

6

The Structure of Personality and Temperament

Gerard Saucier
Jennifer Simonds
University of Oregon

In this chapter we review progress on an important scientific issue—how attributes of personality and temperament can best be organized and structured. After explaining the rationale for doing so, we will discuss some insights that have been gained from studies of person descriptors in diverse languages. We discuss the degree to which one prominent structural model—the Big Five—is indeed an ideal model, and what other structural models are most complementary to this one in terms of strengths. Because inquiries into the structure of attributes depend significantly on how personality and temperament are defined, we begin there.

DEFINING “PERSONALITY” AND “TEMPERAMENT”

For Funder (2001), personality is “an individual’s characteristic patterns of thought, emotion, and behavior, together with the psychological mechanisms—hidden or not—behind those patterns” (p. 2). Funder referred simultaneously to characteristics that are ascribed to individuals, stable over time, and psychological in nature. But this is not the only way to define personality. Allport (1937) reviewed definitions of the concept of personality. He catalogued 50 distinct meanings. The definition he proposed was a synthesis of several of the psychological meanings of concept: “Personality is the dynamic organization within the individual of those psychophysical systems that determine his unique adjustments to his environment” (p. 48).

Definitions make assumptions explicit. The way personality is defined is quite consequential: It affects how variables are selected when studying personality phenomena. Allport’s definition highlights attributes that are seen as residing “within” the individual. Other ways of defining personality emphasize more external types of attributes, such as the role individuals assume or the status they have achieved in society, their external appearance (including their attractiveness), and

the reactions of others to the individuals as a stimulus (i.e., their social stimulus value; see MacKinnon, 1944). Later, we will explore some structural models for personality that include rather than exclude such externally defined attributes.

"Temperament" has usually been defined more narrowly than personality. Allport (1937) defined temperament with reference to "the characteristic phenomena of an individual's emotional nature, including his susceptibility to emotional stimulation, his customary strength and speed of response, the quality of his prevailing mood, and all peculiarities of fluctuation and intensity of mood; these phenomena being regarded as dependent on constitutional makeup and therefore largely hereditary in origin" (p. 54). Some modern theories have retained Allport's emphasis on and limitation to emotional reactivity and activity in their conceptualization of temperament (Goldsmith & Campos, 1982, 1986). However, early influential temperament research emphasized the idea of behavioral "styles" (Thomas, Chess, Birch, Hertzog, & Korn, 1963). Later, Buss and Plomin (1984) placed emphasis on genetics, defining temperament as "inherited personality traits present in early childhood" (p. 84). Rothbart and Bates (1998) took a psychobiological approach, defining temperament as "constitutionally-based individual differences in emotional, motor, and attentional reactivity and self-regulation" (p. 109).

Disagreements about definitions of temperament surround the issue of whether temperament is a purely affective construct, or whether systems related to cognition may be included in the traits that make up the structure of temperament (Goldsmith & Campos, 1982), and on issues regarding the relative influence of biological processes on temperamental variables (Goldsmith, et al., 1987). But virtually all definitions of temperament focus on basic dispositions that underlie and modulate the expression of reactivity, physical activity, emotionality, and sociability; dispositions that are present early in life and influenced, directly or indirectly, by biological factors; and traits that are subject to change due to environment and maturation (McCall in Goldsmith et al., 1987).

The previous review indicates that temperament is a less broad concept than personality. Although aspects of behavior, thought, and affect are widely acknowledged to be reflected in temperament, the emphasis has more often been on the affective elements, and on biologically based traits. There are few recent conceptualizations of temperament that make reference to physique or external appearance.

PARSIMONY IN PERSONALITY MODELS

A survey of the scales in current personality inventories finds scale labels for a bewildering variety of constructs. And if single words potentially referring to personality attributes in modern world languages are the focus, then the situation becomes simply overwhelming: Allport and Odbert (1936), for example, catalogued nearly 18,000 words from *Webster's Second International Dictionary* referring to characteristics that might be used to distinguish one human being from another. There needs to be a more parsimonious summary of this vast domain of concepts.

In the field of personality, there has been a rising wave of interest in the search for a scientifically compelling taxonomy of the huge number of personality attributes. A taxonomy chunks things, systematically dividing phenomena into ordered groups or categories; in other words, it is a way of "chunking" things. A scientific taxonomy helps organize and integrate knowledge and research findings by providing a stan-

standard scientific nomenclature, which facilitates communication and aids in the accumulation of empirical findings.

In taxonomy construction, a variety of procedures might be used to divide (or group) the phenomena under study. The most useful is a class of statistical methods generically referred to as "factor analysis." As noted by Goldberg and Digman (1994), factor analysis can be considered as a variable-reduction procedure, in which many variables are organized by a few factors that summarize the interrelations among the variables.

WHAT MAKES A STRUCTURAL MODEL GOOD?

However, prior to employing factor analysis, one must make a crucial determination—which variables to include in the analysis. One cannot find a dimension or factor without including a domain of variables relevant to it. Variable selection is inevitably guided by the investigator's beliefs about what makes a structural model good. These beliefs involve criteria that can be applied both to variables and factors formed from variables, and tend to focus on criteria from among the following eight alternatives:

1. *Social importance* of the variables or factors, that is, whether they are "shown to interact powerfully with social activities widely regarded as important" (Eysenck, 1991, p. 785).
2. *Comprehensiveness* of the variables or factors (taken as a whole), so that they cover "a wide field, and [are] not restricted to a narrow segment of personality research" (Eysenck, 1991, p. 774).
3. *Reliability and cross-time stability* of the variables or factors. This criterion is important for personality characteristics because they are expected to be relatively consistent across time. However, it is possible that traits themselves have more stability over time than do the correlations between, and thus the structures of, those traits.
4. *Predictive power and validity* of the variables or the factors they form. This criterion is related to social importance, but relies more heavily on specific practical contexts in which personality measures have wide application.
5. *Generalizability across types of data* that one finds for the variables or factors. For example, there should be less interest in a variable or factor found only in self-report data, than in one found to be important also in ratings by knowledgeable others, or in observer data.
6. *Generalizability across cultures and languages* that one finds for the variables or factors. This criterion has often been termed "universality" (Costa & McCrae, 1992, p. 653), or independence of "national, racial and cultural differences" (Eysenck, 1991, p. 784).
7. *Biological or other causal basis* established for the variables or factors. Personality characteristics are known to be moderately heritable (Bouchard, 1994), and heritability indicates biological influences. Therefore, biological bases are prime candidates, but not the only candidates, for important causal factors.
8. *A theory*, plausible and logically consistent, related to personality functioning or dynamics, that is linked to the model. Such a theory might enable researchers fruitfully to derive testable deductions and hypotheses to explain known

phenomena and predict phenomena that are not yet known, without contradicting well-established findings (cf. Eysenck, 1991, p. 774).

Developers of personality measures have long used a wide variety of criteria and combinations of criteria for variable selection. Because of the diversity of criteria employed, the long tradition of packaging structural models into multiscale personality inventories led to little agreement on the most important variables of personality. As of two decades ago, the literature on the structure of personality characteristics was a maelstrom of competing inventories, mostly proprietary, embedded in a mass of mutually isolated research measures. More order has been brought to the field by the lexical approach. This approach, although not perfectly combining or even involving all relevant criteria, has enabled the simultaneous application of most of the major criteria for the goodness of a structural model.

THE BASIS FOR THE LEXICAL APPROACH

Researchers have long recognized (e.g., Allport, 1937; Cattell, 1943; Goldberg, 1981; Norman, 1963) that some of the most basic personality attributes might be discovered from studying conceptions implicit in use of the natural language. If a distinction is highly represented in the lexicon, then it can be presumed to have practical importance. Folk concepts of personality (Tellegen, 1993) provide basic but not exhaustive (necessary but not sufficient) components for a science of personality attributes (Goldberg & Saucier, 1995).

This leads to a key premise of the lexical approach to taxonomy construction: The degree of representation of an attribute in language has some correspondence with the general importance of the attribute in real-world transactions. This premise links semantic representation directly with the *social importance criterion*.

An attribute represented by multiple terms in a language will likely appear as a factor in multivariate analyses. Moreover, if the factor includes terms used with high frequency, then the importance of the factor is underscored. Such factors are but a "starting point" because the lexicon could omit or underemphasize some scientifically important variables, and the meaning of single natural language terms can be vague, ambiguous, or context-dependent (John, Angleitner, & Ostendorf, 1988).

Many variables, and potential factors, might have rich semantic representation and thus satisfy a social importance criterion, so this should not be the sole criterion. Other criteria from among the eight might be fruitfully applied. In the lexical study paradigm, one particular additional criterion has taken on special importance, probably because it is the most demanding and therefore the most potentially efficient in rapidly reducing the field of candidates.

The *cross-cultural generalizability criterion* can adjudicate among competitor taxonomic structures. Structural models derived within one limited population, or a limited sample from that population, are prone to reflect the unique patternings found within that population or sample. Culture-specific patternings may be interesting in their own way. But models that transfer well across populations, and thus across languages and sociocultural settings, are more congruent with the scientific ideals of replicability and generalizability.

If cross-cultural generalizability is taken as a criterion for a good taxonomic structure, it can be applied in either a lenient or a stringent way. The lenient way is to export a set of variables (most often, those represented in a single personality inventory) for use in other populations, and then examine whether or not these pre-selected variables (after translation, if necessary) generate the same factor structure in each new language or culture. If the scales in a personality inventory generate similar factors across populations, it might be argued (McCrae & Costa, 1997) that the structure is widely generalizable. However, this is not a very challenging test. It shows only that when personality measures in a new language are made to conform to the Procrustean specifications of one model, that model can be recovered. A large number of models may be equally exportable and maintain their factor structures across many populations.

A more challenging test is to identify the most salient and important personality concepts within each linguistic/cultural context, derive an indigenous factor structure from those variables, and then examine the extent to which this new structure corresponds to any previously proposed models. A model that could meet this test in any language could be considered far more ubiquitous and universal than a structure that simply met the less demanding imposed test (i.e., showed a high degree of translatability).

The lexical approach involves such an indigenous research strategy. Analyses are carried out separately within each language, using a representative set of native language descriptors, rather than importing selections of variables from other languages (e.g., English). Generally, factors identified by the lexical approach have fared well with respect to the first six criteria, generating a relatively comprehensive set of socially important personality constructs that evidence consistency across time, good predictive validity, and generalizability across differing types of data as well as across cultures. Thus, these factors deserve in-depth consideration.

WHAT HAS BEEN LEARNED FROM NATURAL LANGUAGE PERSONALITY DESCRIPTIONS

The majority of lexical studies of personality descriptors have attempted to test the most widely influential personality model of the last two decades: the Big Five factor structure (Goldberg, 1990, 1993; John, 1990). The Big Five factors are customarily labeled Extraversion, Agreeableness, Conscientiousness, Emotional Stability (or its opposite, Neuroticism), and Intellect (or, in one inventory representation, Openness to Experience). There were signs of the Big Five structure in some studies from an earlier era (as detailed by Digman, 1990; Goldberg, 1993; John, 1990), but its identification in studies of natural language descriptors in English (e.g., Goldberg, 1990) was decisive.

If we value cross-cultural generalizability, however, applicability to one language is not enough. As detailed in more lengthy reviews (Saucier & Goldberg, 2001; Saucier, Hampson, & Goldberg, 2000), lexical studies have yielded structures resembling the Big Five most consistently in languages originating in northern Europe, including German (Ostendorf, 1990) and Polish (Szarota, 1996), as well as English. Although a study in Turkish (Goldberg & Somer, 2000) also found a structure with much resemblance to the Big Five, studies of other non-north-European languages

(e.g., Di Blas & Forzi, 1998; Church, Katigbak, & Reyes, 1998; Church, Reyes, Katigbak, & Grimm, 1997; Szirmak & De Raad, 1994) have led to results that are less clearly supportive. And because a majority of studies have relied exclusively on self-report, the degree of generality of the Big Five in peer ratings is less certain than for self-ratings.

To this point, lexical studies have revealed a great deal about the relative robustness of the Big Five, as well as information about other less well-known candidate models, including some with fewer and some with more factors. The next section discusses the most consistent findings from lexical studies to date by describing models with successively more factors.

What If Only One Factor Is Allowed?

Several lexical studies have reported evidence about factor solutions containing only one factor (Boies, Lee, Ashton, Pascal, & Nicol, 2001; Di Blas & Forzi, 1999; Goldberg & Somer, 2000; Saucier, 1997, 2003b). The findings from these studies have been quite consistent. The single factor contrasts a heterogeneous mix of desirable attributes at one pole with a mix of undesirable attributes at the other pole. This unrotated factor can be labeled *Evaluation*; it involves the contrast between socially desirable and undesirable personal qualities. This one factor structure can be expected to be the most replicable one across languages and cultures based on two principles: (a) The more terms associated with a factor, the more replicable that factor should be, and (b) because the first unrotated factor will have the most terms associated with it, it should be the most ubiquitous factor.

Findings of a single large evaluative factor are no doubt related to a classic finding in psychology. In judgments about the meanings of diverse objects in a wide array of cultural settings, a global evaluation factor (good vs. bad) was typically found to be the single largest factor (Osgood, May, & Miron, 1975). Osgood hypothesized that the ubiquity of this evaluative factor was related to basic evolutionary principles: Our forebears would not have survived if they had not become adapted at a very basic level to any signals of good versus bad objects or events—those to approach versus those to avoid, those leading to pleasure versus those leading to pain (e.g., Can I eat it or will it eat me?). Hawley (chap. 8 in this volume) refers to the evolutionary value of evaluation as a tool for determining an individual's value to a group.

So evaluation may be evolutionarily the first factor. But it is also the first factor to emerge in the cognitions of young children. Whereas older children employ more differentiated trait concepts, younger children rely on global, evaluative inference (Alvarez, Ruble, & Bolger, 2001).

Are Two Factors Better Than One?

Two factor solutions from several lexical studies also suggest a consistent pattern: One factor includes attributes associated with positively valued dynamic qualities and individual ascendancy, whereas the other factor includes attributes associated with socialization, social propriety, solidarity, and community cohesion (Boies et al., 2001; Caprara, Barbaranelli, & Zimbardo, 1997; Di Blas & Forzi, 1999; Digman, 1997; Goldberg & Somer, 2000; Hřebíčková, Ostendorf, Osecká, & Čermák, 1999;

Paulhus & John, 1998; Saucier, 1997, 2003b; Shweder, 1972; White, 1980). Such a factor structure resembles that embodied in the theoretical model of Bakan (1966), who labeled the two factors Agency and Communion. In addition, these two factors may be aligned with some of the other sets of dual personological constructs reviewed by Digman (1997) and Paulhus and John (1998), including Hogan's (1983) distinction between "getting ahead" (Dynamism) and "getting along" (Social Propriety).

This constellation of two factors is also related to the three most ubiquitous dimensions of affective meaning, which include Potency (or Strength) and Activity in addition to Evaluation (Osgood et al., 1975). Whether this correspondence is due entirely to the imposition of universal tendencies in human cognition, or to the natural structure of phenomena "out in the world" remains an open question. In judgments about human targets, Potency and Activity tend to merge into a single dimension that Osgood and his associates called "Dynamism."

As is true of the Big One factor structure, no lexical study has presented evidence to contradict the view that this two factor structure is ubiquitous across languages and cultures. If both the one and two factor structures eventually turn out to be universal, the latter has some advantage, because two factors provide more information than one.

Regularities at the Five Factor Level

As already noted, lexical studies in languages originating in northern Europe (including English) have been supportive of the Big Five, and so has a study in Turkish. But studies in Italian (De Raad, Di Blas, & Perugini, 1997) and Hungarian (Szirmak & De Raad, 1994) found no counterpart to the Intellect factor in five factor solutions. Instead, there were two Agreeableness-related factors, one contrasting peacefulness with aggression and irritability, and the other contrasting humaneness with greed and egotism (cf. Deary, 1996). Extraction of additional factors was necessary to find a factor related to Intellect.

Several lexical studies have included a relatively broad selection of variables, each including terms that could be classified as referring to emotions and moods or as being unusually highly evaluative, and two of these studies included terms referring to physical appearance. Because none of these studies found the Big Five in a five factor solution, it is clear that the appearance of the Big Five as the first five factors is contingent on strictures in variable selection.

Lexical Seven Factor Models

Although not finding the Big Five in five factor solutions, studies with wide variable selection criteria in English and Turkish did find Big Five-like factors in a seven factor solution (Goldberg & Somer, 2000; Saucier, 1997; Tellegen & Waller, 1987). The two additional factors were "Negative Valence" (a factor emphasizing attributes with extremely low desirability and endorsement rates) found in all three studies and either "Positive Valence" (a factor emphasizing vague positive attributes like Impressive and Outstanding; found in Tellegen & Waller, 1987) or Attractiveness (found in the other two studies).

But studies in two other languages with broad variable selection criteria have led to an alternative seven factor structure. The convergences between these studies occurred despite their many differences in methodology. Lexical studies in Filipino (Church et al., 1997, 1998) and Hebrew (Almagor, Tellegen, & Waller, 1995)—languages from unrelated languages and cultures—yielded a highly convergent seven factor structure, although the similarity was obscured by discrepant labels. The English translations of marker adjectives for the Filipino and Hebrew factors have been shown to correspond in a one-to-one way (Saucier, 2003a).

One of these new factors resembles the Negative Valence factor just described. Two of them resemble the Big Five factors of Conscientiousness and Intellect. The other three Big Five factors—Extraversion, Agreeableness, and Emotional Stability—correlate substantially but complexly with the remaining four factors, which map an affective-interpersonal domain (cf. Saucier, 1992). These four can be labeled Gregariousness (or Liveliness), Self-Assurance (or Mettle or Fortitude), Even Temper (Tolerant vs. Temperamental), and Concern for Others (vs. Egotism). Big Five Extraversion is related to Gregariousness and Self-Assurance, Emotional Stability to Self-Assurance and Even Temper, and Agreeableness to Even Temper and Concern for Others.

Similar factors have been obtained from lexical data in English (Saucier, 2003a), and factors found in studies in Italian (e.g., Di Blas & Forzi, 1998) resemble the Multi-Language Seven. However, further replication tests are needed because few studies have used such broad variable selection criteria.

Ashton et al. (2001; cf., Saucier, 2002) presented evidence that more than half of the lexical studies conducted to date yield a consistent pattern in six factor solutions; the analyses involved presuppose removal of the most extremely evaluative terms. This six factor pattern resembles that of the Multi-Language Seven, with the following exceptions: (a) Even Temper is labeled Agreeableness, and Gregariousness is labeled Extraversion; (b) with a narrower variable selection, the three ML7 factors of Negative Valence, Concern for Others, and Self-Assurance become compressed into two (in Ashton and Lee's factors of Honesty and Emotionality).

Saucier's (2003a) analyses of convergences among Filipino, Hebrew, Italian, and English factors, taken together with Ashton and Lee's detection of related recurrent patterns in six factor solutions, suggest that a new competitor alternative to the Big Five is taking shape. There are three exciting features of this new model: Its origin is outside of northern European languages, its cross-cultural generalizability appears likely to exceed that for the Big Five, and it has incrementally greater comprehensiveness as compared to the Big Five.

Limitations

There are important limitations to the body of lexical studies carried out to date, and these can be related to any of the eight criteria. In terms of cross-cultural generalizability, more studies are needed in non-Western settings where the majority of the world's human population resides, and with non-European languages. In terms of generalizability across data types, lexical studies have focused almost entirely on those attributes represented in adjectives, although some attributes may be represented mainly as type nouns (e.g., Hick, Nerd, Slavedriver, Tease) or as attrib-

ute nouns (e.g., Integrity, Mettle); more studies that include attributes represented in nonadjectival forms are needed. In addition, most lexical studies to date have relied exclusively on self-descriptions, a methodology whose use should be supplemented with descriptions by knowledgeable informants.

Are sets of lexical factors comprehensive? Certainly they are more comprehensive than the structural models that came before. The extra content coverage in the NEO inventory after the grafting of two lexical factors (Agreeableness and Conscientiousness) into its structural model (Costa & McCrae, 1985) is a prime reason for its rapid gain in popularity. But, there are clearly dimensions of individual differences that are beyond the Big Five, particularly if we widen the taxonomy to include abilities, social attitudes, or appearance-related characteristics (Saucier & Goldberg, 1998).

Because of their derivation in commonly referenced attribute concepts, lexical factors are guaranteed strong *social importance*, although there is no guarantee that all socially important factors will be richly represented in the lexicon. Lexical factors have already performed a service to the field in enhancing the *comprehensiveness* of personality models; prior to the Big Five, there was little attention to Agreeableness or Conscientiousness. Lexical factors have shown good evidence of *cross-time stability*, and their *predictive validity* (e.g., in work settings) has been a major force behind their rising popularity. As the review indicates, their generalizability across types of data and across cultures has been impressive, and not matched by alternative models. On these six criteria, lexically based factors like the Big Five can be judged as something between adequate and superb. Structural models might be developed that are incrementally better on one or more of these criteria, but improvements are unlikely to be huge. It is the last two criteria that reveal limits of lexical factors.

One of these criteria is biological or other causal basis. The Big Five shows evidence of heritability for all factors in the model, but it is not clear that the factors actually maximize heritability. It may be possible, for example, through analyses of genetic and environmental correlations in suitable data sets, to locate factors whose causal clarity is maximized, located or rotated so as to yield the maximum heritability, or the maximum basis in shared environmental effects. The result would be a model with superior causal clarity. There is no clear evidence that the Big Five correspond closely to the primary lines of genetic or biological influence, or provide the optimal beachhead for inserting personality variables into studies of the brain.

The other criterion is theory. The Big Five (and any other lexical structural model) is inductively and empirically derived, and lacks theoretical underpinnings. There have been attempts to transplant the Big Five into some body of preexisting or ad hoc theory (e.g., McCrae & Costa, 1996; MacDonald, 1995), but it is not clear that these operations have been particularly successful.

Thus, structural models like the Big Five are vulnerable to being superseded by some model with a clearer basis in biology or in some other causal element, and by a model with a stronger basis in theory. There are a number of models that meet this description. Unfortunately, many of them contain only one or two factors, giving them little in the way of comprehensiveness, and have been measured exclusively via self-report methodologies with little attention paid to cross-cultural generalizability. Models that are lacking in so many respects give up more than they gain in comparison with a lexical model like the Big Five.

There are, however, models of temperament that are more comprehensive than this and are not limited to self-report. These models beneficially introduce a longitudinal perspective that is missing from structural models based on ratings of adult targets.

COMPLEMENTARY CONTRIBUTIONS FROM THE STUDY OF TEMPERAMENT

Temperament may be conceptualized as the early-in-a-life framework from which personality develops (Digman, 1994; Rothbart, Ahadi, & Evans, 2000). The nature of temperament as early-appearing dimensions underlying behavior presents challenges different from those faced by personality researchers. Expressions of temperament change during development. The systems that underlie temperament are, like the behavioral indicators of temperament, in the process of development and subject to change throughout the life span (Rothbart, 1989). New systems come "on line" at different stages. For example, the capacity to inhibit approach to new objects does not develop until later in the first year in infants, and once it is in place individual differences in this type of inhibition remain relatively stable (Rothbart & Bates, 1998). As another example, there are differences in elicitors of fear at different ages. Measurement of fear in infancy and early childhood is accomplished through determination of reaction to novel, unusual, or sudden presentation of stimuli (Rothbart, Derryberry, & Hershey, 2000). This type of fear goes away for nearly all children by school age, by which point going to the doctor, sleeping over at another child's house, being in dark rooms, and watching scary movies become major fear-evoking events (Rothbart, Ahadi, Hershey, & Fisher, 2001; Simonds & Rothbart, 2005). But for early adolescents and adults, better indicators of a fearful temperament are anticipation of negative events and possible failures (Capaldi & Rothbart, 1992; Rothbart, Ahadi, & Evans, 2000).

Measurement of temperament relies on different types of data at different ages. In order to measure temperament in infancy and early childhood, parent report is a valuable source of information. At this age assessment through laboratory measurement of behavioral and physiological indicators of underlying temperament is another potential method. But infants and younger children are unable to provide self-reports. Beyond age 4 or 5 years, it is possible for children to report on their own temperament with a fair degree of reliability, and there is some correspondence with parent reports of temperament in the self-reports of children from preschool age to early adolescence (Ellis & Rothbart, 2002; Hwang, 2002; Simonds & Rothbart, 2003). Once children have entered school, teachers become sources of information on temperament and emerging personality (Bramlett, Scott, & Rowell, 2000; Digman & Inouye, 1986; Digman & Shmelyov, 1996) and temperament (e.g., Kurdek & Lillie, 1985). Each rater provides information about a child in a different context: parents in the home, teachers in the school, peers in the social realm, and self-reports provide unique information about a child's internal states. Laboratory measures remain relevant at all ages.

Each method of measurement is controversial because of potential biases unique to it, imperfect agreement between raters, and between questionnaire data and objective laboratory measures (Teglasi, 1998), and the well-known limitations of

self-report data, which apply to self-reports of children as well as adults. Despite arguments that parents are not reliable raters of their own children's temperament (Kagan, 2001), parent report data has shown substantial validity (Rothbart & Bates, 1998) and is the most frequently used method of measuring temperament in infants and children.

An important unresolved problem in temperament research stems from the variety of measurement methods. Although these methods show agreement, much disagreement between methods remains. It is unclear to what degree findings and key models are contingent on utilization of certain methods of measurement and not others. As is true in personality research, the use of a wide variety of measurement instruments makes it difficult to compare results; temperament research has not settled on a consensual structural model.

THE STRUCTURE OF TEMPERAMENT

There are several influential structural models of temperament, each involving a particular way of defining temperament and selecting variables (or creating items). These different theories have given rise to different ideas about the structure of temperament, with some marked similarity, and some remaining differences.

Thomas and Chess' Model of Behavioral Styles

Current models of temperament are strongly influenced by findings from the New York Longitudinal Study (NYLS; Thomas et al., 1963). Through interviewing parents about the behaviors of their 2- to 6-month-old infants, Thomas et al. identified nine dimensions of behavioral styles, or early-appearing patterns of behavior, to describe the observed differences: Activity Level, Approach/Withdrawal (responses to novel stimuli), Adaptability (ease of modification of response to new or altered situations), Mood (negative or positive), Threshold, Intensity (energy level of reaction), Distractibility, Rhythmicity (regularity of sleep, hunger, feeding, and elimination), and Attention Span/Persistence. Martin, Wisenbaker, and Huttunen (1994) conducted a factor analysis of large samples on NYLS-based measures and found five factors: Activity Level, Negative Emotionality, Task Persistence, Adaptability, and Inhibition. Two other factors were found, Rhythmicity and Threshold, but were determined to suffer from lack of relevance past infancy (Rhythmicity) or lack of internal consistency (Threshold).

Questionnaire measures based on the NYLS model have been developed for most age groups (Fullard, McDevitt, & Carey, 1984; Hegvik, McDevitt, & Carey, 1982; McDevitt & Carey, 1978; Windle & Lerner, 1986). NYLS-based infant and child questionnaires use parent reports for measuring temperament.

The NYLS model is based in observable behaviors that appear very early in child development and so have a strong temperament (as contrasted with adult personality) flavor. However, the scales are conceptually overlapping and lack discriminant validity, and there has been difficulty with replicating its nine factors (Martin, Wisenbaker, & Huttunen, 1994; Sanson, Prior, Garino, Oberklaid, & Sewell, 1987). Given these limitations of the NYLS model, other models have attracted attention.

Buss and Plomin's Emotionality-Activity-Sociability (EAS) Model

Buss and Plomin (1975) developed a model of temperament based on heritable dimensions believed to appear in early life: Emotionality, Activity, and Sociability, and Impulsivity (thus, the EASI model). Due to lack of support for the heritability of impulsivity, that dimension was later dropped (Buss & Plomin, 1984).

The EAS model is currently reflected in two measures of temperament: the EAS-III, and the EAS- and NYLS-based parent report Colorado Child Temperament Inventory, which includes a shyness scale in addition to the EAS factors (Buss & Plomin, 1984; Rowe & Plomin, 1977). The dimensions have a theoretical basis as well as evidence of stability (Rende, 1993). However, the model excludes the useful dimension of Positive Affect, and there is doubt about whether the EAS dimensions are actually more heritable than dimensions in other models.

Rothbart and Derryberry's Psychobiological Approach

Rothbart and Derryberry's (1981) model of temperament was influenced by the NYLS dimensions as well as characteristics identified as indicating temperamental variability in animal species (Diamond, 1957) and characteristics identified as heritable in human behavioral genetics research (Rothbart & Bates, 1998). The model was first empirically represented in the Infant Behavior Questionnaire (IBQ; Rothbart, 1981), which included a set of six scales that showed substantial internal consistency and lack of construct overlap: Activity Level, Smiling and Laughter, Fear, Distress to Limitations (Crying or fussing while confined, subject to caretaking actions, or unable to perform desired actions), Duration of Orienting, and Soothability. The IBQ provided the basis for further questionnaires: the revision of the IBQ (IBQ-R; Gartstein & Rothbart, 2003), the Children's Behavior Questionnaire (CBQ; Rothbart et al., 2001), the Temperament in Middle Childhood Questionnaire (TMCQ; Simonds & Rothbart, 2005), Early Adolescent Temperament Questionnaire, Revised (EATQ-R; Capaldi & Rothbart, 1992; Ellis & Rothbart, 2002), and the Adult Temperament Questionnaire (Evans & Rothbart, 2005).

Laboratory assessments can elicit temperament-based behaviors. For instance, latency to, intensity of, and duration of positive reaction (smiling, laughter, positive vocalization) to novel stimuli have been used as laboratory measures of Smiling and Laughter in infancy, a behavioral scale that indicates underlying positive affect. Rothbart, Derryberry, et al. (2000) found low to moderate correspondence between laboratory measures and parent report for a number of scales in infancy, and found that scores on laboratory measures in infancy predict (with medium to large effect sizes) parent ratings of 7-year-old activity and frustration.

Studies using these measures have found two factors that are present from infancy to adulthood: Extraversion/Surgency and Negative Affectivity. Effortful Control appears as a third factor from early childhood through adulthood. Recent evidence indicates that Affiliative tendencies form a factor that is present in infancy (Affiliation/Orienting) and in early adolescence (Affiliation; Putnam, Ellis, & Rothbart, 2001). Moreover, in adulthood, a separate Orienting Sensitivity factor is present (Rothbart, Ahadi, et al., 2000).

Compared to the NYLS and EASI models, this psychobiological model appears to be more comprehensive (Shiner, 1998). It is also more linked to theory regarding physiological mechanisms, and more inclusive of self-control mechanisms. One limitation is the dearth of clear evidence on the stability of its scores across long spans of time; the use of behavioral indicators distinct to each age group on inventories tailored to that age group have so far created some limitations to studying cross-time stability with these measures.

How does this psychobiological model stand with respect to the eight criteria for good structural models? The stability of three main factors from early childhood to adulthood (Extraversion/Surgency, Negative Affectivity, and Effortful Control) indicates some cross-time stability and reliability of these factors of temperament. Evidence for links between temperament factors of Negative Affectivity and Effortful Control and psychopathology (Lonigan & Phillips, 2001; Posner & Rothbart, 2000) indicate some predictive validity and, assuming that psychopathology is a major social problem, also some degree of social importance for the temperament factors. Moderate to strong correspondence between laboratory measures and questionnaire data indicates generalizability across types of data (Rothbart, Derryberry, et al., 2000). Temperament factors of Extraversion/Surgency, Negative Affectivity, and Effortful Control have been found in numerous studies examining temperament across cultures, although some lower level scales were related to different factors in different cultures (Ahadi, Rothbart, & Ye, 1993).

Relative to natural language-based models of the structure of personality, the greatest strengths of this model are on the last two of the eight criteria: biological basis and theory. This model has a theory-driven emphasis on biological systems that underlie the behavioral expression of temperament. It is likely that different genetic alleles predispose to the neural and neurochemical differences that underlie temperament (Rothbart & Bates, 1998). On the other hand, the psychobiological temperament model may not be as strong as lexicon-derived models on criteria like comprehensiveness, predictive validity, and generalizability across cultures and languages.

Generality of Structures of Temperament

Some studies have sought to determine the structure of temperament through use of multiple instruments (Anthony, Lonigan, Hooe, & Phillips, 2002; Lemery, Goldsmith, Klinnert, & Mroczek, 1999; Lonigan & Dyer, 2000). Combining the Children's Behavior Questionnaire (CBQ; Rothbart et al., 2001), Emotionality-Activity-Sociability questionnaire (EAS; Buss & Plomin, 1984), and the Positive Affect Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988), Lonigan and Dyer (2000) found three relatively independent factors, similar to findings of Rothbart et al. (2001): Positive Affectivity/Surgency, Negative Affectivity/Neuroticism, and Effortful Control.

Constructs similar to three main factors of temperament have been identified in models of adult personality. Eysenck's (1967) original model of personality consisted of the two factors of Extraversion and Neuroticism, to which Psychoticism, considered to be a measure of disinhibition (Watson & Clark, 1993), was later added (Eysenck & Eysenck, 1975). Tellegen (1985) put forth a three factor model consisting of Negative Emotional-

ity, Positive Emotionality, and Constraint. Combining models of Eysenck and Tellegen, Clark and Watson (1999) identified three main superfactors of N/NE (Neuroticism/Negative Emotionality), E/PE (Extraversion/Positive Emotionality), and DvC (Disinhibition vs. Constraint). These three factors share similarities with the Extraversion/Surgency, Negative Affectivity, and Effortful Control factors (Rothbart, et al. 2001), although Constraint has fear as a component construct whereas Rothbart et al. group fear with Negative Affectivity and not with Effortful Control. The same three temperament factors correspond to three of the Big Five personality factors: Positive Affectivity/Surgency to Extraversion, Negative Affectivity to Neuroticism/Emotional Stability, and Effortful Control to Conscientiousness (Rothbart, Ahadi, et al., 2000).

Thus, there is considerable overlap between structural models of personality and temperament. Overall, however, important questions remain about the relation of temperament factors to personality factors. Although lexical studies converge with respect to the factors in one and two factor solutions, it is not clear whether these lexical factors converge with one and two factor solutions from temperament measures. Approach (temperament) and Dynamism (lexical) are probably related. As for the lexically emphasized Social Propriety factor, these are attributes related to the effects of socialization, that is, whether a person is relatively well-behaved (e.g., considerate, polite, punctual, patient, honest) or ill-mannered (e.g., thoughtless, rude, negligent, careless, deceitful). Such attributes are certainly important in judgments made about the behavior of children (e.g., by teachers), but have conventionally been thought of as primarily effects of socialization, and perhaps of moral development, not primarily effects of temperament. In other words, Social Propriety (or Morality) has been assumed to be an effect from the outside, whereas Negative Affectivity or even Effortful Control are assumed to be effects from the inside. It appears that this difference in content representation between personality and temperament models may be a direct effect of variable selection: Attributes related to social propriety and morality have not been included in temperament studies, and consequences for structural results follow from this exclusion.

Cloninger, Svrakic, and Przybek (1993) proposed a psychobiological model for adult individual differences that separates temperament from character, with character assumed to represent "effects from outside" (i.e., environment) to a greater degree than temperament. That is, the theory supposes that temperament reflects mainly genetic effects, but character represents mainly environmental effects. However, heritability for the character dimensions (Cooperativeness, Self-directedness, Self-transcendence) does not seem to be markedly lower than that for temperament dimensions (Harm-avoidance, Novelty-seeking, Reward-dependence, Persistence) (Ando et al., 2002). This provides an interesting puzzle for the psychology of personality and temperament: Present psychobiological models have difficulty accounting for individual differences in certain forms of social behavior (i.e., character) that may have analogues in other social species, are important in human societies, and have some degree of genetic basis.

CONCLUSIONS

Recent decades have seen important progress in discerning the structure of personality attributes. At the very broadest level—too broad for many purposes—this

structure appears to have much in common with Osgood's (1962) classic dimensions of affective meaning, which were found in studies of the ways that diverse objects (not just persons) are judged and perceived. At a slightly less broad but more informative level are the well-known Big Five factors. The extent to which the Big Five is optimal at its level in the hierarchy is not fully determined, there being promising new competitor models (Ashton et al., 2004; Saucier, 2003a), with the competitions not really yet begun. It is also not clear whether current models of child temperament structure, which correspond largely but imperfectly with personality structure models derived by studying adults, are ultimately the optimal ones. Factorial equivalence over groups is at present better established for lexically derived factor structures in adults. However, including both childhood and adulthood, factorial equivalence over time is better established for the structure of temperament. Hopefully, in the next decades, research on personality and temperament will converge on a common framework. To enable such a common framework, it is vital that personality and temperament researchers become and remain aware of the strong homologies between the domains they study.

Much remains unresolved, leaving potential for petty disputes about differences between structural models. So it is important to remember that scientific models are, by definition, set out tentatively, subject to the judgment of subsequent evidence. Researchers should bear in mind the criteria by which structural models can be compared—in other words, what makes a structural model “good.” By focusing on these criteria, researchers might keep their eye on the prize—an ultimately optimal structural model—and generate increasing evidence related to overall optimality of structural models.

Future models of temperament and personality will be more comprehensive and more widely generalizable across languages and cultures. These models will not only provide improved prediction of a wide array of useful criteria, but will also include more explicit linkage to the psychological mechanisms that underlie individual differences. The ultimately optimal structural model, in other words, will have both basic science foundations and real-world applications.

ACKNOWLEDGMENTS

Work on this chapter was supported by Grant MH-49227 from the National Institute of Mental Health, U.S. Public Health Service. Some portions of this chapter are derived from a previous contribution (Saucier & Goldberg, 2003) by the first author.

REFERENCES

- Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the U.S. and China: Similarities and differences. *European Journal of Personality*, 7, 359–378.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs*, 47 (1, Whole No. 211).
- Almagor, M., Tellegen, A., & Waller, N. (1995). The Big Seven model: A cross-cultural replication and further exploration of the basic dimensions of natural language of trait descriptions. *Journal of Personality and Social Psychology*, 69, 300–307.
- Alvarez, J. M., Ruble, D. N., & Bolger, N. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development*, 72, 1409–1425.

- Ando, J., Ono, Y., Yoshimura, K., Onoda, N., Shinohara, M., Kanba, S., & Asai, M. (2002). The genetic structure of Cloninger's seven-factor model of temperament and character in a Japanese sample. *Journal of Personality*, *70*, 583–609.
- Anthony, J. L., Lonigan, C. J., Hooe, E. S., & Phillips, B. M. (2002). An affect-based hierarchical model of temperament and its relations with internalizing symptomatology. *Journal of Clinical Child and Adolescent Psychiatry*, *31*, 480–490.
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P. De Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*, 356–366.
- Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man*. Boston: Beacon.
- Boies, K., Lee, K., Ashton, M. C., Pascal, S., & Nicol, A. A. M. (2001). The structure of the French personality lexicon. *European Journal of Personality*, *15*, 277–295.
- Bouchard, T. J. (1994). Genes, environment, and personality. *Science*, *264*, 1700–1701.
- Bramlett, R. K., Scott, P., & Rowell, R. K. (2000). A comparison of temperament and social skills in predicting academic performance in first graders. *Special Services in the Schools*, *16*, 147–158.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality*. New York: Wiley.
- Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Capaldi, D. M., & Rothbart, M. K. (1992). Development and validation of an early adolescent temperament measure. *Journal of Early Adolescence*, *24*, 153–173.
- Caprara, G. V., Barbanelli, C., & Zimbardo, P. G. (1997). Politicians' uniquely simple personalities. *Nature*, *385*, 493.
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*, 476–506.
- Church, A. T., Katigbak, M. S., & Reyes, J. A. S. (1998). Further exploration of Filipino personality structure using the lexical approach: Do the Big Five or Big Seven dimensions emerge? *European Journal of Personality*, *12*, 249–269.
- Church, A. T., Reyes, J. A. S., Katigbak, M. S., & Grimm, S. D. (1997). Filipino personality structure and the Big Five model: A lexical approach. *Journal of Personality*, *65*, 477–528.
- Clark, L. A., & Watson, D. (1999). Temperament: A new paradigm for trait psychology. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 399–423). New York: Guilford.
- Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry*, *50*, 975–990.
- Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*, 653–655.
- Deary, I. J. (1996). A (latent) Big Five personality model in 1915? A reanalysis of Webb's data. *Journal of Personality and Social Psychology*, *71*, 992–1005.
- De Raad, B., Di Blas, L., & Perugini, M. (1997). Two independent Italian trait taxonomies: Comparisons with Italian and between Italian Germanic languages. *European Journal of Personality*, *11*, 167–185.
- Diamond, S. (1957). *Personality and temperament*. New York: Harper.
- Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality descriptors in the Italian language. *European Journal of Personality*, *12*, 75–101.
- Di Blas, L., & Forzi, M. (1999). Refining a descriptive structure of personality attributes in the Italian language: The abridged Big Three circumplex structure. *Journal of Personality and Social Psychology*, *76*, 451–481.

- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 41, pp. 417–440). Palo Alto, CA: Annual Reviews.
- Digman, J. M. (1994). Child personality and temperament: Does the five-factor model embrace both domains? In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 323–338). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Digman, J. M. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256.
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50(1), 116–123.
- Digman, J. M., & Shmelyov, A. G. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology*, 71, 341–351.
- Ellis, L. K., & Rothbart, M. K. (2002). *Revision of the Early Adolescent Temperament Questionnaire*. Manuscript in preparation.
- Evans, D. E., & Rothbart, M. K. (2005). *A hierarchical model of temperament and the Big Five*. Manuscript submitted for publication.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Thomas.
- Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3? Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12, 773–790.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *The Eysenck Personality Questionnaire*. London: Hodder & Stoughton.
- Fullard, W., McDevitt, S. C., & Carey, W. B. (1984). Assessing temperament in one- to three-year old children. *Journal of Pediatric Psychology*, 9, 205–217.
- Funder, D. C. (2001). *The personality puzzle* (2nd ed.). New York: Norton.
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant behavior via the Revised Infant Behavior Questionnaire. *Infant Behavior and Development*, 26, 64–86.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. W. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 141–165). Beverly Hills, CA: Sage.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34.
- Goldberg, L. R., & Digman, J. M. (1994). Revealing structure in the data: Principles of exploratory factor analysis. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 216–242). New York: Springer.
- Goldberg, L. R., & Saucier, G. (1995). So what do you propose we use instead? A reply to Block. *Psychological Bulletin*, 117, 221–225.
- Goldberg, L. R., & Somer, O. (2000). The hierarchical structure of common Turkish person-descriptive adjectives. *European Journal of Personality*, 14, 497–531.
- Goldsmith, H. H., Buss, A., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., Hinde, R. A., & McCall, R. B. (1987). Roundtable: What is temperament? Four approaches. *Child Development*, 58, 505–529.
- Goldsmith, H. H., & Campos, J. J. (1982). Toward a theory of infant temperament. In R. N. Emde & R. J. Harmon (Eds.), *The development of attachment and affiliative systems* (pp. 161–193). New York: Plenum.
- Goldsmith, H. H., & Campos, J. J. (1986). Fundamental issues in the study of early temperament: The Denver Twin Temperament Study. In M. E. Lamb, A. L. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology* (pp. 231–283). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hegvik, R. L., McDevitt, S. C., & Carey, W. B. (1982). The Middle Childhood Temperament Questionnaire. *Developmental and Behavioral Pediatrics*, 3, 197–200.

- Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *Nebraska symposium on motivation* (pp. 336–355). Lincoln: University of Nebraska Press.
- Hřebíčková, M., Ostendorf, F., Osecká, L., & Čermák, I. (1999). Taxonomy and structure of Czech personality-relevant verbs. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (vol. 7, pp. 51–65). Tilburg, The Netherlands: Tilburg University Press.
- Hwang, J. (2002). *Development of a temperament self-report measure for young children*. Unpublished doctoral dissertation, University of Oregon.
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford.
- John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, 2, 171–203.
- Kagan, J. (2001). The structure of temperament. In R. N. Emde & J. K. Hewitt (Eds.), *Infancy to early childhood: Genetic and environmental influences on developmental change* (pp. 45–51). New York: Oxford University Press.
- Kurdek, L. A., & Lillie, R. (1985). The relation between classroom social status and classmate likeability, compromising skill, temperament, and neighborhood social interactions. *Journal of Applied Developmental Psychology*, 6(1), 31–41.
- Lemery, K. S., Goldsmith, H. H., Klinnert, M. D., & Mrazek, D. A. (1999). Developmental models of infant and childhood temperament. *Developmental Psychology*, 35(1), 189–201.
- Lonigan, C. J., & Dyer, S. M. (2000). *Toward a hierarchical affective model of temperament in early childhood*. Unpublished manuscript.
- Lonigan, C. J., & Phillips, B. M. (2001). Temperamental influences on the development of anxiety disorders. In M. W. Vasey & M. R. Dadds (Eds.), *The developmental psychopathology of anxiety* (pp. 60–91). New York: Oxford University Press.
- MacDonald, K. (1995). Evolution, the five-factor model, and levels of personality. *Journal of Personality*, 63, 525–566.
- MacKinnon, D. W. (1944). The structure of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 3–48). New York: Ronald.
- Martin, R. P., Wisenbaker, J., & Huttenen, M. (1994). Review of factor analytic studies of temperament measures based on the Thomas–Chess structural model: Implications for the Big Five. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 323–338). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McCrae, R. R., & Costa, P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford.
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509–516.
- McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3-7 year old children. *Journal of Child Psychology and Psychiatry*, 19, 245–253.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574–583.
- Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, 17, 10–28.
- Osgood, C. E., May, W., & Miron, M. (1975). *Cross-cultural universals of affective meaning*. Urbana: University of Illinois Press.
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit* [Language and personality structure: Toward the validation of the Five-Factor model of personality]. Regensberg, Germany: S. Roderer Verlag.

- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-descriptive styles with basic traits and motives. *Journal of Personality*, 66, 1025–1060.
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12, 427–441.
- Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. In A. Elias & A. Angleitner (Eds.), *Advances/proceedings in research on temperament* (pp. 165–182). Germany: Pabst Scientist Publisher.
- Rende, R. D. (1993). Longitudinal relations between temperament traits and behavioral syndromes in middle childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 287–290.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, 52, 569–578.
- Rothbart, M. K. (1989). Temperament and development. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–247). West Sussex, England: Wiley.
- Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, 78, 122–135.
- Rothbart, M. K., Ahadi, S. A., Hershey, K., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development*, 72, 1394–1408.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.
- Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37–86). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rothbart, M. K., Derryberry, D., & Hershey, K. (2000). Stability of temperament in childhood: Laboratory infant assessment to parent report at seven years. In V. J. Molfese & D. L. Molfese (Eds.), *Temperament and personality development across the life span* (pp. 85–119). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment*, 41, 150–156.
- Sanson, A., Prior, M., Garino, E., Oberklaid, F., & Sewell, J. (1987). The structure of infant temperament: Factor analysis of the Revised Infant Temperament Questionnaire. *Infant Behavior and Development*, 10, 97–104.
- Saucier, G. (1992). Benchmarks: Integrating affective and interpersonal circles with the Big-Five personality factors. *Journal of Personality and Social Psychology*, 62, 1025–1035.
- 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. (2002). Gone too far—or not far enough? Comments on Ashton and Lee (2001). *European Journal of Personality*, 16, 55–62.
- Saucier, G. (2003a). An alternative multi-language structure of personality attributes. *European Journal of Personality*, 17, 179–205.
- Saucier, G. (2003b). Factor structure of English-language personality type-nouns. *Journal of Personality and Social Psychology*, 85, 695–708.
- Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of Personality*, 66, 495–524.
- Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69, 847–879.
- Saucier, G., & Goldberg, L. R. (2003). The structure of personality attributes. In M. Barrick & A. M. Ryan (Eds.), *Personality and work* (pp. 1–29). New York: Jossey-Bass-Pfeiffer.

- Saucier, G., Hampson, S. E., & Goldberg, L. R. (2000). Cross-language studies of lexical personality factors. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 1–36). East Sussex, England: Psychology Press.
- Shiner, R. L. (1998). How shall we speak of children's personalities in middle childhood: A preliminary taxonomy. *Psychological Bulletin*, *124*, 308–332.
- Shweder, R. A. (1972). *Semantic structure and personality assessment*. Unpublished doctoral dissertation, Harvard University.
- Simonds, J., & Rothbart, M. K. (2005). *Temperament in Middle Childhood Questionnaire*. Manuscript in preparation.
- Szarota, P. (1996). Taxonomy of the Polish personality-descriptive adjectives of the highest frequency of use. *Polish Psychological Bulletin*, *27*, 342–351.
- Szirmák, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality*, *8*, 95–118.
- Teglasi, H. (1998). Temperament constructs and measures. *School Psychology Review*, *27*, 564–585.
- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry*, *4*, 122–130.
- Tellegen, A., & Waller, N. G. (1987). *Re-examining basic dimensions of natural language trait descriptors*. Paper presented at the 95th annual convention of the American Psychological Association.
- Thomas, A., Chess, S., Birch, H. G., Hertzog, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.
- Watson, D., & Clark, L. A. (1993). Behavioral disinhibition versus constraint: A dispositional perspective. In D. M. Wenger & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 506–527). Englewood Cliffs, NJ: Prentice-Hall.
- 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.
- Windle, M., & Lerner, R. M. (1986). Reassessing the dimensions of temperamental individuality across the life span: The Revised Dimensions of Temperament Survey (DOTS—R). *Journal of Adolescent Research*, *1*(2), 213–229.
- White, G. M. (1980). Conceptual universals in interpersonal language. *American Anthropologist*, *82*, 759–781.