

Personality Coherence: Moderating Self–Other Profile Agreement and Profile Consensus

Jeremy C. Biesanz and Stephen G. West
Arizona State University

Traditional research on moderator variables in personality has focused on measures of relative consistency. In contrast, using Goldberg's (1992) adjectives representing the Big Five personality traits, the authors examined the applicability of moderator variables to measures of personality coherence. The authors considered 3 traditional moderator variables (interitem variability, construct similarity, and scalability) and one new moderator variable: the temporal stability of response patterns. Across 2 studies, individuals with temporally stable response patterns had higher levels of personality coherence, as measured by self–other profile agreement and informant profile consensus, than did individuals with less temporally stable patterns. By comparison, the normatively based moderator variables did not moderate self–other profile agreement and informant profile consensus. The implications for personality structure and coherence are discussed.

One important focus of personality theories is on those non-physical features that distinguish a person from others and make that person either subtly or strikingly unique. Broad theories of personality (e.g., Allport, 1937; Cattell, 1950; Murray, 1938) all "emphasize the *consistency and coherence* [italics added] of normal personality and view the individual organism as an organized and complexly structured whole" (McAdams, 1997, p. 12). *Consistency* most commonly has been used to refer to the stability of individual differences in behavior across situations, although it may also be used to refer to the stability of these differences over time. *Coherence* refers to the lawful patterning and organization of attributes within an individual. In practice, previous empirical research in personality has focused nearly exclusively on questions of consistency. Researchers have isolated a single dimension of personality—a trait for example—and have asked questions related to individual differences on that dimension (e.g., Do people's relative standing on the dimension predict their relative standing on relevant behavioral measures?). Less frequently, researchers have posed questions about factors that may moderate the degree of consistency; for example, If two people have the same standing on a given dimension, can other characteristics be identified that predict differences between them on consistency among measured behaviors of interest?

Personality, however, is unquestionably more than simple consistency or the sum of different single components or dimensions. The concept of coherence conveys the gestalt that is commonly understood by laypersons to be personality. Coherence in personality refers, in part, to "the feature of personality that is most outstanding—its manifest uniqueness of organization" (Allport, 1955, p. 21). Understanding coherence in personality—the organization and patterning of different attributes of personality within an individual—is one of the major challenges facing the field of personality psychology. In the present article, we focus on coherence and examine whether moderator variables, which have proven useful for understanding consistency in personality, are also useful for understanding coherence in personality.

An Overview of Traditional Research on Moderator Variables

One fundamental task of traditional personality research is to determine how to conceptualize and measure personality so as to better understand and predict behavior. Allport (1961), for example, defined personality as, in part, something internal within persons that determines their characteristic behaviors and thoughts. However, if personality and behavior are theoretically linked, why then, at times, does the empirical relationship between personality and behavior appear to be weak?

The classic studies by Hartshorne and May (1928) are often cited as examples of findings that showed the weak relationship between personality and behavior. In investigating character and honesty in children, Hartshorne and May found that in general, different behaviors related to these presumably stable aspects of personality did not correlate highly. For instance, children's behavior when given an opportunity to steal was not highly correlated with their behavior when given an opportunity to cheat on a test.

Since these studies, a number of researchers have been dissatisfied with the frequent empirical findings of a low to moderate relationship between personality and behavior, what Mischel

Jeremy C. Biesanz and Stephen G. West, Department of Psychology, Arizona State University.

We thank William Graziano, Douglas Kenrick, Steven Neuberg, and Roger Millsap for their valuable comments on previous versions of this article. Portions of this article were presented at the Nags Head Invitational Conference on Personality and Social Behavior, Highland Beach, Florida, June 1998.

Correspondence concerning this article should be addressed to Jeremy C. Biesanz, who is now at the Center for Developmental Science, University of North Carolina at Chapel Hill, 100 E. Franklin Street, CB #8115, Chapel Hill, North Carolina 27599–8115. Electronic mail may be sent to biesanz@email.unc.edu.

(1968) identified as an apparent correlational ceiling of .30. To understand and explain these empirical relationships, researchers have taken several different traditional approaches (Krahe, 1992; West, 1983). One solution, hypothesized by Allport (1937), is that not all personality traits are equally relevant for all persons. As a result, we should expect people to differ reliably in the degree to which measures of their personality can predict their behavior. For example, only some people's behavior should be predicted reasonably well by their level of conscientiousness. In other words, individuals may differ in traitedness: how strongly, if at all, that trait influences each individual's behavior.¹

Understanding personality by considering indicators of traitedness is commonly referred to as a *moderator variable approach*—the level of traitedness is hypothesized to moderate the relationship between personality and behavior. According to this perspective, individuals who are traited in, for example, conscientiousness will have stronger personality-behavior relationships in this domain than will less traited individuals. The moderator variable approach stands in contrast to nomothetic trait-based approaches to personality, which, in their simplest form, assume that all individuals can be characterized as possessing the particular trait under consideration to an equal degree. In the traditional nomothetic view, the central measurement issue is assessment of the individual's position on the underlying trait dimension. However, by reintroducing Allport's (1937) idea that perhaps not all traits are equally relevant to all individuals, researchers face two distinct assessment tasks: (a) assessing each individual's position on the underlying trait dimension and (b) assessing traitedness.

Allport's insightful hypothesis sparked considerable research on how to measure traitedness. One approach has been to ask individuals about their personality trait-behavior relationships and to use those self-reports as indicators of traitedness. To date, there is substantial evidence that self-reports of the *consistency* of personality traits (Bem & Allen, 1974; Cheek, 1982; Kenrick & Stringfield, 1980; Mischel & Peake, 1982; Zuckerman, Bernieri, Koestner, & Rosenthal, 1989; Zuckerman et al., 1988; Zuckerman, Miyake, Koestner, Baldwin, & Osborne, 1991; but see Chaplin & Goldberg, 1984 for a failure to replicate) as well as the *relevance* of personality traits (Zuckerman et al., 1988; Zuckerman et al., 1991) are able to moderate self-other agreement. The self-report of consistency asks individuals to assess how much their behavior related to a specific trait varies from one situation to the next. Relevance asks a slightly different question, How relevant is this trait to their behavior in comparison to other traits? These two self-report measures, although slightly correlated, have largely independent moderating effects on self-other trait-level correlations (Zuckerman et al., 1988).

These self-reported moderators are highly dependent on individuals' inferences about their own past behavior. The use of these self-reports presumes that individuals can successfully perform two different inferential tasks. The first and simpler task is that individuals can self-assess their average standing on the broad trait of interest.² The self-assessment of global trait standing involves examining and aggregating one's levels on behaviors related to each of the specific trait adjectives that compose the broad trait—for example, with respect to the broad trait of conscientiousness, a person may reflect on behaviors related to the specific trait adjectives *punctuality*, *neatness*, *reliability*, *organization*, and so on and integrate these reflections into a single summary composite. The

second and far more complex task is that individuals can self-assess their level of traitedness. This requires that individuals take into account their global trait assessment and then report reliably how the trait is behaviorally expressed across different behaviors that may reflect the trait across different situations. Not surprisingly, this is a very difficult and error-prone task. For example, the test-retest reliability of self-reported consistency is extremely low, indicating little useful and stable variance among individuals' self-ratings of traitedness (see Amelang & Borkenau, 1986).

A separate body of literature has consequently proposed moderators that are less dependent on the quality of individuals' inferential processes. Bem and Allen (1974) examined participant *item response patterns* on a conscientiousness scale as a potential moderator of self-other agreement on conscientiousness. They created an ipsatized measure in which each participant's variability between items on the target dimension of Conscientiousness, relative to that participant's interitem variability on scales measuring other dimensions of personality, was used as a moderator of the relationship between self-reported conscientiousness and conscientiousness-related behaviors. Bem and Allen reasoned that participants with low relative interitem variance would display more consistent levels of conscientiousness across situations than participants with high relative interitem variance. As predicted, individuals with low relative interitem variance had higher trait-behavior correlations than individuals with high relative interitem variance.

Following this general line of reasoning, researchers have proposed other moderator variables that attempt to assess traitedness from an individual's response pattern, namely interitem variability, construct similarity, and scalability (Baumeister & Tice, 1988; Chaplin, 1991; Lanning, 1988). *Interitem variability* (Baumeister & Tice, 1988), following Britt (1993), is computed by simply taking the standard deviation among a person's standardized responses. *Construct similarity* (Chaplin, 1991) is the correlation between a person's response profile and the average response profile. *Scalability* (Lanning, 1988) is the sum total deviations of a person's response profile from the normative profile after adjusting them to have the same mean level. These more recent attempts to measure traitedness, however, all share the theoretical presumption that individuals whose trait profiles' shapes are highly congruent with that of the hypothetical "average" person are most predictable. For example, individuals whose pattern of conscientiousness (e.g., levels of reliability, punctuality, organization) mirrors the pattern of the average person are hypothesized to be more predictable in their behavior than individuals whose pattern is dissimilar to that of the average person.

¹ In keeping with Allport (1937, p. 332), this approach assumes that traitedness is continuous and a matter of degree. This stands in contrast to other approaches (e.g., Baumeister & Tice, 1988), which treat traited and untraited as two distinct types.

² For expository purposes in this article to ensure clarity, we distinguish between broad, global traits such as those identified by the five-factor approach to personality (e.g., Conscientiousness) and more molecular trait adjectives that reflect more specific aspects of broad, global traits (e.g., *punctuality*, *reliability*, *neatness*, *organization*).

Temporal Response Pattern Stability as a Moderator Variable

The logic underlying the moderator variable approach in personality contains a fundamental premise: If two individuals are both at the same location on the latent trait (e.g., have equal levels of conscientiousness), but differ on traitedness, the more traited individual will demonstrate more consistent behavior across situations with respect to that trait than the less traited individual. In this conception, traitedness is treated as a continuous variable on which people vary, as compared to earlier formulations which treated individuals as dichotomously either traited or untraited on a particular personality dimension (e.g., Baumeister & Tice, 1988). If we can thus identify individuals who demonstrate consistent behaviors with respect to a personality trait, we can then infer that this trait is relevant to their personality. Individuals who are traited are expected to show a stronger link between that personality trait and their behavior than less traited individuals.

Ozer (1986) has noted that there are several different ways to conceive of consistency besides nomothetic cross-situational consistency. Individuals can be consistent across different measures of behavior and across time as well as across situations. Allport (1937) suggested an important form of consistency in his commentary on Hartshorne and May's (1928) study that found low cross-situational correlations for honesty in children. Allport (1937) wrote that these low correlations "prove only that children are not consistent *in the same way*, not that they are inconsistent with themselves" (p. 250). If we examine individuals' item response patterns on a scale assessed at a single measurement occasion and ask if they are consistent with themselves, the answer is we do not know—there are insufficient data to make this judgment. Researchers need to know if individuals always behave in the same manner within these situations over time (for detailed review of temporal consistency see Fiske & Rice, 1955). As Allport noted, sampling responses across items does not provide an estimate of the consistency of individuals over time. Responses must be sampled both across items and over time.

This reasoning suggests that instead of comparing individuals' scale response patterns with that of the average person, it may be more useful to compare each person's current response pattern with his or her pattern collected on other measurement occasions. Regardless of the shape of the response pattern, individuals who are consistent in their response patterns across time are expected to be more predictable with respect to that trait and have higher levels of agreement.

Building on Allport's (1937) original idea, Biesanz, West, and Graziano (1998) examined this hypothesis—that individuals whose personality is consistently manifested over time will be more predictable than individuals whose personality is manifested less consistently. Otherwise stated, among individuals having the same overall (mean) level of conscientiousness, those individuals whose patterns of responses on trait adjectives related to Conscientiousness are stable over time were hypothesized to be more predictable in their behavior (i.e., display more consistency) than individuals whose patterns are less temporally stable. Across two studies, Biesanz et al. (1998) found that individuals whose item response patterns were stable over repeated assessments for trait adjectives related to (a) Conscientiousness and (b) Extraversion showed increased mean trait-level agreement with judgments from

their parents and peers on these personality traits. In other words, these individuals had higher levels of relative consistency on these broad traits. In contrast, other proposed moderator variables that compare individuals' response patterns to the "average" individual—the "normative" moderators interitem variability (Baumeister & Tice, 1988), scalability (Lanning, 1988), and construct similarity (Chaplin, 1991)—were not significantly related to self-other agreement on these same traits.

A Person-Centered Approach to Examining Moderation in Personality

To date, virtually all of the research on moderator variables within personality has focused on individual differences with respect to a specified trait. The criterion for judging a moderator variable has almost always been a measure of *relative consistency* (Krahe, 1992; Snyder & Ickes, 1985) such as the correlation on overall (mean) trait-level assessments between different informants. For example, self-other trait-level agreement on Conscientiousness asks, How highly do individuals' self-assessments of their level of conscientiousness correlate with the level reported by a knowledgeable informant? In these studies, the hypothesis is that individuals who are more traited will have higher relative consistency than individuals who are less traited.

By emphasizing and focusing on trait-level relative consistency, researchers have taken a variable-centered approach. The focus is on the consistency between two different informants in their rank ordering of a sample of individuals on a specific trait. The implicit consequence of the variable-centered approach is that the unit of analysis is the personality trait under examination rather than being the specific person under consideration (Magnusson & Torestad, 1993). The variable-centered approach, however, is simply one option for examining agreement. When multiple indicators of broad personality traits are used, researchers can shift the unit of analysis from traits to persons. The person-centered approach (e.g., Block, 1971; Colvin, 1993a; Ozer & Gjerde, 1989) emphasizes person-level agreement. The focus of research shifts to questions about *coherence*—the amount of agreement reached concerning the organization and patterning of different trait adjectives within each specific individual. For instance, person-centered self-other agreement on trait adjectives related to Conscientiousness asks, To what degree does an individual's reported pattern of conscientiousness (e.g., reliability, punctuality, organization, etc.) correlate with the pattern reported by observers of that individual? Similarly, person-centered informant consensus on Conscientiousness asks, How highly do the ratings of different observers correlate with regard to that specific individual's pattern of conscientiousness?

In this person-centered approach, an index of agreement such as a *Q* correlation (see Stephenson, 1952) is determined for each individual to quantify his or her coherence. Computationally the *Q* correlation is the same as the Pearson product-moment correlation commonly used in the variable-centered approach to relative consistency. The distinction between these approaches is that the typical Pearson correlation examines the rank order of individuals and thus is an index of relative consistency. In contrast, the *Q* correlation examines the rank order of attributes within a single individual and thus is an index of coherence.

Little research has been done to examine person-centered self-other agreement or informant consensus or the moderation of these

effects, in part because of the increased methodological complexity of this approach (see Funder & West, 1993). Response profiles have three different components: elevation, scatter, and shape (see Cronbach & Gleser, 1953). *Elevation* is simply the mean-level of response: the arithmetic average across the different profile elements. *Scatter* is the variability around the elevation (e.g., the standard deviation across the response elements). *Shape*—the focus of this article—is what is left in response profiles after equating different profiles on elevation and scatter. When a *Q* correlation is computed to quantify agreement between two profiles for an individual, the resulting correlation coefficient assesses the degree of similarity in the shape of the two profiles. The calculation of the *Q* correlation standardizes each of the two profiles so that their elevation ($M = 0$) and scatter ($SD = 1$) are equated. However, Cronbach (1955) noted that there still remains a component within profiles, which he termed *stereotypic accuracy*, that would potentially enhance the correlation between the two profiles. To illustrate, consider a 3-item profile on three trait adjectives related to Extraversion: *talkative*, *outgoing*, and *bold*. If people in general tend to be more talkative than bold and more bold than daring, then interpreting self–other profile agreement across these three items can be difficult. What part of the agreement is due to the person's unique attributes and personality and what part of the agreement is expected because people in general tend to be more talkative than bold and more bold than daring? Stereotypic accuracy can be removed from indices of profile agreement by the simple procedure of initially standardizing responses within each trait across the entire sample before computing profile agreement. When responses are standardized in this manner, each trait adjective has a mean of 0 and a standard deviation of 1. This procedure removes the normative pattern common to the sample of individuals (the profile composed of the average response to each item) and profile agreement is not artificially inflated because of stereotypic accuracy.

Overview

The present article extends previously published work on temporal response pattern stability as a moderator variable of personality in three important ways. First, across two studies, this article examines the ability of temporal response pattern stability as well as three normative moderators (interitem variability, construct similarity, and scalability) to moderate rater agreement within a person-centered approach.

Second, whereas Study 1 presents a re-analysis of previously published research using all available trait adjective data representing three broad traits, Study 2 expands the scope of traits examined to include each of the broad traits from the five-factor approach to personality (Agreeableness, Conscientiousness, Extraversion, Neuroticism and Openness to Experience). In this article we follow the emerging consensus that natural language descriptors of behavior compose these five broad traits (for historical reviews and theoretical perspectives, see Digman, 1996; John, 1990; Saucier & Goldberg, 1996; Wiggins & Trapnell, 1996, 1997; for empirical evidence see Costa & McCrae, 1988; Fiske, 1949; Goldberg, 1992; Norman, 1963; Tupes & Christal, 1961/1992). These five traits, by virtue of reflecting different behaviors that generally co-occur in persons' descriptions of others, provide a useful framework in which to examine

moderator variables in personality and temporal response pattern stability.

Finally, this article examines both self–other profile agreement and informant profile consensus. Specifically, we hypothesize that overall and for each of the broad traits from the five-factor approach to personality, individuals who have more temporally stable patterns will have higher (a) self–other profile agreement and (b) informant profile consensus than individuals with less temporally stable patterns.

Study 1

Method

Participants

Introductory psychology students ($N = 134$) were recruited to participate in return for partial fulfillment of their introductory psychology class requirements. A total of 111 participants completed the basic study requirements of attending three measurement sessions. Participants were encouraged to bring two acquaintances into the laboratory in exchange for additional credit toward fulfillment of their course requirements. In addition, participants provided consent for obtaining a parental rating via mail. Of the participants completing the basic study requirements, 102 (54 women and 48 men; mean age = 19.68 years, $SD = 2.29$) had ratings from at least two informants, and 79 participants had a parental rating as well as two peer ratings. Note that this study presents a reanalysis of Biesanz et al. (1998; Study 1), using all available data.

Materials

Participants, peers, and parents rated the participant on 57 unipolar trait adjectives developed by Goldberg (1992)—19 for Conscientiousness, 20 for Extraversion, and 18 for Neuroticism. Three trait adjectives proposed by Goldberg (1992) were not included: *Imperturbable* and *haphazard* were frequently not known by participants (cf. Graziano, Jensen-Campbell, Steele, & Hair, 1998); analyses done on a separate large sample indicated that *unexcitable* was not related to the other indicators of neuroticism for self-ratings. All ratings were on a 9-point scale ranging from 0 (*extremely inaccurate*) to 8 (*extremely accurate*). Participants' self-rating instructions were modified from Goldberg (1992) to limit self-assessments of behavior to the previous week. This change encouraged the reporting of more variability in the trait adjectives over time as opposed to general or typical behavior. The specific rating instructions were as follows:

Participant Trait Rating Instructions

Please use this list of common human traits to describe yourself as accurately as possible. Describe yourself as you see yourself at the present time, not as you wish to be in the future. Describe yourself as you were *during this past week*, as compared to other persons you know of the same sex and roughly your same age. Before each trait, please write a number indicating how accurately that trait describes you *for the past week*.

Peers and parents received Goldberg's (1992) standard rating instructions with the participant's name embedded within the instructions. Peers and parents used the same rating scale as participants. The specific rating instructions were as follows:

Peer and Parent Rating Instructions

Please use this list of common human traits to describe [participant's name] as accurately as possible. Describe [participant's name] as you see

[him or her] at the present time, as compared to other persons you know of the same sex and roughly the same age. Before each trait, please write a number indicating how accurately that trait describes [participant's name].

Design and Procedure

Participants completed the self-report inventory three times, at no less than 1-week intervals, in a lecture hall reserved for that purpose. Peers were separated from the participant they rated. In the case of two peers rating the same participant simultaneously, they were separated from each other. Questionnaires were mailed to the parent designated by participants. Peer and parent questionnaires were prefaced with an explanatory cover letter. Peers and parents provided only one assessment of the participant.

Calculating Self-Other Agreement and Informant Consensus

Self-other profile agreement and informant profile consensus were calculated separately both for the set of adjectives corresponding to each broad trait (i.e., across the 19 adjectives for Conscientiousness, the 20 adjectives for Extraversion, and the 18 adjectives for Neuroticism) as well as for the full response profile of 57 adjectives. When appropriate, trait adjectives were reverse-coded prior to any analysis. To reduce stereotypic profile agreement, self, peer, and parental responses for each trait adjective were all separately standardized using the means and standard deviations from the corresponding rating source. More specifically, self-ratings were first aggregated across the three assessments and then standardized using the self-rating data, peer ratings were standardized using peer rating data, and parent ratings were standardized using the parent rating data. Thus, for example, each peer-reported trait adjective would have a mean of 0 and a standard deviation of 1 across all peer ratings.

To quantify person-centered self-other agreement for each participant, the aggregated self-reported profile of trait adjectives for each broad trait (e.g., the 20 trait adjectives for extraversion) was correlated with each of the three informant rating profiles on those same trait adjectives. The three resulting *Q* correlations (i.e., participant with Peer A; participant with Peer B; participant with parent) were then averaged to form a single measure of self-other agreement for that trait. This procedure was performed separately for the trait adjectives corresponding to the broad Big Five traits of Conscientiousness, Extraversion, and Neuroticism. The procedure was then repeated using the full response profile of 57 adjectives.

Person-centered consensus was calculated in a parallel manner. Separately for each broad trait, three pairwise profile correlations (i.e., Peer A with Peer B, Peer A with parent, and Peer B with parent) were computed based on the trait adjectives for Conscientiousness, Extraversion, and Neuroticism. These three *Q* correlations were then averaged into a single composite measure of other-other agreement for that trait. Again, this procedure was also conducted using the full response profile of 57 adjectives.

Moderator Calculations

Values for each of the four moderator variables were calculated separately for each broad trait as well as for the full response profile across traits.

Temporal response pattern stability. To quantify the temporal stability of response patterns, responses were first standardized across persons separately for each assessment (e.g., the Time 1 responses were standardized across participants using the means and standard deviations based only on participant responses at Time 1). The correlation between the pattern of each pair of assessments was then computed within each participant both within each broad trait (e.g., across the 20 trait adjectives for Extraversion) as well as across the full response profile (i.e., the 57 trait adjectives for Conscientiousness and Extraversion). As discussed earlier, standardizing

responses removes the component of stability that is derived from stereotype accuracy (Cronbach, 1955). The three pairwise profile correlations (i.e., Time 1 and Time 2; Time 1 and Time 3; and Time 2 and Time 3) were averaged into a single composite measure of temporal response pattern stability separately for each broad trait and the full response profile. Adequate levels of generalizability were found across these three pairwise profile correlations: for the full response profile ($\rho^2 = .72$), for Conscientiousness ($\rho^2 = .62$), for Extraversion ($\rho^2 = .68$), and for neuroticism ($\rho^2 = .56$; see Shavelson and Webb, 1991).

Normative moderators: Scalability, construct similarity, and interitem variability. A mean participant response profile was computed for each individual by averaging their responses within adjectives across the three administrations. The normative moderators were computed on this mean profile for each individual. Lanning's (1988) scalability index was computed according to the following formula using participants' unstandardized responses:

$$\text{Scalability}_i = \sum |X_{ij} - (X_{.j} + X_{i.} - X_{..})|.$$

Person *i*'s scalability is a function of his or her response to a single trait adjective (X_{ij}), the average response in the full sample to that adjective ($X_{.j}$), person *i*'s mean level across adjectives ($X_{i.}$), and the grand mean across participants and adjectives ($X_{..}$). By multiplying the sum by negative one, scalability is coded such that individuals with higher values are more scalable and thus their profiles resemble more closely the mean response profile.

Construct similarity (Chaplin, 1991) was computed for each participant by calculating the correlation between the participant's unstandardized response profile (after aggregating across the three assessments) and the normative response profile—the average response profile across individuals. Again, higher values of construct similarity represent closer agreement to the normative profile. In research on relative consistency, higher values indicated increased traitedness.

Interitem variability was computed by taking the standard deviation across a participant's responses after standardization. Participant responses on item *j* were converted to *z* scores using the mean and standard deviation from the full sample. Following the procedure in recent studies (Baumeister, 1991; Baumeister & Tice, 1988; Britt, 1993; Chaplin, 1991), the standard deviation as opposed to the variance was used because it results in a less skewed distribution. In research on relative consistency, lower values of interitem variability indicate increased traitedness.

All three normative moderators were calculated separately for Conscientiousness, Extraversion, and Neuroticism. To estimate these moderators across the full range of participants' responses, construct similarity and interitem variability were also calculated using the full response profile across the 57 trait adjectives. Scalability on the full response profile was calculated by averaging the estimates of scalability derived separately for Conscientiousness, Extraversion, and Neuroticism.³

Results

Self-Other Agreement and Informant Consensus

Although our primary interest is in moderator effects, mean levels of profile agreement are initially presented to answer the following question: Did informants agree with each other and participants on profile shape? As shown in Table 1, there were small to moderate levels of self-other agreement and informant

³ Scalability is computed with respect to a specific trait (see Lanning, 1988). If scalability were calculated on the full response profile, mean-level differences among the broad traits would be erroneously incorporated in the scalability index. Averaging the scalability indices from each of the broad traits avoids this problem.

Table 1
Profile Agreement Means (and Standard Deviations) Across Participants for Study 1

Profile agreement	Full profile	Broad trait		
		CO	EX	NE
Self-other agreement	.18 (.18)	.10 (.21)	.15 (.22)	.08 (.18)
Informant consensus	.17 (.14)	.12 (.21)	.11 (.20)	.06 (.17)

Note. All mean correlations presented were transformed by means of a Fisher's r -to- z before averaging. N ranges from 73 to 79. CO = Conscientiousness; EX = Extraversion; NE = Neuroticism.

consensus for the full profile across participants. The broad traits of Conscientiousness, Extraversion, and Neuroticism offer views of profile agreement on a more focused level; across participants, the magnitude of self-other agreement and informant consensus was small but reliable for these three traits.

Although the magnitude of profile agreement presented in Table 1 is small to moderate, these average correlations are almost all statistically significant. The indices of profile agreement across participants are independent and thus can be combined and examined meta-analytically. Accounting for the range of sample sizes and the varying number of items that compose the profile correlations, average correlations greater than .07 are significant at $p < .01$ (see Shoda et al., 1994 for a parallel analysis).

Moderator Variables

Participants were consistent with themselves over time; the moderate to large values of temporal stability in Table 2 indicate that participants report the same patterns of responses across time. The results for temporal stability demonstrate that there is agreement based only on profile shape; recall that two large components of agreement—mean level and the normative pattern—have been removed by standardizing responses and are not represented in these correlations. The impact of the normative pattern on profile correlations based on unstandardized responses can be observed in the large mean value of construct similarity. Across participants, construct similarity—the correlation between a participant's unstandardized responses and the normative pattern—is moderate to large and uniformly slightly higher than the values for temporal stability.

The mean values of interitem variability and scalability are not easily interpretable because they are not in a commonly used metric. However, it is clear that interitem variability and scalability are not constrained by floor or ceiling effects, respectively (0 being the potential lower limit for interitem variability and the upper limit for scalability). Note that a value of 0 for interitem variability (because it is based on standardized scores), construct similarity, and scalability indicates that a person's response profile corresponds perfectly with the average response profile.

Moderating Profile Agreement

Did individuals with more temporally stable response patterns have higher levels of profile self-other agreement and profile consensus? Consider first the results for the full response profile

presented in rows 1 and 2 of Table 3. As predicted, individuals with temporally stable response patterns had significantly higher levels of self-other profile agreement and informant profile consensus. Thus, for individuals who have temporally stable response patterns, informants agreed with each other and with the individual on the shape of that pattern of behavior more than individuals with less temporally stable response patterns. At the more focused trait level, temporal stability on the traits of Extraversion and Neuroticism—but not Conscientiousness—had small to moderate relationships with self-other agreement and informant consensus on these traits. It is worth noting that with the modest sample size present in Study 1 ($N \cong 75$), the statistical power to detect an effect of the expected small to moderate size ($\rho = .20$) was only .41.

Are the traditional normative moderators of interitem variability, scalability, and construct similarity associated with agreement and consensus? The results for these proposed moderators were less encouraging. Although the relationship between interitem variability and self-other agreement on the full response profile was significant, this effect was opposite of the predicted direction. As originally formulated (Bem & Allen, 1974; see also Baumeister & Tice, 1988; Britt 1993), individuals with lower interitem variability were predicted to have increased agreement. The marginally significant relationship between construct similarity and self-other agreement on extraversion provided the only evidence of the predicted relationship for the traditional normative moderators out of 24 correlations presented in Table 3.

Discussion

The results from Study 1, although encouraging for temporal stability as a moderator variable, were less encouraging for the traditional normative moderators. However, nonsignificant results obtained under conditions of low statistical power must be interpreted with caution. Consequently, a second study was conducted to (a) examine temporal stability and the traditional normative moderators with greater statistical power and (b) expand the breadth of the traits examined to encompass the full five-factor approach by including Agreeableness and Openness to Experience.

Table 2
Moderator Variable Means (and Standard Deviations) for Study 1

Moderator variable	Full profile	Broad trait		
		CO	EX	NE
Temporal stability	.42 (.21)	.32 (.23)	.41 (.30)	.36 (.26)
Interitem variability	.88 (.22)	.69 (.21)	.76 (.21)	.77 (.21)
Scalability	-15.3 (3.2)	-13.4 (4.07)	-16.2 (4.78)	-16.6 (4.84)
Construct similarity	.55 (.27)	.36 (.22)	.42 (.30)	.54 (.35)

Note. Temporal stability and construct similarity are within-person correlations and potentially range from -1.0 to +1.0. Interitem variability ranges from 0 upward, whereas scalability ranges from 0 downwards, with higher numbers indicating greater scalability. N ranges from 73 to 79. CO = Conscientiousness; EX = Extraversion; NE = Neuroticism.

Table 3
Moderator Variable Correlations With Self-Other Agreement and Informant Consensus for Study 1

Profile agreement	Moderator variable			
	Temporal stability	Interitem variability	Scalability	Construct similarity
Full profile agreement				
Self-other agreement	.31**	.21†	-.04	.10
Informant consensus	.27*	.04	.05	.01
Trait-level profile agreement				
Self-other agreement				
Conscientiousness	-.10	-.04	.01	.07
Extraversion	.22†	.11	-.19	.22†
Neuroticism	.22*	.13	-.17	-.13
Informant consensus				
Conscientiousness	-.02	.04	-.01	-.10
Extraversion	.11	.09	-.11	.01
Neuroticism	.21†	.11	-.12	-.14

Note. *N* = 73 to 79.
† *p* < .10 (marginally significant). * *p* < .05. ** *p* < .01.

Study 2

Method

Participants

Introductory psychology students (*N* = 345) were recruited to participate in return for partial fulfillment of their introductory psychology class requirements. A total of 305 participants completed the basic study requirements of attending three measurement sessions. Participants were encouraged to bring two acquaintances into the laboratory in exchange for additional credit toward fulfillment of their course requirements. In addition, participants provided consent for obtaining a parental rating via mail. Of the participants completing the basic study requirements, 295 (202 women and 93 men; mean age = 19.57 years, *SD* = 2.91) had ratings from at least two informants and 233 participants had a parental rating as well as two acquaintance ratings. The statistical power in Study 2, given an *N* of 200, was sufficient at .81 to detect an effect of the expected small to moderate size (*p* = .20).

Materials

Participants, peers, and parents rated the participant on 97 unipolar trait adjectives—20 for Agreeableness, 19 for Conscientiousness, 20 for Extraversion, 18 for Neuroticism, and 20 for Openness to Experience (Goldberg, 1992). As in Study 1, peers and parents provided only one assessment of the individual and three trait adjectives proposed by Goldberg (1992)—*imperturbable*, *haphazard*, and *unexcitable*—were not included. The instructions for participants and informants were identical to those of Study 1.

Design and Procedure

The design and procedure exactly paralleled those of Study 1. Participants completed the self-report inventory three times, at no less than 1-week intervals, in a lecture hall reserved for that purpose. Peers were separated from the participant they rated. In the case of two peers rating the same participant simultaneously, they were separated from each other. Questionnaires were mailed to the parent designated by participants.

Self-other agreement, informant consensus, and moderator variables were calculated in the manner described in Study 1 for both the full

response profile across the 97 adjectives and separately for each of the five broad traits. Generalizability for the three measures of temporal response pattern stability was again adequate for the full response profile ($\rho^2 = .83$ for the full response profile; ρ^2 ranged from .66 to .73 for temporal stability calculated on the five broad traits).

Results

Self-Other Agreement and Informant Consensus

As we found in Study 1, did participants and informants in this study agree with each other on profile shape? Overall, the levels of profile agreement corresponded very closely to those from Study 1. As shown in Table 4, there were small to moderate levels of self-other agreement and informant consensus for the full profile across participants.

The magnitude of self-other agreement and informant consensus was generally small within each of the broad Big Five traits. Although these mean levels of agreement are small to moderate in size, they are statistically significant (*p* < .05). As in Study 1, levels of agreement across participants are independent and thus can be combined and examined meta-analytically. Average correlations presented within Table 4 greater than .047 are significant at *p* < .01. Visual confirmation of this analysis is presented in Figure 1. The majority of participants have peer-peer profile consensus levels greater than 0 for both the full profile and for each of the broad Big Five traits. However, there is considerable variability across participants in the magnitude of peer-peer profile consensus.

Before interpreting these small to moderate average correlations presented in Table 4, we performed a check on the standardization procedure used to remove possible artifactual sources of correlation. Standardizing responses before computing profile correlations, as was done in the current studies, theoretically reduces artifactual agreement due to shared meaning (i.e., stereotypic

Table 4
Profile Agreement Means (and Standard Deviations) Across Participants for Study 2

Profile agreement	Full profile	Broad trait				
		AG	CO	EX	NE	OP
Self-other agreement						
Self-Peer	.18 (.16)	.07 (.20)	.12 (.22)	.14 (.22)	.12 (.22)	.12 (.22)
Self-Parent	.17 (.21)	.09 (.28)	.12 (.28)	.16 (.28)	.11 (.30)	.11 (.28)
Informant consensus						
Peer-Peer	.16 (.21)	.06 (.26)	.11 (.27)	.10 (.27)	.10 (.28)	.10 (.28)
Peer-Parent	.14 (.16)	.07 (.21)	.06 (.20)	.11 (.20)	.07 (.20)	.08 (.19)

Note. *N* ranges from 149 to 219. AG = Agreement; CO = Conscientiousness; EX = Extraversion; NE = Neuroticism; OP = Openness to Experience.

agreement due to the normative response profile).⁴ In contrast, using *unstandardized* responses, Blackman and Funder (1998) reported average correlations of $r = .16$ for profile consensus and $r = .18$ for self-other profile agreement between full response profiles based on ratings of *different* individuals who were randomly paired (for similar levels in a different context, see Pelham, 1993). As a check on whether the standardization procedure removed this artifactual source of correlation, the expected level of profile agreement was examined for ratings of

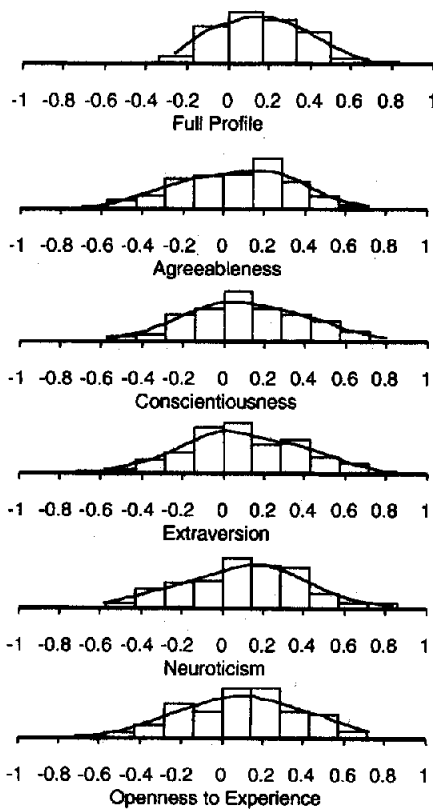


Figure 1. Coherence among peer ratings: Summary of peer-peer profile agreement for Study 2. Values on the x-axis are peer-peer correlations. The histogram represents actual data and the height represents the proportion of cases. The superimposed kernel density plot (the curve) estimates the underlying distribution.

different participants by participants, peers, and parents. Following Blackman and Funder's procedure, we calculated the average correlation between self-ratings of different individuals, peer ratings of different individuals, and parental ratings of different individuals using the *standardized* responses. Separately for each of these three sources of ratings, we created 511 random pairings. We then computed profile correlations based on both the full profile and each of the five broad traits. Across self-, peer, and parental ratings, the mean profile correlations for full profile and broad traits ranged from $-.019$ to $.019$ with a standard error of $.01$ and thus were not significantly different from zero. This check strongly supports the interpretation that the mean levels of self-other profile agreement and informant profile consensus presented in Table 3 reflect levels of coherence that are unique to participants and are not inflated by stereotype accuracy or produced artifactually.⁵

Moderator Variables

Across participants, the mean levels of the moderator variables were very similar to those obtained in Study 1 (see Table 5). Replicating Study 1, the mean levels of temporal stability and construct similarity were, on average, moderate to large. The mean levels of interitem variability and scalability corresponded closely to those from Study 1 as well.

⁴ Although responses were standardized separately based on the rating source, there was substantial agreement among participants, peers, and parents on the normative profile. For example, across the 97 trait adjectives, the correlation between the normative profiles derived from peers and parents was $r(95) = .93, p < .001$. Similarly, the normative profile based on self-reports correlated extremely highly with that of peers, $r(95) = .94, p < .001$, as well as parents, $r(95) = .88, p < .001$.

⁵ For comparison purposes, this same analysis was repeated based on *unstandardized* responses for another 511 random pairings of participant, peer, and parental ratings. The average level of profile agreement for ratings of different participants ranged from small to moderate: Full profile (self $\bar{r} = .23$, peer $\bar{r} = .21$, parent $\bar{r} = .33$); Agreeableness (self $\bar{r} = .17$, peer $\bar{r} = .14$, parent $\bar{r} = .21$); Conscientiousness (self $\bar{r} = .07$, peer $\bar{r} = .03$, parent $\bar{r} = .09$); Extraversion (self $\bar{r} = .15$, peer $\bar{r} = .12$, parent $\bar{r} = .22$); Neuroticism (self $\bar{r} = .24$, peer $\bar{r} = .16$, parent $\bar{r} = .25$); Openness to Experience (self $\bar{r} = .23$, peer $\bar{r} = .20$, parent $\bar{r} = .19$).

Table 5
Moderator Variable Means (and Standard Deviations) for Study 2

Moderator variable	Full profile	Broad trait				
		AG	CO	EX	NE	OP
Temporal stability	.46 (.17)	.33 (.21)	.34 (.23)	.40 (.22)	.40 (.22)	.41 (.22)
Interitem variability	.90 (.25)	.71 (.23)	.72 (.26)	.73 (.24)	.78 (.25)	.79 (.28)
Scalability	-16.7 (4.7)	-14.4 (5.3)	-16.0 (6.1)	-17.7 (6.3)	-18.9 (6.5)	-17.3 (6.7)
Construct similarity	.49 (.22)	.43 (.29)	.29 (.27)	.38 (.22)	.45 (.11)	.45 (.25)

Note. Temporal stability and construct similarity transformed by means of Fisher's *r*-to-*z* before averaging. *N* ranges from 193 to 219 for each of the broad traits because of missing data. For the full profile, 149 participants had complete data across all five broad traits for both self-reports and informant reports. AG = Agreement; CO = Conscientiousness; EX = Extraversion; NE = Neuroticism; OP = Openness to Experience.

Moderating Profile Agreement

Full profile analyses. Table 6 presents the moderator analyses based on the full profile of 97 trait adjectives. Consistent with Study 1, participants with more temporally stable response patterns had significantly higher full profile self-other agreement and informant consensus than participants with less temporally stable patterns. The larger sample size in Study 2 also permitted finer grained analyses differentiating among informants. Temporal stability was more strongly associated with self-peer profile agreement and peer-peer profile agreement than with self-parent or peer-parent agreement. For the traditional normative moderators of interitem variability, scalability, and construct similarity (presented in columns 2, 3, and 4, respectively), the association with full profile agreement and consensus was either nonsignificant or in the direction opposite of prediction. Lower interitem variability, higher scalability and higher construct similarity were predicted to be related to agreement and consensus.⁶

Trait-level profile analyses. On a more focused level, the next question we examined was, Is temporal stability, calculated solely within a broad trait, associated with profile agreement and consensus for that same trait? Table 7 presents the correlations between trait-level temporal stability and profile agreement and consensus for that same trait. Note that these analyses represent a more stringent test of the hypothesis than the analyses using the full response profile. Averaging meta-analytically across traits, high trait-level temporal stability was more significantly and positively associated with self-peer agreement, self-parent agreement, and peer-peer consensus, than was low trait-level stability.⁷

This meta-analysis combining across traits complements the examination of the relationship between temporal stability and agreement and consensus based on the full response profile. Analyses based on the full response profile are composed of consensus or self-other agreement based on both the organization and patterning of the attributes *within* broad traits for a person (e.g., within the broad trait Conscientiousness, she is more organized than neat) as well as differences *between* their broad traits (e.g., in general, she is more conscientious than extraverted). Computing temporal stability, self-other agreement, and consensus separately within each broad trait and then averaging across broad traits allows one to isolate effects based on *within* broad trait coherence and remove the influence of

across broad trait coherence. Thus this meta-analysis indicates that temporal stability is related to agreement and consensus within the specific attributes that compose broad Big Five traits. Note that although the relationship between temporal stability and self-other profile agreement for Conscientiousness in Study 2 appears substantially larger than that of Study 1, 95% confidence intervals around these effect sizes do overlap.

For the traditional normative moderators, no effects on the trait level were significant in the predicted direction. Indeed, as in the full profile analysis, several were significant in the direction opposite of prediction. The marginally significant positive relationship between construct similarity and self-other agreement for Extraversion observed in Study 1 was not replicated. Indeed, this nonsignificant relationship was, in fact, negative in Study 2.

Personality Correlates of Temporal Stability and Profile Consensus

What are the personality correlates of temporal stability and informant consensus based on the full response profiles? Examination of the mean trait level (i.e., the mean trait level of Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness to Experience) derived from self-reports and informant reports when compared with profile measures revealed several

⁶ When self-other profile agreement and consensus are computed on unstandardized profiles, their relationship with construct similarity is strong and positive. Given the negative relationship between construct similarity and consensus shown in Table 6, we conclude that the relationship observed based on unstandardized profiles is artifactual and likely driven by stereotype accuracy (see Cronbach, 1955).

⁷ In research on consistency, researchers (e.g., Paunonen & Jackson, 1985) have expressed concern that mean trait-level extremity may account for demonstrated moderator effects. In the present context, this hypothesis may be represented as an interaction between a quadratic trait-level term and temporal stability (see Aiken & West, 1991, pp. 69-70). Tests of this effect for both the full response profiles and the trait-level profiles found no evidence for this hypothesis for either self-other profile agreement or informant profile consensus.

Table 6
Full Profile Moderator Variable Correlations With Self-Other Agreement and Informant Consensus for Study 2

Full profile (<i>N</i> = 149)	Moderator variable			
	Temporal stability	Interitem variability	Scalability	Construct similarity
Self-other agreement	.28**	.13	.04	-.01
Self-peer agreement	.30***	.16*	-.03	-.02
Self-parent agreement	.14†	.02	.13	.01
Informant consensus	.17*	.14†	-.02	-.11
Peer-peer consensus	.27**	.18*	-.08	-.02
Peer-parent consensus	.06	.08	.02	-.14†

† $p < .10$ (marginally significant). * $p < .05$. ** $p < .01$. *** $p < .001$.

interesting findings.⁸ First, informant profile consensus was negatively related to self-ratings of the participant's level of Extraversion and Agreeableness, $r(147) = -.21, p < .01$, and $r(147) = -.18, p < .04$, respectively. Second, informant ratings of Extraversion were associated with lower temporal stability, $r(147) = -.27, p < .01$. No other relationship with mean trait level based on full response profiles approached statistical significance. Note that the present analyses differ from that of previous research (e.g., Colvin, 1993b) in two ways: (a) profile measures are based on standardized responses, and (b) separate reporting sources are used for mean trait-level and profile measures. In sum, extraverted individuals had both lower temporal stability and lower informant consensus levels relative to less extraverted individuals.

General Discussion

Across both studies and as predicted, temporal stability moderated profile agreement. Participants with temporally stable response patterns had higher informant profile consensus and self-other profile agreement than participants with less temporally stable response patterns. As the temporal stability of participants' patterns of responses increased, informants were more likely to agree with each other on the participants' profile of trait adjectives. At the same time, informants were also more likely to agree with participants' self-reported patterns of responses. These results, in sum, strongly support the interpretation of temporal stability as a personality moderator variable.

The current studies present the first evidence, of which we are aware, concerning the existence and moderation of measures of personality coherence for both the full response profile and within the broad traits from the Big Five. On average and across individuals, there was reliable self-other profile agreement and informant profile consensus at even the focused level of broad traits. Such profile agreement and consensus represents a coherence of personality—an organization and patterning of attributes that is unique to the individual and not shared with people in general. Profile agreement and consensus do not inform us about a person's standing on a trait (i.e., mean level) but rather inform us about agreement and consensus on the ordering of traits and the attributes of traits within an individual.

Beyond the existence of coherence within the broad traits of the Big Five is the moderation of this coherence. Traditionally, moderator analyses have examined factors that influence mean-level

agreement across individuals (e.g., agreement on how conscientious a person is). In contrast, moderation of profile agreement asks what factors influence agreement on the relative ordering of different attributes within a person (e.g., for the trait of Conscientiousness, agreement on whether a person is more reliable than organized). Examining the moderation of profile agreement thus provides data directly relevant to the existence, nature, and manifestation of coherence in personality. In the present studies, the clearest results are for the full profiles. To understand this, recall that the full response profile contains information on the relative ordering of the broad Big Five traits (e.g., whether, in general, a person is more extraverted than conscientious) as well as the relative ordering of the specific attributes that compose each of these traits. The more stringent examination of profiles within each of the Big Five—which do not contain information about the relative ordering of the broad Big Five—reveals both the existence and moderation of profile agreement and consensus for the attributes within these broad traits. Understanding the existence and moderation of coherence within each of the Big Five may require careful analysis of the facets or subcomponents of these broad traits. In recent work, Saucier and Ostendorf (1999) identified subcomponents of each of the Big Five traits and thus provided an empirical basis for future research in this direction.

No significant support was obtained for the traditional normative moderators of interitem variability, scalability, and construct similarity. Indeed, the significant results for interitem variability in the direction opposite of prediction (i.e., higher interitem variability associated with self-other agreement and informant consensus) appear to be inconsistent with Baumeister and Tice's (1988) proposal that interitem variability is an indicator of traitedness. To understand these results, it is worth reconsidering the logic under-

⁸ We report only analyses between mean trait-level and profile measures based on *separate* reporting sources (i.e., self-reported trait level with informant profile consensus; informant-reported trait level with temporal stability). We note that although there are no significant correlations between mean trait-level (for both self and informant reports) and self-other profile agreement, there is substantial evidence of curvilinearity in the relationship between mean trait-level and self-other profile agreement. However, because the mean trait level (for either reporting source) is a direct component of self-other profile agreement, interpretation of these relationships is difficult and potentially artifactual.

Table 7
Trait-Level Temporal Stability Correlations With Self-Other Agreement and Informant Consensus for Study 2

Profile agreement	Average across traits	Broad trait				
		AG	CO	EX	NE	OP
Self-other agreement	.16***	.05	.25***	.12†	.22**	.16*
Self-peer agreement	.14***	.07	.24***	.11	.16*	.13†
Self-parent agreement	.10*	.00	.10	.09	.16*	.12†
Informant consensus	.06†	-.05	.07	.13†	.14*	.00
Peer-peer consensus	.09*	.03	.08	.10	.12†	.13†
Parent-peer consensus	.01	-.09	.04	.11	.09	-.10

Note. The average across traits was tested meta-analytically using Rosenthal and Rubin's (1986) procedure for dependent effect sizes. *N* ranges from 193 to 219. AG = Agreement; CO = Conscientiousness; EX = Extraversion; NE = Neuroticism; OP = Openness to Experience.

† $p < .10$ (marginally significant). * $p < .05$. ** $p < .01$. *** $p < .001$.

lying the development of the traditional normative moderator variables. These moderator variables were developed to answer questions of relative consistency. The fundamental premise of the traditional normative moderator variables is that the mean trait level—*elevation* in Cronbach and Gleser's (1953) terminology—is the only meaningful aspect of personality measured within a response profile. However, the present studies show that after correcting for the normative response profile, there is self-other profile agreement and informant profile consensus for each of the broad Big Five traits. Meaningful and measurable aspects of personality remain within response profiles even after removing the mean trait-level and the normative profile. Because the normative moderators classify components of personality beyond mean trait level as indicators of "error" or untraitedness, it is not entirely surprising that these proposed moderators do not function as hypothesized in the present context.

Extraversion, Temporal Stability, and Informant Consensus

Among the broad traits from the five-factor approach, a person's level of extraversion is a very visible and salient aspect of his or her personality. After seeing only very thin slices of a person's behavior—behavior that lasts mere seconds—there is consensus among observers' ratings of *mean* levels of extraversion (Kenny, Horner, Kashy, & Chu, 1992). The present study, in contrast, demonstrates the apparently paradoxical finding that increased extraversion is associated with decreased *profile* consensus. Recall, moreover, that increased extraversion was also associated with decreased temporal stability, which in turn, was associated with informant profile consensus. Close examination of the behavioral correlates of extraversion provides some insight into the underlying behavioral processes that may account for these relationships.

Theorizing has emphasized extraversion's interpersonal nature (e.g., John, 1990). Research examining extraversion (e.g., Emmons & Diener, 1986; Emmons, Diener, & Larsen, 1986) has shown that, not surprisingly, individuals high on this construct are very socially active when free to choose their situations. Brown and Moskowitz (1998) recently replicated and extended this basic finding. Among a community sample of working adults who

participated in experience and event sampling, extraverted individuals attended more events and had both more and varied social partners during their nonwork time than less extraverted individuals. This increased social activity had an interesting consequence: Social activity was associated with increased variability in interpersonal behaviors throughout the day. The present finding of a negative relationship between extraversion and temporal stability is consistent with this empirical relationship.

That extraversion predicts more varied social partners and activities as well as variability in interpersonal behaviors suggests two potential mechanisms through which extraversion could be associated with decreased profile consensus. These behavioral correlates of extraversion correspond to two factors within Kenny's (1991) weighted-average model of interpersonal consensus: overlap and consistency. *Overlap* refers to the extent that different informants observe a target at the same time. If extraverts vary their social partners over time, these partners should have less overlap, which should adversely impact their agreement. Similarly, as extraverts are more variable in their interpersonal behaviors, this implies that they are necessarily less consistent in their behaviors, which the weighted-average model predicts leads to decreased consensus. In sum, these results reinforce the distinction between mean-level and profile analyses and how conclusions drawn from one source may not generalize to the other. Informants easily reach consensus that a person is, in general, highly extraverted, but that same high level of extraversion apparently makes it more difficult to reach consensus on that extraverted individual's coherence—the ordering of his or her attributes that compose extraversion and other aspects of personality.

Personality Coherence Revisited

Careful consideration of the different possibilities of how trait relevance may be related across traits raises questions about the nature of personality coherence, broadly construed. Understanding the nature of personality coherence—that "people respond consistently across some contexts and display distinctive patterns of variation across others" (Cervone & Shoda, 1999a, p. 27)—is a fundamental goal of personality psychology, essential to better prediction of behavior, and the primary focus of this article. That most people display at least modest levels of coherence in their

behavior across situations, time, and different behaviors is not in doubt. The strong levels of temporal stability in people's self-reported patterns of behavior as well as self-other profile agreement and informant profile consensus are all empirical manifestations of personality coherence. What is currently being debated is the origins and nature of that coherence.

Recent research (e.g., Shoda et al., 1994) and theorizing (Cervone & Shoda, 1999b) has emphasized that people display idiosyncratic and stable situation-behavior profiles that, in essence, are "behavioral signatures" (Mischel & Shoda, 1995, p. 246). Stable situation-behavior profiles are simply manifestations of the familiar person by situation interaction (Shoda, Mischel, & Wright, 1993, p. 1029) that have been extensively discussed (e.g., Endler & Magnusson, 1976). Cross-situational behavioral profiles are but one of many lenses through which personality coherence may be viewed and examined (Ozer, 1986). In this article, we present another different—but not mutually exclusive—perspective. Within the broad trait categories from the Big Five, individuals display idiosyncratic and stable average trait profiles on which knowledgeable informants agree. This is simply a manifestation of the less frequently examined person by response class interaction. To understand personality coherence, both of these perspectives must ultimately be carefully examined and considered.

Summary and Conclusion

Consistency and coherence are hallmarks of personality. Traditionally, researchers have focused on relative consistency in attempting to understand and measure personality. The dissatisfaction among some researchers with the empirical manifestation of measures of consistency led to the development of moderator variables that attempt to identify individuals who are more or less consistent than others in a given domain. The present studies indicate that the moderator variable approach is also applicable to measures of coherence. Individuals differ in how coherent measures of their personality are across time, which in turn influences how the coherence in their personality is perceived and understood by others.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Allport, G. W. (1955). *Becoming: Basic considerations for a psychology of personality*. New Haven, CT: Yale University Press.
- Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.
- Amelang, M., & Borkenau, P. (1986). The trait concept: Current theoretical considerations, empirical facts, and implications for personality inventory construction. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaire* (pp. 7–34). Berlin, Germany: Springer-Verlag.
- Baumeister, R. F. (1991). On the stability of variability: Retest reliability of metatraits. *Personality and Social Psychology Bulletin*, 17, 633–639.
- Baumeister, R. F., & Tice, D. M. (1988). Metatraits. *Journal of Personality*, 56, 571–598.
- 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, 506–520.
- Biesanz, J. C., West, S. G., & Graziano, W. G. (1998). Moderators of self-other agreement: Reconsidering temporal stability in personality. *Journal of Personality and Social Psychology*, 75, 467–477.
- Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology*, 34, 164–181.
- Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.
- Britt, T. W. (1993). Metatraits: Evidence relevant to the validity of the construct and its implications. *Journal of Personality and Social Psychology*, 65, 554–562.
- Brown, K. W., & Moskowitz, D. S. (1998). Dynamic stability of behavior: The rhythms of our interpersonal lives. *Journal of Personality*, 66, 105–134.
- Cattell, R. B. (1950). *Personality: A systematic theoretical and factual study*. New York: McGraw-Hill.
- Cervone, D., & Shoda, Y. (1999a). Beyond traits in the study of personality coherence. *Psychological Science*, 8, 27–32.
- Cervone, D., & Shoda, Y. (1999b). Social-cognitive theories and the coherence of personality. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization* (pp. 3–33). New York: Guilford Press.
- Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality*, 59, 143–178.
- Chaplin, W. F., & Goldberg, L. R. (1984). A failure to replicate the Bem and Allen study of individual differences in cross-situational consistency. *Journal of Personality and Social Psychology*, 47, 1074–1090.
- Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology*, 43, 1254–1269.
- Colvin, C. R. (1993a). Childhood antecedents of young-adult judgment. *Journal of Personality*, 61, 611–635.
- Colvin, C. R. (1993b). "Judgable" people: Personality, behavior, and competing explanations. *Journal of Personality and Social Psychology*, 64, 861–873.
- Costa, P. T., Jr., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, 55, 258–265.
- Cronbach, L. J. (1955). Process affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, 52, 177–193.
- Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, 50, 456–473.
- Digman, J. M. (1996). The curious history of the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 1–20). New York: Guilford Press.
- Emmons, R. A., & Diener, E. (1986). Situation selection as a moderator of response consistency and stability. *Journal of Personality and Social Psychology*, 51, 1013–1019.
- Emmons, R. A., Diener, E., & Larsen, R. J. (1986). Choice and avoidance of everyday situations and affect congruence: Two models of reciprocal interactionism. *Journal of Personality and Social Psychology*, 51, 815–826.
- Endler, N. S., & Magnusson, D. (Eds.). (1976). *Interactional psychology and personality*. New York: Halsted Press.
- Fiske, D. W. (1949). Consistency of the factorial structure of personality ratings from different sources. *Journal of Abnormal and Social Psychology*, 44, 329–344.
- Fiske, D. W., & Rice, L. (1955). Intra-individual response variability. *Psychological Bulletin*, 52, 217–250.
- Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality*, 61, 457–476.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Graziano, W. G., Jensen-Campbell, L. A., Steele, R. G., & Hair E. C.

- (1998). Unknown words in self-reported personality: Lethargic and provincial in Texas. *Personality and Social Psychology Bulletin*, 24, 893–905.
- Hartshorne, H., & May, M. A. (1928). *Studies in the nature of character: Vol. I. Studies in deceit*. New York: Macmillan.
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.
- Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98, 155–163.
- Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, 62, 88–97.
- 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.
- Krahé, B. (1992). *Personality and social psychology: Towards a synthesis*. London: Sage.
- Lanning, K. (1988). Individual differences in scalability: An alternative conception of consistency for personality theory and measurement. *Journal of Personality and Social Psychology*, 55, 142–148.
- Magnusson, D., & Torestäd, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology*, 44, 427–452.
- McAdams, D. P. (1997). A conceptual history of personality psychology. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 4–40). San Diego, CA: Academic Press.
- 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.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102, 246–268.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- 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.
- Ozer, D. J. (1986). *Consistency in personality: A methodological framework*. Berlin, Germany: Springer-Verlag.
- Ozer, D. J., & Gjerde, P. F. (1989). Patterns of personality consistency and change from childhood through adolescence. *Journal of Personality*, 57, 483–508.
- Paunonen, S. V., & Jackson, D. N. (1985). Idiographic measurement strategies for personality and prediction: Some unredeemed promissory notes. *Psychological Review*, 92, 486–511.
- Pelham, B. W. (1993). The idiographic nature of human personality: Examples of the idiographic self-concept. *Journal of Personality and Social Psychology*, 64, 665–677.
- Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, 99, 400–406.
- Saucier, G., & Goldberg, L. R. (1996). The language of personality: Perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 21–50). New York: Guilford Press.
- Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, 76, 613–627.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.
- Shoda, Y., Mischel, W., & Wright, J. C. (1993). The role of situational demands and cognitive competencies in behavior organization and personality coherence. *Journal of Personality and Social Psychology*, 65, 1023–1035.
- Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology*, 67, 674–687.
- Snyder, M., & Ickes, W. (1985). Personality and social behavior. In E. Aronson & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 2, pp. 248–305). New York: Random House.
- Stephenson, W. (1952). Some observations of the Q technique. *Psychological Bulletin*, 49, 483–498.
- Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60, 225–251. (Original work published 1961)
- West, S. G. (Ed.). (1983). Personality and prediction: Nomothetic and idiographic approaches [Special Issue]. *Journal of Personality*, 51(3).
- 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* (pp. 88–162). New York: Guilford Press.
- Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the Big Five. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 737–766). San Diego, CA: Academic Press.
- Zuckerman, M., Bernieri, F., Koestner, R., & Rosenthal, R. (1989). To predict some of the people some of the time: In search of moderators. *Journal of Personality and Social Psychology*, 57, 279–293.
- Zuckerman, M., Koestner, R., DeBoy, T., Garcia, T., Maresca, B. C., & Sartoris, J. M. (1988). To predict some of the people some of the time: A reexamination of the moderator variable approach in personality theory. *Journal of Personality and Social Psychology*, 54, 1006–1019.
- Zuckerman, M., Miyake, K., Koestner, R., Baldwin, C. H., & Osborne, J. W. (1991). Uniqueness as a moderator of self-peer agreement. *Personality and Social Psychology Bulletin*, 17, 385–391.

Received July 22, 1999

Revision received February 18, 2000

Accepted February 18, 2000 ■