Journal of Adolescent Research
http://jar.sagepub.com

Latent-Variable Confirmatory Factor Analysis of the Adolescent Temperament Questionnaire

Lawrence M. Scheier, Robin J. Casten and William Fullard

Journal of Adolescent Research 1995; 10; 246
DOI: 10.1177/0743554895102004

The online version of this article can be found at:
http://jar.sagepub.com/cgi/content/abstract/10/2/246

Published by:
SAGE Publications
http://www.sagepublications.com

Latent-Variable Confirmatory Factor Analysis of the Adolescent Temperament Questionnaire

Lawrence M. Scheier
Cornell University Medical College
Robin J. Casten
William Fullard
Temple University

Latent-variable confirmatory factor analysis was used to examine the dimensional structure of adolescent temperament from a sample of 436 adolescents. The nine-dimensional model proposed by A. Thomas and S. Chess in 1977 was extended by including a bidimensional structure of Mood (including both positive and negative affectivity) and a dimension of Ego Control tapping resiliency and flexibility. All 11 dimensions were statistically reliable and a superior fit was obtained with a correlated, rather than an orthogonal, model. Results indicated that the factors of Threshold, Intensity, and Distractibility may not be developmentally consistent nor conceptually homogeneous. Moreover, model fit statistics underscored that a single nomothetic model could not adequately account for the variability underlying the temperamental styles of these adolescents. Separate primary models reflecting temperamental "systems" of Cognitive-Diligence, Sociability/Resilience, and Vigor/Mobility also were tested. Findings suggest that extension of temperament assessment from infancy and childhood to adolescence is developmentally appropriate, although further elucidation of temperamental styles consistent with adult personality is warranted.

Temperament long has been recognized as an important component of socioemotional development in early life. A substantial influence in the American temperament research community comes from the pioneering work of Thomas and Chess and their colleagues (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Hertzig, & Korn,

This research was supported in part by a grant to Barbara Schraeder, Ph.D. by the National Institute of Mental Health (RD1315-7). An earlier version of this article was presented at the 101st American Psychological Association National Convention, Toronto, Canada. The authors appreciate the suggestions for revision by the anonymous reviewers. We also are grateful for the support and participation of students and staff at the Germantown Academy, Philadelphia, PA and Haddonfield Middle School of New Jersey. Robin Casten is currently at the Philadelphia Geriatric Center, Polisher Research Institute, Philadelphia, PA.

Journal of Adolescent Research, Vol. 10 No. 2, April 1995 246-277

1963). Based on their longitudinal study of development, the New York Longitudinal Study, Thomas and Chess elucidated the structure and clinical significance of temperament. Their contribution also included establishing a theoretical framework from which to understand linkages between temperament and individual differences in psychological development.

One distinguishing component of the research of Thomas and Chess (1977) is that they defined temperament as the stylistic component or *how* of behavior, rather than emphasizing motivational or ability-related causes or functions. Their nine dimensions describe individual differences with respect to sensory threshold; intensity of mood expression; distractibility to extraneous stimuli; task persistence; adaptability to the requirement of change; approach/withdrawal from new situations, demands, or environments; positive and negative mood; rhythmicity or regularity of biological functions; and activity level (Chess & Thomas, 1991).

Some Essential Questions Regarding the Appropriateness of Temperamental Constructs for Adolescence

Two distinct problems are associated with applying conceptualizations of infant temperament to older individuals. First, the notion that adolescence is a developmental bridge between childhood and adulthood in which temperamental traits "meld" into personality has been extended theoretically, but not rigorously tested empirically (Buss, 1989; Prior, Crook, Stripp, Power, & Joseph, 1986). Several differing conceptual positions have been extended that address the correspondence between temperament and personality, in part suggesting that temperament might be a subset of personality (Berger, 1982; Hofstee, 1991; Strelau, 1983). Although adolescence is often characterized by the emergence of new and different socializing forces (i.e., the shift from parental to peer influence), adolescence may represent a transitional interface between temperamental styles and personality, a finding that has received some empirical support (Prior et al., 1986; Windle, 1989). If this is the case, then assessment of temperament in adolescence represents an important bridge from which to better appreciate and understand personality development between late childhood and adulthood.

In addition, several researchers have suggested that attention needs to be paid to the operationalization of key temperament constructs (Hubert, Wachs, Peters-Martin, & Gandour, 1982). Hubert et al. have suggested that studies concerned with the factorial validity of infant and early childhood assessments have not provided clear empirical support nor connections to their respective theoretical orientations. In particular, they cite the lack of evidence for consistently replicable marker factors (i.e., factors that can be tied to

theoretical constructions) as "further indication of the weak operational status of current approaches to measuring infant and child temperament" (Hubert et al., 1982, p. 580). Although this comment is mostly reserved for problems associated with studies of early temperament, similar methodological and conceptual problems are recognized with older individuals as well (Windle, 1988).

Several intriguing directions are suggested by this latter point. First, most studies have relied on exploratory data analysis to examine the structure of temperament. Tanaka and Huba (1984) have commented that, among other things, structures obtained by exploratory factor analysis may be influenced by choice of rotation (their treatment of this problem notably was constrained to analyses of affective measures). Thus any different number of primary or secondary structures can be obtained depending on whether an oblique or orthogonal rotation was implemented. The lack of appreciable factor congruence for models of both infant and adolescent temperament presents an important problem for researchers interested in establishing conceptual linkages between temperament and personality literatures (e.g., Angleitner & Riemann, 1991).

Second, the psychometric adequacy of most temperament assessments have been based primarily on the use of observed measured variables, which are usually construed at the item level or as composite indexes. A shortcoming of this approach is the inclusion of "measurement error," which attenuates reliabilities and correlations (Bollen, 1989). To a large degree, these statistical and psychometric biases directly influence interpretations of dimensional structures and may even affect generalizability. Third, many attempts to develop psychometrically adequate assessments of temperament in adolescence have embarked from a similar empirical vantage point as did Thomas and Chess (1977). As a result, the lack of a clear and consistent theoretical approach for operationalizing key constructs can only hamper validation efforts (Skinner, 1986).

Temperament and Personality in Adolescence

According to lifespan theorists, adolescence is a transitional period marked by several developmental milestones. Foremost among these are formation of an identity and development of cognitive self-regulatory skills (Erikson, 1968; Keating, 1990; Marcia, 1980). Adolescence represents a period in which emphasis is placed on consolidation of earlier—and the emergence of new—social, emotional, and cognitive skills as well as the experience of rapid physical growth associated with puberty (Simmons & Blyth, 1987). During adolescence, there is evidence of reorganization of the

self and rapid transitions in personality (Eccles, et al., 1989). Furthermore, essential linkages between cognitive abilities and emotion become solidified and may manifest themselves as individual differences or behavioral styles.

Ego Control and Ego Resiliency

During this prolonged period of individuation and increasing cognitive self-regulation, many disparate facets of growth and development blend together into a coherent whole or self (Blos, 1979; Josselson, 1980). Specifically, this is a period marked by the onset of formal operational thought (representing more complex and abstract thinking skills), peaking and diminishing of egocentric thinking, recognition of the self in the context of a social milieu, and establishment of a mature identity (Erikson, 1968; Seltzer, 1989). During adolescence, mechanisms of cognitive and affective impulse control (i.e., rigidity or flexibility) become accentuated and linked to a host of behaviors, including resiliency, academic competence, and mental health (Funder & Block, 1989; Rutter, 1987; Windle, 1991; Windle et al., 1986).

Although some theorists have characterized adolescence as a period of personality disequilibrium (see Rutter, Graham, Chadwick, & Yule, 1976, for a discussion), others have noted the marked crystallization of personality features and renewed emphasis on the role of self in society (e.g., Erikson, 1968). Thus, given the tendency during this age period toward unification of behavioral styles into a more coherent self, an important question is whether models of temperament that favor distinct (i.e., orthogonal) stylistic components are appropriate for characterizing adolescent temperament. To some degree, models that support the coalescing of behavioral styles are suggesting a stronger connection between cognitive self-monitoring functions and affectively based emotional systems.

Current Assessments of Temperament in Adolescence

Several researchers have tried to bridge the operationalization of key temperamental constructs derived from infants to older aged samples corresponding to the years 12 through 18. For example, Windle and his colleagues (Windle, 1992; Windle & Lerner, 1986) have developed the Dimensions of Temperament Survey-Revised (DOTS-R), a 54-item questionnaire for assessing temperament both in adolescents and young adults. The DOTS-R both extends and refines the nine-dimensional model originally proposed by Thomas and Chess (1977). Although certain dimensions in the DOTS-R correspond with the nine-dimensional model proposed by Thomas and Chess (i.e., Approach/Withdrawal, Distractibility, Mood, Persistence), others are

not present (Adaptability, Intensity, Threshold), are greatly modified (Rhythmicity and Activity were divided into multiple subscales), or added (Flexibility-Rigidity). Notwithstanding, Windle (1992) obtained well-fitting primary and second-order latent-variable models using confirmatory techniques, albeit the constructs specified in his model departed somewhat from the exact nine-dimensional structure advocated by Thomas and Chess (additionally, their primary factor structure was hypothesized to be orthogonal).

Hegvik, McDevitt, and Carey (1982) developed the Middle Childhood Temperament Questionnaire (MCTQ) based on the nine temperament categories proposed by Thomas and Chess. The MCTQ is a 99-item Likert-type parental rating scale designed for children 8 through 12 years of age (the upper range overlaps with early adolescence). Item content analysis of the MCTQ shows many of the items to be task or situation specific (i.e., Persistence: "stays with homework until finished") and other items conceptually broad enough to perhaps load on more than one factor (i.e., "reacts strongly to a disappointment or failure" reflected Mood, whereas "shows strong reactions when pleasantly surprised" reflected Intensity). Results from factor analyses of this scale have only partially confirmed the nine dimensions in both United States and German samples (Czeschlik, 1992; McClowry, Hegvik, & Teglasi, 1993). For example, Czeschlik reported an overlap of 65% for marker variables from the nine dimensions reported by Hegvik et al. (1982) and suggested that an eight-factor solution might best describe the structure of the German version of the MCTQ. Based on their findings, Czeschlik concluded that only partial support could be garnered for the nine-factor model, because she obtained two new factors (Persistence and Organization) and reported substantial disparity in item correspondence from the original MCTQ. Likewise, McClowry et al. (1993), using the MCTQ, found support for only a five-factor model.

Finally, Capaldi and Rothbart (1992) developed and validated the Early Adolescent Temperament Questionnaire (EATQ). The EATQ departed somewhat from the nine-dimensional model of Thomas and Chess and incorporated the dimensional structure proposed by Derryberry and Rothbart (1988) for adult temperament characteristics. Items in the EATQ were rewritten to be age appropriate and tapped three general constructs of emotionality, reactivity, and activity or self-regulation (these were subsequently decomposed into subconstructs). As a result, the 92-item EATQ included 11 scales assessing sensitivity (i.e., threshold to detect stimulation and somatic arousal); autonomic reactivity (i.e., physical reactions related to tension, stress, excitement); motor activation (behavior related to somatic arousal); fear, irritability, shyness, sadness (all four scales tapping negative emotionality); high- and low-intensity pleasure (both scales assessing positive emo-

tionality); activity level (i.e., impulsive and controlled behavior); and attentional control (i.e., focusing and shifting), the latter two scales tapping self-regulation.

Through principal axis factor analysis, the authors obtained three conceptual dimensions corresponding to negative aspects of temperament, positive aspects of temperament, and behavioral inhibition (the first two dimensions were replicated on a second and slightly larger sample). Overall, the downward extension of the adult version of the EATQ required extensive modification (scales were either eliminated, collapsed, or did not replicate on the adolescent sample). Many of the scales in the EATQ are psychometrically sound; however, they are not conceptually unidimensional (sadness, fear, attention, motor activation, and irritability loaded together on a single dimension), making it difficult to map to other valid measures or theoretical positions.

Conceptualization of Mood: Positive and Negative Affectivity

In addition to the methodological and conceptual problems already mentioned, a recent literature examining self-rated adult mood provides evidence that mood is best conceptualized as bidimensional, rather than bipolar and unidimensional as previously considered (e.g., Warr, Barter, & Brownbridge, 1983; Watson & Tellegen, 1985; Watson, Clark, & Carey, 1988; Watson, Clark, & Tellegen, 1988). The two constructs of mood include negative affectivity, composed of negative affective states such as distressed, nervous, anxious, or angry, and positive affectivity, reflecting pleasurable engagement and enthusiasm (e.g., carefree, excited, lively, and content). Recognition that the structure of mood is perhaps best conceptualized as bidimensional is primarily based on adult samples, although some evidence has confirmed the bidimensionality of mood in children (King, Ollendick, & Gullone, 1991; Wolfe et al., 1987). Contrary to current conceptualizations, Thomas and Chess have suggested that mood is a bipolar unidimensional construct, with negative mood reflected by the absence of positive affect.

Importance of the Current Study

To address these important theoretical and methodological concerns, previous research was extended in several ways: First, findings were included from the adult literature pertaining to the structure of mood. Second, a dimension of ego control tapping resiliency and flexibility was included, which was hypothesized as essential toward obtaining a more complete picture of temperament during adolescence. Third, latent-variable confirma-

tory factor analysis was used, which permitted a more statistically rigorous empirical evaluation of several alternative conceptualizations of temperament (Bentler, 1978, 1980).

METHOD

Sample and Procedures

Data were obtained from a sample of 436 adolescents attending two schools, one in suburban New Jersey and a private nonsecular school in suburban Philadelphia. Based on their grade equivalents, students ranged from 12 to 18 years of age. Of the respondents, 11.5% were in 7th and 8th grades, 55.5% were in 9th and 10th grades, and 33.0% were in 11th and 12th grades. The study originally had been intended to equally balance students between the junior and senior high school grade levels; however, because the study was conducted during the spring, senior high school students at the suburban Philadelphia school were unavailable for assessment (they participate in an off-campus spring project). The small percentage of early middle school students (7th and 8th grades) included was inadvertently produced by teachers surveying classes with respondents below the required grade levels. Overall, 182 of 330 students from the suburban Philadelphia school were surveyed, and 254 of 545 students in the New Jersey school were surveyed. There was no follow-up procedure to acquire data from students absent on the day of administration. Fifty-five percent of the total sample were male and the racial composition was predominantly White.

A passive consent procedure was used in both schools, entailing a letter sent home to each parent describing the goals of the study and the content of the Adolescent Temperament Questionnaire (ADTQ). Parents and students both were informed of the confidential nature of the data collected in the study and encouraged to discuss the study with school administrators if any questions existed. Two parents in one school refused to allow their adolescents to participate. Students responded to the questionnaire in homeroom, study periods, or during lunch recess within their school. The ADTQ was administered jointly by the authors with the assistance of school personnel (teachers). All teachers had been informed, through memo and discussions conducted during faculty meetings, regarding the study protocols and research goals. In general, these schools were quite favorable toward school-based research and had informal institutional review boards, which facilitated implementation of this study.

Instrumentation

Item construction for the 70-item ADTQ was primarily based on the nine-dimensional Thomas and Chess (1977) model of temperament. Items were written to address various broad content domains, rather than specific situational tendencies. In addition, to avoid any confusion between the "stylistic" component of behavior as opposed to "frequency" of behavior, item content emphasized intensity, rather than how often a person engaged in the activity. For instance, rather than phrasing an item, "I often fidget when doing quiet tasks," which reflects Activity and taps frequency of behavior, the item was phrased as, "I fidget during quiet activities," which taps the "how" of behavior.

Overall, 11 primary latent constructs were hypothesized, including Adaptability, Approach/Withdrawal, Activity, Rhythmicity, Threshold, Intensity, Persistence, Distractibility, Positive Mood, Negative Mood, and Ego Control. Response formats for all items ranged from 1 = *never* through 4 = *always*, with the exception of the mood items, which were scored from 1 = *never true* through 4 = *always true*. To limit the influence of response set, some items were intentionally reverse coded.

Six items were used to reflect a construct of Adaptability (e.g., "I am comfortable with change" and "I can adjust to any situation"). In general, Adaptability taps willingness to accept change and ability to adjust to changes in the surrounding environment. A latent factor of Approach/Withdrawal was reflected by four items (e.g., "It's easy for me to talk to people that I don't know" and "I enjoy interacting with other people"). Approach/Withdrawal taps both a willingness to approach people socially and an inquisitiveness toward strange environments (e.g., "It's hard for me to go to unfamiliar places").

A latent factor of Activity was reflected by four items tapping general activity level (e.g., "I sit quietly while waiting" and "I fidget during quiet activities"). Items tapping reflective thought also were included (e.g., "I daydream and think about things not related to what I am doing") in an effort to tap emergent cognitive activities consistent with this period of adolescent development. A latent factor of Threshold was reflected by three items tapping both sensitivity toward changes in other people (e.g., "I notice when people change something physically about themselves") and detection of climatic and environmental change (e.g., "I easily notice changes in the outdoor temperature"). Intensity was reflected by five items; four of which tapped the emotional reactions and content (e.g., "I react strongly when I am surprised") and one that tapped intensity of movement ("I don't make a lot of noise when moving around").

A latent factor of Persistence was reflected by seven items (e.g., "I stick with one thing until it is done" and "When working on a task, I give up when I feel frustrated"). Items tapping frustration and inability to stay on task were reverse coded. To distinguish persistence from a more specific (and school-related) task orientation, only items assessing persistence as a general behavioral style were used. A latent factor of Distractibility was reflected by four items (e.g., "I have trouble concentrating when I am upset" and "I have difficulty switching from one activity to another"). As with Persistence, items that might encourage situation-specific answers were avoided, whereas items that tapped a general dimension of Distractibility were included.

A latent factor of Rhythmicity was reflected by five items tapping diurnal cycles based on sleep and hunger (e.g., "I get tired the same time every night") and daily routines (e.g., "I have the same morning routine every day"). Finally, separate latent factors of Negative and Positive Mood were hypothesized tapping, respectively, negative affectivity (e.g., irritable, nervous, and distressed) and positive affectivity (e.g., enthusiastic, cheerful, and lively). These items were selected from word lists associated with a number of circumplex models that have demonstrated empirically that these adjectives are reliable indicators of negative and positive mood.

RESULTS

Gender and Grade Differences in Temperament

Table 1 contains summary descriptive statistics and item-scale correlations for the items included in the analyses. Although most of the item-scale correlations are modest (average item-scale correlation is .35 across the 11 scales), several are notably low; however, any further discussion of these associations is addressed following the confirmatory factor analyses (CFA).

Point-biserial correlations between gender and composite scale scores are contained in the right-hand corner of Table 1. Overall, the largest significant mean difference based on gender was for Approach/Withdrawal, which only accounted for 3% of the variance. The average absolute mean difference (i.e., percentage of variance) across all composite scales based on gender was 2.3%. Although mean gender differences across a majority of the composite scales were ever so slight, it is still worth noting that males reported they were more active, had lower thresholds for sensory stimuli, were more adaptable to changing situations, approached strange situations/people more readily, reacted more intensely, were more flexible and resilient to change, and had higher levels of both positive and negative affect.

Using analysis of variance mean differences were examined for the same composite scores by grade, contrasting early (7th and 8th), middle (9th and 10th), and late adolescence (11th and 12th), and testing Gender × Grade interactions. A significant interaction was found only for Approach/Withdrawal, $F(1, 435) = 6.96$, $p < .01$, with older males (junior/seniors) having the highest mean overall. Main effects by grade were obtained for Persistence, $F(1, 435) = 3.97$, $p < .05$, with early adolescents having the highest mean overall and differing significantly from older adolescents (11th and 12th grades). The absence of pronounced mean differences and few observed significant interactions supported the decision to analyze the combined data. Moreover, the use of pooled sample data produces more reliable and stable estimates when using confirmatory techniques (Tanaka, 1987).

Confirmatory Factor Analysis

The first model tested the adequacy of constraining all the items to load on a single dimension. This model represents the lowest level of dimensionality in the series of "restricted" confirmatory models and is usually regarded as the baseline model against which all other multidimensional models can be statistically contrasted.[1]

The unidimensional model poorly fit the data, $\chi^2(1430, N = 436) = 4144.56$, $p < .001$, Comparative Fit Index (CFI) = .452, and had a χ^2/df ratio exceeding 2.90. For the most part, standardized factor loadings were significant with the exception of three intensity items, one indicator of activity, three rhythmicity items, and one indicator of negative mood, all of which were nonsignificant. Overall, this model was not expected to provide the best fit. However, it represents a good starting point in the analytic process to establish the multidimensional nature of temperament.

Next, the nine-dimensional model proposed by Thomas and Chess (1977) was tested with a few minor modifications. First, a 10th factor was added to capture Ego Control (i.e., flexibility and resilience). Items reflecting both positive and negative mood were constrained to load on a single bipolar construct of Mood. This model substantially gained in the CFI and overall model fit as compared to the previous model, $\chi^2(1385, N = 436) = 3148.9$, $p < .001$; CFI = .644, $\chi^2/df = 2.27$. The significant nested difference ($\Delta\chi^2$) between the 10-dimensional and unidimensional models reinforces the improvement gained with a multidimensional structure, $\Delta\chi^2(45, N = 436) = 995.7$, $p < .001$.

Items were then separated conceptually according to whether they reflected negative and positive mood (and constrained accordingly to load

TABLE 1: Summary Descriptive Statistics and Psychometric Properties for Temperament Items and Composite Scales

Latent Construct and Measured Variable	*Mean*	*Item-Scale Correlation*	SD	*Skew*	*Kurtosis*	*Mean Gender Difference*[a] r_{pb}
Adaptability (4)[b]	2.91		.44	−.53	1.50	−.05
Comfortable with change	2.80	.42	.69	−.30	.13	
Can adjust to any situation	2.95	.40	.63	−.40	.76	
Adjust quickly to changes	3.00	.39	.58	−.50	1.75	
Can easily compromise	2.90	.19	.80	−.53	.06	
Approach/Withdrawal (4)	3.11		.53	−.52	.43	−.17***
Easy to talk to strangers	2.75	.50	.84	−.20	−.58	
Hard to go to unfamiliar places	1.99	.33	.81	.59	.00	
Enjoy interacting with others	3.45	.48	.66	−1.03	.92	
Have trouble making friends	1.76	.43	.73	.79	.57	
Activity (4)	2.53		.55	.34	−.02	−.04
Sit still while in car	2.72	.34	.82	−.52	−.12	
Sit quietly while waiting	2.62	.39	.86	−.24	−.56	
Fidget during quiet activities	2.53	.42	.81	.15	−.52	
I daydream and think about things not related to what I am doing	2.94	.23	.84	−.19	−.94	
Threshold (3)	3.14		.47	−.26	.22	−.11*
Notice when people change something about themselves	3.23	.23	.65	−.47	.18	
Can easily detect changes in people's voices	3.06	.17	.74	−.40	−.26	
Easily notice changes in the outdoor temperature	3.13	.15	.75	−.55	−.06	
Persistence (7)	2.84		.44	−.50	2.06	.00
Stick with one thing until it is done	2.79	.47	.67	−.38	.35	
Can return to work when interrupted	2.91	.41	.68	−.45	.54	
Give up on task when I feel frustrated	2.34	.44	.81	.45	−.20	

Finish all projects that I begin	2.99	.49	.68	−.43	.44	
See things through to the end	2.89	.51	.64	−.87	1.83	
Don't give up when working on a difficult task	2.66	.42	.80	−.05	−.51	
Keep trying when things don't work out	3.00	.47	.64	−.53	1.13	
Intensity (5)	2.71		.43	.19	.18	−.10*
React strongly when surprised	2.71	.19	.71	.02	−.36	
Yell or scream loudly when in pain	2.27	.18	.83	.42	−.27	
I am a soft-spoken person	2.27	.20	.91	.18	−.80	
Don't make a lot of noise when moving around	2.53	.24	.79	−.06	−.42	
Laugh loud when something strikes me as funny	3.35	.16	.73	−.93	.47	
Distractibility (4)	2.36		.43	−.12	.07	−.05
Finish a task even when distractions are present	2.67	.27	.70	−.27	−.03	
Have trouble concentrating when upset	3.04	.23	.78	−.57	.02	
Can easily do two things at once	2.77	.28	.78	−.33	−.19	
Have difficulty switching from one activity to another	1.86	.13	.59	.45	1.58	
Rhythmicity (5)	2.65		.54	−.03	.06	.03
Like having structure in my life	2.87	.27	.82	−.43	−.26	
Get tired same time every night	2.37	.43	.85	.15	−.57	
Get hungry same time everyday	2.62	.41	.87	−.19	.60	
Have the same morning routine	2.94	.38	.91	−.49	−.60	
Am alert and peppy same time(s) everyday	2.45	.38	.87	.15	−.64	
Positive mood (6)	2.99		.50	−.36	.64	−.14**
Enthusiastic	3.08	.57	.70	−.40	.02	
Cheerful	3.02	.55	.68	−.41	.37	
Lively	3.08	.57	.75	−.43	−.27	
Content	2.84	.32	.77	−.42	−.04	

(continued)

TABLE 1: Continued

Latent Construct and Measured Variable	*Mean*	*Item-Scale Correlation*	SD	*Skew*	*Kurtosis*	*Mean Gender Difference*[a] r_{pb}
Positive mood (6)						
Good natured	3.23	.45	.74	−.73	.26	
Carefree	2.70	.41	.92	−.13	−.85	
Negative Mood (7)	2.21		.46	.88	1.63	.04
Irritable	2.25	.39	.65	1.14	1.53	
Nervous	2.22	.47	.75	.72	.51	
Worrisome	2.31	.48	.85	.48	−.32	
Distressed	2.15	.48	.74	.59	.49	
Anxious	2.61	.31	.83	.17	−.69	
Angry	2.12	.44	.65	1.07	2.17	
Panicky	1.82	.50	.79	.97	.91	
Ego Control (6)	3.13		.43	−.59	1.53	−.08
Like to explore new things	3.22	.37	.68	−.61	.45	
Having goals is important	3.31	.32	.80	−1.08	.77	
There are many ways to see things	3.23	.31	.74	−1.03	.97	
Express my feelings outwardly	2.93	.20	.85	−.29	−.74	
Bounce right back after setback	3.02	.42	.66	−.45	.67	
Can handle stressful things in life	2.98	.34	.78	−.72	.54	

NOTE: $N = 436$.

a. Males coded (2) and females coded (1), so that negative correlations indicate females with higher means.

b. Number of items in scale.

*$p < .05$; **$p < .01$.

separately on dimensions of Positive and Negative Mood). The model fit indexes, $\chi^2(1375, N = 436) = 2805.20$, $p < .001$; CFI = .71, $\chi^2/df = 2.04$, and significant improvement on the 10-dimensional model, $\Delta\chi^2(10, N = 436) = 343.7$, $p < .001$, substantiated the conceptual distinction between elements of positive and negative mood. The two dimensions of mood correlated –.67, which is appreciably less than unity, and extends support to the notion that mood is best conceptualized as bidimensional (e.g., Warr et al., 1983; Watson et al., 1988). Error-free factor intercorrelations from the 11-dimensional model are contained in the lower triangle of Table 2. The upper triangle of this same table also contains correlations among the psychometric (observed) scales formed for each dimension.

The matrix of associations among linear composites is presented with caution, because their interpretation is tenuous for two reasons: (a) Current research with the ADTQ is considered exploratory even in light of the use of confirmatory techniques for model testing; and (b) the use of CFA techniques is intended to delineate "hypothetical" constructs mapped conceptually with those proposed earlier by Thomas and Chess (1977). Therefore, linear scales are not to be regarded as indicative of greater amounts of the "dimension" in question, because most of the scales are considerably heterogeneous.

Figure 1 depicts the factor structure and standardized parameter loadings for the 11-dimensional model. As depicted, all factor loadings based on this model were significant. For the most part, factor composition was fairly uniform and moderately high, indicating the psychometric soundness of the hypothesized factors. Several exceptions were observed, including two items on Intensity, "I react strongly when I am surprised" and "When I am in pain, I yell or scream loudly," both which had comparatively lower loadings (–.243 and –.231, respectively). Likewise, "I have difficulty switching from one activity to another" loaded poorly on Distractible (.257), as did "I am the kind of person who can easily compromise" on Adaptability (.257). Importantly, these relatively lower loadings reflect two features: (a) that the specific item is not as strong an indicator of the hypothesized dimension given the shared variances of the remaining items reflecting that specific factor; and (b) the latent factor is more than likely heterogeneous in composition, especially given the statistical significance, albeit differential loading, of each of the indicators (i.e., on inspection of the zero-order correlation matrix all items within-scale were significantly correlated).

Finally, much prior research using exploratory factor analysis relied on variance maximization procedures, which entails extracting orthogonal factors as opposed to oblique rotations with correlated factors. To adequately test the hypothesis that the 11 dimensions are orthogonal, a model that posited

TABLE 2: Composite Scale (upper triangle) and Latent Factor (lower triangle) Intercorrelations From the 11-Factor Confirmatory Factor Analysis Model of Adolescent Temperament

Factor	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1	(.576)*[a]	−.08[b]	.21***	−.36***	−.14**	.03	.31***	−.12*	.06	.17***	−.09
2	.20*	(.335)	−.11*	.25***	.24***	.24***	.09	.07	.21***	.01	.23***
3	−.36***	−.51***	(.433)	−.38***	−.29***	−.16***	.11*	−.01	−.18***	.13**	−.31***
4	.51***	.49***	−.85***	(.740)	.35***	.23***	−.14**	.13**	.27***	−.17***	.44***
5	.18*	.54***	−.84***	.55***	(.569)	.44***	.01	.02	.47***	−.23***	.46***
6	−.07	.53***	−.49***	.37***	.68***	(.649)	.28***	−.02	.46***	−.20***	.47***
7	.68***	−.18	−.11	.23**	−.09	−.64***	(.383)	−.06	.22***	.04	.13**
8	.17*	.11	.08	.16*	.03	−.06	.14	(.623)	.09	.07	.04
9	.03	.43***	−.49***	.40***	.61***	.72***	−.45***	.10	(.733)	−.26***	.48***
10	−.21**	−.01	.31***	−.23***	−.35***	−.32***	.04	.10	−.33***	(.726)	−.26***
11	.17*	.49	−.83***	.67***	.80***	.68***	−.24**	.04	.76***	−.49***	(.589)

NOTE: F1 = Activity; F2 = Threshold; F3 = Distractibility; F4 = Persistence; F5 = Adaptability; F6 = Approach/Withdrawal; F7 = Intensity; F8 = Rhythmicity; F9 = Positive Mood; F10 = Negative Mood; F11 = Ego Control.

a. Numbers in parentheses are internal consistency estimates computed by Werts, Linn, and Jöreskog's (1974) formula that adjusts for latent factor variances (and are nearly identical to Cronbach alphas).

b. Scale correlations may be in opposite direction to latent factor correlations because items have been reverse coded (i.e., "more of") in scales, whereas CFA models produce factor loadings, which can be both negative and positive.

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

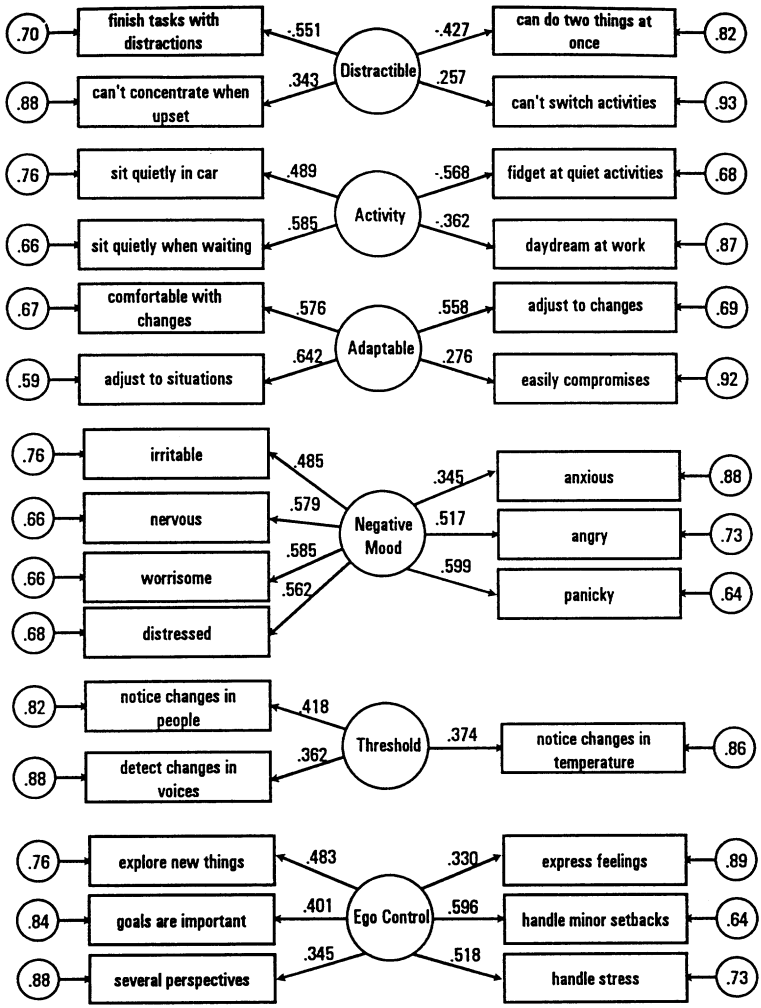
correlated factors was compared with an uncorrelated version. The orthogonal version of this model substantially degraded the fit compared with the correlated model, $\chi^2(1430, N = 436) = 3923.98, p < .001$; CFI = .497, [$\Delta\chi^2(55, N = 436) = 1118.78, p < .001$].

Distinguishing Systems of Temperament and Higher-Order Factor Structures

It is important to note that none of the primary models achieved an adequate model fit ($p > .05$), underscoring that other equally plausible model structures might apply to these data. The current research examined several conceptual improvements to earlier views of temperament and contrasted statistically these improvements using nested hierarchical tests. Notwithstanding, the model fit statistics derived from the "best" 11-factor model were less than optimal and indicated several research directions worth pursuing. First, a closer inspection of the residual matrices and modification indexes provided by the LaGrange Multiplier (LM) test (Chou & Bentler, 1990) corresponding to the 11-factor model indicated that some reparameterization might enhance the overall model fit.[2]

Several minor additions to a pared-down and smaller version of the current model might improve the fit indexes and be defensible statistically (given the reduced number of estimated parameters), as well as conceptually. Testing this aspect of the ADTQ represents an important step toward scale refinement for two essential reasons: (a) because redundant items could then be eliminated making the ADTQ more efficient, and (b) the fine-tuning and model enhancement process could reveal more about the conceptual overlap between possible dimensional "systems" of temperament that have been regarded previously as distinct (especially in younger ages), but that are related more than likely developmentally in older youth. Therefore, prior to conducting any second-order tests (factor intercorrelations among the 11 primary indicated the necessity of testing the hypothesis that the structure of temperament is generated by one or more higher-order constructs), a series of restricted primary models were tested followed by an hierarchical test of a second-order structure.

Based on an examination of the error-free factor intercorrelations from the 11-dimensional model, a model was posited that included primary factors of Adaptability, Approach/Withdrawal, Positive Mood, Negative Mood, and Ego Control. The aggregate profile among these factors underscored that the temperamental styles of many youths could be described as "Sociable and Resilient." An adequate fit for this five-factor model was achieved, $\chi^2(314,$



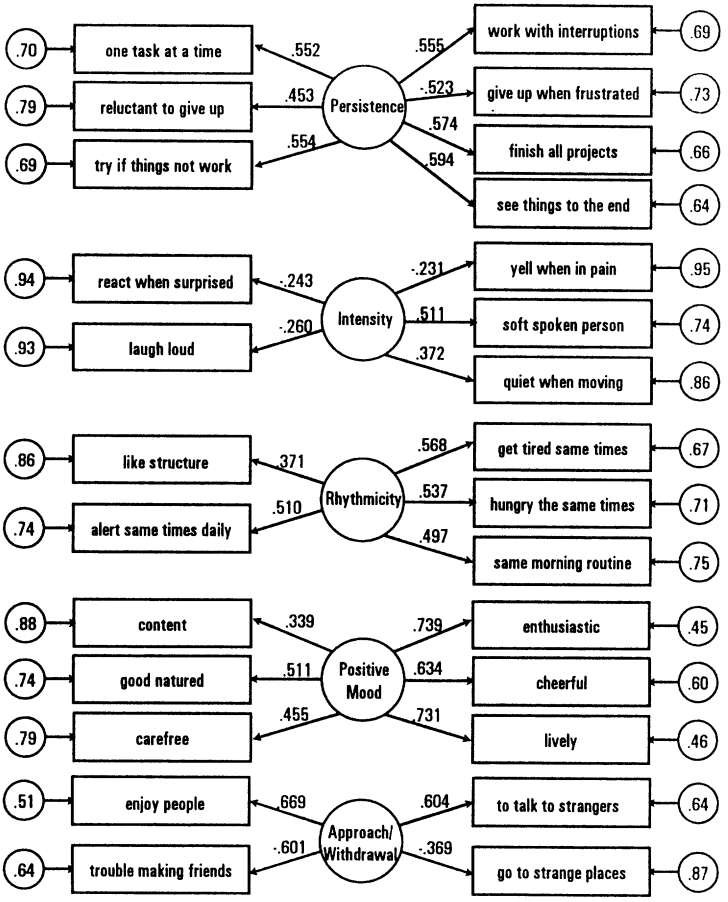


Figure 1: Confirmatory Factor Analysis Model Depicting 11-Factor Model of Temperament

NOTE: Large circles are latent constructs, rectangles are measured variables. Small circles with unidirectional arrows are residual variables (variances). Parameter estimates are standardized and significance levels are based on critical ratios. All loadings significant at $p < .001$.

$N = 436) = 809.21, p < .001$; CFI = .806, root mean square residual (RMSR) = .04, $\chi^2/df = 2.58$, although the significant p value (.001) and the χ^2/df greater than 2.0 indicated that some modifications (i.e., correlated disturbances) might capture more residual variance (albeit not a large amount because the RMSR was small [.04], indicating that little meaningful covariation remained).

Based on the stepwise multivariate LM test, a number of residual covariances were then added. Table 3 contains correlations among the residual variances from the restricted five-factor model (i.e., expressing a temperamental style of Sociability/Resilience). Following the addition of 20 residual covariances, a model fit threshold was achieved that is considered to be satisfactory, $\chi^2(294, N = 436) = 546.63, p < .001$; CFI = .901, $\chi^2/df = 1.86$ (Bentler, 1990; Marsh, Balla, & McDonald, 1988).

The addition of the covariances among the residual terms provides a more refined analysis of specific temperamental styles after statistically controlling for the associations among the general constructs of temperament. In the case of the first restricted analysis, the majority of the 20 covariances were cross-construct (only 5 were within-construct). These adjustments primarily indicated the need for specific (and perhaps stylistic) relationships between positive and negative mood, adaptability, approach/withdrawal, and ego control (resilience and flexibility).

The same type of analysis was conducted for a second group of factors determined to be moderately associated based on the factor intercorrelations from the 11-factor model. In this model, potential associations among residual variances corresponding to the primary factors of Distractibility, Threshold, and Persistence were examined (based on the empirical modification indexes provided by the LM test). Together these factors are thought to reflect a dimension of "Cognitive-Diligence," a system that should be conceptually distinct from the more affectively laden system reflected by "Sociability/Resilience."

This model also initially fit adequately, $\chi^2(74, N = 436) = 205.75, p < .001$; CFI = .843, $\chi^2/df = 2.78$, and all the parameter loadings were significant. In the restricted model, Persistence and Distractibility were strongly correlated ($r = .83$), and the remaining primary factor correlations were more moderate (Distractibility and Threshold = .49; Persistence and Threshold = .49, $p < .001$, respectively). A total of six residual covariances were added based on the LM test, and these also are contained in Table 3. With the addition of these few covariances, a substantial improvement in the model fit was achieved, $\chi^2(68, N = 436) = 121.75, p < .001$; CFI = .936, $\chi^2/df = 1.79$. Again, five of the six model additions represented cross-construct relationships further

TABLE 3: Correlated Uniquenesses From Tripartite System of Temperament

Item Content (factor) Covariance	*Standardized Residual*
Sociability-Resilience System Model	
Irritable (NM), angry (NM)	.35***
Easy to talk to unfamiliar people (AW), express feelings outwardly (EC)	24***
Irritable (NM), can easily compromise (AD)	−.15**
Many ways to see things (EC), distressed (NM)	.23***
Good natured (PM), can easily compromise (AD)	.17**
Handle stressful things in life (EC), adjust to any situation (AD)	.18**
Handle stressful things in life (EC), bounce back after setback (EC)	.19**
Easy to talk to unfamiliar people (AW), not go to unfamiliar places (AW)	−.18***
Having goals is important part of life (EC), worrisome (NM)	.17**
Angry (NM), kind of person who can easily compromise (AD)	−.20**
Cheerful (PM), irritable (NM)	−.15**
Carefree (PM), comfortable with change (AD)	.16**
Like to explore new things (EC), hard to go to unfamiliar places (AW)	−.20***
Anxious (NM), lively (PM)	.16**
Worrisome (NM), nervous (NM)	.17**
Lively (PM), nervous (NM)	−.15**
Like to explore new things (EC), easy to talk to unfamiliar people (AW)	.15**
Like to explore new things (EC), enjoy interacting with people (AW)	.15**
Many ways to see things (EC), like to explore new things (EC)	.12*
Content (PM), worrisome (NM)	−.12*
Cognitive-Diligence System Model	
I stick with one thing (P), doesn't work out keep trying (P)	−.22***
Easily do two things at once (D), easily detect voice changes (T)	.15**
Can finish tasks even with distractions (D), can return to task (P)	.30***
Easily notice temperature changes (T), see things through to end (P)	.18**
Can't switch from one activity to another (D), easily notice changes in temperature	−.16**
Finish projects I begin (P), easily do two things at once (D)	−.16**
Vigor-Mobility System Model	
Fidget during quiet activities (AC), daydream (AC)	.19**
Don't make noise moving around (I), like having structure (R)	.17**
Alert and peppy same time every day (R), don't make noise moving (I)	.15**
Get hungry same time (R), daydream (AC)	.13*
Laugh loud when things are funny (I), daydream (AC)	.12**

NOTE: AC = Activity; AD = Adaptability; AW = Approach/Withdrawal; D = Distractibility; EC = Ego Control; I = Intensity; NM = Negative Mood; PM = Positive Mood; P = Persistence; R = Rhythmicity; T = Threshold.
*$p < .05$; **$p < .01$; ***$p < .001$.

underscoring the moderate associations among items selected to reflect these constructs. With respect to these additions, two points are worth noting. First, the size of the associations are relatively small and require further cross-replication to assure their stability. Second, in comparison to the first restricted model, where 20 such additions were required to surpass the .90 CFI criteria, only a very few additions were needed in the current model to achieve a similar benchmark.

A third systems model captured the relations among the primary factors of Intensity, Rhythmicity, and Activity. This temperamental system was construed to capture a component of physiological arousal or "Vigor/Mobility" of these youth. Overall, the initial model fit was adequate, $\chi^2(74, N = 436) = 137.87$, $p < .001$; CFI = .871, $\chi^2/df = 1.86$, and indicated that perhaps few additions would be required to achieve the .90 criteria. Five such additions were made and a final model achieved, $\chi^2(69, N = 436) = 98.70$, $p < .05$; CFI = .940, $\chi^2/df = 1.43$.

Finally, to obtain a clearer picture of the associations among all three temperamental systems, a second-order structure that integrated the three separate primary models was tested (Rindskopf & Rose, 1988). In the second-order model, each of the factors corresponding to the restricted primary models was hypothesized to be an "indicator" of a second-order factor and three higher order factors were posited.[3] Factor intercorrelations and standardized parameter loading corresponding to the second-order model are depicted in Figure 2.

Consistent with the fit of the 11-factor primary model, the fit of the second-order model, $\chi^2(1416, N = 436) = 2962.09$, $p < .001$; CFI = .688, $\chi^2/df = 2.09$, indicated that some modifications (i.e., correlated disturbances) might capture more residual variance. Factor loadings from the second-order structure were significant for all the primary factors with the exception of Rhythmicity. Associations among the three second-order constructs were small to moderate in size, ranging from a low of –.11 between Vigor/Mobility and Sociability/Resilience (marginally significant $p < .06$) to a high of .73 between Sociability/Resilience and Cognitive-Diligence ($p < .001$). The second-order model yielded additional information on the convergent and divergent nature of several dimensions of temperament. First, the association between Cognitive-Diligence and Sociability/Resilience was quite large in magnitude, indicating that perhaps these constructs might well collapse into a single dimension, reflecting environmental (i.e., school), social-interpersonal, and emotional facets of a general adaptability.

A second indication of the purity, statistical, and conceptual composition of the second-order constructs comes from an inspection of the standardized

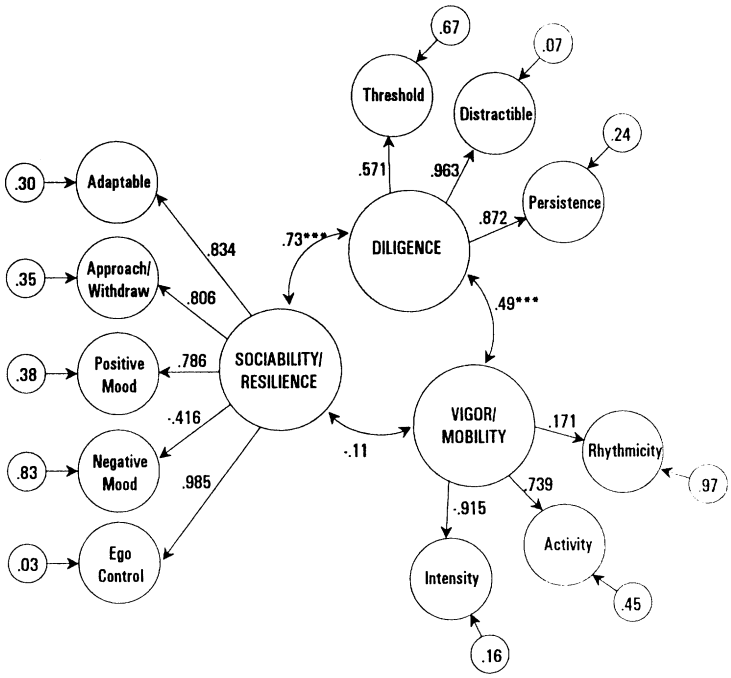


Figure 2: Confirmatory Factor Analysis Model Depicting Second-Order Latent Constructs of Sociability/Resilience, Diligence, and Vigor/Mobility
NOTE: Large circles with depict higher order constructs, smaller circles indicate primary constructs, and small circles numbers are disturbance terms (residual variances). Measured variables have not been included in figure for purposes of clarity, but can be found in Figure 1. All standardized factor loadings significant $p < .001$.
*$p < .05$; **$p < .01$; ***$p < .001$.

loadings for these dimensions. For instance, the factor loadings of both Negative Mood (constrained to load on Sociability/Resilience) and likewise, Rhythmicity (constrained to load on Vigor/Mobility) were appreciably lower than the remaining primary constructs constrained to load on these higher order dimensions. On the other hand, Ego Control reflected a large part of the variance for Sociability/Resilience and to a lesser degree Adaptability and Approach/Withdrawal.

Vigor/Mobility was most strongly reflected by Intensity (–.915), although the disparity between the loadings was not as great for Activity (.739) as it was for Rhythmicity (.171, $p > .05$). Perhaps for this age group, biological

and diurnal rhythms are less influential on overall activity and intensity of actions than for younger ages where observation of these cycles by significant caretakers provides reliable indications of behavioral style.

Notably, the standardized parameter loading for Threshold (.571) constrained to load on the second-order factor of Cognitive-Diligence was substantially lower in magnitude than both Distractibility (.963) and Persistence (.872), indicating that perhaps the cognitive undertones of Threshold sensitivity influences (i.e., interacts with) task diligence and persistence; however, some independent variance remains unique to Threshold.

With respect to the nested model tests, a comparison of the primary and second-order models (both models without residual covariances to preserve the hierarchical comparison) were significantly different [$\Delta\chi^2(41, N = 436) = 156.89, p < .001$], the decrement in incremental variance (the CFI for the 11-factor model is 71.1%, whereas the CFI for the higher order model is 68.8%), and the χ^2/df for the two respective models (2.09 for the higher order model and 2.04 for the 11-dimensional model) indicated that the second-order model did not gain any explanatory power over the 11-dimensional structure, although it is considerably more parsimonious and provides a clearer picture of the associations among the respective temperamental "systems."

DISCUSSION

Temperament researchers who adopt the position of developmental consistency suggest that structures of temperament can be extended across different age periods. The results of this study partially support this position and extend the notion beyond the middle childhood years to include the adolescent years. Furthermore, these analyses confirm that temperament in adolescence is best conceptualized as multidimensional. These multiple dimensions were moderately related and provided additional evidence for a higher order dimensional structure. Moreover, the specification of broad-band dimensions of personality as exemplified by the separate temperamental systems including Cognitive-Diligence (i.e., task orientation), Vigor/Mobility (i.e., activity and intensity), and Sociability/Resilience (i.e., flexibility, adaptability, and social extraversion) map nicely with reported findings in infant (e.g., Matheny, 1980; Rothbart & Derryberry, 1981), childhood (McClowry et al., 1993), and adult literatures (Buss & Plomin, 1975, 1984).

In their original research, Thomas and Chess (1977) used a cluster of similarly related items to describe the "easy child," exemplified as moder-

ately active, approachable, positive in mood, flexible, and adaptable. Mapping between the Thomas and Chess item level cluster analysis and the second-order factors derived in the current study, empirical confirmation of this clinically meaningful structure was obtained, albeit at a much higher level of abstraction.

Because the original summary objectives included confirming the psychometric properties of an adolescent temperament scale, along with testing and refining several theoretical constructions in this age group, these two goals are discussed accordingly. First, only partial replication of the nine-dimensional structure suggested by Thomas and Chess (1977) was obtained. That is, the dimensions of (sensory) Threshold, Intensity, and Distractibility were not entirely homogeneous and low in reliability. Likewise, error variances for the items comprising these factors were relatively large in magnitude, underscoring the lack of conceptual purity in these factors. In a preliminary set of exploratory purposes/confirmatory techniques analyses, distributions among the original six distractibility items did not overlap sufficiently enough to support a single latent factor and substantial covariation was observed in the residual matrix.

A closer examination of zero-order associations among the distractibility items indicated that perhaps two or more finer aspects of distractibility could be extracted distinguished by external stimuli (i.e., "If someone is playing music, it is hard for me to study"), internal stimuli (i.e., "I have trouble concentrating when I am upset"), activities that do not involve social cognition (i.e., "I have difficulty switching from one activity to another"), and those that involve active social cognition (i.e., "It is easy for people to get my attention when I am busy concentrating").

A similar pattern underlying the Intensity factor also was observed. Some items tapped intensity of emotional reactions to specific situations (i.e., "I react strongly when I am surprised"), whereas other items measured a more general property of intensity (i.e., "I am a soft-spoken person"), not attributed to a specific emotion or situation. Strelau (1983) conducted laboratory trials in which he examined physiological indicators along with self-report measures of temperament and reported that many facets of temperament were stimuli and/or sensory dependent (i.e., he observed strong differentiation between thresholds for auditory and visual stimuli). Although others have reported similar low internal consistency estimates for the same factors based on much younger ages (e.g., Toddler Temperament Scale) (Fullard, McDevitt, & Carey, 1984), it is unclear whether this is an artifact of measurement (i.e., the fewer numbers of items comprising these dimensions) or a result of some cognitive-affective regulatory mechanism that differentially influences receptivity to external stimuli during or before this critical age period.

Based on the current results, the poor psychometric properties for these few dimensions (i.e., Threshold, Intensity, and Distractibility) are attributed to a lack of conceptual purity and item heterogeneity. In effect, certain components of adolescent temperament are developmentally consistent with structures derived from infant samples; however, based on the current findings, more variegation exists in adolescent temperament. Such observed heterogeneity in behavioral style is consistent with developmental change, growth, and increased use of cognitive self-regulatory processes (Keating, 1990). For instance, those factors that did not replicate entirely in the current research were observed to be tapping either sensory- and situation-specific temperamental traits or encompassing both cognitive and emotion-focused styles of behavior. In light of these findings, future studies of temperamental characteristics in adolescence may want to adopt a more age-specific approach that includes formulating items that can differentiate between general expression of behavioral style and situation-specific styles.

In addition to the psychometric concerns addressed by these data, there are also several developmental considerations that need to be addressed. First, these findings underscore the appropriateness of developmental linkages between early appearing individual differences in temperament and later personality structures. For example, Buss and Plomin (1984) defined the adult temperamental category of Activity with items reflecting vigor, endurance, and intensity of participation, items corresponding to the second-order factor (and temperamental system) of Vigor/Mobility in the current study. Likewise, Rothbart and her colleagues have identified self-regulation as an important structural component of temperament in children and infants (Rothbart & Derryberry, 1981; Rothbart & Posner, 1985). Accordingly, they conceptualized self-regulation as consisting of elements of approach/withdrawal, excitement, stimulus seeking, and positive reactivity. In the current study, this self-regulative component of temperament is exemplified by the second-order factor of Sociability/Resilience and by the moderate association between Sociability/Resilience and Cognitive-Diligence (task orientation), and likewise between Cognitive-Diligence and Vigor/Mobility (the latter reflecting activity and intensity). It would appear based on these and other reported temperamental structures that although these structures evolve somewhat across the lifespan, their basic organization remains stable from birth to adulthood.

Second, inclusion of items that reflect ego control, flexibility, and resilience underscores the importance of a dimension (or set of related constructs) that reflects an emerging socioemotional component of individual differences during adolescence. Block and Block (1980) suggested that the differing levels of boundary permeability associated with ego control delineate behav-

ioral styles and accordingly may overlap considerably with temperamental traits. In addition, ego control and self-monitoring skills become more pronounced and essential components of behavior during adolescence. Perhaps, by incorporating a dimension of ego control the true covariation among constructs previously construed as independent has been specified correctly (Thomas & Chess, 1977). In essence, the conceptualization of temperament as composed of distinct behavioral styles may be appropriate for infants for whom certain temperamental attributes may remain somewhat independent components in their behavioral repertoire (i.e., behavior is regulated through separate and primitive systems). However, the same behavioral styles may be more unified or interrelated for older ages.

Additional support for this claim is provided by the large percentage of shared construct variance between the second-order factors of Cognitive-Diligence and Sociability/Resilience and to a somewhat lesser degree between Cognitive-Diligence and Vigor/Mobility. Thus the view that temperament reflects solely the emotional undertones of behavior is not entirely supported by these data. Cognitive-Diligence reflects cognitive undertones (i.e., task persistence and attentional focus); Vigor/Mobility primarily reflects activity and intensity of engagement, the latter not limited to emotional qualities. Moreover, Positive Mood was a significant and fairly strong indicator of Sociability/Resilience, reinforcing that greater general adaptability is associated with positive mood (i.e., being enthusiastic, cheerful, lively, content, and good natured). Taken together, the indicators (primary factors) of Sociability/Resilience reflect a positive emotional tone and balance with which to engage in life, the indicators of Cognitive-Diligence reflect cognitive task orientation, and the indicators of Vigor/Mobility reflect calmness and behavioral quiescence.

Additionally, certain key features of temperament, which may form a conceptual bridge between early and later life, have been identified. For example, the consistently strong loadings of four of the five primary factors comprising the second-order factor of Sociability/Resilience reinforce the primacy of these characteristics in adolescence and their apparent extension into adulthood. Any number of differing models of adult personality contain some element of sociability (i.e., extraversion or agreeableness) in the conceptualization (i.e., Costa & McCrae, 1992; Digman, 1989; Eysenck & Eysenck, 1969) and the models are largely composed of very similar characteristics to those reflected by Positive Mood, Approach/Withdrawal, Adaptability, and Ego Control. Because the selection and content of many of the items used in the current analyses may seem quite specific to the Thomas and Chess nine-dimensional model, the relative strengths of the loadings for the second-order factor of Sociability/Resilience and its replication in

older aged samples supports the developmental consistency of this temperamental dimension.

Importantly, in the current study the choice of nomenclature for all three second-order factors is closely linked to the Block and Block (1980) conceptualization of ego-control and ego-resiliency. That is, Block and Block delineated the continuum of ego control as a set of operating characteristics (i.e., behavioral styles), which included regulation of impulse, delay of gratification, modulation of action and affect, adaptability to changing environmental circumstances, and vulnerability to environmental distractors. Likewise, Block and Block characterized ego-resilient individuals by their levels of resourceful adaptation, flexibility, management of stress, cognitive and emotional engagement, capability for processing multiple stimuli, and organization. Many, if not all, of these characteristics are encompassed by the second-order factors, reinforcing the strong ties between temperament and personality development.

Finally, the structure of Mood was recast as bidimensional (capturing both negative and positive affectivity), a finding that has garnered much support from the adult literature on mood and affect (Warr et al., 1983; Watson et al., 1988). The moderate correlation between these constructs in the first-order model and their respectively divergent loadings in the second-order model supports the contention that negative and positive emotionality are somewhat distinct components during this age period, although they share substantial portions of variance (Goldsmith & Rothbart, 1991).

Several limitations to the current research are worth noting. First, the sample of youths is relatively homogeneous with respect to race and socioeconomic status. The current study focused on temperamental structures for a particular age period, but did not address potential ethnic or socioeconomic differences. Others who have addressed this empirical question did not find marked differences in behavioral style based on gender or social class in younger aged samples (Fullard, Simeonsson, & Huntington, 1989; Persson-Blennow & McNeil, 1981). Notwithstanding, the validity of the current findings and their generalizability may be hampered by a lack of representativeness.

In the current study, a heterogeneous age grouping was included that spanned early-middle to late adolescence, and these ages were coalesced under the rubric of a single developmental period. Considering adolescence as a single developmental period may gloss over evolving patterns of development that could be captured by finer age discrimination. It is unfortunate that many lifespan theorists consider adolescence to be so all encompassing both agewise and experientially, especially in light of the many complex and multifaceted changes (i.e., physical, cognitive, social) that occur during this

age period. Just as there is no single developmental task that can adequately characterize adolescence, perhaps no single model of temperament can characterize the emotional, cognitive, and social stylistic components that evolve during adolescence. By necessity, future replications may want to pattern and test models at each chronological age level, rather than proposing a single model for such a wide developmental period. In addition, the current research indicates that different systems or styles of temperament may be distinctly evolving during this critical age period, corresponding to affective and cognitive processing and physiological arousal. Future research on adolescent temperament could address these systems or at least begin to impose some differentiation on the items that tap these underlying stylistic tendencies.

Finally, in the current study no attempt has been made to validate the ADTQ using criterion-related validity (both concurrent and predictive) (Goldsmith, Rieser-Danner, & Briggs, 1991). By necessity, further refinement of the ADTQ necessitates establishing its heuristic utility via a network of relationships with clinically appropriate and psychometrically valid assessments (Windle, 1989). Perhaps future research may examine parent, teacher, or friends' ratings, all of which may ensure greater ecological validity of adolescent behavioral styles.

NOTES

1. Several model fit criteria were used to evaluate the fit and nested fit of the various models. These included (a) χ^2 degree of freedom ratio (optimally less than 2.0), (b) p value associated with the χ^2 ($p > .05$), and (c) the Comparative Fit Index (Bentler, 1990), indicating the amount of covariation accounted for in the sample data by the hypothetical model. Benchmarks for this latter statistic with small models are generally considered adequate approaching .90.

2. In general, post hoc model fitting includes relaxing constraints (i.e., allowing complex factor loadings rather than maximizing simple structure) or adding residual covariances (both within and between constructs), the latter which captures elements of both method (i.e., item wording) and shared construct variance. To obtain a more fine-grained analysis of potential enhancements to confirmatory and structural models, the EQS statistical software program (Bentler, 1989) provides a multivariate stepwise procedure indicating where relaxing specific constraints might improve the fit (by decreasing the χ^2 sufficiently enough for each degree of freedom change, and likewise removing substantively meaningful covariation from the residual matrix).

This type of model enhancement must be approached with caution, especially because a recent series of simulation analyses have demonstrated empirically the fragility of these modifications with small sample sizes (e.g., less than 2,000, Bollen, 1990; MacCallum, Roznowski, & Necowitz, 1992). Accordingly, in the initial runs that compared statistically the different model conceptualizations no attempt was made to "fit" the final models through

reparameterization (i.e., adding residual covariances). These models contained excessively large numbers of parameters (i.e., factor covariances, factor loadings, residual variances) and given the small sample size likely would capitalize on chance (MacCallum et al., 1992). In addition, the modification indexes from the 11-factor model indicated well over 222 univariate constraints (residual covariances set to zero) that could be relaxed (estimated as true covariances) to improve the overall model fit. Likewise, the multivariate version of the modification indexes suggested 166 modifications ($p < .05$) before the model χ^2 would drop substantially and produce a better fitting model. Moreover, many of these suggested modifications were inconsistent with substantive theory.

3. We are grateful to the anonymous reviewer who pointed out that in our earlier submission, we had allowed both primary constructs of Rhythmicity and Negative Mood to correlate freely with the second-order factors, which does not accurately address their "true" validity relationships. A more effective approach is to allow these primary constructs to load on the second-order factors, and by the very nature of their respective parameter loadings determine empirically the structure of temperament. Although the relative fit of the model as currently presented was virtually identical to the model that allowed the two specific primary factors to correlate with the second-order constructs, it is difficult to identify and clarify the precise nature of the relationship between the freestanding primary factor and second-order factor, because the second-order factor is multiply determined (several primary constructs are hypothesized to be statistically caused by the second-order dimension). Therefore, at the reviewer's suggestion, it is more meaningful to constrain the primary factor to load on the second-order factor and determine its conceptual validity accordingly.

REFERENCES

Angleitner, A., & Riemann, R. (1991). What can we learn from the discussion of personality questionnaires for the construction of temperament inventories? In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 191-204). New York: Plenum.

Bentler, P. M. (1978). The interdependence of theory, methodology, and empirical data: Causal modeling as an approach to construct validation. In D. B. Kandel (Ed.), *Longitudinal research on drug use: Empirical findings and methodological issues* (pp. 267-302). Washington, DC: Hemisphere.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31*, 419-456.

Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Berger, M. (1982). Personality development and temperament. In *Temperamental differences in infants and young children* (pp. 176-190). Ciba Foundation Symposium 89, London: Pitman.

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 13, pp. 39-101). Hillsdale, NJ: Lawrence Erlbaum.

Blos, P. (1979). *The adolescent passage: Developmental issues*. New York: International Universities Press.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A. (1990). Overall fit in covariance structure models: Two types of sample size effects. *Psychological Bulletin, 107*, 256-259.

Buss, A. H. (1989). Temperaments as personality traits. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 49-58). New York: Wiley.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Lawrence Erlbaum.

Capaldi, D. M., & Rothbart, M. K. (1992). Development and validation of an early adolescent temperament measure. *Journal of Early Adolescence, 12*, 153-173.

Chess, S., & Thomas, A. (1991). Temperament and the concept of goodness of fit. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 15-28). New York: Plenum.

Chou, C. -P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange multiplier, and Wald tests. *Multivariate Behavioral Research, 25*, 115-136.

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

Czeschlik, T. (1992). The Middle Childhood Temperament Questionnaire: Factor structure in a German sample. *Personality and Individual Differences, 13*, 205-210.

Derryberry, D., & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality and Social Psychology, 55*, 958-966.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57*, 195-214.

Eccles, J., Wigfield, A., Flanagan, C., Miller, C., Reuman, D., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: Relations and changes in early adolescence. *Journal of Personality, 57*, 283-310.

Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.

Eysenck, H. J., & Eysenck, S. B. (1969). *Personality structure and measurement*. London: Routledge & Kegan Paul.

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.

Fullard, W., Simeonsson, R. J., & Huntington, G. S. (1989). Sociocultural factors in temperament. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 523-536). New York: Wiley.

Funder, D. C., & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology, 57*, 1041-1050.

Goldsmith, H. H., Rieser-Danner, L. A., & Briggs, S. (1991). Evaluating convergent and discriminant validity of temperament questionnaires for preschoolers, toddlers, and infants. *Developmental Psychology, 27*, 566-579.

Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249-272). New York: Plenum.

Hegvik, R. L., McDevitt, S. C., & Carey, W. B. (1982). Middle Childhood Temperament Questionnaire. *Developmental and Behavioral Pediatrics, 3*, 197-200.

Hofstee, W.K.B. (1991). The concepts of personality and temperament. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives in theory and measurement* (pp. 177-188). New York: Plenum.

Hubert, N. C., Wachs, T. D., Peters-Martin, P., & Gandour, M. J. (1982). The study of early temperament: Measurement and conceptual issues. *Child Development, 53*, 571-600.

Josselson, R. (1980). Ego development in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 133-210). New York: Wiley.

Keating, D. P. (1990). Adolescent thinking. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 54-89). Cambridge, MA: Harvard University Press.

King, N. J., Ollendick, T., & Gullone, E. (1991). Negative affectivity in children and adolescents: Relations between anxiety and depression. *Clinical Psychology Reviews, 11*, 441-459.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*, 490-504.

Marcia, J. (1980). Identity in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159-187). New York: Wiley.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

Matheny, A. P. (1980). Bayley's Infant Behavior Record: Behavioral components and twin analyses. *Child Development, 51*, 1157-1167.

McClowry, S. G., Hegvik, R. L., & Teglasi, H. (1993). An examination of the construct validity of the Middle Childhood Temperament Questionnaire. *Merrill-Palmer Quarterly, 39*, 279-293.

Persson-Blennow, I., & McNeil, T. F. (1981). Temperament characteristics of children in relation to gender, birth order, and social class. *American Journal of Orthopsychiatry, 51*, 710-714.

Prior, M., Crook, G., Stripp, A., Power, M., & Joseph, M. (1986). The relationship between temperament and personality: An exploratory study. *Personality and Individual Differences, 7*, 875-881.

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research, 23*, 51-67.

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.

Rothbart, M. K., & Posner, M. I. (1985). Temperament and the development of self-regulation. In L. C. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 93-123). New York: Plenum.

Rutter, M. (1987). Temperament, personality, and personality disorder. *British Journal of Psychiatry, 150*, 443-458.

Rutter, M., Graham, P., Chadwick, O., & Yule, W. (1976). Adolescent turmoil: Fact or fiction. *Journal of Child Psychology and Psychiatry, 17*, 35-56.

Seltzer, V. C. (1989). *The psychosocial worlds of the adolescent: Public and private*. New York: Wiley.

Simmons, R., & Blyth, D. (1987). *Moving into adolescence: Impact of pubertal change and school context*. New York: de Gruyter.

Skinner, H. A. (1986). Toward the integration of classification theory and methods. *Journal of Abnormal Psychology, 90*, 68-87.

Strelau, J. (1983). *Temperament, personality, activity*. New York: Academic Press.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58*, 134-146.

Tanaka, J. S., & Huba, G. J. (1984). Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology, 46*, 621-635.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. G. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.

Warr, P., Barter, J., & Brownbridge, G. (1983). On the independence of positive and negative affect. *Journal of Personality and Social Psychology, 44*, 644-651.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depression disorders. *Journal of Abnormal Psychology, 97*, 346-353.

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.

Werts, C. F., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement, 34*, 25-33.

Windle, M. (1988). Psychometric strategies of measures of temperament: A methodological critique. *International Journal of Behavioral Development, 11*, 171-201.

Windle, M. (1989). Temperament and personality: An exploratory interinventory study of the DOTS-R, EASI-II, and EPI. *Journal of Personality Assessment, 53*, 487-501.

Windle, M. (1991). The difficult temperament in adolescence: Associations with substance use, family support, and problem behaviors. *Journal of Clinical Psychology, 47*, 310-315.

Windle, M. (1992). Revised dimensions of temperament survey (DOTS-R): Simultaneous group confirmatory factor analysis for adolescent gender groups. *Psychological Assessment, 4*, 228-234.

Windle, M., Hooker, K., Lernerz, K., East, P. L., Lerner, J. V., & Lerner, R. M. (1986). Temperament, perceived competence, and depression in early- and late-adolescents. *Developmental Psychology, 22*, 384-392.

Windle, M., & Lerner, R. M. (1986). Reassessing the dimensions of temperamental individuality across the lifespan: The Revised Dimensions of Temperament Survey (DOTS-R). *Journal of Adolescent Research, 1*, 213-230.

Wolfe, V. V., Finch, A. J., Saylor, C. F., Blount, R. L., Pallmeyer, T. P., & Carek, D. J. (1987). Negative affectivity in children: A multitrait-multimethod investigation. *Journal of Consulting and Clinical Psychology, 55*, 245-250.

Requests for reprints should be addressed to Lawrence M. Scheier, Institute for Prevention Research, Cornell University Medical College, 411 East 69th St., Kips Bay Room 201, New York, NY 10021.