of their friends.

Although the magnitude of correlations seen in Table 4 compares favorably with most in the literature (see McCrae, 1982, for a review), and although virtually all exceed the .3 barrier sometimes thought to set a limit to validity coefficients in personality research, it is also true that there was room for considerable difference of opinion between raters with regard to the same subject.

Agreement Between Self-Reports and Ratings

Agreement among raters was only one piece of evidence for consensual validation. It could be argued that shared stereo-

types account for some or all of the agreement among peers (Bourne, 1977). A more rigorous test would compare ratings with self-reports, because it is unlikely that any of the artifacts affecting either of these sources would be shared (McCrae, 1982). The top half of Table 5 presents the correlations between averaged peer ratings and self-reports for each of the five factors. With the exception of conscientiousness among subjects having only a single rater, all the correlations were statistically significant and many were substantial in magnitude.

The bottom half of Table 5 gives corresponding correlations for the NEO Personality Inventory for subjects with complete data. Although several correlations were small when only a single rater was used, they increased considerably in magnitude when multiple raters were averaged. Only the Agreeableness scale failed to show the utility of aggregating raters.

Table 4
Intraclass Correlations for Peer Ratings

Number of peer raters	N ^a	Factor				
		N	E	O	A	C
Adjective-factor scores						
2	146	30	59	65	43	37
3	267	38	37	37	44	36
4	248	30	42	41	36	41
NEO Personality Inventory						
2	142	53	52	51	38	47
3	270	30	38	39	38	38
4	252	31	43	40	28	40

Note. N = Neuroticism. E = Extraversion. O = Openness. A = Agreeableness. C = Conscientiousness. All correlations are significant at $p < .01$. Decimal points are omitted.

^a Refers to number of raters.

Finally, we considered divergent as well as convergent validation of the five factors across instruments and observers. To simplify presentation of the data, the four subsamples were combined by calculating an average peer rating (standardized within subsample) for each adjective-factor and questionnaire scale. Table 6 presents the correlations of mean peer ratings on both adjective factors and questionnaire scales with self-reports on the same variables. Convergent correlations, given in boldface, were invariably larger than divergent correlations, markedly so for all cases except those involving the self-report questionnaire Agreeableness–Antagonism scale. Because there was good agreement across observers when the agreeableness adjective-factor scores were used, it could be inferred that the problem here lay with the questionnaire scale and not with the construct. This was expected given the low reliability of the preliminary Agreeableness scale used in self-reports. For all five dimensions, the median validity coefficient was .44. When examined by sex, convergent correlations ranged from .19 to .58 for men (median = .35) and from .17 to .56 for women (median = .48).

Discussion

Convergence Across Observers and Instruments

This research examined the correspondence between assessments of five major personality dimensions among peer ratings and between peer ratings and self-reports, using both adjective factors and questionnaire scales. The results are straightforward, showing convergent and discriminant cross-observer and cross-instrument validation for all five factors.

The magnitude of the correlations—generally .4 to .6—deserves some comment, because it was larger than typically reported (e.g., Borgatta, 1964). In part, the higher agreement may be due to reliable and well-constructed measures and, in part, to the nature of the raters. On the whole, raters were very well acquainted with the subjects they rated, having seen them frequently in a variety of circumstances over a period of many

years. As Table 5 shows, aggregating across raters also tended to increase agreement. However, as Kammann et al. (1984) noted, there are limits to the improvements in accuracy offered by aggregating. Although the averaged ratings may reflect more accurately the consensus of how the individual is viewed, they may always diverge to some extent from the individual's phenomenological view of himself or herself. Given the qualifications of the raters in this study, it seems likely that the correlations seen here will be near the ceiling for self–other agreement.

It is also worth pointing out that ratings and self-reports differed in another respect as well. When raw scores on the NEO Inventory were compared, ratings were approximately one-half standard deviation higher on extraversion, and one-third lower on neuroticism, than were self-reports. Separate norms would thus be needed to make self-reports and ratings comparable.

The Nature of the Five Factors

These methodological considerations lay the groundwork for the equally important question of substance. A growing body of research has pointed to the five-factor model as a recurrent and more or less comprehensive taxonomy of personality traits. Theorists disagree, however, on precisely how to conceptualize the factors themselves. It seems useful at this point to review each of the factors and attempt to define the clear elements as well as disputed aspects. The factors in Table 3, which so closely parallel factors found in self-reports and which show such clear evidence of convergent and discriminant validity across observers and instruments, can form a particularly useful guide to the conceptual content of the dimensions of personality.

Neuroticism versus emotional stability. There is perhaps least disagreement about neuroticism, defined here by such terms as worrying, insecure, self-conscious, and temperamental. Al-

Table 5
Convergent Correlations Between Self-Reports and Peer Ratings for Adjective-Factor Scores and NEO Personality Inventory

No. of peer raters	N ^a	Self-reports				
		N	E	O	A	C
Adjective-factor scores						
1	49	33*	29*	46***	41**	25
2	72	53***	62***	46***	42***	30**
3	85	59***	45***	54***	55***	49***
4	61	51***	48***	52***	59***	50***
NEO Personality Inventory						
1	45	26	37*	53***	20	21
2	68	47***	60***	53***	34**	33**
3	81	43***	46***	62***	35**	50***
4	54	51***	56***	67***	24	58***

Note. N = Neuroticism. E = Extraversion. O = Openness. A = Agreeableness. C = Conscientiousness. Decimal points are omitted.

^a Refers to targets, not to raters.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 6
Correlations of Self-Reports With Mean Peer Ratings for Adjective Factors and Questionnaire Scales

Mean peer rating	Self-reports									
	Adjective factors					NEO Personality Inventory				
	N	E	O	A	C	N	E	O	A	C
Adjective factors										
N	50***	00	02	05	-10	38***	06	08	01	-09
E	19**	48***	01	09	-07	08	40***	16*	04	-03
O	01	-01	49***	-01	-08	02	11	43***	-06	11
A	-05	-14*	-18**	49***	-20***	-08	-26***	-02	28***	-19**
C	-09	-08	-12*	-08	40***	-11	-02	-09	11	40***
NEO Personality Inventory										
N	44***	-03	00	-03	-15*	42***	02	02	-11	-14*
E	06	45***	16**	00	06	-04	47***	25***	02	02
O	07	08	45***	13*	-07	03	13*	57***	02	13*
A	-06	-11	-15*	45***	-10	-12	-25***	-03	30***	-12
C	-11	-05	-10	-09	39***	-14*	-02	-08	08	43***

Note. N = Neuroticism. E = Extraversion. O = Openness. A = Agreeableness. C = Conscientiousness. *N* = 255 to 267. Convergent correlations are shown in boldface. Decimal points are omitted.

* $p < .05$. ** $p < .01$. *** $p < .001$.

though adjectives describing neuroticism are relatively infrequent in English (Peabody, 1984), psychologists' concerns with psychopathology have led to the development of innumerable scales saturated with neuroticism. Indeed, neuroticism is so ubiquitous an element of personality scales that theorists sometimes take it for granted.

A provocative view of neuroticism is provided by Tellegen (in press), who views it as negative emotionality, the propensity to experience a variety of negative affects, such as anxiety, depression, anger, and embarrassment. Virtually all theorists would concur in the centrality of negative affect to neuroticism; the question is whether other features also define it. Tellegen himself (in press) pointed out that his construct of negative emotionality has behavioral and cognitive aspects. Guilford included personal relations and objectivity in his emotional health factor (Guilford, Zimmerman, & Guilford, 1976), suggesting that mistrust and self-reference form part of neuroticism. We have found that impulsive behaviors, such as tendencies to overeat, smoke, or drink excessively, form a facet of neuroticism (Costa & McCrae, 1980), and *impulse-ridden* is a definer of the neuroticism factor in self-reports, although not in ratings. Others have linked neuroticism to irrational beliefs (Teasdale & Rachman, 1983; Vestre, 1984) or to poor coping efforts (McCrae & Costa, 1986).

What these behaviors seem to share is a common origin in negative affect. Individuals high in neuroticism have more difficulty than others in quitting smoking because the distress caused by abstinence is stronger for them. They may more frequently use inappropriate coping responses like hostile reactions and wishful thinking because they must deal more often with disruptive emotions. They may adopt irrational beliefs like self-blame because these beliefs are cognitively consistent with the negative feelings they experience. Neuroticism appears to include not only negative affect, but also the disturbed thoughts and behaviors that accompany emotional distress.

Extraversion or surgency. Sociable, fun-loving, affectionate,

friendly, and talkative are the highest loading variables on the extraversion factor. This is not Jungian extraversion (see Guilford, 1977), but it does correspond to the conception of H. J. Eysenck and most other contemporary researchers, who concur with popular speech in identifying extraversion with lively sociability.

However, disputes remain about which elements are central and which are peripheral to extraversion. Most writers would agree that sociability, cheerfulness, activity level, assertiveness, and sensation seeking all covary, however loosely. But the Eysencks have at times felt the need to distinguish between sociability and what they call impulsiveness (S. B. G. Eysenck & Eysenck, 1963; Revelle, Humphreys, Simon, & Gilliland, 1980). Hogan (1983) believed that the five-factor model was improved by dividing extraversion into sociability and assertiveness factors. In Goldberg's analyses, surgency (dominance and activity) were the primary definers of extraversion, and terms like warm-cold were assigned to the agreeableness-antagonism factor. Tellegen (in press) emphasized the complementary nature of neuroticism and extraversion by labeling his extraversion factor positive emotionality.

These distinctions do seem to merge at a high enough level of analysis (H. J. Eysenck & Eysenck, 1967; McCrae & Costa, 1983a), and sociability—the enjoyment of others' company—seems to be the core. What is essential to recall, however, is that liking people does not necessarily make one likable. Salesmen, those prototypic extraverts, are generally happier to see you than you are to see them.

Openness to experience. The reinterpretation of Norman's culture as openness to experience was the focus of some of our previous articles (McCrae & Costa, 1985a, 1985b), and the replication of results in peer ratings was one of the purposes of the present article. According to adjective-factor results, openness is best characterized by original, imaginative, broad interests, and daring. In the case of this dimension, however, questionnaires may be better than adjectives as a basis for interpretation

and assessment. Many aspects of openness (e.g., openness to feelings) are not easily expressed in single adjectives, and the relative poverty of the English-language vocabulary of openness and closedness may have contributed to confusions about this domain (McCrae & Costa, 1985a). We know from questionnaire studies that openness can be manifest in fantasy, aesthetics, feelings, actions, ideas, and values (Costa & McCrae, 1978, 1980), but only ideas and values are well represented in the adjective factor. Interestingly, questionnaire measures of openness give higher validity coefficients than do adjective-factor measures—indeed, the correlation of .57 between the self-reported NEO Openness scale and the peer-rated NEO Openness scale is the highest of those shown in Table 6.

Perhaps the most important distinction to be made here is between openness and intelligence. Open individuals tend to be seen by themselves and others as somewhat more intelligent, and there are correlations of .30 between psychometric measures of intelligence and openness. However, joint factor analyses using Army Alpha intelligence subtests and either adjectives (McCrae & Costa, 1985b) or NEO Inventory scales (McCrae & Costa, 1985a) show that intelligence scales define a factor clearly separate from openness. Intelligence may in some degree predispose the individual to openness, or openness may help develop intelligence, but the two seem best construed as separate dimensions of individual differences.

Agreeableness versus antagonism. As a broad dimension, agreeableness–antagonism is less familiar than extraversion or neuroticism, but some of its component traits, like trust (Stark, 1978) and Machiavellianism (Christie & Geis, 1970), have been widely researched. The essential nature of agreeableness–antagonism is perhaps best seen by examining the disagreeable pole, which we have labeled antagonism. As the high-loading adjectives in Table 3 and the items in Table 2 show, antagonistic people seem always to set themselves against others. Cognitively they are mistrustful and skeptical; affectively they are callous and unsympathetic; behaviorally they are uncooperative, stubborn, and rude. It would appear that their sense of attachment or bonding with their fellow human beings is defective, and in extreme cases antagonism may resemble sociopathy (cf. H. J. Eysenck & Eysenck's, 1975, psychoticism).

An insightful description of antagonism in its neurotic form is provided by Horney's account of the tendency to move against people (1945, 1950). She theorized that a struggle for mastery is the root cause of this tendency and that variations may occur, including narcissistic, perfectionistic, and arrogant vindictive types. Whereas some antagonistic persons are overtly aggressive, others may be polished manipulators. The drive for mastery and the overt or inhibited hostility of antagonistic individuals suggests a resemblance to some formulations of Type A personality (Dembroski & MacDougall, 1983), and systematic studies of the relations between agreeableness–antagonism and measures of coronary-prone behavior should be undertaken.

Unappealing as antagonism may be, it is necessary to recognize that extreme scores on the agreeable pole may also be maladaptive. The person high in agreeableness may be dependent and fawning, and agreeableness has its neurotic manifestation in Horney's self-effacing solution of moving toward people.

Antagonism is most easily confused with dominance. Amelang and Borkenau (1982), working in German and apparently

unaware of the Norman taxonomy, found a factor they called *dominance*. Among its key definers, however, were Hartnäckigkeit (*stubbornness*) and Erregbarkeit (*irritability*); scales that measure agreeableness and cooperation defined the opposite pole in their questionnaire factor. Clearly, this factor corresponds to antagonism. In self-reports (McCrae & Costa, 1985b), submissive–dominant is a weak definer of extraversion; in Table 3, from the peers' point of view, it is a definer of antagonism. The close etymological relationship of *dominant* and *domineering* shows the basis of the confusion.

Agreeableness–antagonism and conscientiousness–undirectedness are sometimes omitted from personality systems because they may seem too value laden. Indeed, the judgment of character is made largely along these two dimensions: Is the individual well or ill intentioned? Is he or she strong or weak in carrying out those intentions? Agreeableness–antagonism, in particular, has often been assumed to be an evaluative factor of others' perceptions rather than a veridical component of personality (e.g., A. Tellegen, personal communication, March 28, 1984).

However, the fact that a trait may be judged from a moral point of view does not mean that it is not a substantive aspect of personality. The consensual validation seen among peers and between peer-reports and self-reports demonstrates that there are some observable consistencies of behavior that underlie attributions of agreeableness and conscientiousness. They may be evaluated traits, but they are not mere evaluations.

Conscientiousness versus undirectedness. Conscientious may mean either governed by conscience or careful and thorough (Morris, 1976), and psychologists seem to be divided about which of these meanings best characterizes the last major dimension of personality. Amelang and Borkenau (1982) labeled their factor self-control versus impulsivity, and Conley (1985) spoke of impulse control. This terminology connotes an inhibiting agent, as Cattell (Cattell, Eber, & Tatsuoka, 1970) recognized when he named his Factor *G superego strength*. A conscientious person in this sense should be dutiful, scrupulous, and perhaps moralistic.

A different picture, however, is obtained by examining the adjectives that define this factor. In addition to conscientious and scrupulous, there are a number of adjectives that suggest a more proactive stance: hardworking, ambitious, energetic, persevering. Digman and Takemoto-Chock (1981) labeled this factor *will to achieve*, and it is notable that one of the items in the questionnaire measure of conscientiousness, "He strives for excellence in all he does," comes close to the classic definition of need for achievement (McClelland, Atkinson, Clark, & Lowell, 1953).

At one time, the purposefulness and adherence to plans, schedules, and requirements suggested the word *direction* as a label for this factor, and we have retained that implication in calling the opposite pole of conscientiousness *undirectedness*. In our view, the individual low in conscientiousness is not so much uncontrolled as undirected, not so much impulse ridden as simply lazy.

It seems probable that these two meanings may be related. Certainly individuals who are well organized, habitually careful, and capable of self-discipline are more likely to be able to adhere scrupulously to a moral code if they choose to—al-

though there is no guarantee that they will be so inclined. An undirected individual may have a demanding conscience and a pervasive sense of guilt but be unable to live up to his or her own standards for lack of self-discipline and energy. In any case, it is clear that this is a dimension worthy of a good deal more empirical attention than it has yet received. Important real-life outcomes such as alcoholism (Conley & Angelides, 1984) and academic achievement (Digman & Takemoto-Chock, 1981) are among its correlates, and a further specification of the dimension is sure to be fruitful.

Some personality theorists might object that trait ratings, in whatever form and from whatever source, need not provide the best foundation for understanding individual differences. Experimental analysis of the psychophysiological basis of personality (H. J. Eysenck & Eysenck, 1984), examination of prototypic acts and act frequencies (Buss & Craik, 1983), psychodynamic formulations (Horney, 1945), or behavioral genetics (Plomin, DeFries, & McClearn, 1980) provide important alternatives. But psychophysiological, behavioral, psychodynamic, and genetic explanations must eventually be related to the traits that are universally used to describe personality, and the five-factor model can provide a framework within which these relations can be systematically examined. The minor conceptual divergences noted in this article suggest the need for additional empirical work to fine-tune the model, but the broad outlines are clear in self-reports, spouse ratings, and peer ratings; in questionnaires and adjective factors; and in English and in German (Amelang & Borkenau, 1982; John, Goldberg, & Angleitner, 1984). Deeper causal analyses may seek to account for the structure of personality, but the structure that must be explained is, for now, best represented by the five-factor model.

References

- 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 neuroticism]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 119–146.
- Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 81, 88–104.
- Block, J. (1977). Advancing the psychology of personality: Paradigmatic shift or improving the quality of research? In D. Magnusson & N. S. Endler (Eds.), *Personality at the cross-roads: Current issues in interactional psychology* (pp. 37–63). Hillsdale, NJ: Erlbaum.
- Bond, M. H., Nakazato, H., & Shiraishi, D. (1975). Universality and distinctiveness in dimensions of Japanese person perception. *Journal of Cross-Cultural Psychology*, 6, 346–357.
- Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioral Science*, 9, 8–17.
- Bourne, E. (1977). Can we describe an individual's personality? Agreement on stereotype versus individual attributes. *Journal of Personality and Social Psychology*, 35, 863–872.
- Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, 39, 214–227.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, 90, 105–126.
- Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *The handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Christie, R., & Geis, R. L. (Eds.). (1970). *Studies in Machiavellianism*. New York: Academic Press.
- Conley, J. J. (1985). Longitudinal stability of personality traits: A multi-trait–multimethod–multioccasion analysis. *Journal of Personality and Social Psychology*, 49, 1266–1282.
- Conley, J. J., & Angelides, M. (1984). *Personality antecedents of emotional disorders and alcohol abuse in men. Results of a forty-five year prospective study*. Manuscript submitted for publication.
- Costa, P. T., Jr., & McCrae, R. R. (1978). Objective personality assessment. In M. Storandt, I. C. Siegler, & M. F. Elias (Eds.), *The clinical psychology of aging* (pp. 119–143). New York: Plenum Press.
- Costa, P. T., Jr., & McCrae, R. R. (1980). Still stable after all these years: Personality as a key to some issues in adulthood and old age. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life span development and behavior* (Vol. 3, pp. 65–102). New York: Academic Press.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Dembroski, T. M., & MacDougall, J. M. (1983). Behavioral and psychophysiological perspectives on coronary-prone behavior. In T. M. Dembroski, T. H. Schmidt, & G. Blumchen (Eds.), *Biobehavioral bases of coronary heart disease* (pp. 106–129). New York: Karger.
- 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.
- Edwards, A. L., & Klockars, A. J. (1981). Significant others and self-evaluation: Relationships between perceived and actual evaluations. *Personality and Social Psychology Bulletin*, 7, 244–251.
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37, 1097–1126.
- Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research*, 18, 197–218.
- Eysenck, H. J., & Eysenck, M. (1984). *Personality and individual differences*. London: Plenum Press.
- Eysenck, H. J., & Eysenck, S. B. G. (1967). On the unitary nature of extraversion. *Acta Psychologica*, 26, 383–390.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: EDITS.
- Eysenck, S. B. G., & Eysenck, H. J. (1963). On the dual nature of extraversion. *British Journal of Social and Clinical Psychology*, 2, 46–55.
- Funder, D. C. (1980). On seeing ourselves as others see us: Self–other agreement and discrepancy in personality ratings. *Journal of Personality*, 48, 473–493.
- Goldberg, L. R. (1980, May). *Some ruminations about the structure of individual differences: Developing a common lexicon for the major characteristics of human personality*. Paper presented at the meeting of the Western Psychological Association, Honolulu, HI.
- 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. (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. (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.
- Guilford, J. P. (1977). Will the real factor of extraversion–introversion please stand up? A reply to Eysenck. *Psychological Bulletin*, 84, 412–416.
- Guilford, J. S., Zimmerman, W. S., & Guilford, J. P. (1976). *The Guil-*

- ford-Zimmerman Temperament Survey handbook: Twenty-five years of research and application.* San Diego, CA: EdITS.
- Haggard, E. A. (1958). *Intraclass correlation and the analysis of variance.* New York: Dryden Press.
- Hogan, R. (1983). Socioanalytic theory of personality. In M. M. Page (Ed.), *1982 Nebraska Symposium on Motivation: Personality—current theory and research* (pp. 55–89). Lincoln: University of Nebraska Press.
- Horney, K. (1945). *Our inner conflicts.* New York: Norton.
- Horney, K. (1950). *Neurosis and human growth.* New York: Norton.
- 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. J. C. Bonarius, G. L. M. van Heck, & N. G. Smid (Eds.), *Personality psychology in Europe: Theoretical and empirical developments.* Lisse, Switzerland: Swets & Zeitlinger.
- Kammann, R., Smith, R., Martin, C., & McQueen, M. (1984). Low accuracy in judgments of others' psychological well-being as seen from a phenomenological perspective. *Journal of Personality, 52,* 107–123.
- Kenrick, D. T., & Dantchik, A. (1983). Interactionism, idiographics, and the social psychological invasion of personality. *Journal of Personality, 51,* 286–307.
- Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review, 87,* 88–104.
- Kline, P., & Barrett, P. (1983). The factors in personality questionnaires among normal subjects. *Advances in Behaviour Research and Therapy, 5,* 141–202.
- Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist, 36,* 276–289.
- Lorr, M., & Manning, T. T. (1978). Higher-order personality factors of the ISI. *Multivariate Behavioral Research, 13,* 3–7.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive.* New York: Appleton-Century-Crofts.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology, 43,* 293–303.
- McCrae, R. R., & Costa, P. T., Jr. (1983a). Joint factors in self-reports and ratings: Neuroticism, extraversion, and openness to experience. *Personality and Individual Differences, 4,* 245–255.
- McCrae, R. R., & Costa, P. T., Jr. (1983b). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51,* 882–888.
- McCrae, R. R., & Costa, P. T., Jr. (1985a). Openness to experience. In R. Hogan & W. H. Jones (Eds.), *Perspectives in personality: Theory, measurement, and interpersonal dynamics* (Vol. 1). Greenwich, CT: JAI Press.
- McCrae, R. R., & Costa, P. T., Jr. (1985b). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49,* 710–721.
- McCrae, R. R., & Costa, P. T., Jr. (1986). Personality, coping, and coping effectiveness in an adult sample. *Journal of Personality, 54,* 385–405.
- McGowan, J., & Gormly, J. (1976). Validation of personality traits: A multicriteria approach. *Journal of Personality and Social Psychology, 34,* 791–795.
- Mischel, W. (1968). *Personality and assessment.* New York: Wiley.
- Mischel, W., & Peake, P. K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review, 89,* 730–755.
- Morris, W. (Ed.). (1976). *The American Heritage dictionary of the English language.* Boston: Houghton Mifflin.
- 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. (1969). "To see ourselves as others see us!": Relations among self-perceptions, peer-perceptions, and expected peer perceptions of personality attributes. *Multivariate Behavioral Research, 4,* 417–443.
- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4,* 681–691.
- Peabody, D. (1984). Personality dimensions through trait inferences. *Journal of Personality and Social Psychology, 46,* 384–403.
- Plomin, R., DeFries, J. C., & McClearn, G. E. (1980). *Behavior genetics: A primer.* San Francisco: Freeman.
- Revelle, W., Humphreys, M. S., Simon, L., & Gilliland, K. (1980). The interactive effect of personality, time of day, and caffeine: A test of the arousal model. *Journal of Experimental Psychology: General, 109,* 1–31.
- Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18–38.
- Shock, N. W., Greulich, R. C., Andres, R., Arenberg, D., Costa, P. T., Jr., Lakatta, E. G., & Tobin, J. D. (1984). *Normal human aging: The Baltimore Longitudinal Study of Aging* (NIH Publication No. 84-2450). Bethesda, MD: National Institutes of Health.
- Small, S. A., Zeldin, R. S., & Savin-Williams, R. C. (1983). In search of personality traits: A multimethod analysis of naturally occurring prosocial and dominance behavior. *Journal of Personality, 51,* 1–16.
- Stark, L. (1978). Trust. In H. London & J. E. Exner, Jr. (Eds.), *Dimensions of personality* (pp. 561–599). New York: Wiley.
- Teasdale, J. D., & Rachman, S. (Eds.). (1983). Cognitions and mood: Clinical aspects and applications [Special issue]. *Advances in Behaviour Research and Therapy, 5,* 1–88.
- Tellegen, A. (in press). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders.* Hillsdale, NJ: Erlbaum.
- Tupes, E. C., & Christal, R. E. (1961). Recurrent personality factors based on trait ratings. *USAF ASD Technical Report* (No. 61–97).
- Vestre, N. D. (1984). Irrational beliefs and self-reported depressed mood. *Journal of Abnormal Psychology, 93,* 239–241.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412.
- Wrobel, T. A., & Lachar, D. (1982). Validity of the Weiner subtle and obvious scales for the MMPI: Another example of the importance of inventory-item content. *Journal of Consulting and Clinical Psychology, 50,* 469–470.

Received September 7, 1984

Revision received January 18, 1985 ■