

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, Department of Psychology

Psychology, Department of

6-1-2002

What Do We Assess When We Assess a Big 5 Trait? : A Content Analysis of the Affective, Behavioral, and Cognitive Processes Represented in Big 5 Personality Inventories

Lisa M. Pytlik Zillig

University of Nebraska-Lincoln, lpytlikz@nebraska.edu

Scott H. Hemenover

Kansas State University, shemenov@ksu.edu

Richard A. Dienstbier

University of Nebraska-Lincoln, rdienstbier2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychiatry and Psychology Commons](#)

Pytlik Zillig, Lisa M.; Hemenover, Scott H.; and Dienstbier, Richard A., "What Do We Assess When We Assess a Big 5 Trait? : A Content Analysis of the Affective, Behavioral, and Cognitive Processes Represented in Big 5 Personality Inventories" (2002). *Faculty Publications, Department of Psychology*. 19. <https://digitalcommons.unl.edu/psychfacpub/19>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

What Do We Assess When We Assess a Big 5 Trait? A Content Analysis of the Affective, Behavioral, and Cognitive Processes Represented in Big 5 Personality Inventories

Lisa M. Pytlik Zillig

University of Nebraska–Lincoln

Scott H. Hemenover

Kansas State University

Richard A. Dienstbier

University of Nebraska–Lincoln

What are personality traits? Are all “broad” traits equally broad in the constructs they encompass and in the pervasiveness of their effects? Or are some traits more or less affective, behavioral, or cognitive in nature? The present study examined these issues as they applied to the Big 5 traits of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Expert and novice raters judged the extent to which items from four popular Big 5 inventories contain behavioral, cognitive, or affective components. Traits and inventories were then compared in terms of their relative assessment of these components. Results indicate convergence among inventories but remarkable differences between traits. These findings have implications for the conceptualization and assessment of traits and suggest directions for future research.

What are traits? What are the Big 5 traits? Throughout the past decade, there has been a growing consensus that individual differences in personality may be parsimoniously described by a hierarchical system composed

of three to seven major traits, and among these approaches, the five-factor models¹ have gained distinct prominence (John & Srivastava, 1999; Pervin, 1994). The Big 5 traits—Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness—emerged from decades of research and have been celebrated for their ability to simplify an otherwise overwhelming number of traits (Hofstee, 1994; John, 1990; McCrae & Costa, 1987), their cross-cultural applicability (McCrae & Costa, 1997b), and their ability to predict health-relevant and other outcomes (e.g., Emmons, 1995). Although the adequacy of the five-factor model of traits has been debated (e.g., Block, 1995; Pervin, 1994), several Big 5 trait measures are currently in wide use (Widiger & Trull, 1997). Furthermore, although researchers often debate the emphases associated with each model or measure, proponents of the various Big 5 conceptualizations sometimes unite and proclaim, “Despite differences in emphasis and interpretation . . . there is agreement among all these investigators that they are addressing the same phenomenon” (Costa & McCrae, 1992a, p. 653).

Authors’ Note: The first author’s last name is “Pytlik Zillig,” not “Zillig.” When citing this article, please cite as follows: Pytlik Zillig, Hemenover, & Dienstbier, 2002. Portions of this research were presented at the 72nd annual meeting of the Midwestern Psychological Association, May 2-4, 2000. The authors would like to thank Jennifer Linder, Tamara Davenport, Amy Osberg, and Amber Struebel for portions of data entry and preparation of tables; Audrey Jones, Alisha Maul, and Charles Hardy for their assistance with the literature review; Dave Hansen, Scott Roesch, Theresa Dethlefs, Valerie Meier, Sherri Lowrey, and Karen Smith for expertly rating items; and the undergraduates who also volunteered to rate items. We would also like to thank those who have contributed comments, suggestions, references, and/or permission to use copyrighted items: Scott Roesch, Greg Schraw, Gustavo Carlo, Amy N. Spiegel, Oliver P. John, Calvin P. Garbin, Rick Bevins, and Psychological Assessment Resources (PAR). A special thanks to Brian L. Pytlik Zillig, librarian extraordinaire; space constraints do not permit a listing of the many ways he contributed to the production of this article. Correspondence concerning this article should be addressed to Lisa M. Pytlik Zillig, Department of Psychology, 238 Burnett Hall, University of Nebraska–Lincoln, Lincoln, NE 68588-0308; e-mail: lpytlik@unlserve.unl.edu.

But what, exactly, are these traits? Understanding what a trait is—that is, what one is describing when one invokes a trait construct or references a trait score—has important implications for building theory and interpreting research. However, even among those who agree on the importance of traits—and of the Big 5 in particular—there remains much debate regarding their conceptualization. Here, we review some of those debates and argue for the utility of examining operational definitions of traits, especially the extent that they assess affects, behaviors, and cognitions (hereafter ABCs or A, B, and C). This analysis offers a different perspective on both the general nature of traits and the nature of individual Big 5 traits.

The Nature of Traits in General

Historically, traits have been defined primarily in terms of overt behavior (e.g., as reviewed by Pervin, 1994). However, emphases on behavior, cognition, and affect have changed in recent decades. Perhaps at least partly due to two rare content analyses revealing traits to be operationalized by both covert and overt responses (Angleitner, John, & Löhr, 1986; Werner & Pervin, 1986), recent trait definitions have reduced their behavioral emphases. Thus, MacDonald (1995) describes traits as “motivational systems with an affective core” (p. 525), whereas other researchers define traits even more broadly in terms of “stylistic and habitual patterns of cognition, affect and behavior” (Emmons, 1989, p. 32; see also Johnson, 1997; Winter, John, Stewart, Klohnen, & Duncan, 1998, p. 232).

Presently, the scope of the trait domain—inclusive of cognitions, affects, and overt behaviors—seems to be widely but perhaps ambivalently accepted, without much empirical or theoretical consideration. For example, Pervin (1994) inquired about the acceptance and implications of such a widened definition of the trait domain but reported that his concerns were answered nonchalantly, with at least one prominent researcher suggesting that the domain of traits has always included such varied indicators. Nonetheless, ongoing debates about whether traits should subsume motives provide an exception to the otherwise unquestioned expansion of the trait domain, suggesting a potential boundary between traits and at least some other psychological constructs (Allport, 1937; Pervin, 1994; Winter et al., 1998). In addition, although not often explicated, current distinctions between similar Big 5 models sometimes imply differential emphases on As, Bs, and Cs between models. For example, Saucier and Goldberg (1996) state that they emphasize the phenotypical aspects of the Big 5 traits, suggesting a corresponding emphasis on observable trait expressions (behaviors), whereas McCrae

and Costa (1997a, 1999) emphasize the genotypical bases of the Big 5, suggesting a greater emphasis on covert trait expressions such as cognitions and affects (see Johnson, 1997). These debates over the constructs that traits may or must subsume, exclude, or emphasize are related to debates concerning traits as descriptive or explanatory (e.g., Epstein, 1994; Funder, 1994) because, as is sometimes noted, a construct cannot both describe and explain another construct (Bandura, 1999; Cervone & Shoda, 1999).

The Nature of Individual Big 5 Traits

Perhaps, however, the extent to which traits include ABC components is not a question of “general” relevance but instead varies by individual trait. Although broad traits such as the Big 5 have not been distinguished in ABC terms, debates concerning the facets (narrow traits) that are central to or subsumed within each broad trait often imply different ABC emphases. For example, Extraversion is sometimes related primarily to sociability (i.e., a cognitive/affective preference for and enjoyment of other people; McCrae & Costa, 1987), to surgency (behavioral dominance and achievement seeking; e.g., Goldberg, 1992), or to positive affect (Lucas, Diener, Grob, Suh, & Shao, 2000; Watson & Clark, 1997). Similarly, whereas McCrae and Costa (1997a) regard Openness as “a broad constellation of traits with cognitive, affective, and behavioral manifestations” (p. 832), alternative conceptualizations of Openness emphasize cognitive intellect (Goldberg, 1992), absorption (Tellegen & Atkinson, 1974), or culture (Tupes & Christal, 1992). Wiggins and Trapnell (1996) suggest that the “intellect” view of Openness emphasizes agency and the “mind,” whereas the “Openness to Experience” view emphasizes communal matters of the “heart” (p. 144). These disputes over central trait content have taken place at the theoretical and empirical levels. Here, we add a rare examination of these issues at the operational level.

Are ABC Distinctions and Emphases Important?

At the level of narrow traits, ABC dimensions are invoked when traits are distinguished as stylistic, intentional, motivational, emotional, interpersonal, temperamental, as primarily reflecting desires and preferences or processes (e.g., Allport, 1937; Buss, 1989; Hirschberg, 1978; Wiggins, 1997), and even as specifically A, B, or C (Johnson, 1997). ABC dimensions are also sometimes invoked to describe a trait’s impact—for example, explicating the relationships between the aggression-related constructs of anger (A), violence (B), and hostility (C) (Martin, Watson, & Wan, 2000)—or to ensure complete coverage of a domain (e.g., explicating the ABC components of attitudes or interpersonal

al impacts) (Wiggins & Trapnell, 1996). Thus, ABC constructs often play a central role in descriptions of causal relationships (e.g., designating the roles for As and Cs in models explaining Bs) (Bandura, 1999; Mischel & Shoda, 1998; Weiner, 1996) and for differentiating assessments and interventions appropriate to As versus Bs versus Cs (e.g., Lazarus, 2000; O'Keefe & Berger, 1999; Palmer, 1997). Thus, ABC distinctions seem important, at least at the level of narrow traits.

Are ABC distinctions also important at the level of broad traits and at the level of trait assessment? Because people often infer internal states from overt behavior, it is possible that differences in ABC emphases matter little. Johnson (1997) claims that "in the process of assessment, the genotypic/phenotypic distinction disappears" (p. 74) and cites his findings of very little difference in the validity of genotypic (cognitive and affective, according to Johnson) versus phenotypic (behavioral) items for assessing traits. However, Johnson also notes that his results varied somewhat between individual traits and that when assessing Extraversion, behavioral indicators were more valid than cognitive-affective indicators. Similarly, Angleitner et al.'s (1986) content analysis of trait inventories found certain reliable differences between traits. Neuroticism was typically measured with items assessing covert feelings, cognitions, and physical sensations, whereas Extraversion was assessed with more items referencing behaviors. If explicated further and for other broad traits, such differences between traits may have implications for how each trait is individually conceptualized and best assessed and (eventually perhaps) for clarifying what each trait can or should describe, explain, and predict.

The Present Study

So little empirical attention has been paid to understanding "what" it is that personality inventories really measure that Fiske (1986) suggested that inventories—the tools perhaps most often used by personality researchers—should not be used as operational definitions in experimental research. In this study, we examined inventories and further considered Pervin's (1994) question about the number and nature of constructs included in the trait domain. Is each Big 5 trait assessed by items reflecting similar balances of the ABC dimensions? Are inventories similar in how they assess Big 5 traits? To answer these questions, we selected four Big 5 inventories and asked judges to rate the extent to which ABCs were represented in each (and every) item. We also heeded warnings (e.g., Johnson, 1997) that test takers and constructors may not similarly interpret items and asked both experts and novices to provide item ratings.

Although other approaches, traits, and dimensions may be usefully and importantly investigated (e.g., Osgood, Suci, & Tannenbaum, 1957), our focus on the ABC dimensions and on the "operationalization" of Big 5 traits has certain advantages. First, A, B, and C form basic dimensions that have been reliably distinguished in prior content analyses (Werner & Pervin, 1986), although our approach is unique in allowing ABC blends to be reflected in our rating method (characterizing items by percentages of ABCs). Second, reflection on the Big 5 traits has implications for how these widely referenced traits have been conceptualized and measured. Third, examination of trait operationalizations provides a relatively efficient opportunity for researchers to re-evaluate the correspondence between conceptual and operational definitions. Inventories are often iteratively revised, not only to explicate the emerging theoretical definition of the construct but also based on scale properties and theory concerning assumed relationships between multiple constructs. Because inventory items are chosen in part based on empirical data reflecting their ability to fit a model, predict expected outcomes, or contribute to scale reliability, revised inventories can come to reflect implicit characteristics of the construct. To the extent that a set of inventories converges on the constructs of interest, an examination of those inventories may explicate those implicit features of the constructs (Werner & Pervin, 1986).

Hypotheses and expectations.

This study was intended to be exploratory and descriptive. However, we expected that the shared intuitions of inventory constructors should sometimes lead to substantial convergence between the four inventories, with regard to specific Big 5 traits. In particular, we expected our results to show the general agreement on the relevance of (negative) affect to Neuroticism, and the relevance of cognitive phenomena to Openness, as reviewed above.

However, we also anticipated that several factors might reduce the agreement between inventories selected in the present study. For example, the included inventories were selected to represent varying conceptualizations of the Big 5, each was derived from different initial item pools, and each used different scale construction methods for different purposes (for reviews, see John & Srivastava, 1999; Widiger & Trull, 1997). Variations in format and length also may reduce agreement. Adjective inventories may allow for more variability in item interpretations, and shorter inventories may differentially weight the prototypical aspects of the traits, which may in turn relate to different ABC components. Thus, replication across instruments, despite these differences, will allow for added confidence in the results.

METHOD

Personality Inventory Item Sample

Our item sample included 476 items from four popular trait inventories with established validity. To choose inventories for this study, we conducted a search of the PsychLit database for the keywords *five factor model*, or *Big 5*, and *personality traits*. This search, including articles and book chapters from 1980 to August 1998, resulted in more than 600 citations. Examination of the citations revealed that popular Big 5 inventories vary widely in length, format, theoretical origin, and usage. Within the titles, abstracts, and other fields of the citations, the NEO (FFI, PI-Ror PI) (Costa & McCrae, 1985, 1992b) was mentioned more than 400 times, whereas other specific inventories were mentioned no more than 50 times. Therefore, we included the 240-item NEO-PI-R in this study and the appendix emphasizes detailed NEO results. Various adjective checklists were also very popular and were mentioned by name approximately 100 times. In this study, we included the 100-item Unipolar Adjective Trait Descriptors (ATD) (Goldberg, 1992) and the 92-item Revised Interpersonal Adjective Scales (IASR-B5) (Trapnell & Wiggins, 1990). Although not as frequently mentioned, the 44-item Big Five Inventory (BFI) (John & Srivastava, 1999) also was included because its use of short phrases provided a compromise between decontextualized adjectives versus statements. As suggested by Werner and Pervin (1986), we obtained ratings for the entire population of items from each inventory.

Design

The units of analysis in this study were individual inventory items and the dependent variables were percentages reflecting the ABC content of each item. The items were grouped by specific personality inventory (NEO-PI-R, ATD, IASR-B5, and BFI) and trait (N, E, O, A, C), and the data were primarily analyzed using 4 (inventory) \times 5 (trait) factorial ANOVAs. However, to explore the generalizability of the findings across inventories varying in length, we also present descriptives for the subsets of items comprising the 60-item NEO-FFI (the short form of the NEO-PI-R) and 40-item ATD-Mini (i.e., the Mini-Markers scales) (Saucier, 1994).

Participant Raters for Item Assessment

We followed the lead of previous researchers (Angleitner et al., 1986; Werner & Pervin, 1986) and assessed item content using relatively few primary raters. The three present authors provided the primary expert ratings. Although agreement between three expert raters would provide evidence of

internal reliability, we wished to also establish that those ratings would correspond with the ratings of both other experts and nonexperts. Thus, a second group consisted of six other experts, including two PhD-level psychologists (one in clinical and one in social/personality psychology) and four advanced graduate students (each with 5 to 8 years of graduate-level training in cognitive, clinical, or social/personality psychology).

The third rating group consisted of 30 student raters from undergraduate sophomore- and junior-level psychology courses who volunteered for extra course credit. However, the data of three students were dropped due to failure to follow instructions (e.g., not assigning percentages summing to 100%). Of the remaining 27 students, 21 were women, 2 were men, and 4 did not report their gender. Their mean age was 22 years (range 20 to 39); 11 were psychology majors, 12 reported other majors, and 4 did not report their major.

Procedures

To rate the items in the same contexts in which they are normally encountered, all items were presented in the sequence of their inventories. To establish between-group reliabilities, all three classes of raters rated the 240 items from the NEO-PI-R. Only the three primary expert raters rated the remaining 236 inventory items. Items included in the short-form inventories were rated only once by each rater in the context of the long inventory. However, 26 adjective items overlapped between the ATD and the IASR-B5 and were therefore rated twice (once in the context of each inventory).

Raters were provided with instructions, a copy of items from the relevant inventories, and answer sheets. For convenience and to reduce fatigue, raters were instructed to take frequent breaks and to complete the ratings at home. The instructions described the purpose of the study and defined the ABC dimensions in a manner that paralleled definitions used in prior research (e.g., Werner & Pervin, 1986) and that would be understood by both experts and laypersons. Behavior was defined as overt and directly observable actions, including both active (e.g., bike-riding) and passive (e.g., watching television) behaviors, but not including strictly mental events (e.g., thinking). In contrast, thoughts/cognitions were defined as thoughts, beliefs, patterns, or modes of thinking. Affective feeling/emotion constructs were described as internal, motivational, and evaluative, valenced states, including patterns of feelings, emotions, "feeling-like" states, and preferences. Although two other component categories were also initially included, the extremely low use of the "bodily" and "other" categories led us to focus entirely on the ABC ratings.

Raters assigned percentages reflecting the extent that each inventory item assessed each component (ABC, physiological, and other). The percentages across components were required to add to 100% for each item. As an example of our rating procedures, consider the following item from the BFI: "I see myself as someone who . . . is generally trusting." For this item, the ratings averaged across the three primary experts were 37% A, 12% B, and 52% C. Those responses suggest the importance of both cognitive and affective processes when considering "trust." In contrast, average ratings for the BFI item "I see myself as someone who . . . generates a lot of enthusiasm" were 50% A, 50% B, and 0% C.

DATA ANALYSES AND RESULTS

Rater Accuracy

The technical accuracy of the primary and other expert ratings was very high. There were no missing ratings, and of the ratings by the primary raters (476 items × 3 primary raters), only one set of component ratings (for a single item) did not sum to exactly 100%. Of the other expert ratings (240 NEO-PI-R items × 6 raters), less than .2% reflected similar obvious rater errors. The error rate of the 27 undergraduate student raters was only .4%. These errors were minor (sums ranging from 90% to 110%).

Computation of the Composite Rating Scales

To summarize the rating data, we computed mean A, B, and C component ratings for each item within the appropriate classes of raters. To summarize the A, B, and C components

reflected in each trait within each inventory, mean ratings were then computed across the relevant items.

Reliability and Generalizability of Composite Ratings

To examine the internal consistency of each composite rating scale, Hoyt's intraclass reliability coefficients, mathematically equivalent to Cronbach's alpha (Crocker & Algina, 1986), were computed. As shown in the top half of Table 1, the intraclass reliabilities computed for the composite A, B, and C scales using ratings of the 240 NEO-PI-R items were high for each class of raters. The lower half of Table 1 summarizes the estimated alphas for the three primary raters across items from each of the other long-form inventories. Consistent with the results specific to the NEO-PI-R, interrater reliability for the primary raters was high and indicative of good internal consistency for each of the individual inventories.

The top half of Table 2 displays the average interrater correlations for each group of raters. Agreement was highest among primary experts and lowest among students. This pattern was expected for a number of reasons. First, the primary experts (i.e., the three authors) worked together to define ABC components, resulting in stronger and more salient shared conceptualizations. Second, the primary raters were all social/personality psychologists, whereas the other experts included psychologists from other areas. Third, as numbers of raters increase, the potential for different opinions and ratings also increases, especially when the heterogeneity of one's sample is intentionally increased along with sample size. The greatest heterogeneity of opinion (and of conscientious effort) was expected among the undergraduate students.²

TABLE 1: Intraclass Reliabilities Computed for Each of the Primary Rating Dimensions by Measure and Rating Group

Rating Group (N raters) or Measure (N items)	Affect	Behavior	Cognition	Mean
Intraclass reliabilities by rater group				
All experts (9)	.94	.95	.92	.93
Primary experts (3)	.89	.90	.87	.89
Other experts (6)	.89	.90	.85	.88
Students (27)	.92	.94	.91	.92
Intraclass reliabilities by measure				
NEO-PI-R (240)	.89	.90	.87	.89
BFI (44)	.90	.87	.89	.89
ATD (100)	.91	.91	.89	.90
IASR-B5 (92)	.94	.90	.92	.92
All items (476)	.90	.90	.88	.89

NOTE: BFI = Big Five Inventory, ATD = Unipolar Adjective Trait Descriptors, IASR-B5 = Revised Interpersonal Adjective Scales. Rater group reliabilities are based only on data from the NEO-PI-R (the inventory rated by all groups). Measure reliabilities are based only on data from the primary experts (the group rating all inventories).

TABLE 2: Average Interrater Correlations and Intergroup Composite Correlations for Affective, Behavioral, and Cognitive Rating Data

Rating Groups	Rating Dimension		
	Affect	Behavior	Cognition
Average interrater correlations			
Primary experts (n = 3)	.74 (.75)	.75 (.75)	.70 (.72)
Other experts (n = 6)	.59	.62	.52
Students (n = 27)	.28	.36	.26
Intergroup composite correlations			
Primary experts–other experts	.89	.90	.86
Primary experts–students	.86	.90	.84
Other experts–students	.85	.91	.82

NOTE: Values in parentheses for the primary raters are the average interrater correlations based on all items rated from all inventories (N = 476 items). Other average interrater correlations were based on ratings of the NEO-PI-R data (N = 240 items). All intergroup composite correlations are significant at p < .001 and were based on composite ratings for 240 NEO-PI-R items, computed as described in the text.

As shown in the bottom half of Table 2, correlations between composite ratings (for the NEO-PI-R data) calculated from primary ratings, other expert ratings, and student ratings indicated relatively high ordinal correspondence. This suggests that the ratings by the primary raters are ordinally representative of ratings made by the other two groups of raters. To assess possible differences in means between the three groups of raters, we conducted three univariate mixed factorial ANOVAs: 3 (rater groups) \times 5 (traits), using the A, B, and C composites as separate dependent variables (DVs).³ These univariate tests indicated a significant Rater \times Trait Factor interaction for each ABC component, $F_s(8, 470) > 8$, $ps < .001$, indicating that the differences between the primary experts, other experts, and student ratings depended on the trait factor. Further examination of these data (see the appendix) revealed that although each class of raters ordered the traits similarly on each ABC dimension, the primary experts tended to magnify the differences between traits on the ABC dimensions, relative to the students and the other experts. For example, the primary experts rat-

ed Neuroticism and Extraversion items (generally rated as substantially affective) as significantly higher than students on the affective component, but they rated Agreeableness and Conscientiousness items (judged as low on the affective dimension by all raters) lowest on the affect dimension. The tendency of the primary raters to magnify the differences between items and traits likely reflects greater sensitivity to ABC differences.

Differences Between Traits in ABC Content

Our central questions concerned whether ABC components are equally represented across traits and whether there is consistency between major Big 5 trait inventories. The mean component ratings averaged across items for each trait factor are listed by inventory in Table 3 and illustrated in Figure 1. To facilitate comparisons between facet and trait ABC composition, and to clarify potential sources of differences between the NEO (the most frequently used inventory according to our literature review) and other inventories, the mean ABC ratings for each trait facet assessed by the NEO-PI-R are listed in the appendix.

TABLE 3: Item Means (and standard deviations) by Rating Dimension, Trait Factor, and Personality Inventory Measure

Measure	N	E	O	A	C
Behavioral ratings					
ATD	17.1 ^a (16.4)	64.2 ^{bc} (14.3)	19.3 ^a (21.1)	50.1 ^b (24.9)	67.3 ^c (18.2)
ATD-Mini	13.5 ^a (7.8)	61.5 ^{bc} (13.6)	14.6 ^a (15.3)	43.5 ^b (22.8)	66.5 ^c (10.1)
IASR-B5	15.1 ^a (7.3)	53.5 ^{bc} (16.7)	19.4 ^a (17.1)	39.9 ^b (16.7)	62.3 ^c (19.6)
NEO-PI-R	19.9 ^{ab} (32.1)	44.8 ^c (36.9)	15.5 ^a (28.7)	38.7 ^{bc} (39.3)	70.2 (30.8)
NEO-FFI	0.3 ^a (1.0)	38.1 ^b (33.6)	17.5 ^{ab} (37.3)	52.0 ^b (41.1)	85.8 (12.3)
BFI	13.3 ^a (14.5)	72.1 ^b (20.3)	14.2 ^a (11.2)	60.4 ^b (34.5)	67.6 ^b (19.2)
Overall	17.1 ^b (24.4)	52.9 ^a (30.3)	16.9 ^b (23.6)	43.4 ^a (33.5)	67.8 (25.4)
Cognitive ratings					
ATD	12.3 ^{ab} (9.0)	7.7 ^a (10.1)	70.4 (20.8)	15.6 ^{ab} (12.1)	23.8 ^b (15.3)
ATD-Mini	14.0 ^{ab} (9.3)	3.1 ^a (4.9)	76.7 (12.7)	12.8 ^a (7.7)	27.7 ^b (11.8)
IASR-B5	14.6 ^a (12.8)	14.4 ^a (9.9)	73.8 (21.0)	14.3 ^a (6.2)	31.1 (21.1)
NEO-PI-R	12.2 ^{ab} (20.6)	7.4 ^a (13.8)	48.0 ^c (36.9)	43.0 ^c (36.9)	24.2 ^b (29.2)
NEO-FFI	6.4 ^a (10.4)	5.9 ^a (15.4)	49.3 ^c (35.6)	37.4 ^{bc} (35.7)	13.1 ^{ab} (11.7)
BFI	7.7 ^a (15.3)	5.2 ^a (5.5)	75.0 (24.1)	22.8 ^a (24.9)	29.8 ^a (19.3)
Overall	12.3 ^a (16.7)	8.49 ^a (12.06)	60.6 (32.3)	30.20 ^b (31.1)	26.1 ^b (24.4)
Affective ratings					
ATD	70.6 (16.3)	28.2 ^b (16.5)	10.3 ^a (11.1)	34.3 ^b (22.3)	8.8 ^a (14.5)
ATD-Mini	72.5 (6.6)	35.4 ^b (14.9)	8.8 ^a (5.7)	43.7 ^b (24.4)	5.8 ^a (4.7)
IASR-B5	70.3 (11.4)	32.1 (18.0)	6.83 ^a (7.8)	45.9 (20.0)	6.6 ^a (3.8)
NEO-PI-R	67.9 (67.9)	47.1 ^b (38.1)	35.3 ^{ab} (33.3)	18.3 ^a (23.3)	5.5 (14.6)
NEO-FFI	93.3 (10.4)	55.9 ^c (32.7)	31.5 ^{bc} (32.7)	10.5 ^{ab} (12.4)	1.1 ^a (2.2)
BFI	75.3 (75.3)	22.7 ^a (18.2)	10.8 ^a (17.0)	16.9 ^a (14.1)	2.6 ^a (4.7)
Overall	69.6 (27.0)	38.3 (31.2)	21.9 ^a (27.8)	26.4 ^a (24.2)	6.2 (12.4)

NOTE: N = Neuroticism, E = Extraversion, O = Openness, A = Agreeableness, C = Conscientiousness, ATD = Unipolar Adjective Trait Descriptors, ATD-Mini = the Mini-Markers scales, IASR-B5 = Revised Interpersonal Adjective Scales, NEO-FFI = the short form of the NEO-PI-R, BFI = Big Five Inventory. Common superscripts indicate nonsignificant pairwise differences within a row, as determined using Tukey's honestly significant difference (HSD) follow-ups (all one-way ANOVAs for each inventory revealed $F_s > 7.76$, $ps < .001$, uncorrected). Overall means were computed across all relevant individual items. Therefore, the overall means more heavily weight scales with more items. (Short-form items were not included twice in the computation of the overall means.) Numbers of items by scale were as follows: ATD: 20 items per trait factor; ATD-Mini: 8 items per trait factor; IASR-B5: 16 items each for Agreeableness and Extraversion, 20 items for each of the other trait factors; NEO-PI-R: 48 items per trait factor; NEO-FFI: 12 items per trait factor; BFI: 8 items each for Neuroticism and Extraversion, 10 items for Openness, 9 items each for Agreeableness and Conscientiousness.

To examine differences in ABC component ratings between inventories and between traits, we conducted three separate univariate factorial ANOVAs: 4 (long-form inventories) × 5 (trait factors), using each ABC component rating composite as a dependent variable.⁴ For these omnibus analyses, only the composite ratings made by the primary raters were used as dependent variables because only the primary raters rated all items from all instruments. Also, because items in short-

form inventories overlapped with those in the long-form inventories, we focused primarily on differences between long-form inventories. However, Table 3 also indicates the results of analyses conducted for each individual inventory (long or short form). Those one-way ANOVAs were all significant at the $p < .001$ level (uncorrected), and Tukey's honestly significant difference (HSD) follow-ups were used to determine pairwise differences between Big 5 traits within inventories.

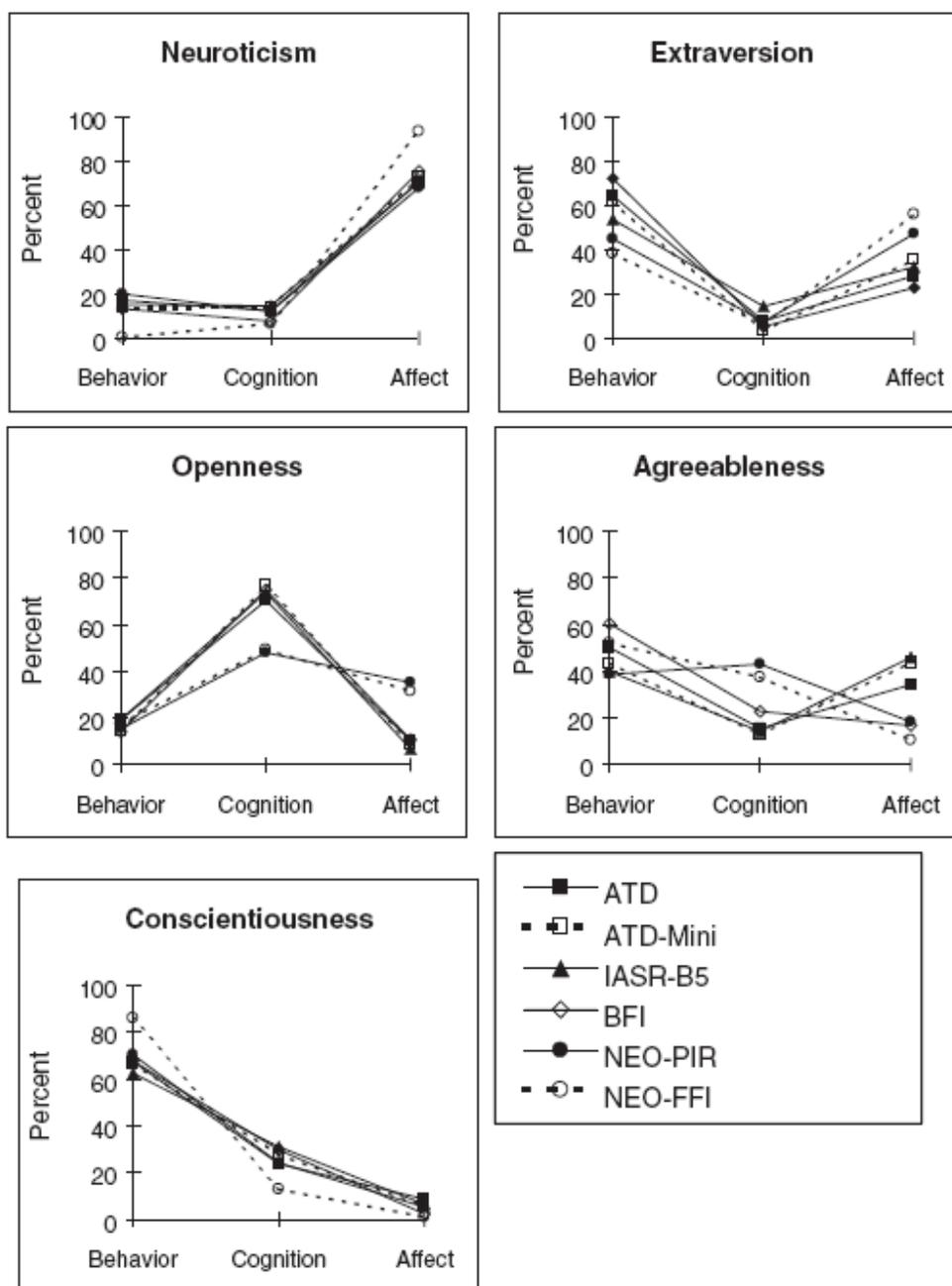


Figure 1 The relative importance of affective, behavioral, and cognitive components to each of the Big 5 traits, as determined by primary expert ratings of inventory items.
 NOTE: ATD = Unipolar Adjective Trait Descriptors, ATD-Mini = the Mini-Markers scales, IASR-B5 = Revised Interpersonal Adjective Scales, BFI = Big Five Inventory, NEO-FFI = the short form of the NEO-PI-R.

Behavioral ratings. The univariate factorial ANOVA revealed a main effect of trait factor, $F(4, 456) = 63.15, p < .001$, but no significant main effect of inventory and no Trait Factor \times Inventory interaction. Pairwise follow-up analyses conducted using Tukey's HSD procedures (and collapsing the data across inventories) revealed that Conscientiousness was rated significantly higher than all other traits on the behavioral dimension (see Table 3). Extraversion and Agreeableness were next highest and not significantly different from one another. Finally, Neuroticism and Openness were significantly less behavioral than the other three trait dimensions.

Cognitive ratings. Both a main effect of trait, $F(4, 456) = 73.41, p < .001$, and an Inventory \times Trait interaction, $F(12, 456) = 4.91, p < .001$, were found. One-way ANOVA follow-ups revealed a significant trait effect for all four major inventories as well as for the two short-form inventories. Tukey's HSD follow-ups (see Table 3) revealed that Openness included significantly more cognitive content than all other traits when assessed with the ATD, ATD-Mini, IASR-B5, and BFI inventories. Openness also contained the most cognitive content when assessed with the NEO-PI-R and NEO-FFI inventories but not significantly more than was included in Agreeableness assessed by those inventories. In addition, all inventories tended to assess Extraversion and Neuroticism using items with low average amounts of cognitive content.

Affective ratings. ANOVA analysis of affect ratings revealed a main effect for traits, $F(4, 456) = 93.39, p < .001$, and an Inventory \times Trait interaction, $F(12, 456) = 4.82, p < .001$. One-way follow-ups again revealed significant differences between traits for each long- and short-form inventory. Pairwise examination of differences between traits using Tukey's HSD follow-ups (see Table 3) revealed that Neuroticism was significantly higher in affective content than the other traits, regardless of which long- or short-form inventory was being analyzed. Furthermore, Conscientiousness was always lowest in affect and nearly always rated as significantly less affective than Extraversion (with the exception of the BFI) and Agreeableness (with the exception of the BFI and NEO-FFI).

Consistency across inventories. As noted above, the pattern of differences between traits was quite robust across inventories. However, additional one-way ANOVA analyses revealed that the greatest differences among long-form inventories occurred when assessing cognitive and affective components of Agreeableness, $F_s(3, 89) = 6.87$ and $7.90, ps < .001$, for cognitive and affective ratings respectively, and Openness, respective $F_s(3, 94) = 5.51$ and $9.22, ps < .01$. Tukey's HSD follow-ups revealed that the NEO-PI-R in-

cluded significantly less cognitive content in Openness and significantly more cognitive content in Agreeableness than the two adjective inventories. In addition, the NEO-PI-R included greater assessment of affect for the Openness dimension than did each of the other three inventories and significantly less affect for Agreeableness than did either the ATD or IASR-B5. In general, the pattern of means obtained from the short-form measures (the NEO-FFI and the ATD-Mini) closely resembled the pattern displayed by the corresponding long-form measure (see Figure 1). However, at times, data from the short-form measures resulted in more extreme ABC estimates. For example, of all the inventories, items from the NEO-FFI, the BFI, and the ATD-Mini resulted in the highest affect estimates for trait Neuroticism. The NEO-FFI and the BFI also provided the lowest affect estimates for trait Conscientiousness.

DISCUSSION

The present study offers a different perspective on the nature of traits in general, and the nature of the Big 5 traits in particular, by asking what levels of A, B, and C are included within the operational definitions of each of the Big 5 traits and by asking whether various prominent inventories assess those traits similarly. Our results showed that the overall differences in ABC content between inventories were either nonsignificant or were subsumed by Trait \times Inventory interactions. However, within each inventory, different Big 5 traits were typically operationalized and described by different levels of ABC components. All of the inventories examined converged on that finding, although they reflect different formats and were constructed from different perspectives by different theorists using different methods for somewhat different purposes. The most striking features of these operational definitions of traits were the associations of behavioral content with Conscientiousness (and secondarily, Extraversion and Agreeableness), of cognitive content with Openness (followed by Agreeableness and Conscientiousness), and of affective content with Neuroticism (followed by Extraversion and Agreeableness). Next, we compare the observed operational definitions with conceptual views of the Big 5 traits.

Conceptual and Operational Depictions of Big 5 Traits

Conscientiousness. Neuroticism and Conscientiousness were the two factors on which we found most agreement across inventories. Consistent with McCrae and Costa's (1987) proactive conceptualization of Conscientiousness, there was convergence across inventories that Conscientiousness is dominated by behavior (68% overall). Further-

more, consistent with Conscientiousness including “being governed by conscience” (McCrae & Costa, 1987, p. 88), each inventory included at least a modest cognitive component (26% overall).

Neuroticism. McCrae and Costa (1987) stated, “virtually all theorists would concur in the centrality of negative affect to Neuroticism” (p. 87). Indeed, we found that Neuroticism scales were dominated by affective items (70% overall). But McCrae and Costa also noted that Neuroticism includes “the disturbed thoughts and behaviors that accompany emotional distress.” Consistent with that additional emphasis, the NEO-PI-R had the least affective emphasis (68%). However, the NEO-PI-R’s emphases on behavior (20%) and cognition (12%) are still small, with the behavioral component almost solely due to the inclusion of the facet of impulsivity (see the appendix). The shortened version of the Neuroticism scale in the FFI minimizes the contributions of behavior and cognition even further, assessing Neuroticism with 93% affectivity. Thus, although Neuroticism has been related to cognitions such as irrational beliefs (e.g., Barlow, 1988), the operationalization of Neuroticism is dominated by items assessing negative affect.

Openness. Consistent with descriptions of (and several complaints about) Openness as a cognitive disposition (e.g., Revelle, 1995; Wiggins & Trapnell, 1996), each scale assessing Openness included more cognitive than affective or behavioral content. However, consistent with a dual emphasis on cognition and motivation (McCrae & Costa, 1997a), the NEO inventories diverged from others in their inclusion of a substantial (35%) affective component. In fact, the NEO inventories emphasize affect somewhat over behavior, whereas the other inventories give a slight precedence to behavior over affect. Costa and McCrae’s (1992b) relative emphasis on affect and Goldberg’s emphasis on cognition (in the ATD) are consistent with Wiggins and Trapnell’s (1996) claim that Openness to Experience versus intellect conceptualizations of this trait factor emphasize matters of the “heart” and “mind,” respectively. However, the NEO actually emphasizes both cognition and affect over behavior, perhaps consistent with descriptions of FFM as assessed by the NEO-PI-R and NEO-FFI as generally focused on genotypes rather than phenotypes (Saucier & Goldberg, 1996).

Extraversion. For Extraversion, the lack of cognitive content is striking, as is the fact that each inventory includes substantial emphasis on both affective and behavioral components. Across inventories, behavior tended to be emphasized to a greater extent (53%) than affect (38%). The predominance of behavioral content was most accentuated in the BFI, the inventory constructed with the aim of representing the

most prototypical aspects of each trait. Those observations contrast with the view that positive affect is the “core” of Extraversion (e.g., Lucas et al., 2000; Watson & Clark, 1997). The NEO inventories include the greatest percentage of affective content, achieving that affective emphasis through the inclusion of the facets of positive affect and warmth (see the appendix). Consistent with an emphasis on agency and dominance within Extraversion, the IASR-B5 and the ATD included somewhat more behavioral content.

Agreeableness. Both the dyadic-interactional and the lexical perspectives of the Big 5 traits assign great breadth (second only to Extraversion) to Agreeableness, and John and Srivastava (1999) note that great breadth allows different researchers to emphasize different aspects of the domain. Perhaps this explains why there was the least agreement among inventories over the conceptualization of Agreeableness. A recurring theme for each of the other traits was that the NEO inventories would sometimes diverge, in varying extremes, from the adjective assessments and the BFI inventory of prototypical trait indicators would generally most closely resemble the adjective inventories. However, the pattern for Agreeableness was more complex. The BFI accentuated the behavioral component of Agreeableness more than other inventories while converging with the NEO inventories in portraying Agreeableness as relatively low in affective content (see Figure 1 and Table 3).

Implications and Directions for Future Research

The present study distinguished between rather broad classes of responses and therefore may be seen as a rough (but robust) guide to the content of the scales included. It is interesting that despite certain explicit theoretical distinctions often made between different instantiations of the Big 5 traits, so much convergence across inventories was found with regard to the inventories used to assess the Big 5 traits. For example, whereas Saucier and Goldberg (1996) emphasize phenotypic descriptions of the Big 5, overall they do not rely on behavioral descriptors any more than other scale constructors. Similarly, whereas McCrae and Costa are somewhat unique in their conceptual attempts to explicitly reference all three ABC indicators of Big 5 traits (e.g., McCrae & Costa, 1987), their operationalization of the traits in the NEO instruments, similar to the other inventories included in this study, tends to emphasize different ABC components depending on the trait under consideration.

These findings of differential ABC emphases by broad trait may have practical implications for interpreting research. For example, finding the strongest correlations between cognitive dependent variables and the Big 5 traits of Openness and Agreeableness (e.g., Langston & Sykes, 1997) may sim-

ply reflect the fact that measures of these traits include more cognitive items. However, relations between cognition, affect, and behavior are multiple and complex, and many factors influence correlations found in research, including overlap in operationalizations, but also including actual causal relationships, measurement error, and so on. Therefore, future research is needed to determine the extent to which the current ABC content imbalances actually contributes to inflated or deflated trait-outcome correlations.

More important, however, these findings raise a host of other questions: Why are different traits assessed by different sorts of indicators? Do the differences found reflect differences in the sorts of constructs that are the “best” indicators of the various traits? Do the varying A versus B versus C emphases reflect on “core” aspects of the trait? Or do the differential emphases reflect bias in the initial item pools or (as might be suggested by the lexical hypothesis) differences in the import of the social impact of A, B, and C?

Assume for the moment that there is some very basic core or reality to Big 5–level traits, that the ABC dimensions are highly meaningful constructs for assessing that core, and that the operational definitions of traits on ABC dimensions in major inventories reasonably reflect those underlying latent traits. Given those assumptions, our findings suggest that abstract arguments (and conceptual definitions of traits, such as found in personality texts) about the basic nature of traits may miss the mark. The Big 5 traits seem to be very different from each other in basic dimensions of structure and substance, not merely in which facets they subsume. Abstract arguments about whether “traits” should include motivation or be conceptualized as behavioral dispositions rather than as affective or cognitive in nature, and so forth, are largely irrelevant if these major broad traits are substantially different in underlying substance and structure.

In addition, if the Big 5 traits are vastly different from each other in underlying structure and substance, then different traits may require different types of measurement models and instruments (as previously noted by Hirschberg, 1978). For example, if observers are better judges of behavior and self-reports are most accurate for more covert thoughts and feelings (as argued by Johnson, 1997), then—judging from the current findings—peer reports may be more accurate assessments of Conscientiousness, whereas self-reports may be more accurate for assessing Neuroticism and Openness. Finer-grained analyses aimed at identifying the types or subsets of ABC constructs related to individual traits might be particularly useful for further reflection on and refinement of trait assessments (Hirschberg, 1978). For example, the cognitive items used to assess Agreeableness may include more belief statements, whereas the cognitive items assessing other traits may focus more on cog-

nitive processes (e.g., the tendency to be inclusive in one’s thinking [Openness]).

However, we remain conscious that the most obvious limitation to this study is that although it describes how traits are currently measured, it leaves unanswered the question of how traits should be measured. If we are right in identifying the importance of ABC constructs for defining the structure of traits, but wrong in assuming that the operational definitions of traits in the inventories we have studied reflect the underlying traits, then the missing components, such as the cognitive components of Extraversion and Neuroticism and the affective component of Conscientiousness, suggest that more balanced inventories need to be developed.

NOTES

1. Here we use the terms *Big 5* and *five-factor model* interchangeably to refer to several different instantiations of models that have five factors of similar types (consistent with usage by McCrae & Costa, 1997b; Revelle, 1995). Our choice of Costa and McCrae’s (1992b) labels to describe these was based on their interpretability and their common use in the literature. For clarity, we distinguish between various specific five-factor models by referencing specific theorists and instruments commonly associated with each.

2. Note that it appeared that eight students used idiosyncratic rating methods; for example, uniformly (across items) over- or underweighting certain components. Without these eight students, the average interrater student correlations increase to .43, .55, and .39 for ABC components, respectively. However, the variability added by these eight students was mostly random and therefore changed the composite (mean) student ratings only slightly.

3. We avoid reliance on and emphasis of multivariate tests due to the nonnormality often associated with ipsative data (Greer & Dunlap, 1997). However, our multivariate tests nonetheless always revealed statistical significance. In this case, a similar 3×5 mixed-factorial MANOVA conducted using the behavioral and cognitive ratings as dependent variables (DVs) also revealed a significant Rater \times Trait Factor interaction, Wilks’s Lambda = .71, approximate $F(16, 938) = 10.76, p < .001$, as well as main effects of factor, Wilks’s Lambda = .50, approximate $F(8, 468) = 23.87, p < .001$, and rater, Wilks’s Lambda = .86, approximate $F(4, 938) = 17.83, p < .001$.

4. In addition, a 4 (inventory) \times 5 (trait factor) \times 2 (rating: cognitive, feeling) mixed-factorial MANOVA revealed a three-way interaction, $F(12, 456) = 6.72, p < .001$, as well as a Factor \times Dimension two-way interaction, $F(4, 456) = 80.29, p < .001$, and a main effect of factor, $F(4, 456) = 48.06, p < .001$.

APPENDIX
Mean Percentage Affective, Behavioral, and Cognitive Components for NEO-PI-R Traits and Facets,
as Determined by Primary Expert (PEX), Other Expert (OEX), and Student (Stu) Ratings

Trait or Trait Facet	Affect			Behavior			Cognition		
	PEX	OEX	Stu	PEX	OEX	Stu	PEX	OEX	Stu
Neuroticism	<u>67.9</u> ^{ab}	<u>54.9</u> ^{ac}	<u>41.9</u> ^{bc}	<u>19.9</u>	<u>18.9</u>	<u>21.8</u>	<u>12.9</u> ^{ab}	<u>23.6</u> ^{ac}	<u>30.5</u> ^{bc}
Anxiety	81.9	51.9	46.4	0.8	2.6	6.3	17.3	39.1	39.7
Angry hostility	88.1	67.2	43.8	10.4	10.1	22.8	1.5	21.9	27.3
Depression	79.2	75.3	55.9	0.0	5.8	10.8	20.8	18.9	29.0
Self-consciousness	90.8	75.1	50.7	5.8	12.1	18.8	3.3	12.6	25.7
Impulsiveness	18.3	21.8	19.4	77.5	56.7	52.0	4.2	15.1	22.0
Vulnerability	49.2	38.0	35.2	25.0	25.9	20.1	25.8	34.1	39.5
Extraversion	<u>47.1</u> ^b	<u>39.6</u>	<u>32.5</u> ^b	<u>44.8</u> ^b	<u>42.1</u> ^c	<u>35.8</u> ^{bc}	<u>7.4</u> ^{ab}	<u>16.2</u> ^{ac}	<u>25.8</u> ^{bc}
Warmth	57.9	49.5	41.1	30.6	25.9	28.4	11.5	24.6	26.0
Gregariousness	61.0	63.8	41.3	32.9	28.7	27.1	6.0	7.6	27.5
Assertiveness	6.7	9.6	16.3	78.8	70.8	49.1	10.4	19.6	30.3
Activity	20.4	7.2	19.5	77.9	70.5	53.2	1.7	16.4	20.8
Excitement seeking	69.8	56.6	37.9	25.8	29.1	30.4	4.4	11.8	24.4
Positive emotions	66.8	50.7	39.2	22.6	27.8	26.4	10.5	17.5	26.1
Openness	<u>35.3</u>	<u>34.1</u>	<u>30.6</u>	<u>15.5</u>	<u>19.3</u>	<u>18.7</u>	<u>48.0</u>	<u>46.1</u>	<u>46.9</u>
Fantasy	19.2	19.6	19.5	8.5	16.4	19.0	72.3	64.1	57.7
Aesthetics	60.6	54.8	39.1	12.3	18.3	16.1	22.9	25.1	39.4
Feelings	77.9	71.8	60.5	0.0	5.9	9.2	22.1	21.0	25.5
Actions	19.2	22.0	19.3	66.7	60.0	47.6	14.2	18.0	28.7
Ideas	31.0	32.7	22.3	3.8	10.6	12.9	61.9	56.7	62.8
Values	3.8	4.0	22.9	1.7	4.4	7.3	94.6	91.7	67.1
Agreeableness	<u>18.3</u> ^b	<u>21.2</u>	<u>25.9</u> ^b	<u>38.7</u> ^{ab}	<u>31.9</u> ^a	<u>28.5</u> ^b	<u>43.0</u>	<u>47.5</u>	<u>42.1</u>
Trust	19.0	16.2	28.7	3.3	10.4	10.7	77.7	72.7	57.4
Straightforwardness	19.6	15.5	21.3	56.7	44.7	36.4	23.8	39.8	38.0
Altruism	8.5	9.6	17.6	62.2	49.6	41.2	29.3	40.8	37.1
Compliance	10.4	20.1	21.3	73.3	49.4	48.4	16.3	30.3	25.8
Modesty	21.9	31.5	32.3	35.6	25.1	23.5	42.5	43.4	41.2
Tender-mindedness	30.6	34.3	34.2	0.8	7.8	10.7	68.5	58.1	53.5
Conscientiousness	<u>5.5</u> ^b	<u>5.8</u> ^c	<u>14.8</u> ^{bc}	<u>70.2</u> ^b	<u>64.0</u> ^c	<u>50.5</u> ^{bc}	<u>24.9</u> ^b	<u>29.7</u>	<u>31.5</u> ^b
Competence	4.2	9.6	19.5	52.5	41.7	32.9	43.3	48.8	45.3
Order	12.5	9.7	13.0	69.2	69.2	56.6	18.3	21.2	26.9
Dutifulness	0.0	1.0	14.6	94.2	80.0	60.8	5.8	15.8	21.4
Achievement	15.2	11.0	18.6	71.7	64.9	52.7	13.1	24.1	25.0
Self-discipline	0.8	2.3	13.3	89.0	72.9	56.3	10.2	24.8	26.8
Deliberation	0.4	0.8	9.7	45.0	55.4	43.9	54.6	43.8	43.9

NOTE: Percentages may not add to exactly 100 due to use of the bodily and other component ratings by a few raters. Common superscripts indicate significant ($p < .01$) differences between raters on an A, B, or C rating component, determined using paired *t* tests (significance was tested only for the major traits). Underlined values are mean ratings for each broad trait, averaged across the six narrower trait facets listed below each broad trait.

REFERENCES

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

Angleitner, A., John, O. P., & Löhr, F.-J. (1986). It's *what* you ask and *how* you ask it: An itemmetric analysis of personality questionnaires. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 61-108). Berlin: Springer-Verlag.

Bandura, A. (1999). Social cognitive theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 154-196). New York: Guilford.

Barlow, D. H. (1988). *Anxiety and its disorders*. New York: Guilford.

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

Buss, A. H. (1989). Personality as traits. *American Psychologist*, *44*, 1378- 1388.

Cervone, D., & Shoda, Y. (1999). Beyond traits in the study of personality coherence. *Current Directions in Psychological Science*, *8*, 27-32.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO personality inventory (NEO-PI) manual*. Odessa, FL: Psychological Assessment Resources.

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

Costa, P. T., Jr., & McCrae, R. R. (1992b). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Holt, Rinehart & Winston.

- Emmons, R. A. (1989). Exploring the relations between motives and traits: The case of narcissism. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 32-44). New York: Springer-Verlag.
- Emmons, R. A. (1995). Levels and domains of personality: An introduction. *Journal of Personality*, *63*, 341-364.
- Epstein, S. (1994). Trait theory as personality theory: Can a part be as great as the whole? *Psychological Inquiry*, *5*, 120-122.
- Fiske, D. W. (1986). The trait concept and the personality questionnaire. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 35-46). Berlin: Springer-Verlag.
- Funder, D. C. (1994). Explaining traits. *Psychological Inquiry*, *5*, 125-127.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*, 26-42.
- Greer, T., & Dunlap, W. P. (1997). Analysis of variance with ipsative measures. *Psychological Methods*, *2*, 200-207.
- Hirschberg, N. (1978). A correct treatment of traits. In H. London (Ed.) *Personality: A new look at metatheories* (pp. 45-68). New York: John Wiley.
- Hofstee, W.K.B. (1994). Will the true trait theorist please stand up? *Psychological Inquiry*, *5*, 134-137.
- John, O. P. (1990). The "Big 5" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford.
- John, O. P., & Srivastava, S. (1999). The Big 5 trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford.
- Johnson, J. A. (1997). Units of analysis for the description and explanation of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 73-93). San Diego, CA: Academic Press.
- Langston, C. A., & Sykes, W. E. (1997). Beliefs and the Big 5: Cognitive bases of broad individual differences in personality. *Journal of Research in Personality*, *31*, 141-165.
- Lazarus, A. A. (2000). Multimodal replenishment. *Professional Psychology: Research & Practice*, *31*, 93-94.
- Lucas, R. E., Diener, E., Grob, A. Suh, E. M., & Shao, L. (2000). Crosscultural evidence for the fundamental features of Extraversion. *Journal of Personality and Social Psychology*, *79*, 452-468.
- MacDonald, K. (1995). Evolution, the five-factor model, and levels of personality. *Journal of Personality*, *63*, 525-567.
- Martin, R., Watson, D., & Wan, C. K. (2000). A three factor model of trait anger: Dimensions of affect, behavior, and cognition. *Journal of Personality*, *68*, 869-897.
- McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of a five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81-90.
- McCrae, R. R., & Costa, P. T., Jr. (1997a). Conceptions and correlates of openness to experience. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 825-847). San Diego, CA: Academic Press.
- McCrae, R. R., & Costa, P. T., Jr. (1997b). Personality trait structure as a human universal. *American Psychologist*, *52*, 509-516.
- McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139-153). New York: Guilford.
- Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, *49*, 229-258.
- O'Keefe, E. J., & Berger, D. S. (1999). *Self-management for college students: The ABC approach* (2nd ed.). Hyde Park, NY: Partridge Hill Publishers.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Palmer, S. (1997). Modality assessment. In S. Palmer & G. McMahon (Eds.), *Client assessment: Professional skills for counselors* (pp. 134-167). London: Sage Ltd.
- Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, *5*, 103-113.
- Revelle, W. (1995). Personality processes. *Annual Review of Psychology*, *46*, 295-328.
- Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, *63*, 506-516.
- Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21-50). New York: Guilford.
- Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology*, *83*, 268-277.
- Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the Big 5 dimensions of personality. *Journal of Personality and Social Psychology*, *59*, 781-790.
- Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, *60*, 225-251.
- Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 767-793). San Diego, CA: Academic Press.
- Weiner, B. (1996). Searching for order in social motivation. *Psychological Inquiry*, *7*, 199-216.
- Werner, P. D., & Pervin, L. A. (1986). The content of personality inventory items. *Journal of Personality and Social Psychology*, *51*, 622-628.
- Widiger, T. A., & Trull, T. J. (1997). Assessment of the five-factor model of personality. *Journal of Personality Assessment*, *68*, 228-250.
- Wiggins, J. S. (1997). In defense of traits. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 97-115). San Diego, CA: Academic Press.
- Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 88-162). New York: Guilford.
- Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward and integration of two traditions in personality research. *Psychological Review*, *105*, 230-250.