

Toward a Structure- and Process-Integrated View of Personality: Traits as Density Distributions of States

William Fleeson
Wake Forest University

Three experience-sampling studies explored the distributions of Big-Five-relevant states (behavior) across 2 to 3 weeks of everyday life. Within-person variability was high, such that the typical individual regularly and routinely manifested nearly all levels of all traits in his or her everyday behavior. Second, individual differences in central tendencies of behavioral distributions were almost perfectly stable. Third, amount of behavioral variability (and skew and kurtosis) were revealed as stable individual differences. Finally, amount of within-person variability in extraversion was shown to reflect individual differences in reactivity to extraversion-relevant situational cues. Thus, decontextualized and noncontingent Big-Five content is highly useful for descriptions of individuals' density distributions as wholes. Simultaneously, contextualized and contingent personality units (e.g., conditional traits, goals) are needed for describing the considerable within-person variation.

The purpose of this article is to identify aspects of individuals' everyday trait-relevant behavior over time that are relevant to and explainable by personality psychology. The presumption is that, because the same individual behaves differently on different occasions, an individual's behavior over time forms a distribution; the central proposal is that the entire distribution and its several components are relevant to—and to be explained by—personality psychology. Specifically, on the basis of a view of individuals as actively reacting to context (Allport, 1937; Brown & Moskowitz, 1998; Cantor, 1990; Diener, Larsen, & Emmons, 1984; Mischel, 1968; Mischel & Peake, 1982; Mischel & Shoda, 1998; Moskowitz, 1982; Nesselroade, 1988, 1991; Revelle, 1995), it is proposed, first, that the average individual routinely and regularly expresses all levels of all traits and that this within-person variability is predictable as individual differences in reactions to situational cues. Second, on the basis of work concerning the predictability of behavior (e.g., Diener & Larsen, 1984; Epstein, 1979; Moskowitz, 1982), it is proposed that although single behaviors are less predictable, the mean of the distribution is among the most predictable variables in psychology. Third, on the basis of theory that variability is itself a stable individual-differences characteristic (Fiske, 1961; Larsen, 1989; Murray, 1938), it is proposed that parameters beyond the mean are also meaningful aspects of personality. That is, it is useful for personality to be conceived of as density distributions as wholes rather than as only one aspect of the distributions. To evaluate these claims empirically, three experience-sampling studies are presented in which many individ-

uals described their current behavior several times per day for 2–3 weeks.

Directly assessing individuals' everyday trait-relevant behavior over time is intended as important for at least three reasons. Most important, it addresses a basic question about the nature of personality: How is trait content manifest in everyday behavior, and do individuals differ in such manifestation? If personality psychology is going to say something about behavior with its constructs, we must know how that behavior is patterned and where individual differences occur in those patterns (Barrett & Pietromonaco, 1997; Borkenau & Liebler, 1995; Borkenau & Ostendorf, 1998; Botwin, 1989; Buss & Craik, 1983; Diener & Larsen, 1984; Emmons, Diener, & Larsen, 1986; Larsen, 1989; Revelle, 1995; Schwartz, Neale, Marco, Shiffman, & Stone, 1999). A second reason this is important is that considerable debate in personality psychology depends on the actual nature of behavioral distributions (e.g., Block, 1995; Cervone, 1991; Pervin, 1994). Specifically, the usefulness of structural or noncontingent approaches to explaining behavior depends on the stability of aggregated means, whereas the usefulness of process or contingent approaches to explaining behavior depends on the amount of within-person variability to be explained. With both Epstein (1994) and Mischel and Shoda (1998) calling for compromise and integration between the structural and the process approaches to personality, it appears the field is prepared for the end of this debate. This article proposes one way to integrate the two factious approaches to personality by demonstrating that trait concepts are inclusive of both impressive levels of within-person stability and impressive levels of within-person variability.

The third reason for assessing everyday trait-relevant behavior is that it may reveal additional aspects of everyday behavior over time as relevant to personality. For example, it has long been theorized that the amount of variability itself is a stable, individual-differences characteristic (Fiske, 1961; Larsen, 1989; Murray, 1938), but it has not been investigated empirically for everyday behavior. The present article operationalizes such concepts as

I thank Anne Fleeson, Jason Fors, Alexandra Freund, Batja Mesquita, Benjamin Peterson, Daniel Green, Elizabeth Laney, and Adriane Malanos for their help with this line of work. I also thank the Max Planck Institute for Human Development, Center for Lifespan Development (Director, Paul B. Baltes) for support during the writing of this article.

Correspondence concerning this article should be addressed to William Fleeson, Department of Psychology, Wake Forest University, Winston-Salem, North Carolina 27109. Electronic mail may be sent to fleesoww@wfu.edu.

parameters of the density distributions and is thus able to empirically investigate their stability and validity as well.

This article focuses in particular on the variability and stability of Big-Five-relevant behavior. Empirically demonstrating such phenomena for the Big Five is important for at least two reasons. On the one hand, it would be trivially easy to choose behaviors that show very little or very much stability and variability (Diener & Larsen, 1984); what is important is to investigate the amount of variability and stability of specifically that content of behavior that the leading structural theory implies should be stable. On the other hand, this investigation is useful for further testing and elaboration of the viability (usefulness) of the Big Five. Specifically, this article tests whether there is sufficient predictability (stability) in everyday manifestations of the Big Five traits to justify their usefulness as descriptions of individuals' behavior.

Individual Differences as Density Distributions of States

Determining what in behavior is to be explained requires measuring an extensive sample of several individuals' everyday behavior, using items that allow (a) quantification of variability on a meaningful scale and (b) direct translation of trait concepts into behavior. The present studies accomplish this by assuming that, just as an individual can be characterized by a level on a trait, representing the degree to which that individual expresses the trait (e.g., a 5 on a 7-point Extraversion dimension represents a moderately extraverted individual), a behavior can be characterized by a level on the same dimension, representing the degree to which that behavior expresses the trait (e.g., a 5 on 7-point Extraversion dimension represents a moderately extraverted behavior). That is, states—short-term, continuous, concrete ways of acting, feeling, and/or thinking (Cattell, Cattell, & Rhymer, 1947; Fridhandler, 1986; Nesselroade, 1988; Patrick & Zuckerman, 1977; Spielberger, Lushene, & McAdoo, 1977)—can be described with the same content and scales as are traits.

Once it is accepted that behavior can be characterized as trait-relevant states, it is easy to see that over time, one individual's behavior forms a distribution. Figure 1 shows two extreme but possible natures of state distributions. The left possibility is one of low within-person variability. The three depicted individuals overlap very little; although each individual occasionally enacts states along the entire dimension, individuals typically act in a narrow and characteristic way. Thus, the central tendency is an adequate description of such distributions, and the remaining variance can be dismissed as error. The right extreme is one of sizeable within-person variation, depicting the same three individuals. In this case, each distribution is wide, such that each individual routinely and regularly manifests all levels of the trait, and, as a result, the primary feature of the figure is overlap (i.e., individuals are highly similar in how they behave). Nonetheless, there are differences between individuals in the locations (central tendency or level) and also in the sizes and shapes of the distributions.

Structural approaches are intended to describe what an individual is like in general and, thus, tend to emphasize the mean alone and dismiss the remainder of the distribution as small or meaningless; process approaches are intended to describe how individuals differentially react to situations and, thus, tend to emphasize the variability within the distribution, dismissing individuals' *unique* distributional parameters or contours as unpredictable (Mc-

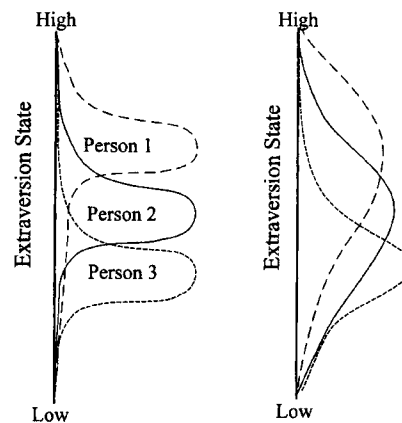


Figure 1. Two extreme but possible distributions of states over time. The two figures differ primarily in the amount of within-person variability in the manifestation of traits. As a consequence, the right three individuals overlap considerably in their manifestations of traits in everyday behavior, and the distribution mean is a less apt description of each individual than for the left possibility.

Adams, 1995; Pervin, 1994). The proposal of this article is that the most adequate characterizations of behavior focus on the entire distribution. That is, personality psychology should conceive of behavior as consisting of density distributions of states.

Five empirical outcomes would support the shift toward a focus on entire density distributions for adequate description of individual differences in behavior. First, the wider the typical individual's distribution, the more of any given individual's behavior left undescribed by his or her mean alone. Second, single states would not be stable (predictable), as predicting an individual's precise momentary location within the distribution would be inefficient. Third, means of the distributions nonetheless would be highly stable, because the distribution is the reliable characteristic of individuals' behavior and the location (e.g., mean) is an important feature of a distribution. Fourth, at least one other parameter of distributions would also be stable. That is, if the distribution is the stable characteristic of individuals, some mean-independent way of describing that distribution must show stability. Fifth, within-person behavioral variability must be meaningful (predictable) and, preferably, differentially so for different individuals. If such variability were only error (capricious or meaningless), it would not be worth describing. In sum, if individual differences in behavior are best described as density distributions, a large amount of behavioral variability will be present within the typical individual, individual differences in distribution parameters will be highly stable, and within-person variability will be meaningful. The remainder of the introduction elaborates each of these three ideas in turn.

How Much Within-Person Variability Is Present in Personality-Relevant Behavior?

Within-person variability describes the same individual acting differently on different occasions, and the first hypothesis of the proposed model is that the amount of within-person variability should be sizeable, near the high end of expectations. In the

present studies, participants described their states several times per day for 2–3 weeks using adjectives rated on 7-point scales, and within-person variability is defined as the standard deviation across the resulting large number of states (i.e., how much the states belonging to one person differ from each other).

The amount of variation within the average individual is first located between the minimum extreme of little variation and the maximum extreme of total variation across the entire sample of states (across and within individuals). Although it is trivially obvious that within-person variation is greater than zero, the closer within-person variation is to total variation, the closer individuals are to perfect and complete overlap in actual manifestation of traits in everyday behavior. Thus, the distance from total variation describes the amount of information gained by knowing which individual is acting.

In addition to being compared against total variation, within-person variation is compared against two other observed variations: between-person variation in traits, and within-person variation in affect. Because the amount of variability between people is sufficient to conclude that different individuals have different personalities, comparing the amount of variability within the average individual against the amount of variability between individuals describes how close within-person variability is to the magnitude sufficient for concluding that one individual's behavior represents different personalities. Within-person variation in affect is a useful comparison standard, because affect is so variable that it is conceived of primarily as a state and only secondarily and recently as a trait (e.g., see the Ekman & Davidson, 1994, volume) and because several studies have quantified the amount of within-person variability in affect (e.g., Cattell, 1973; Eid & Diener, 1999; Eckenrode, 1984; Larsen, 1987; Larson, Csikzentmihalyi, & Graef, 1980; Zelenski & Larsen, 2000). Thus, comparing within-person variability in behavior against within-person variability in affect indicates whether behavior is variable enough to justify it also being studied as a variable state concept. Another advantage to affect is that it is collected with the same scale and the same method as is behavior variability, so that any biases introduced by the method should affect each estimate of variability equivalently and, thus, not be relevant to comparisons between them.

Personality psychologists have long assumed there is within-person behavioral variability (e.g., Allport, 1937; Fiske, 1961; Murray, 1938); the purpose of this article is to find out how much there is. Existing theory and research cannot tell us whether the amount is closer to the left or to the right extreme in Figure 1. However, the larger the degree of within-person variability, the more need there is for research that examines within-person processes of personality; emphasizes flexibility, adaptiveness, and responsivity; and explains behavior as (differential) responses to situations, even in Big Five content. That is, the greater the within-person variability, the greater is the opportunity for integration of two approaches to personality, making quantification of the amount of within-person variability an important goal.

Are Individual Differences in Behavioral Manifestations of Traits Stable (Predictable)?

Individual differences in states are important to the usefulness of the trait concept. That is, one way trait descriptions of individuals may acquire usefulness is by referring to the way the trait is

manifest in those individuals' daily behavior. For example, describing an individual as moderately agreeable acquires usefulness when that individual in fact behaves in a moderately agreeable way. The second, third, and fourth hypotheses of the proposed model concern the existence of such individual differences in stable manifestation of traits.

Because behavioral regularities provide usefulness for the trait concept, it has always been an important task in personality psychology to examine the degree to which behavior shows regularities. The current view is based on empirical work demonstrating distressingly low levels of predictability of one behavior from another. At best, behavior on one occasion predicts behavior on another occasion at around the .30 level (see Mischel, 1968, for a review). However, the current view also proposes highly stable and predictive individual differences in the distributions of states, as long as these differences are conceived of at the distributional level. Epstein (1979) and, later, Moskowitz (1982) and Diener and Larsen (1984) were able to show impressive levels of predictability by examining averaged behavioral aggregates. Rather than predict one behavior from one other behavior, Epstein (1979) successfully predicted the aggregate of several behaviors from the aggregate of several other behaviors. Thus, individual differences in everyday behavior are present and evident. The present studies build on this work by testing whether aggregated means are stable when specifically Big-Five-relevant behavior is assessed and when it is assessed with adjectives. It is predicted that such means will be highly stable and, thus, that noncontingent Big Five descriptions of individuals can be highly useful.

The current view, however, differs from the traditional conception of behavioral distributions as a mean or tendency plus some error. In contrast, I propose that personality psychologists focus on the entire distribution, that it is meaningful as a whole, and that the mean is only one parameter of such distributions. One test of this view is whether at least one distributional parameter other than the mean also shows regularity (Berdie, 1969; Larsen, 1989). The average tendency describes the location of the bulk of the individual's states. For example, a person with a mean of 6 on a 7-point Extraversion scale is not seen as someone who acts highly extraverted but rather as one who acts in the upper half of the Extraversion dimension slightly more frequently than in the lower half. Another distributional property is its size (standard deviation), which describes the diversity of an individual's behavior—how differently he or she acts from moment to moment. Finally, shape (skew and kurtosis) describes the frequency and direction of an individual's extreme behaviors. Although it is a decades-old idea that people have characteristic ranges of variability in behavior (Fiske, 1961; Murray, 1938), it has been studied mainly with mood, affect, or emotion (e.g., Eid & Diener, 1999; Larsen & Diener, 1987; Larson et al., 1980). It is not known whether the amount of everyday behavioral variability within persons is a stable individual-differences characteristic. Stability of a distributional parameter in addition to the mean, combined with the large size of the average distributions, would suggest that the entire distribution become the focus of individual differences.

Is Within-Person Variability Meaningful? The Role of Situational Cues

Paying attention to within-person variability in behavior is worthwhile only if that variability is meaningful (predictable) rather than measurement error (unpredictable). Much within-person variance in behavior is likely to be a response to variability in relevant situational cues (e.g., people are more extraverted in large groups than in small groups): Because situational cues vary in everyday behavior, behavior varies as well. Explaining situational cues' joint influence with personality on behavior is an important mission for personality psychologists. However, it is likely to be particularly interesting to personality psychologists, because individuals are likely to differ in their sensitivity or reactivity to such cues (e.g., some individuals may not increase or may even decrease their level of Extraversion with the size of the group). These differences have been theorized as Person \times Situation interactions (Magnusson & Endler, 1977) and as conditional, if-then traits (Shoda, Mischel, & Wright, 1994; Thorne, 1989). An additional consequence of individual differences in sensitivity to cues may be individual differences in the amount of within-person variability (Brown & Moskowitz, 1998). Generalized responsiveness to cues (e.g., self-monitoring; Snyder, 1974; or affect intensity; Larsen, Diener, & Emmons, 1986) would lead to heightened variability in all traits. However, if different cues are relevant to different traits (the size of the group is less likely to be a cue for Conscientiousness than for Extraversion), being sensitive to the cues for a particular trait would result in the individual acting more variably on that trait, in particular. That is, a person whose actions differ from each other on a trait may be a person who responds strongly to the momentary cues for that trait.

Testing these ideas requires identifying the relevant situational cues for a given trait and then assessing them along with behavior. Doing so for all traits is clearly beyond the scope of this article, but it is possible to do so for Extraversion as an illustration. I use time of day and the number of present others as situational cues, because they both are intuitively powerful for Extraversion and have been shown to be relevant to Extraversion (Blake, 1971; Brown & Moskowitz, 1998; Eysenck, 1981; Larsen, 1985; Larsen & Kasimatis, 1990; less is known about their relevance for other traits, although see Rusting & Larsen, 1998). Four predictions are made in support of the hypothesis that sensitivity to cues and individual differences in sensitivity provide within-person variance with psychological meaning. First, if within-person variability represents meaningful variance in behavior rather than capricious or random responding, then it will vary significantly with time of day and with number of present others, but different traits will respond differently to these situational cues. Second, if personality is at least somewhat expressed in differential sensitivity to cues, individuals will differ in how their trait levels react to the cues. Third, if sensitivity to cues is trait specific, variability will not be strongly correlated across traits. Finally, if individual differences in the amount of variability in a trait represent sensitivity to that trait's cues, then variability of Extraversion will be predictable from reactivity of Extraversion to time and to number of others (the same will not hold true for other traits, unless time and number of others turn out to be relevant for another trait as well).

The following three experience-sampling studies assess the average individual's distribution across time as well as individual

differences in such distributions. The proposed model predicts that (a) density distributions evince sizeable and meaningfully predictive variability, near the high end of possibilities; (b) means of density distributions are highly stable; and (c) additional parameters of density distributions (size and shape) are also stable characteristics of individuals. Beyond adding a useful fifth empirical demonstration that aggregated means are highly stable, the present studies are the first to directly and quantifiably assess the manifestation of trait content in an extended period of the typical individual's everyday behavior, test the stability of size and shape properties of behavioral state distributions, and examine stability and predictability of the specifically Big-Five-relevant properties of behavior. These tasks are important for discovering how, in fact, people differ from each other in manifestations of traits in daily life, for contributing to the integration of process and structure within the very trait concept itself, and for empirically identifying additional individual differences in everyday behavior.

Study 1

Method

Participants. Forty-six students at a small southeastern university participated in the experiment in partial fulfillment of the requirements for an introductory psychology course.

Procedure. Five times per day for 13 days, participants described how they had been acting and feeling during the previous hour. These reports were completed on a regular schedule, every 3 hrs (noon, 3 p.m., 6 p.m., 9 p.m., and midnight), and took about 1–2 min to complete. Reports were completed on Palm Pilots (U.S. Robotics, Los Angeles, CA), hand-held computers about the size of a calculator. Each question was printed on a small screen, and participants responded by pressing a number with a plastic stylus. Every 2 days, participants downloaded their data, and those who missed a download were contacted. The last day of the study, participants also completed a standard Big Five and affect inventory (the Positive and Negative Affect Scale [PANAS]; Watson, Clark, & Tellegen, 1988).

The first report occurred during a 45-min introductory session. The procedure was explained, questions were answered, and participants chose code names to make their anonymity salient. The unique nature of this study, that it investigated a complete picture of 2 weeks of each individual's life, was stressed as well as that it was important that they complete as many reports as honestly as possible. At the end of the introductory session, participants were invited to withdraw for partial credit if they felt the study would be too intrusive.

The response rate was within normal range for such experience-sampling studies. Of the possible 65 reports per participant, the mean was 49.6 reports (76%), and the median was 53.5 reports (82.3%), with a range of 13–63 reports. Participants had been instructed to miss a report if it would be a major inconvenience to complete (e.g., while driving, during an exam, during practice, while sleeping). Participants were also told they could complete a report up to 3 hrs later than the scheduled time but to nonetheless describe the scheduled hour. Reports were also missed because participants forgot or had computer problems. In the interest of maintaining data quality, I excluded completed reports if they did not meet strict criteria. First, reports that contained four or more missing values or at least 85% identical responses were excluded (e.g., reports in which a participant responded "1" to every item—leaving these reports in would artificially inflate within-person variability). Second, an advantage to Palm Pilots is that they surreptitiously record the date and time of completion. Thus, all events completed at least 1 hr earlier or 3 hrs later than the scheduled time were excluded, guaranteeing that all reports were completed close in time to the described behavior. In total, 342 of the 2,281 reports were excluded for one of these reasons (15%).

Materials. The daily reports were the same format as traditional, adjective-based Big Five and affect scales, with the exception that rather than describing themselves in general, participants described their behavior and emotion during the previous 1 hr (e.g., "During the previous hour, how well does 'talkative' describe you?"). Each Big Five and affect factor was represented by four items (Extraversion: talkative, energetic, assertive, adventurous; Agreeableness: cooperative, trustful, rude, warm; Conscientiousness: organized, undependable, hardworking, responsible; Emotional Stability: perturbable, insecure, optimistic, vulnerable; Intellect: intelligent, philosophical, inquisitive, creative; Positive Affect: excited, enthusiastic, proud, alert; Negative Affect: distressed, irritable, nervous, guilty). The Big Five is appropriately assessable with a large variety of adjectives (Goldberg, 1992); for this study, I chose adjectives that (a) loaded on the correct factor in Goldberg (1992), either alone or as part of a bipolar item, or in De Raad, Hendriks, and Hofstee (1994); (b) together represented the breadth of a factor; (c) were easily used to describe behavior; and (d) contained no emotion words, especially for Extraversion and Emotional Stability (to avoid redundancy with the affect scales). In addition, *perturbable* and *inquisitive* were used rather than their opposites for ease of participant understanding. For affect, eight representative items were chosen from the PANAS. Items were presented in Goldberg's (1992) opaque order: The five traits were cycled through one adjective per cycle in the above-listed order, followed by alternating negative and positive affect adjectives. All adjectives were responded to on scales ranging from 1 to 7, with higher numbers meaning that the adjective was more descriptive. Scale scores for the five traits and two affects were computed for each report by taking the mean of the corresponding four items (after reversing). Thus, each participant produced about 45 descriptions of how well his or her behavior in 45 different 1-hr periods was describable by each of the Big Five factors and how he or she was feeling during those same 1-hr periods. Differences across these descriptions represent variability in states, and similarities across the descriptions represent stability in states. Participants also indicated whether 0, 1–3, 4–10, or 10 or more others were present.

After the daily study, participants completed standard Big Five and affect scales, describing what they are like in general. The Big Five scale consisted of the same 20 adjectives (except that *imperturbable* and *uninquisitive* were used in this assessment). Affect was measured with the same eight items.

Reliability was calculated across all included reports and was found to be similar to reliability of Big Five traits and affect scales in previous work: for Extraversion, Cronbach's $\alpha = .72$; for Agreeableness, $\alpha = .66$; for Conscientiousness, $\alpha = .68$; for Emotional Stability, $\alpha = .62$; for Intellect, $\alpha = .69$; for Positive Affect, $\alpha = .80$; and for Negative Affect, $\alpha = .74$.

Results and Discussion

Results are divided into three parts. First, the distribution of states across time within the average individual is described. Specifically, the amount of variability is quantified and then compared with several other variability quantities. Second, individual differences in density distributions are assessed by testing split-half stabilities of several distributional parameters. Third, the role of sensitivity to situational cues in producing within-person variability is tested.

Amount of variability. The first section addresses how much the behavior of one individual changes from hour to hour. I attempt to locate the amount of variability between the one extreme view that there is little within-person variability and the other extreme view that individuals overlap completely in the expression of traits in their daily behavior. I assessed the latter extreme by taking the standard deviation, separately for each trait and affect, across the entire data set of states (thus, it is the total and individual-ignorant variation in states); the left bars in Figure 2 depict these standard deviations.

Within-person variation was assessed by calculating seven standard deviations per participant, one for each trait and affect factor, with each such standard deviation representing the amount the individual varies from hour to hour in how he or she manifests a given trait. The second set of bars in Figure 2 shows the averages across individuals of these standard deviations; thus, they represent the magnitude of the average participant's across-time state variability. It is not surprising that the five trait within-person variabilities are larger in magnitude than zero; however, they are

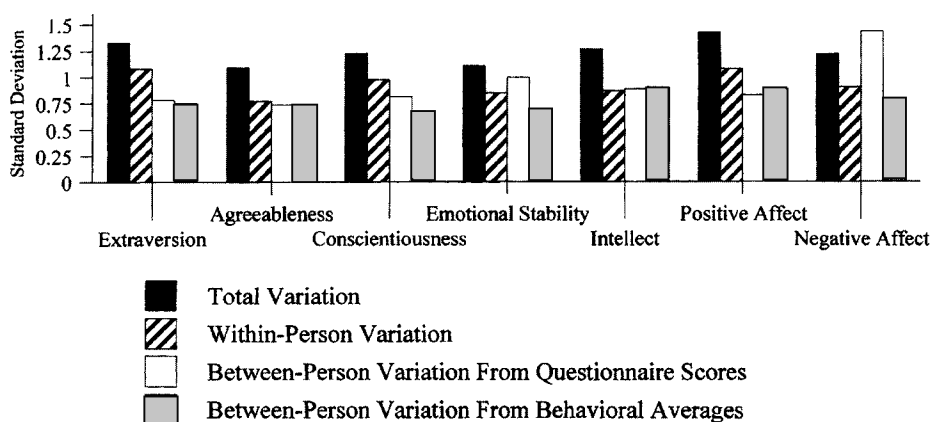


Figure 2. Within-person variation and four comparisons. Within-person variation is the average of 46 within-person standard deviations; thus, it represents how much the average individual's states differ over time. Total variation is the standard deviation calculated across all experience-sampling reports, thus representing how much states differ from each other (ignoring the actor). Between-person variation is the standard deviation across trait levels as assessed in a standard questionnaire or across behavioral mean levels from the experience sampling. As predicted, within-person variation is close to the total, about the same as between-person variation in traits, and nearly as much as within-person variation in affect.

relatively close to the total variability in the sample and, thus, near the high end of possibilities and to complete overlap across individuals.

Two additional ways of evaluating the magnitude of within-person variability are provided in Figure 2. First, trait standard deviations are almost equal to affect standard deviations, and some are larger. A within-subject one-way analysis of variance (ANOVA) revealed that the seven standard deviations differ from each other, $F(6, 270) = 15.16, p < .001$ (any pairwise differences greater than .09 are significantly different from each other at $p < .05$). Most important, Extraversion was more variable than Negative Affect, all other traits were less variable than Positive Affect, and Agreeableness was the only trait less variable than was Negative Affect. Second, Figure 2 also depicts between-person variation in traits, represented by how much individuals differ from each other in their average levels. This study provides two ways to calculate average levels: scores on standard questionnaires, and averages of states across the 2 weeks. The rightmost bars in Figure 2 depict standard deviations across these two measures of average levels. F tests on differences between two variances revealed that within-person variance in Extraversion was greater than between-person variance in Extraversion questionnaire scores, $F(41, 43) = 1.90, p < .05$, and greater than between-person variance in Extraversion behavioral averages, $F(41, 45) = 2.06, p < .05$; that within-person variation in Conscientiousness was greater than between-person variance in Conscientiousness behavioral averages, $F(41, 45) = 2.15, p < .01$; and that within-person variation did not differ from between-person variation for any other trait. In sum, individuals differ from themselves over time at

least as much as they differ from each other at the average level, and more so for the traits of Extraversion and Conscientiousness.

The first purpose of this article is to describe the average individual's distributions of states over a short time period. Figure 3 shows normal distributions calculated with the means and standard deviations observed for the average individual for each trait. The width of these distributions makes clear that the average individual routinely and regularly manifests all levels of Extraversion, Conscientiousness, and Intellect and most levels of Agreeableness and Emotional Stability in his or her daily behavior. In sum, one individual's behavior varies from hour to hour over a 2-week period close to the maximum extreme possible, almost as much as affect varies from hour to hour and at least as much as individuals differ from each other. Such variability is near the high end of expected possibilities.

Individual differences in state distributions. The second purpose of this article is to identify the individual differences in behavior that are available to be described in a nomothetic manner. The basic assumption is that the individual differences in behavior that are available to be described are those that show some regularity over time—thus, those that are correlated across two independent assessments (i.e., split-half reliability). The central claim is that the individual differences that show regularity are the distributions of states, as opposed to single states or to only one parameter of the distributions. If this claim is correct, then (a) single states are not highly correlated across independent assessments, (b) locations of the distributions (means) are highly correlated across assessments, and (c) at least one other distributional parameter is also highly correlated across assessments.

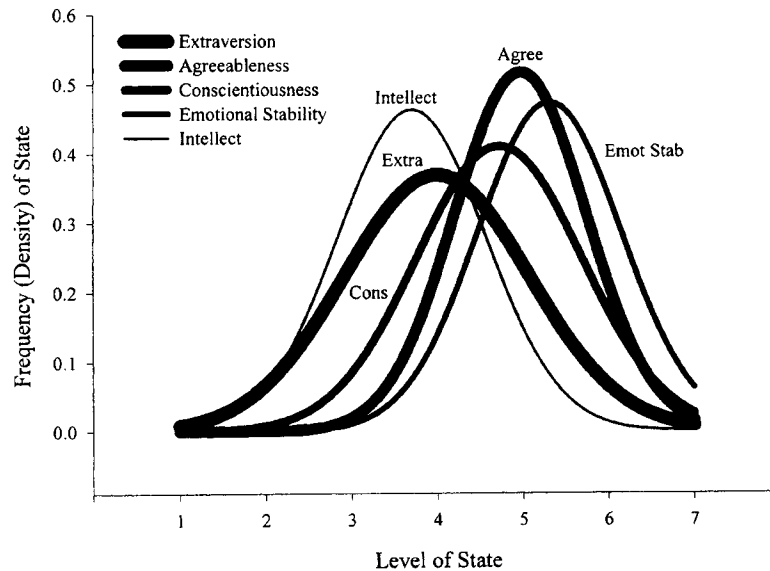


Figure 3. The average individual's distributions of states over 2 weeks: Within-person variability is near the high end of expectations. How the typical individual acts over a 2-week period: Normal distributions calculated with the means and standard deviations observed within the typical individual. The larger standard deviations (for Extraversion, 1.08; for Conscientiousness, 0.97) show that variability cannot be much larger on a 7-point scale while maintaining normality and that individuals must largely overlap with each other in their behavior. Conditional and contextual personality units are needed to describe this variation. Extra = Extraversion; Agree = Agreeableness; Cons = Conscientiousness; Emot Stab = Emotional Stability.

First, two reports were randomly selected for each participant (with the constraints that none were within the first five reports and that the two reports were at least six reports apart from each other) and correlations were computed for each of five traits across those two reports (i.e., from one randomly selected hour to another). This was repeated for 20 trials, and the average of the 20 resulting correlations for each trait are presented in the top line of Table 1 (because participants differed in the number of completed reports, different numbers of participants contributed to each trial; note also that the standard errors are all less than 0.05, suggesting that additional trials would be unlikely to substantially change the results). These correlations are slightly higher than the traditional .30 limit but are low enough to indicate that levels of single states do not exhibit sufficient regularity to be the focus of personality descriptions.

Second, the across-subjects mean and standard deviation for each of the within-subject parameters are shown in Table 1 (skews greater than ± 2.6 and kurtoses greater than ± 2.1 were wind-sorized). The between-subjects distributions were relatively normal, indicating that such parameters can be treated as individual-differences variables (kurtoses were less normal because of outliers).

Finally, regularity at the distribution level was tested, starting with the mean. Each participant's reports were randomly split into two approximately equal halves; means for each half were calculated for each trait, and correlations were calculated between the two halves. The Location Stability line of Table 1 shows the average correlations obtained from 10 such trials (the standard errors are all under 0.02, suggesting that additional trials would be unlikely to substantially change the results). These correlations approach 1.00, showing that the location of an individual's distribution is highly predictable. For example, the .90 correlation for Extraversion means that an individual's average level of Extraversion in half the data was nearly identical to his or her average level

of Extraversion in the other half of the data. (Stepping up the single-state correlations with the Spearman-Brown prophecy formula predicted values similar to these empirically observed values.)

Figure 4 shows two example scatterplots from single trials, one predicting Extraversion on a single occasion from Extraversion on another single occasion, and the second predicting averaged Extraversion in one half of the data from averaged Extraversion in the other half of the data. The important point in these figures is that the correlations are based on the entire range of possible means. That is, individuals differed from each other in the location of their distributions, and such differences were matched precisely in two independent halves of the data.

However, to show that it is the distribution that is stable and not only the average tendency, it is important to show that at least one other parameter of the distributions also shows regularity. Each participant's data were randomly divided into two approximately equal-size halves; standard deviations, skews, and kurtoses were computed for each half; and the parameters were correlated across halves (skews greater than ± 2.6 were recoded as ± 2.6 to avoid outliers dominating the results; extreme kurtoses were recoded to ± 2.1). Ten such trials were repeated, and the averages of the correlations across the 10 trials are shown in Table 1. The size of an individual's distribution was clearly a reliable characteristic of the individual, although not as reliable as was the location. Skew demonstrated some regularity, although not much more than did single states. Kurtosis is of questionable status: Independent assessments were positively correlated, but only slightly. It may be that kurtosis requires more reports for a stable assessment, or it may be that kurtosis is not an important individual difference (the low correlations for kurtosis also point out that aggregation does not artificially create stability—only when the characteristic is a stable aspect of the individual does aggregation produce large correlations). In sum, average tendency is a highly regular char-

Table 1
Density Distributions and Their Parameters as Personality Characteristics

Characteristic	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect
Single state stability	.29	.48	.28	.36	.54
Location (mean)					
<i>M</i>	3.99	4.97	4.72	5.31	3.70
<i>SD</i>	0.75	0.74	0.66	0.68	0.89
Stability	.90	.94	.87	.90	.94
Size (standard deviation)					
<i>M</i>	1.08	0.77	0.97	0.84	0.86
<i>SD</i>	0.24	0.23	0.27	0.28	0.25
Stability	.59	.58	.67	.55	.58
Shape (skew)					
<i>M</i>	-0.16	-0.34	-0.28	-0.89	-0.07
<i>SD</i>	0.57	0.78	0.69	0.76	0.73
Stability	.41	.58	.53	.41	.44
Shape (kurtosis)					
<i>M</i>	-0.19	0.19	-0.11	0.56	0.44
<i>SD</i>	0.90	0.90	0.87	1.06	0.92
Stability	.37	.28	.36	.20	.10

Note. Mean and standard deviation are across-subjects statistics calculated on the within-subjects parameters. Mean describes the typical participant score on the parameter, and standard deviation describes how much participants differ from each other in the parameter. Stability describes correlations computed from one randomly selected hour to another from the same individual or from one randomly sampled half to the other half belonging to the same individual; stability is thus similar to split-half reliability.

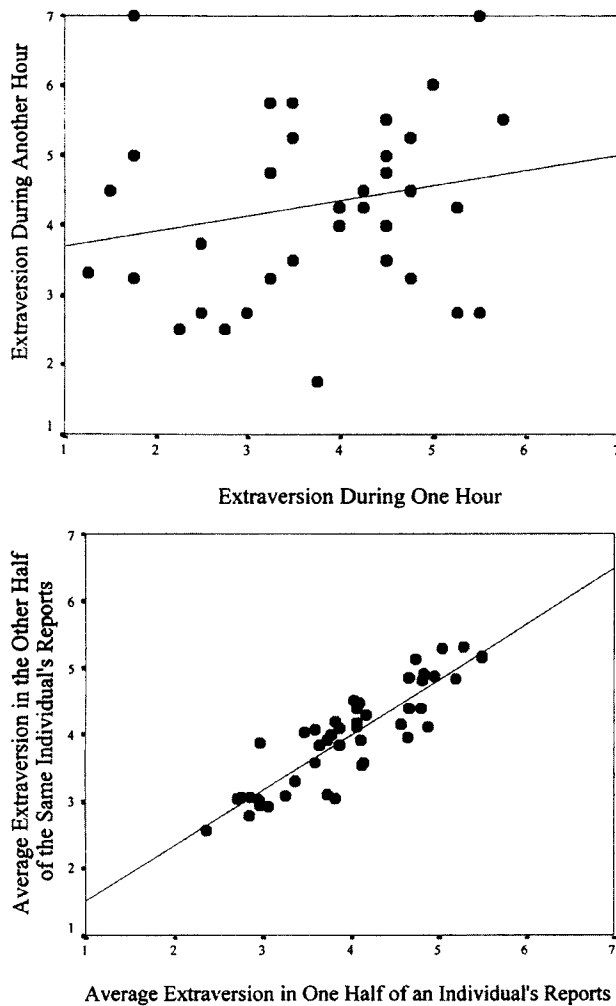


Figure 4. The left panel shows that Extraversion level during one randomly selected hour only slightly predicts Extraversion level of the same individuals during one other randomly selected hour. The right panel shows that average Extraversion calculated over one randomly selected half of the reports strongly predicts average Extraversion of the same individuals calculated over the other half of their data. Furthermore, these correlations are not due to unusual distributions nor to outliers.

acteristic of individuals' behavior, and more than average tendency is required to adequately describe all that is regular about individuals' behavior.

Discriminant validity of parameters. Individuals who were more variable, skewed, or kurtotic on one trait tended to be slightly more variable, skewed, or kurtotic, respectively, on other traits. Across-trait correlations of standard deviations ranged from $-.08$ to $.70$, with an average $r = .38$; across-trait correlations of skewness ranged from $-.12$ to $.50$, average $r = .28$; and across-trait correlations of kurtosis ranged from $.04$ to $.42$, average $r = .20$. This positive manifold in variability of responding, as is true of individual differences in affect variabilities (Eid & Diener, 1999; Larsen & Diener, 1987), likely represents both meaningful person-level tendencies to behave more or less variably for all behavioral

contents (e.g., generalized situational cue sensitivity) as well as less meaningful method variance or individual differences in response styles (e.g., the tendency to use extreme numbers).

Although these across-trait correlations are substantially smaller than even the split-half, same-trait correlations reported above, it is important to establish that behavioral regularities represent more than response styles and that they are trait specific. Thus, the above stability analyses were repeated for 10 more trials, except that the standard deviation of a given trait in one half of the data was predicted simultaneously from the standard deviations of all five traits in the other half of the data. The resulting unique standardized betas, shown in Table 2, describe the stability of standard deviations that is independent of generalized response styles. The pattern of stronger correlations along the diagonals shows that removing generalized response styles had little effect on same-trait stabilities yet eliminated the across-trait correlations. Thus, individual differences in parameters represent more than response styles, and participants responded discriminatively to the different traits. Skew and kurtosis stabilities dropped somewhat more when generalized response styles were removed: Skew same-trait unique betas ranged from $.28$ to $.47$, average $r = .37$, and kurtosis same-trait unique betas ranged from $.05$ to $.33$, average $r = .22$. However, these were higher than across-trait unique betas: skew = $-.18$ to $.29$, average $r = .06$, kurtosis = $-.11$ to $.14$, average $r = .05$.

In sum, although the level of single states was not highly predictable (and less so a characteristic of the individual), individual differences in density distributions were stable and highly predictable. Combined with the typical size of such distributions, such stability suggests that the individual differences in daily manifestations of traits may be best thought of as individual differences in density distributions.

Situational cues and the meaning of within-person variability. The next section of analyses illustrates the role of trait-relevant cues in producing within-person variability. Time of day and number of present others are used as situational cues because they are intuitively powerful for Extraversion. Table 3 shows the results of separate ANOVAs for each trait, with time of day and number

Table 2
Discriminant Validity of Distribution Sizes
(Standard Deviations)

Predictor from one half of data	Dependent variable from other half of data				
	Extra	Agreeable	Cons	Emot Stab	Intellect
Extra	.44	.03	.10	.06	.18
Agreeable	-.02	.49	.04	.02	.02
Cons	.21	.13	.58	-.04	.03
Emot Stab	.02	.04	-.03	.55	-.05
Intellect	.10	.00	.05	-.12	.48

Note. Results of five multiple regressions. In each regression, the standard deviation of all five traits computed over one half of the data predicted the standard deviation for one trait computed over the other half of the data. Table entries are resulting unique predictiveness of each trait's size in one half to each trait's size in the other half (averaged across 10 trials of randomly splitting the data). Larger numbers along the diagonal indicate discriminant validity. Extra = Extraversion; Agreeable = Agreeableness; Cons = Conscientiousness; Emot Stab = Emotional Stability.

Table 3
Within-Person Variation and Responsiveness to Situations

Source of variation	Extra	Agreeable	Cons	Emot Stab	Intellect
Time of day	.01***	.003*	.004*	.00	.01***
No. of present others	.03***	.005***	.00	.00	.00
Individual differences in the effect of time	.07*	.05	.07	.05	.05
Individual differences in the effect of others	.07***	.05***	.07***	.06**	.04**
Between-subjects variation	.19*** (.30***)	.27*** (.47***)	.23*** (.31***)	.23*** (.39***)	.32*** (.50***)
Remaining within-subject variation	.48 (.70)	.40 (.53)	.53 (.69)	.49 (.61)	.39 (.50)

Note. Results of five analyses of variance, one per trait. Table entries are η^2 s squared for the corresponding effect. Individual differences in effects are the Participant \times Effect interactions. Totals are less than 100% because of slight overlap among independent variables. There is no error term against which to evaluate the significance of within-subject variation. Parentheses contain proportions obtained with participants as the only independent variable. Extra = Extraversion; Agreeable = Agreeableness; Cons = Conscientiousness; Emot Stab = Emotional Stability.

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

of others as fixed between-subjects factors and subjects as a random factor (note that the error term for participant and for the interaction are underestimated because of nonindependence of cases). The significant effects of time of day and number of others show that sensitivity to situation cues is, indeed, one source of within-person variation in behavior. Second, the effects of time and of others differed across traits (in whether they affected the trait, how much they affected the trait, and in the pattern of effect). Third, large amounts of variance were accounted for by the interactions between participant, time, and others, meaning that individual differences in sensitivity to cues (i.e., conditional personality units) should be fruitful for explaining within-person variation. Finally, the results echo Figure 2 in showing that within-person variance remained high even when variance due to time of day and others being present was removed. Together, these results provide validity evidence for within-person variation in behavior: Participants reported more extraverted behavior as time of day and number of others increased, suggesting that such ratings truly reflected extraverted behavior.

If within-person variability in behavior is partially due to sensitivity to trait-relevant cues, then individual differences in within-person variability may be due to individual differences in sensitivity to trait-relevant cues. For each participant separately, a regression predicted state Extraversion from time of day. The resulting betas indicated how much each participant varied his or her Extraversion with time of day. As predicted, these betas correlated with standard deviations of Extraversion, $r = .40$, $p < .01$, meaning that the more an individual's Extraversion varied with time of day, the more his or her Extraversion varied in total. However, time of day is not a relevant cue for all traits, so a similar result did not obtain for all traits. For each trait, a beta was obtained for each participant indicating how much that participant increased the given trait with time of day. Corresponding betas did not predict overall variability for Agreeableness, $r = .14$, $p = .37$, nor for Conscientiousness, $r = .05$, $p = .73$, but did so negatively

for Emotional Stability, $r = -.37$, $p < .05$, and positively for Intellect, $r = .32$, $p < .05$. The meaning of the negative relationship for Emotional Stability is unclear, but it clearly indicates that the positive correlation for Extraversion is not a statistical artifact nor a logical necessity. The positive relationship for Intellect suggests that time of day may be a relevant cue for Intellect as well as for Extraversion (however, see Study 2 results). Similar betas were computed for the responsiveness of Extraversion to number of present others and, as predicted, were correlated with individual differences in Extraversion standard deviations, $r = .38$, $p < .01$. Thus, the more individuals increased their Extraversion with the number of present others, the more variable they were in Extraversion in general. Corresponding betas were also correlated with standard deviations for Agreeableness, $r = .29$, $p < .05$, and Intellect, $r = .37$, $p < .05$, but not for Emotional Stability, $r = -.13$, nor for Conscientiousness, $r = .23$. It is plausible that the number of present others is relevant to Agreeableness and to Intellect as well as to Extraversion.

Study 2: Standardizing the Definitions

Study 1 showed that within-person state distributions are characterized by a sizeable amount of variability yet the parameters of the distributions are highly stable. One advantage to using Big Five adjectives is that even such trait concepts are shown to (a) need explanations of within-person variability, and (b) not be threatened in their usefulness by the existence of within-person variability, because of the equally large degree of distributional stability. However, one possible criticism of Study 1 is that the large degree of stability was an artifact of idiosyncratic adjective definitions (Dunning & McElwee, 1995; Goldberg & Kilkowski, 1985). That is, two different definitions of the same adjective may result in 2 participants rating identical behaviors at different points on the scale. Over several occasions, this would result in two different average ratings on that adjective for the 2 individuals, which would in turn result in apparent stability of those average ratings. Thus, stable means may have reflected not stability in state means but rather stability in definition of the adjective. In Study 2, participants memorized standard definitions for every adjective. Thus, differences in mean levels are less likely to reflect differences in definitions and more likely to reflect differences in states (Goldberg & Kilkowski, 1985).

A second purpose to Study 2 is to provide more reliable estimates of the especially labile parameters of shape. Thus, participants completed 3 weeks of reports rather than the 2 weeks in Study 1. If somewhat lower stabilities of individual differences in skew and kurtosis were due to limited numbers of reports, Study 2 should show stronger stabilities of these parameters.

Study 2 also used different and more adjectives. First, this addresses whether the results of Study 1 were due to unusual adjectives (Saucier, 1997). Second, this provides a more reliable indicator of each of the Big Five traits. Finally, Study 2 allowed a replication of the time-of-day analyses from Study 1.

Method

Participants. Twenty-nine students at a small southeastern university participated in the experiment in partial fulfillment of the requirements for an introductory psychology course. One participant was dropped for com-

pleting only eight reports, and another participant completed no valid reports.

Procedure. Five times per day for 20 or 22 days, participants described how they were acting during the previous 3 hrs. Ten of the participants used Palm Pilots as in Study 1; 18 completed the study on paper. The latter participants carried pads of paper on which were printed the 25 adjectives in the same order as appeared on the Palm Pilots. The response rate was again within normal range for such daily studies. Of the possible 110 reports per participant, the mean was 71.4 reports (65%), and the median was 75.5 reports (69%), with a range of 25–93. Reports that contained more than two missing values, had at least 80% identical responses, were completed at least 2 hrs earlier than scheduled, or were completed on the day after scheduled were excluded. In total, 227 of the 2,000 reports (11.4%) were excluded for one of these reasons.

In an extension of the introductory session, participants memorized definitions for each of the 25 adjectives. Participants studied the definitions and then wrote them verbatim from memory. Any definitions that were not precisely correct were studied and tested again, for 30 min or until the participants correctly remembered all 25 definitions (7 of the 29 participants did so before 30 min passed).

Materials. Materials were identical to Study 1, with three exceptions. First, no affect measures were included, because of the increased number of trait adjectives. Second, participants described their behavior during the previous 3 hrs rather than the previous 1 hr. Third, rather than four items per factor, traits were assessed with a new set of five items per factor (Extraversion: talkative, assertive, shy, bold, energetic; Agreeableness: rude, cooperative, warm, trustful, kind; Conscientiousness: organized, careful, steady, conscientious, undependable; Emotional Stability: relaxed, imperturbable, irritable, nervous, insecure; Intellect: uninquisitive, bright, artistic, unreflective, imaginative). I chose items that (a) loaded on the correct factor in Goldberg (1992), (b) together represented the breadth of the factor, (c) were easily used to describe behavior, and (d) were amenable to the types of definitions described below.

Adjective definitions were modified from dictionary definitions (*Random House College Dictionary*, 1975; *Webster's Ninth New Collegiate Dictionary*, 1986) to meet five criteria: (a) they were easy to understand, (b) they were concrete, (c) they did not deviate from the meaning of the

adjective's Big Five factor, (d) they did not use the same words as another item or definition, and (e) they were short. For example, the definitions of trustful and insecure were "willing to believe or rely on others" and "uncertain about your abilities or worth," respectively. The important aspect of the definitions is not their exact content but rather that all participants were using the same ones.

Reliability was calculated across all included reports and, except for Intellect, was found to be reasonably high: for Extraversion, Cronbach's $\alpha = .75$; for Agreeableness, $\alpha = .78$; for Conscientiousness, $\alpha = .74$; for Emotional Stability, $\alpha = .75$; and for Intellect, $\alpha = .51$.

Between-subjects assessment. A different group of 394 introductory psychology students completed a standard assessment of the Big Five, using the same items but indicating how much each item described them in general. These data were used only to assess the amount of variability between individuals in the Big Five.

Results and Discussion

Amount of variability. In this case, all between-person variances calculated from questionnaires were larger than were the within-person variances; $F(393, 64)$ tests on differences between two variances were significant, $p < .05$, for all but Emotional Stability, $p > .20$. However, within-person variance did not differ from between-person variance calculated from behavioral means, all $ps > .20$ —Emotional Stability within-person variance was marginally larger than was between-person variance from behavioral means, $F(64, 26) = 1.96$, $p = .06$. Thus, within-person variation was again close to the total possible and to the amount between individuals.

Stability of parameters. The central question in Study 2 is whether large amounts of stability are evident even when definitions are standardized across participants, and I assess this with the same method of correlating parameters across randomly split halves. Table 4 shows the average correlations across 10 trials. The results were similar to those of Study 1,

Table 4
Density Distributions and Their Parameters as Personality Characteristics, Using Standardized Definitions

Parameter	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect
Location (mean)					
<i>M</i>	4.40	5.01	4.71	4.83	4.36
<i>SD</i>	0.71	0.67	0.71	0.61	0.60
Stability	.95	.97	.97	.93	.96
Size (standard deviation)					
<i>M</i>	0.73	0.60	0.62	0.84	0.62
<i>SD</i>	0.20	0.21	0.20	0.28	0.16
Stability	.80	.78	.80	.85	.72
Shape (skew)					
<i>M</i>	-0.07	-0.52	-0.31	-0.56	-0.13
<i>SD</i>	0.46	0.59	0.55	0.45	0.45
Stability	.41	.50	.52	.44	.31
Shape (kurtosis)					
<i>M</i>	0.01	0.54	0.16	0.00	0.21
<i>SD</i>	0.73	0.88	0.78	0.95	0.69
Stability	.05	.15	.28	.45	.17

Note. Mean and standard deviation are across-subjects statistics calculated on the within-subjects parameters. Mean describes the typical participant's score on the parameter, and standard deviation describes how much participants differ from each other in the parameter. Stability describes correlations computed from one randomly selected hour to another from the same individual or from one randomly sampled half to the other half belonging to the same individual; thus, stability is similar to split-half reliability.

with individual differences in location correlating close to 1.00, individual differences in size meeting standard reliability criteria, and stability of windsorized skew lower but positive. Windsorized kurtosis showed very weak stability, suggesting that uniform definitions may have prevented some of the extreme responses. Furthermore, parameters were distributed across participants relatively normally, showing a reasonable spread of individual differences in the parameters.

Discriminant validity of standard deviations was improved over that of Study 1. As in Study 1, stability analyses were repeated predicting parameters for one half of the data from parameters for the other half of the data, except using all five traits' parameters as independent variables. Table 5 shows the unique standardized betas resulting from 10 such trials. The higher diagonal correlations indicate that removing any potential response bias did not appreciably reduce stabilities but did eliminate across-trait correlations. However, skew and, especially, kurtosis showed less discriminant validity, with same-trait unique betas averaging .29 for skew and .20 for kurtosis. In sum, location and size were highly predictable and stable even when adjective definitions were standardized across participants.

Explaining within-person variability in Extraversion from time of day. For each trait, a 5 (time of day) × 27 (participants) ANOVA revealed the results depicted in Table 6. As in Study 1, behavior varied systematically with time of day, Extraversion was the trait most strongly related to time of day in a consensual manner, and the majority of the effect was manifest in systematic individual differences in reactions to time of day. Thus, within-person variability in Extraversion likely reflects variability in extraverted behavior, and explaining such variability is a potentially fruitful area for future personality research.

Analyses to illustrate a process approach to convergent validity also replicated Study 1. Specifically, the extent to which individuals' Extraversion responded linearly to time of day predicted how variable they were in Extraversion, $r(27) = .34$, $p < .05$, one-tailed. Conscientiousness also showed a significant relationship, $r = .38$, $p = .05$, but the other traits did not:

Table 5
Discriminant Validity of Distribution Sizes (Standard Deviations): Standardized Definitions

Predictor from one half of data	Dependent variable from other half of data				
	Extra	Agreeable	Cons	Emot Stab	Intellect
Extra	.68	.12	-.06	.08	-.01
Agreeable	.05	.69	.12	.06	.08
Cons	.01	.02	.70	.01	.08
Emot Stab	.11	.02	.02	.68	.11
Intellect	.07	.04	.08	.14	.61

Note. Results of five multiple regressions. In each regression, the standard deviation of all five traits computed over one half of the data predicted the standard deviation for one trait computed over the other half of the data. Table entries are resulting unique predictiveness of each trait's size in one half to each trait's size in the other half (averaged across 10 trials of randomly splitting the data). Larger numbers along the diagonal indicate discriminant validity. Extra = Extraversion; Agreeable = Agreeableness; Cons = Conscientiousness; Emot Stab = Emotional Stability.

Table 6
Proportion of Total Variation Due to Time of Day and Participant

Source of variation	Emot Stab				
	Extra	Agreeable	Cons	Emot Stab	Intellect
Between-subjects variation	.45*** (.47***)	.52*** (.55***)	.53*** (.56***)	.30*** (.31***)	.40*** (.44***)
Time of day	.02***	.003*	.00	.01***	.00
Individual differences in the effect of time	.04*	.03*	.04**	.06***	.04*
Remaining within-subject variance	.47 (.53)	.41 (.45)	.40 (.44)	.62 (.69)	.51 (.56)

Note. Results of five analyses of variance, one per trait. Table entries are etas squared for the corresponding effect. Individual differences in the effect of time are the Participant × Time interactions. Totals are less than 100% because of slight overlap among independent variables. There is no error term against which to evaluate significance of within-subject variance. Parentheses contain proportions obtained with participants as the only independent variable. Extra = Extraversion; Agreeable = Agreeableness; Cons = Conscientiousness; Emot Stab = Emotional Stability. * $p < .05$. ** $p < .01$. *** $p < .001$.

Agreeableness, $r = .24$, $p > .20$; Emotional Stability, $r = .36$, $p = .06$; Intellect, $r = .03$, $p > .80$. Extraversion's results from Study 1 were replicated, but no other trait showed similar results across the two studies, supporting the theory that time of day is a relevant situational feature for Extraversion but is less so for the other traits and that individual differences in Extraversion sensitivity to time of day predict individual differences in Extraversion variability.

Study 3: Standardizing the Scale

Study 2 shows that the large degrees of distributional stability are not an artifact of adjective definitions. Study 2 also shows, by using more and a different set of adjectives, that neither the impressive stability nor the impressive variability described in Study 1 was due to an unusual item selection. Finally, Study 2 replicates the convergent validity analyses concerning Extraversion standard deviations: More diversity in Extraversion behavior is partly a function of responding more strongly to Extraversion-relevant situational cues.

Study 3 attempts to address an artifactual explanation for the large degree of within-person variability. Specifically, the magnitude of variability may have been an artifact of idiosyncratic scale usage. The scales in Studies 1 and 2, as is recommended for Big Five scales (Goldberg, 1992), anchored the highest value, 7, with an adjective and did not anchor the lowest value, 1. Thus, participants may have interpreted the lower anchor so as to produce more variability than is actually present. For example when 7 was anchored with *talkative*, participants may have interpreted 1 to mean their own personal minimum level of talkativeness. At least two results in Study 1 argue against this possibility. First, traits differed significantly in the amount of within-person variation; it is difficult to explain why such a bias would differ across traits. Second, participants differed from each other reliably and discriminatively in their mean levels and in their degrees of variability,

indicating that they adjusted their use of rating scales to the specific content of the items.

Nonetheless, the purpose of Study 3 is to more directly address possible idiosyncratic scale definitions by using bipolar adjectives, with each pole anchored by the antonym of the opposite pole (e.g., anchoring 1 with *silent* when 7 is *talkative*). Such bipolar scales require participants to locate their responses by reference to two opposite behaviors, and, thus, participants are less free to adjust scales to produce more variability. In addition, Study 3 used still another set of adjectives, to further test the breadth of items for which the findings hold.

Method

Participants. Thirty students at a small southeastern university participated in the experiment in partial fulfillment of the requirements for an introductory psychology course.

Procedure. Five times per day for 21 days, participants described how they were acting and feeling during the previous 3 hrs. Twelve of the participants used Palm Pilots, as in Study 1; 18 completed the study on paper, as in Study 2. Additionally, prior to the experience-sampling portion of the study, participants completed a standard assessment of the Big Five, using the same adjectives but describing the way they are in general.

The response rate was again within normal range for such daily studies. Of the possible 105 reports per participant, the mean was 73.5 reports (70%), and the median was 76 reports (72.4%), with a range of 40–112 reports (1 participant continued longer than necessary). Again, strict criteria were used to cull reports. Reports that contained more than four missing values, had at least 80% identical responses, were completed at least 0.5 hr earlier than scheduled, or were completed on the day after scheduled were excluded. In total, 210 of the 2,204 reports (9.5%) were excluded for one of these reasons.

Materials. Materials were identical to Study 2, with three exceptions. First, each item was bipolar, with the poles antonyms of each other. Second, participants described how they acted since the last time they completed a report. Third, traits were assessed with the following six items per factor (Extraversion: introverted–extraverted, energetic–unenergetic, bold–timid, talkative–silent, unenthusiastic–enthusiastic, unassertive–as-

sertive; Agreeableness: cooperative–uncooperative, rude–polite, warm–cold, stingy–generous, kind–unkind, trustful–distrustful; Conscientiousness: irresponsible–responsible, disorganized–organized, careless–thorough, negligent–conscientious, undependable–reliable, hardworking–lazy; Emotional Stability: guilt free–guilt ridden, secure–insecure, nervous–at ease, relaxed–tense, unemotional–emotional, envious–not envious; Intellect: unintelligent–intelligent, unimaginative–imaginative, reflective–unreflective, curious–uninquisitive, sophisticated–unsophisticated, uncreative–creative). I chose items that (a) loaded on the correct factor in Goldberg (1992), (b) together represented the breadth of the factor, and (c) were easily used to describe behavior. All adjectives were responded to on scales ranging from 1 to 7, with 1 meaning that the leftmost adjective was more descriptive and 7 meaning that the rightmost adjective was more descriptive. Reliability was calculated across all included reports and, as expected given the greater number of items per trait, was found to be reasonably high: for Extraversion, Cronbach's $\alpha = .78$; for Agreeableness, $\alpha = .83$; for Conscientiousness, $\alpha = .86$; for Emotional Stability, $\alpha = .76$; and for Intellect, $\alpha = .68$.

Results and Discussion

Amount of variability. Five standard deviations were calculated for each participant (each representing the amount of within-person variability for one trait), and the middle bars in Figure 5 show the averages of these standard deviations. The left bars show total variation in the sample, and the right bars show between-person variation in traits calculated from the standard questionnaire and from means of each participant's behavior. As can be seen, within-person variability was again close to the maximum possible and to the amount that individuals differ from each other. Although within-person variance was slightly smaller than in Study 1, so were total variance and between-person variance, suggesting that this was a function of the particular adjectives used. *F* tests on differences between two variances revealed that within-person variation in Extraversion was greater than between-person variation of Extraversion behavior means, $F(65, 29) = 2.61$, $p < .01$, but that within-person variance did not differ from between-person variance for any other trait. In sum, within-person variability

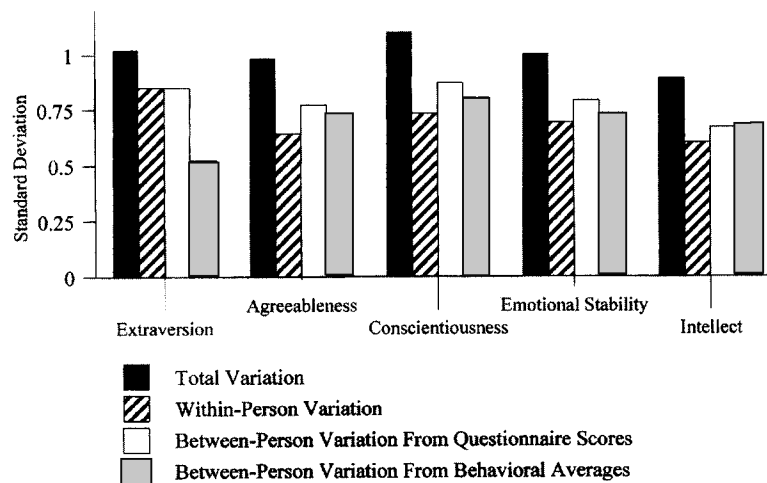


Figure 5. Within-person variation and three comparisons, using bipolar scales. Study 3 shows that sizeable within-person variation is not due to “stretching” of unipolar scales. Even when both ends of the scale are anchored by opposite extremes (e.g., 1 = *silent*, 7 = *talkative*), within-person variation is close to the total variation and nearly equal to between-persons variation.

was near the high end of expectations, even when scales were anchored on both ends to prevent artificially variable responses.

Stability of individual differences in distributional parameters. Parameter stability analyses also replicated Studies 1 and 2. Average correlations obtained from 10 trials, correlating each individual's average states across two randomly selected halves of the data, again approached 1.00, showing that the location of the distribution was highly predictable and stable: Extraversion, $r = .91$; Agreeableness, $r = .98$; Conscientiousness, $r = .97$; Emotional Stability, $r = .97$; Intellect, $r = .97$. Stabilities for the standard deviations approached those for the location of the distributions: Extraversion, $r = .85$; Agreeableness, $r = .86$; Conscientiousness, $r = .90$; Emotional Stability, $r = .83$; Intellect, $r = .87$. Skew and kurtosis stabilities improved: average skew stability, $r = .49$; average kurtosis stability, $r = .32$.

Discriminant validity. Again, standard deviations showed strong discriminant validity between traits. Same-trait unique predictiveness of standard deviations ranged from .56 to .83, whereas across-trait unique predictiveness ranged from $-.11$ to .18. Controlling for all other traits' skews, average skew stability remained at .44, and controlling for all other traits' kurtoses, average kurtosis stability remained at .31.

General Discussion

This article is concerned with how Big Five content (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect) is manifest in individuals' everyday behavior. The findings support the general proposal that the most adequate descriptions of behavior focus on the entire distributions of behavior. That is, the amount of variability is too large and meaningful to be ignored, average tendencies are highly stable descriptions of individuals, and the size and shape of the distributions are also reliable individual differences characteristics. Such distributions, as depicted, for example, in Figure 3, lay out a map of the terrain to be explored and explained by personality psychologists. Specifically, personality psychologists need to explain both the differences between individuals in the parameters of the distributions and the causes of the considerable variability within the distributions.

One reason that directly describing the behavior patterns is important is that the structural or trait approach and the process approach have continued to be at odds. The structural approach emphasizes broad tendencies that are manifest in stable and situation-independent behavioral averages. The process approach emphasizes laws relating situational conditions to individuals' behavioral reactions. The present studies are the first to support both approaches simultaneously, in the same data involving everyday behavior. The findings show that not only is the room for each approach large but the two approaches do not conflict with each other: The structural approach can focus on correlates of highly reliable means, and the process approach can explain the plentiful deviations from these means. The theoretical proposition that allowed this integration is that individuals' behavior is best conceived of as density distributions.

Implications for Decontextualized and Noncontingent Trait Descriptions of Individuals

In a structural approach, emphasis is on what a person is like in general; that is, on the decontextualized and noncontingent trait

unit (e.g., "he or she is moderately agreeable"). Such a personality unit (a) is intended to describe the content of an individual's behavior (e.g., "he or she behaves moderately agreeably"; Fleeson, Zirkel, & Smith, 1995), and (b) explicitly and intentionally does not refer to the conditions or antecedents of the individual's behavior but rather describes the individual per se. One question of this article is whether such trait concepts are viable—that is, whether it is possible to describe individuals' behavior in general without having to state when or under what conditions the individual behaves that way. The Big Five content of behavior was chosen because it is likely the strongest contender for an affirmative answer.

The results show that noncontingent descriptions of individuals can be highly accurate and useful descriptions of individuals' behavioral distributions. The mean of the distribution was nearly as stable as is possible ($r_s > .90$) and describes the area of the dimension in which a person most frequently acts. For example, an individual with a mean of 5 and a standard deviation of 1 might act in the 3 range about 11% of the time, in the 4 range about 28% of the time, in the 5 range about 43% of the time, and in the 6 range about 11% of the time. The findings also reveal new parameters of the distributions as stable and, thus, as characteristics of individuals. The standard deviation was nearly as stable as was the mean, and the skew and kurtosis of the distributions showed some but not impressive stability. Size and shape, like the mean of a person's behavioral distribution, have inherent interest and face validity as descriptions of individuals, descriptions that originate in behavior itself. The standard deviation describes how differently the person acts from moment to moment. For example, a standard deviation of 1.5 would extend the above frequencies out toward the extremes, with such an individual acting in the 3 range about 17% of the time, in the 4 range about 17% of the time, in the 5 range about 24% of the time, and in the 6 range about 18% of the time, making this individual even more different from moment to moment.

Beyond the inherent interest and face validity of individual differences in variability of behavior, findings also demonstrate the psychological meaning of variability. First, the results provide validity evidence that individual differences in standard deviations represented genuine individual differences in how variable individuals are on the traits rather than, for example, some kind of response bias. Specifically, behavior ratings varied with situational cues as they should, and discriminant validity results demonstrated that the same individual differed across traits in amount of variability, meaning that individuals distinguished between item contents in indicating the variability in their behavior. Second, variability was shown to reflect individual differences in reactivity to trait-relevant cues: The more reactive an individual was to the appearance of relevant cues, the more variable he or she was. Specifically, individual differences in Extraversion variability were predictable from individual differences in Extraversion sensitivity to time of day and to the number of others present. It is important for this latter result that other traits' sensitivity to these cues did not predict variability for those traits; this means that the result was not a trivial result of the method but was specifically predicated on the relevance of the cue to the particular trait.

It is possible that other factors may also influence how variable an individual is on a given trait, such as the dynamics of goal pursuit (Cantor & Kihlstrom, 1987), cyclic and inertial behavior patterns (Brown & Moskowitz, 1998; Fleeson, 2001; Larsen, 1987;

Larsen & Kasimatis, 1990), adaptive strategy learning (Siegler & Shipley, 1995), biological or hormonal rhythms (Haus, Lakatua, Swoyer, & Sackett-Lundeen, 1983), and the beginning stages of long-term change (Nesselroade, 1988). Future research is needed to identify cues for other traits and to investigate other potential influences on individual differences in within-person variability.

The results also show limitations of noncontingent descriptions. The severity of these limitations depends primarily on where, between the two extremes depicted in Figure 1, the width of the typical individual's distribution was observed to be. First, because noncontingent personality descriptions are constant within a person, they cannot describe variation within a person. Thus, the observed sizeable within-person variability implies that noncontingent personality descriptions fail to describe a large portion of behavior. Second, the large within-person variability suggests that any one particular parameter of a distribution is misleading when it is offered as a relatively complete description of an individual rather than as only one parameter of a distribution. Furthermore, each aspect of the distribution can be represented by several statistics. For example, the mean is only one possible summary of location, and location may be better summarized in some cases by the interquartile range, the median, or some other central tendency indicator. That is, the mean does not have special status as the primary feature of the distribution. At the same time, these results do not deny the usefulness of the mean as a fast and frugal summary of an individual's state distribution; rather, these results point out that the mean is only a fast and frugal summary.

It is important to acknowledge that not all noncontingent trait concepts are intended to describe how an individual behaves (Saucier & Goldberg, 1996) and that scores on standard trait assessments may reflect some other trait concept. Thus, this article is not about what standard trait assessment scores mean; rather, it is about whether and how individuals' behavior can be meaningfully described with Big Five trait concepts, and, therefore, I began by describing the typical individual's distribution of trait contents across several behaviors. Individual differences in such distributions are then offered as potential descriptives for trait concepts that do intend to refer to behavior.

Amount of Within-Person Variance and Implications for Contextual and Contingent Personality Units

A second primary implication of the findings is that the large degree of within-person variance (a) demonstrates that the typical individual is highly diverse, flexible, and responsive in his or her behavior and (b) calls for more research incorporating within-person variance into models of traits and of personality. Others have made similar points (Brown & Moskowitz, 1998; Cervone, 1999; Larsen, 1989; Mischel & Peake, 1982; Mischel & Shoda, 1998; Nesselroade, 1988; Thome, 1995). What is unique about the current demonstration is that it (a) directly quantifies the amount of within-person variability, (b) shows within-person variability even in the Big Five relevant properties of behavior, and (c) shows that within-person variability coexists comfortably with within-person stability. This result is within the range of expected possibilities (i.e., some personality psychologists would have expected that trait-relevant behavior is as variable within person as is affect and at least as variable as personality is between persons). However, it is at the high end of expectations. Many psychologists would have

expected within-person standard deviations to be nearer the low end of possibilities, closer to the .25 to .40 range. The fact that such a wide range of results is equally expectable, with important differential consequences for the usefulness of various personality units, for the importance of density distributions, for the role of situational cues, and for the integration of personality psychology, makes it essential that it be known where the actual value is. It is much easier for personality psychology to say something about behavior with our constructs when we know how those constructs are manifest in everyday behavior.

Where does all of this within-person variance come from? First, the results suggest that it is not solely measurement error: The reliabilities of the states found in all three studies were reasonably high, states varied predictably with time of day and with number of present others, and the amount of within-person variance showed stability. Rather, much of the source of within-person variance is likely to be variation in the situation, and the large amount of within-person variance shows how flexible individuals are in responding to situational presses, even as they maintain their distributional properties over time. Finally, much of the within-person variance was found to be due to person-situation interactions—that is, to individual differences in how they reacted behaviorally to the same situational cues. How much variation is due to situation and how much to Situation \times Person interactions will differ by situation and by behavior. Regardless, such variation is not a threat to the usefulness of the trait concept. Thus, the person-situation debate would be better focused on the causes of specific behaviors rather than on the causes of behavior in general.

These findings support the usefulness of a process approach and of conditional personality units. First, principles relating situational cues to behavior reactions are necessary for explaining the variability within people. Second, contextual or contingent personality units such as conditional traits (Cervone, 1999; Murtha, Kanfer, & Ackerman, 1996; Shoda et al., 1994; Thome, 1989, 1995), goals (Fleeson & Cantor, 1995), or expectancies (Mischel & Shoda, 1998) are needed to explain the Person \times Situation interactions. An additional implication of these findings, therefore, is the availability of a promising new method for investigating such Person \times Situation interactions (Endler & Parker, 1992).

Using Adjectives to Characterize Trait-Relevant Behavior (States)

The present method deviates from traditional assessments of behavior by using adjective ratings rather than categorizations. Typically, researchers assess behavior by designating several categories of behavior and then obtaining reports of whether or not a given behavior falls into that category (Borkenau & Ostendorf, 1987; Buss & Craik, 1983; Gosling, John, Craik, & Robins, 1998; Moskowitz, 1982). For example, "introduced oneself to strangers" would be a category of behaviors, and an individual's score would be the number of times that individual introduced him- or herself to strangers. The present studies use adjective ratings of behaviors rather than categorizations, because that is the most direct method of assessing Big Five content in behavior and because it allows quantification of the variability of trait manifestations within one individual's behavior as well as comparison with other variabilities. That is, if one is willing to assume that behaviors can vary along the same dimensions as do traits (that behaviors express the

same content as do traits) and that such variation is a meaningful aspect of the behavior, then states are the manifestations of trait content in daily behavior (e.g., being talkative is the manifestation of the trait of talkativeness). Then, diversity of trait content manifestations within one individual is conceivable, as the variability of states across time.

However, there is a trade-off between knowledge of the specific actions performed and knowledge of the trait-relevant meaning of those actions, and adjectives sacrifice specificity for meaning. Adjective ratings sacrifice specificity in that a given rating on an adjective—for instance, a 5 on assertiveness—could correspond to any number of actions. What adjective ratings gain is clarity in the trait-relevant meaning of the behavior. For example, whether introducing oneself to strangers is gregarious (at a party), conscientious (at a conference), or assertive (when the strangers are not interested) depends on various factors. Adjective ratings directly assess the degree to which the behavior expresses a given trait and thereby captures its meaning. For the present purposes—assessing trait manifestation in daily behavior—the trait-content meaning of an action is of central concern.

A related issue is that the adjectives were self-reported (of course, both adjectives and categorizations can be assessed with self-report or with observer report). Self-report potentially suffers from inaccuracy and social desirability (Borkenau & Ostendorf, 1987; Gosling et al., 1998). For example, one might find a way to interpret an irresponsible behavior as a responsible one. However, in these studies, variability was large on all scales, both between and within individuals, meaning that participants were often willing to admit being, for example, irresponsible, rude, or lazy. In fact, the main consequence of socially desirable responding would have been a reduction in the amount of within-person variability, working against one of the primary hypotheses of this study. For example, both Gosling et al. (1998) and Borkenau and Ostendorf (1987) found that Agreeableness-relevant self-reports were more prone to social desirability than were Extraversion-relevant self-reports, and in the present studies Agreeableness evinced less within-person variation than did Extraversion. However, those two studies may not apply to the present research, as both evaluated the accuracy in self-reporting (a) between-person differences in (b) occurrence of specific actions, and the present studies were concerned with across-time differences in adjective ratings of states. Regardless, the current results demonstrated convergent and discriminant validity to such reports, lowering the likelihood that self-report biases were important. Future research on the accuracy of multiple on-line adjectival self-reports clearly would facilitate the investigation of trait concepts' relationships to everyday behavior.

Conclusion

Personality descriptions of individuals are predicated on knowing how such descriptions are manifest in everyday behavior, and the primary purpose of this article is to empirically observe just how Big Five trait contents are manifest in everyday behavior. It was found that conceiving of behavior as density distributions over time rather than as single behaviors reveals the manifestation of Big Five content as being both conditional and nonconditional. A large degree of variability does not deny the stability of means, and the stability of means does not dismiss the large degree of vari-

ability. The density distribution conception also reveals that additional aspects of behavior patterns, especially the amount of variability an individual shows, are stable and psychologically meaningful individual differences variables. Such findings will hopefully contribute to supporting collaborative work on process and structure in the field of personality.

References

- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Henry Holt.
- Barrett, L. F., & Pietromonaco, P. R. (1997). Accuracy of the five-factor model in predicting perceptions of daily social interactions. *Personality and Social Psychology Bulletin*, *23*, 1173–1187.
- Berdie, R. F. (1969). Consistency and generalizability of intraindividual variability. *Journal of Applied Psychology*, *53*, 35–41.
- Blake, M. J. F. (1971). Temperament and time of day. In P. Colquhoun (Ed.), *Biological rhythms and human performance* (pp. 109–148). London: Academic Press.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.
- Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality*, *63*, 1–25.
- Borkenau, P., & Ostendorf, F. (1987). Retrospective estimates of act frequencies: How accurately do they reflect reality? *Journal of Personality and Social Psychology*, *52*, 626–638.
- Borkenau, P., & Ostendorf, F. (1998). The Big Five as states: How useful is the Five-Factor Model to describe intra-individual variations over time? *Journal of Research in Personality*, *32*, 202–221.
- Botwin, M. D. (1989). Structure of act-report data: Is the five-factor model of personality recaptured? *Journal of Personality and Social Psychology*, *56*, 988–1001.
- Brown, K. W., & Moskowitz, D. S. (1998). Dynamic stability of behavior: The rhythms of our interpersonal lives. *Journal of Personality*, *66*, 105–134.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, *90*, 105–126.
- Cantor, N. (1990). From thought to behavior: "Having" and "doing" in the study of personality and cognition. *American Psychologist*, *45*, 735–750.
- Cantor, N., & Kihlstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice Hall.
- Cattell, R. B. (1973). *Personality and mood by questionnaire*. San Francisco: Jossey-Bass.
- Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psycho-physiological source traits in a normal individual. *Psychometrika*, *12*, 267–288.
- Cervone, D. (1991). The two disciplines of personality psychology: Review of *Handbook of Personality: Theory and Research*. *Psychological Science*, *2*, 371–377.
- Cervone, D. (1999). Bottom-up explanation in personality psychology: The case of cross-situational coherence. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization* (pp. 303–341). New York: Guilford.
- De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. B. (1994). The Big Five: A tip of the iceberg of individual differences. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 91–109). Hillsdale, NJ: Erlbaum.
- Diener, E., & Larsen, R. J. (1984). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. *Journal of Personality and Social Psychology*, *47*, 871–883.
- Diener, E., Larsen, R. J., & Emmons, R. A. (1984). Person * Situation

- interactions: Choice of situations and congruence response models. *Journal of Personality and Social Psychology*, 47, 580–592.
- Dunning, D., & McElwee, R. O. (1995). Idiosyncratic trait definitions: Implications for self-description and social judgment. *Journal of Personality and Social Psychology*, 68, 936–946.
- Eckenrode, J. (1984). Impact of chronic and acute stressors on daily reports of mood. *Journal of Personality and Social Psychology*, 46, 907–918.
- Eid, M., & Diener, E. (1999). Intraindividual variability in affect: Reliability, validity, and personality correlates. *Journal of Personality and Social Psychology*, 76, 662–676.
- Ekman, P., & Davidson, R. J. (1994). *The nature of emotion*. New York: Oxford University Press.
- 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., & Parker, J. D. A. (1992). Interactionism revisited: Reflections on the continuing crisis in the personality area. *European Journal of Personality*, 6, 177–198.
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37, 1097–1126.
- Epstein, S. (1994). Trait theory as personality theory: Can a part be as great as the whole? *Psychological Inquiry*, 5, 120–122.
- Eysenck, H. J. (1981). General features of the model. In H. J. Eysenck (Ed.), *A model for personality* (pp. 1–37). New York: Springer-Verlag.
- Fiske, D. W. (1961). The inherent variability of behavior. In D. W. Fiske & S. R. Maddi (Eds.), *Functions of varied experience* (pp. 326–354). Homewood, IL: Dorsey Press.
- Fleeson, W. (2001). Behavioral inertia and explaining within-person variability. Unpublished manuscript, Wake Forest University.
- Fleeson, W., & Cantor, N. (1995). Goal relevance and the affective experience of daily life: Ruling out situational explanations. *Motivation and Emotion*, 19, 25–57.
- Fleeson, W., Zirkel, S., & Smith, E. E. (1995). Mental representations of trait categories and their influences on person perception. *Social Cognition*, 13, 365–397.
- Fridhandler, B. M. (1986). Conceptual note on state, trait, and the state–trait distinction. *Journal of Personality and Social Psychology*, 50, 169–174.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, 48, 82–98.
- Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology*, 74, 1337–1349.
- Haus, E., Lakatua, D. J., Swoyer, J., & Sackett-Lundeen, L. (1983). Chronobiology in hematology and immunology. *American Journal of Anatomy*, 168, 467–517.
- Larsen, R. J. (1985). Individual differences in circadian activity rhythm and personality. *Personality and Individual Differences*, 6, 305–311.
- Larsen, R. J. (1987). The stability of mood variability: A spectral analytic approach to daily mood assessments. *Journal of Personality and Social Psychology*, 52, 1195–1204.
- Larsen, R. J. (1989). A process approach to personality psychology: Utilizing time as a facet of data. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 177–193). New York: Springer-Verlag.
- Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality*, 21, 1–39.
- Larsen, R. J., Diener, E., & Emmons, R. A. (1986). Affect intensity and reactions to daily life events. *Journal of Personality and Social Psychology*, 51, 803–814.
- Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology*, 58, 164–171.
- Larson, R. W., Csikszentmihalyi, M., & Graef, R. (1980). Mood variability and the psychosocial adjustment of adolescents. *Journal of Youth and Adolescence*, 9, 469–490.
- Magnusson, D., & Endler, N. S. (1977). *Personality at the crossroads: Current issues in interactional psychology*. Hillsdale, NJ: Erlbaum.
- McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, 63, 365–396.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W., & Peake, P. K. (1982). Beyond deja vu in the search for cross-situational consistency. *Psychological Review*, 89, 730–755.
- Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, 49, 229–258.
- Moskowitz, D. S. (1982). Coherence and cross-situational generality in personality: A new analysis of old problems. *Journal of Personality and Social Psychology*, 43, 754–768.
- Murray, H. A. (1938). *Explorations in personality: A clinical and experimental study of fifty men of college age*. New York: Oxford University Press.
- Murtha, T. C., Kanfer, R., & Ackerman, P. L. (1996). Toward an interactionist taxonomy of personality and situations: An integrative situational–dispositional representation of personality traits. *Journal of Personality and Social Psychology*, 71, 193–207.
- Nesselroade, J. R. (1988). Some implications of the trait–state distinction for the study of development over the life span: The case of personality. In P. B. Baltes, D. L. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. 8, pp. 163–189). Hillsdale, NJ: Erlbaum.
- Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 92–105). Washington, DC: American Psychological Association.
- Patrick, A. W., & Zuckerman, M. (1977). An application of the state–trait concept to the need for achievement. *Journal of Research in Personality*, 11, 459–465.
- Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, 5, 103–113.
- Random House College Dictionary* (Rev. ed.). (1975). New York: Random House.
- Revelle, W. (1995). Personality processes. *Annual Review of Psychology*, 46, 295–328.
- Rusting, C. L., & Larsen, R. J. (1998). Diurnal patterns of unpleasant mood: Associations with neuroticism, depression, and anxiety. *Journal of Personality*, 66, 85–103.
- Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology*, 73, 1296–1312.
- 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.
- Schwartz, J. E., Neale, J., Marco, C., Shiffman, S. S., & Stone, A. A. (1999). Does trait coping exist? A momentary assessment approach to the evaluation of traits. *Journal of Personality and Social Psychology*, 77, 360–369.
- Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intra-individual stability in the organization and patterning of behavior: Incorporating psycho-

- logical situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology*, 67, 674–687.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526–537.
- Spielberger, C. D., Lushene, R. E., & McAdoo, W. G. (1977). Theory and measurement of anxiety states. In R. B. Cattell & R. M. Dreger (Eds.), *Handbook of modern personality theory* (pp. 239–253). Washington, DC: Hemisphere.
- Thorne, A. (1989). Conditional patterns, transference, and the coherence of personality across time. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 149–159). New York: Springer-Verlag.
- Thorne, A. (1995). Juxtaposed scripts, traits, and dynamics of personality. *Journal of Personality*, 63, 593–616.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070.
- Webster's Ninth New Collegiate Dictionary* (9th ed.). (1986). Springfield, MA: Merriam-Webster.
- Zelenski, J. M., & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from experience sampling data. *Journal of Research in Personality*, 34, 178–197.

Received December 14, 1999

Revision received January 12, 2001

Accepted January 12, 2001 ■