

The Development of the *School-Age Temperament Inventory*

Sandra Graham McClowry
New York University

The purpose of this study was to develop the *School-Age Temperament Inventory (SATI)* as a parental report of children ages 8–11 years. Based on previous studies of temperament in children, four empirically derived dimensions were proposed: task persistence, negative reactivity, approach/withdrawal, and energy. Parent informants included 435 mothers and 228 of their spouses. Principal factor analysis with varimax rotation supported, in general, the empirically derived dimensions. The validity and reliability of the SATI was then evaluated. Cronbach's alphas were .85 to .90. Correlations between maternal and paternal reports were .51 to .68. Test-retest reliability was .80 to .89. The impact of gender and age was minimal. Convergent validity was evaluated by comparing the SATI with another temperament questionnaire designed for preschool children. Although continued development is recommended, the SATI appears to have adequate validity and reliability for use in research and structured parenting programs.

Temperament refers to the stylistic component of behavior that an individual generally exhibits across a variety of settings. An increasing number of studies have reported that temperament is related to developmental outcomes, both positive and negative (Chess & Thomas, 1984; Keogh, 1986; McClowry et al., 1994; Sanson, Oberklaid, Pedlow, & Prior, 1991). In spite of the large number of reported studies, questions remain regarding the measurement of the construct (Institute of Medicine, 1989). Although parental reports are the most frequently used technique in the temperament field, their validity and reliability remains a source of frequent debate (Goldsmith & Rothbart, 1991; Hubert, Wachs, Peters-Martin, & Gandour, 1982; McClowry, Hegvik, & Teglasi, 1993; Slabach, Morrow, & Wachs, 1991).

This study was funded by the National Institute for Nursing Research, National Institute of Health (1 R15 NR03059-01). The author wishes to express appreciation to Jack Bates, Roy Martin, Mary Rothbart, Ron Seifer, and Jim Victor for serving as content validity experts and to Ann Sanson for her review of this paper. Correspondence may be sent to Sandee McClowry, School of Education, Division of Nursing, New York University, 429 Shimkin Hall, New York, NY 10002.

Merrill Palmer Quarterly, July 1995, Vol. 41, No. 3, pp. 271–285.
Copyright © 1995 by Wayne State University Press, Detroit, MI 48201

A variety of definitions, conceptualizations, and approaches to measure temperament have created a great deal of plurality in the field (Goldsmith et al., 1987). Such multiplicity can be construed as advantageous if used to further clarify temperament and its components. Four example, consistent findings across studies, even when based on different conceptualizations, serve as a source of validation that may ultimately lead toward more theoretical unification. Indeed, researchers in recent reports have remarked that an integration of temperament theories is emerging, on the basis of studies in which the construct validity of existing questionnaires has been evaluated (McClowry et al., 1993; Rothbart & Mauro, 1990).

In studies of parental reports of infants, negative affect, positive affect, activity, rhythmicity, and attention-span/persistence have consistently emerged as factors and appear to be key dimensions of temperament during infancy (Rothbart & Mauro, 1990). Because the development of children differs from infants, a separate examination of reports focusing on questionnaires for children is warranted. Based on a review of existing measures of child temperament and standard psychometric procedures, a new generation of empirically derived temperament instruments could be created. Consistent findings across a variety of instruments might provide further insight into the dimensions of childhood temperament. The purpose of this study was to develop the School-Age Temperament Inventory (SATI).

Literature review

Ten item-based factor analytic studies, using a variety of temperament questionnaires, have included children older than 3 years old. The construct validity of the Middle Childhood Temperament Questionnaire (MCTQ) was examined in two studies. The MCTQ was developed by Hegvik, McDevitt, and Carey (1982) as a parental report of children who are 8 to 12 years old. McClowry et al. (1993) re-examined the data collected in the original development of the tool and compared them with data from two other studies. Czeschlik (1992) factor analyzed data collected on a German translation of the questionnaire.

Four studies focused exclusively on young children. Presley and Martin (1994) examined the Temperament Assessment Battery for Children that Martin (1988) developed for preschool children. Another study of preschool children was conducted by Finegan, Niccols, Zacher, and Hood (1989). They modified the Child Characteristics Questionnaire created by Lee and Bates (1985) and then tested the factor structure of a preschool version. The rest involved questionnaires constructed by

Thomas and Chess (1977). A parental report, the Childhood Temperament Questionnaire, was studied by Sanson, Smart, Prior, Oberklaid, and Pedlow (1994) and focused on children 3–8 years old. The Teacher Temperament Questionnaire was examined in two studies (Baker & Velicer, 1982; Keogh, Pullis, & Cadwell, 1982).

Other studies incorporated subjects across a wider range of ages. Buss and Plomin (1975) included subjects from 1 to 9 years of age when they developed the EASI questionnaire. Later, in order to merge the dimensions of the EASI with those identified by Thomas and Chess (1977), Rowe and Plomin (1977) involved children at ages 5 months to 6 years. The questionnaire that resulted was the Colorado Childhood Temperament Inventory. With an even greater span of years, Windle and Lerner (1986) developed the Dimensions of Temperament Survey (DOTS-R). Their participants ranged from preschool age to early adulthood.

Although these researchers used a variety of instruments with subjects of assorted ages, four factors have consistently emerged. Each study had one factor that described the social responses of children. Names for it included approach/withdrawal (Czeschlik, 1992; McClowry et al., 1993; Windle & Lerner, 1986), inhibition (Presley & Martin, 1994), extroversion (Baker & Velicer, 1982), and sociability (Buss & Plomin, 1975; Rowe & Plomin, 1977; Sanson et al., 1994). Keogh et al. (1982) found a related but broader factor, labeled personal-social flexibility, which merged positive mood and adaptability items. Finegan et al. (1989) also had a broader factor called negative adaptation and affect, which included items that described only negative responses to people and novel situations.

In each of the studies a factor emerged related to the negative expression of affect. The various names of the factors included negative reactivity (McClowry et al., 1993), negative emotionality (Presley & Martin, 1994), emotionality (Buss & Plomin, 1975; Rowe & Plomin, 1977), interpersonal affect (Baker & Velicer, 1982), mood (Czeschlik, 1992; Windle & Lerner, 1986), inflexibility (Sanson et al., 1994), reactivity (Keogh et al., 1982), and difficult (Finegan et al., 1989).

A factor focusing on persistence was reported in all of the studies. McClowry et al. (1993) named it task persistence whereas others simply labeled it persistence (Czeschlik, 1992; Presley & Martin, 1994; Sanson et al., 1994). Additional names for it were compliance (Baker & Velicer, 1982), attention span–persistence (Rowe & Plomin, 1977), task orientation (Keogh et al., 1982; Windle & Lerner, 1986), and persistent/unstoppable (Finegan et al., 1989). Buss and Plomin (1975) described a factor called impulsivity, which consisted of items that reflected self-control.

An additional factor related to activity was reported across the studies, but with less conceptual clarity. Although Baker and Velicer (1982) and Finegan et al. (1989) did not find an activity factor, others did (Buss & Plomin, 1975; Presley & Martin, 1994; Rowe & Plomin, 1977) and Windle and Lerner (1986) reported two: one related to sleep, whereas the other included more general activity items. In both McClowry et al. (1993) and Czeschlik (1992), activity items merged with those that were intended to measure intensity. Activity combined with mood in another study (Sanson et al., 1994). Keogh et al. (1982) described a broader factor of task orientation that consisted of activity, persistence, and distractibility items. Thus, although activity seems to be part of the expression of childhood temperament, it appears to have elements that are broader than just motor activity. This may be because as most children get older they are socialized to channel their energy so that their parents observe behaviors of which motor activity is only a part. The term *energy* may better describe how school-age children expend and exhibit their propensity toward activity.

Based on these studies, four empirically derived dimensions of school-age temperament were chosen as the conceptualization for the instrument: approach/withdrawal, task persistence, negative reactivity, and energy. The following research questions were addressed in this study in order to assess the reliability and validity of the SATI: (a) Does the SATI support the four empirically derived dimensions of negative reactivity, task persistence, approach/withdrawal, and energy? (b) Does the SATI demonstrate adequate validity and reliability?

Instrument Development

The investigator prepared 83 items to represent the four dimensions of temperament. To avoid respondent bias, approximately half of the items were reverse ordered.

Content validity was assessed in the manner described by Lynn (1986). Five temperament experts, with extensive instrumentation experience, were provided with definitions of the dimensions and asked to evaluate whether the generated items were relevant to the intended dimensions and whether they were developmentally appropriate for children this age. A content validity index (CVI) consisting of a Likert-type scale from *irrelevant* (1) to *extremely relevant* (4) was used to rate the relevancy of the items. A similar process was used to assess developmental appropriateness: *inappropriate* (1) to *extremely appropriate* (4). Items that rated a 3 or 4 on both indexes by all of the experts were retained. Otherwise, the items were either omitted or revised, and reevaluated by the experts after at least 2 weeks. After the generated items were

evaluated, the same experts were asked if any areas were omitted from the instrument. Suggested items were subjected to the same CVI process.

Once the content validity stage was completed, the resulting 63 items were placed in a usable form and arranged in sequence. To avoid respondent bias, the items from the four dimensions were randomly placed throughout the questionnaire. Consistent with most other temperament questionnaires, a Likert-type scale from *never* (1) to *always* (5) was used. The readability of the items was also evaluated. The SATI was assessed as below a sixth-grade level on the *RightWriter* program (Que Software, 1992).

DESIGN AND METHODS

Participants

A total of 435 mothers and 228 of their spouses served as parent informants for this study. Their children were from 8 to 11 years and averaged 9.91 years of age ($SD = 0.91$). Approximately half (51%) of the children were boys ($n = 221$) and the remainder ($n = 214$) were girls. The average age of the mothers was 38.6 years ($SD = 4.9$) and ranged from 27 to 60 years. Their husbands' ages ranged from 23 to 62 years with an average age of 40.8 ($SD = 5.3$). Most of the children (75%) lived with both of their parents. An additional 10% lived in blended families with their mothers and stepfathers and the remaining 10% lived with their mothers who were single parents. The rest of the subjects declined to describe their family configuration. The socioeconomic status of the families, as identified on the Four Factor Index by Hollingshead (1975), were 5% unskilled, 12% semiskilled, 22% clerical, 40% technical, and 21% professional. The vast majority (89%) of the children were reported by their mothers to be Caucasian. An additional 6% of the children were African American, 2% Hispanic, 2% Asian, and 1% Native American.

Instruments

The SATI and standard demographic information were collected from all participants. To examine convergent validity, the parent version of the Temperament Assessment Battery for Children—Revised (TABC-R), an instrument that is conceptually similar to the SATI, but is intended for younger children, also was administered. The TABC was developed by Martin (1988) to measure 3- to 7-year-olds. The revised version of the parent form of the instrument, the TABC-R (Presley & Martin, 1994) was used in this study. It consists of 39 items from five dimensions: negative emotionality, inhibition, adaptability, activity, and persistence.

Procedure

A power analysis was first conducted to determine the numbers of subjects necessary for the various portions of data analyses (Cohen, 1977; Knapp & Campbell-Heider, 1989). For conducting the factor analysis, at least 430 mothers were deemed necessary, with far fewer subjects required for the other parts of the study. A pool of mothers was then sought to provide an adequate amount of data for the factor analysis. To reduce respondent burden and cost, mothers were randomly selected from the pool when additional data were needed.

Parental informants were recruited from three school districts in New England by sending letters of invitation to approximately 3,700 mothers of children in regular third- to fifth-grade classrooms. Mothers whose children were between the ages of 8 and 11 years and who were willing to participate were asked to return an enclosed, stamped postcard and to indicate whether the child's father was also willing to take part in the study. A complimentary movie ticket was promised to each parent who completed the questionnaires.

A total of 614 nonduplicate postcards were returned by families who had a child between 8 and 11 years old. Data packets including consent forms, the respective instruments, and stamped return envelopes were sent to the parents' homes. Separate return envelopes were included when both parents agreed to participate ($n = 376$). Parents who had more than one eligible child were instructed to focus on a particular one who was randomly selected by the investigator.

All parents were asked to complete the SATI and all mothers were requested to provide demographic information. In order to evaluate the convergent reliability of the SATI, a randomly selected number of mothers ($n = 120$) from the pool also received the TABC-R in their packet.

When needed, standard reminder postcards and telephone calls were employed. The response rate for the mothers was 76% and 64% for the fathers. Thirty-two families, however, were removed from the study because only the fathers participated ($n = 9$), the mothers provided incomplete data ($n = 10$), or the child was in a special education classroom ($n = 13$). Finally, a second SATI was sent after 4-6 months to 78 mothers who were randomly selected from the pool to examine test-retest reliability, and 91% complied.

RESULTS

The first research question was whether the SATI supported the four empirically derived temperament dimensions of negative reactivity, task

persistence, approach/withdrawal, and energy. Maternal data only ($n = 435$) were submitted to a series of principal factor analyses with orthogonal (varimax) rotations. This method of factor analysis was chosen because it capitalizes on the common variance of the variables so that the possibility of distinct and separate clusters is optimized (Zeller & Carmines, 1980). When the Kaiser-Guttman criterion was used six factors were obtained. Based on the scree test, four of the factors were retained. The criterion used for selection was that the factor loading was $\geq .50$ on one factor with at least a .20 spread on the others. The total amount of explained variance was 72%.

Table 1 illustrates the varimax rotated items. Each of the items loaded on just one of the pre-established conceptualized dimensions. The first factor, Negative Reactivity, described the intensity and frequency with which the child expresses negative affect. The second factor, Task Persistence, depicted the degree of self-direction that a child exhibits in fulfilling tasks and other responsibilities. The third factor, Approach/Withdrawal, portrayed the child's initial response to new people and situations. The last factor included only items from the energy dimension. However, the items did not reflect a combination of intensity and activity as did the factor in McClowry et al. (1993). Instead, the items were related exclusively to large motor activity. Thus, the factor was named Activity.

The 38 items that were selected through this process were again subjected to a principal factor analysis with a varimax solution. With use of the same criteria, all items remained on the same factor with only slight variations in the order of the loadings and in the loadings. The total amount of explained variance for the four factors when only the 38 items were considered was 89%. A copy of the factor loadings is available from the author.

The second research question concerned whether the SATI demonstrated adequate validity and reliability and was assessed in several ways. Scores were obtained by simply summing the items and then dividing by the respective number of items. The means and standard deviations for the dimensions as reported by the mothers and fathers were similar and are reported in Table 2. Higher scores indicate that the child is high in negative reactivity, is task persistent, has a tendency to withdraw in new situations, and is highly active.

Reliability was also assessed by examining the Cronbach's alphas for mothers and fathers. The alphas were similar and ranged from .85 to .90 (Table 2). Correlations between maternal and paternal reports were between .51 and .68 and are shown in Table 3. Test-retest correlations from maternal data ($n = 71$) after 4 to 6 months were between .80 and .89 and are listed in Table 3.

Table 1. Varimax Rotated Items of the School-Age Temperament Inventory

Item Stem	Factor Structure			
	Negative Reactivity	Persistence	Approach	Activity
Factor 1: Negative Reactivity				
Gets mad when mildly criticized	.72	-.13	.13	.11
When angry, yells or snaps at others	.71	-.12	.07	.06
Moody when corrected for misbehavior	.69	-.10	.11	.01
Responds intensely to disapproval	.67	-.16	.04	.18
Gets upset when he/she can't find something	.66	-.13	.16	.17
Makes loud noises when angry	.63	-.18	.03	.16
Reacts strongly to a disappointment	.62	-.18	-.03	.20
When he/she disagrees, speaks quietly ^a	.62	-.24	-.01	.13
Has off days when he/she is moody	.56	-.13	.18	.02
Gets very frustrated	.56	-.16	.16	.15
Gets upset when there is a change in plans	.52	-.18	.24	.11
Gets angry when teased	.51	-.05	.25	.12
Factor 2: Task Persistence				
Has difficulty completing assignments ^a	-.21	.74	-.07	-.14
Does not complete homework ^a	-.11	.74	-.05	-.05
Stays with homework until finished	-.07	.71	-.03	-.09
Remembers to do homework	-.08	.68	-.04	-.07
Returns to responsibilities	-.21	.66	-.03	-.17
Switches from one activity to another ^a	-.25	.60	-.03	-.24
Gets frustrated with projects and quits ^a	-.34	.60	-.16	-.10
Goes back to the task at hand	-.14	.60	.01	-.15
Quits routine household chores ^a	-.19	.54	.03	-.15
When an activity is difficult, gives up ^a	-.28	.52	-.16	-.10
Leaves own projects unfinished ^a	-.17	.51	-.02	-.03
Factor 3: Approach/Withdrawal				
Is shy with adults he/she doesn't know	.04	.10	.78	-.07
Bashful when meeting new children	.12	.02	.72	-.01
Seems uncomfortable at someone's house	.21	.02	.66	.03
Approaches children his/her age ^a	.03	-.07	.66	-.16
Avoids new guests	.06	-.04	.66	-.06
Seems nervous or uncomfortable in a new situation	.18	-.01	.65	.13
Moves right into a new place ^a	.08	-.04	.63	-.01
Prefers to play with someone he/she knows	.12	-.01	.56	-.08
Smiles or laughs with new adults	.07	-.08	.53	-.05
Factor 4: Activity				
Runs to where he/she wants to go	.12	-.13	.00	.70
Runs when entering or leaving	.05	-.28	.01	.68
Runs or jumps when going down	.16	-.16	.07	.67
Seems to be in a hurry	.24	-.13	-.03	.60
Walks quietly in the house ^a	.21	-.27	.04	.58
Bursts loudly into the room	.30	-.16	-.01	.56

^a Reverse coded before factor analyzed.

Table 2. Means, Standard Deviations, and Cronbach's Alphas of the School-Age Temperament Inventory Dimensions by Mothers and Fathers

Dimensions	Mothers (n = 435)			Fathers (n = 228)		
	M	SD	Alphas	M	SD	Alphas
Negative reactivity	3.09	0.74	.90	3.08	0.65	.88
Task persistence	3.60	0.72	.90	3.52	0.67	.90
Approach/withdrawal	2.47	0.74	.88	2.59	0.66	.85
Activity	2.73	0.80	.85	2.84	0.74	.85

Table 3. Correlations Between Mothers and Fathers and Test-Retest Correlations on the School-Age Temperament Inventory Dimensions

Dimensions	Parents' Correlations	Test-Retest Correlations
Negative reactivity	.51**	.89**
Task persistence	.68**	.81**
Approach/withdrawal	.56**	.82**
Activity	.53**	.80**

** $p \leq .01$.

Validity was further examined by exploring whether gender and age influence the expression of school-age temperament. Table 4 lists the means and standard deviations of the dimensions by gender and age, respectively.

To examine the effect of age and gender on the dimensions of temperament, separate analyses of variance (ANOVAs) for the four factors were conducted. Each ANOVA included two between-subject variables, age (four levels) and gender (two levels) and their interactions. When effects were significant, planned post hoc analyses were examined.

Significant age, $F_{3,434} = 3.95$, $p < .01$, and gender, $F_{1,434} = 13.90$, $p < .01$, main effects were found on activity without any interactive effects. Boys were significantly more active than girls, and based on the Bonferroni post hoc comparison, 8-year-old children were significantly more active than those who were 10 ($p < .05$) and 11 years old ($p < .05$).

A main effect of gender, $F_{1,434} = 19.29$, $p < .01$, was found only on task persistence, indicating that boys were less task persistent than girls. There were no significant effects or interactions on negative reactivity and approach/withdrawal.

Stepwise multiple regression was then used to determine the magnitude of the identified effects on school-age temperament. For activity, the contribution of gender (dummy coded) was 4%. The children's age

Table 4. Means and Standard Deviations of the School-Age Temperament Inventory Dimensions in Relation to Child Gender and Age

<i>Dimension</i>	<i>Negative Reactivity</i>	<i>Task Persistence</i>	<i>Approach/Withdrawal</i>	<i>Activity</i>
<i>Girls</i>				
<i>M</i>	3.06	2.53	3.77	2.56
<i>SD</i>	0.70	0.77	0.68	0.76
<i>Boys</i>				
<i>M</i>	3.12	2.41	3.45	2.89
<i>SD</i>	0.65	0.70	0.72	0.81
<i>Both Genders</i>				
<i>Age/Dimension</i>				
8 Years				
<i>M</i>	3.12	3.58	2.40	3.13
<i>SD</i>	0.70	0.70	0.65	0.95
9 Years				
<i>M</i>	3.14	3.54	3.61	3.66
<i>SD</i>	0.69	0.67	0.71	0.72
10 Years				
<i>M</i>	3.06	3.61	2.46	2.65
<i>SD</i>	0.74	0.74	0.71	0.82
11 Years				
<i>M</i>	3.09	3.66	2.47	2.66
<i>SD</i>	0.80	0.74	0.77	0.79

added an additional 2%. Gender also explained 5% of the variance in task persistence. These findings are summarized in Table 5.

Convergent validity of the SATI was evaluated by comparing it with the TABC-R. As shown in Table 5, the SATI dimensions that were conceptually similar to those in the TABC-R showed strong correlations. Specifically, negative reactivity on the SATI and negative emotionality on the TABC-R correlated .71. The SATI approach/withdrawal factor correlated .87 with the TABC-R. The correlation between the activity factors was .73. The SATI task persistence correlated $-.67$ with the TABC-R persistence. The difference in the direction of the SATI and TABC-R is attributable to how the items are scored. Whereas high scores on the SATI indicate that the child is task persistent, high scores on the TABC-R suggest that the child is low in task persistence.

In addition to the correlations between SATI and TABC-R dimensions that are conceptually similar, others were also found. The TABC-R has an adaptability factor that the SATI does not. Its highest correlation with the SATI factors was $-.40$. The SATI task persistence correlated $-.60$ with the TABC-R activity factor, indicating that low task persistence was

Table 5. Regression Equations for Examining Gender and Age on Childhood Temperament

Variable	Cumulative R ²	R ² Change	Standardized Beta Weights	F for R ² Change
Activity				
Gender	.04	.04	-.20	18.73**
Age	.06	.02	-.14	8.60**
Task persistence				
Gender	.05	.05	.22	23.54**

***p* < .01.**Table 6.** Correlations between Maternal Reports of the School-Age Temperament Inventory (SATI) and the Temperament Assessment Battery for Children—Revised (TABC-R) (*n* = 89)

SATI Dimensions	TABC-R Dimensions				
	Negative Emotionality	Inhibition	Adaptability	Activity	Persistence
Negative reactivity	.71**	ns	-.40**	.45**	-.28**
Task persistence	-.45**	ns	.25*	-.60**	-.67**
Approach/ withdrawal	.23**	.87**	ns	ns	ns
Activity	.47**	ns	.29**	.73**	-.46**

p* ≤ .05. *p* ≤ .01.

associated with high activity. The rest of the significant correlations were between .23 and .47.

DISCUSSION

In general, the SATI appears to have adequate reliability and validity. The empirically derived dimensions that served as the conceptualization for the questionnaire were supported, for the most part, by principal factor analysis. Although some of the initial items did not meet the selection criteria, those that did loaded together on their respective hypothesized dimension. The one exception to that statement involved the fourth factor, which was expected to depict a dimension of energy more expansive than just motor activity. However, the items that emerged on that factor pertained only to motor activity. Consequently, the factor was

named Activity. The names and definitions of the other three factors remained consistent with the way they were proposed.

Another strength of the SATI is evident in its high explained variance whether interpreted as 72% or 89%. Previous studies of existing instruments have been reported to explain approximately one third of the variance (Bohlin, Hagekull, & Lindhagen, 1981; Garside et al., 1975; Prior, Sanson, & Oberklaid, 1989; Rowe & Plomin, 1977). The higher explained variance of the SATI may be due to its conceptualization that was empirically derived from previous studies, thus having the advantage of being a second- or third-generation temperament measure.

The strong correlations between the conceptually similar dimensions on the SATI with those on the TABC-R is a demonstration of convergent validity even though the TABC-R was not developmentally appropriate for the children in this study. The items on the SATI were intentionally worded to represent situations and responses that were representative of school-age children. Still, the two instruments seem to capture similar dimensions. Consistency of dimensions across different developmental ages is a necessary precursor for conducting longitudinal studies in order to discern how temperament is related to outcomes such as competencies or behavioral disturbances.

Other evidence of the SATI's reliability was found. The Cronbach's alphas for the maternal and paternal reports were adequate as were the test-retest correlations. One of the reasons that reliability is evidenced may be developmental. Children between the ages of 8–11 years may be more consistent in their behavior, thus allowing parents to be more reliable when reporting it. Likewise, the impact of age and gender were minimal, a finding consistent with Sanson et al. (1994) and Presley and Martin (1994).

When maternal and paternal reports were correlated, they ranged from .51 to .68. This is generally better than many of studies that found correlations between parental reports of $-.05$ to $.73$ (Hinde & Tobin, 1986; Keogh, 1986; Lyon & Plomin, 1981). But, as Reich and Earls (1987) maintain, achieving perfect consensus between parents is less important than understanding why informants differ. The biases inherent in parental reports may be a sign of psychometric weakness or may be indicative of systematic error requiring further clarification. They may also reflect that children's behavior differs across situations. Further evaluation of the SATI is needed so that particular types of biases can be acknowledged and, when possible, minimized. The use of observational techniques or physiological measurements are two possibilities. The use of self-reports from the children is another.

Further development of the SATI is also warranted to explore its use with culturally diverse samples. Like many temperament instruments, the SATI was normed on a convenience sample of mostly white, middle-class families and thus could be culturally biased. A study is currently underway to determine whether the psychometric estimates of this instrument are stable when used by a sample with more ethnic diversity and with a larger proportion of children from less advantaged homes.

This report describes the development of the School-Age Temperament Inventory. An assessment of its current psychometric properties supports that the SATI has more than adequate validity and reliability, but its utility is still to be tested. In research, the SATI could be employed to identify how the temperament of school-age children acts as a risk or as a protective factor in psychological development. The SATI also could be used to explore the transaction between child temperament and various parenting styles. Clinically, the SATI might be useful in prevention programs aimed at teaching parents effective temperament-based management strategies.

REFERENCES

- BAKER, E. H., & VELICER, W. F. (1982). The structure of reliability of the Teacher Temperament Questionnaire. *Journal of Abnormal Child Psychology*, 10, 531-546.
- BOHLIN, G., HAGEKULL, B., & LINDHAGEN, K. (1981). Dimensions of infant behavior. *Infant Behavior and Development*, 4, 83-96.
- BUSS, A. H., & PLOMIN, R. (1975). *A temperament theory of personality development*. New York: Wiley.
- CHESS, S., & THOMAS, A. (1984). *Origins and evolution of behavior disorders*. New York: Guilford.
- COHEN, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- CZESCHLIK, T. (1992). The Middle Childhood Temperament Questionnaire: Factor structure in a German sample. *Personality and Individual Differences*, 13, 205-210.
- FINEGAN, J., NICCOLS, A., ZACHER, J., & HOOD, J. (1989). Factor structure of the *Preschool Characteristics Questionnaire*. *Infant Behavior and Development*, 12, 221-227.
- GARSIDE, R. F., BIRCH, H., SCOTT, D. M., CHAMBERS, S., KOLVIN, I., TWEDDLE, E. G., & BARBER, L. M. (1975). Dimensions of temperament in infant school children. *Journal of Child Psychology and Psychiatry*, 16, 219-231.
- GOLDSMITH, H. H., BUSS, A. H., PLOMIN, R., ROTHBART, M. K., THOMAS, A., CHESS, S., & MCCALL, R. (1987). Roundtable: What is temperament? Four approaches. *Child Development*, 58, 505-529.

- GOLDSMITH, H. H., & ROTHBART, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In A. A. Angleitner & J. Strelau (Eds.), *Explorations in temperament: International perspectives on theory and measurement*. New York: Plenum Press.
- HEGVIK, R. L., MCDEVITT, S. C., & CAREY, W. B. (1982). *The Middle Childhood Temperament Questionnaire*. *Journal of Developmental and Behavioral Pediatrics*, 3, 197-200.
- HINDE, R. A., & TOBIN, C. (1986). Towards a more open construct. In G. A. Kohnstamm (Ed.), *Temperament discussed*. Lisse: Swets & Zeitlinger.
- HOLLINGSHEAD, A. B. (1975). *Four Factor Index of Social Status*. New Haven, CT: Department of Sociology, Yale University.
- HUBERT, N. C., WACHS, T. D., PETERS-MARTIN, P., & GANDOUR, M. J. (1982). The study of early temperament: Measures and conceptual issues. *Child Development*, 53, 571-600.
- INSTITUTE OF MEDICINE. (1989). *Research on children and adolescents with mental, behavioral and developmental disorders: Mobilizing a national initiative*. Washington, DC: National Academy Press.
- KEOGH, B. K. (1986). Temperament and schooling: Meaning of "Goodness of fit"? In J. V. Lerner & R. M. Lerner (Eds.), *Temperament and social interaction during infancy and childhood: New directions for child development*. San Francisco: Jossey-Bass.
- KEOGH, B. K., PULLIS, M. E., & CADWELL, J. (1982). A short form of the Teacher Temperament Questionnaire. *Journal of Education Measurement*, 19, 323-328.
- KNAPP, T. R., & CAMPBELL-HEIDER, N. (1989). Numbers of observations and variables in multivariate analyses. *Western Journal of Nursing Research*, 11, 634-641.
- LEE, C. L., & BATES, J. E. (1985). Mother-child interaction at age 2 years and perceived difficult temperament. *Child Development*, 56, 1314-1325.
- LYNN, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35, 382-386.
- LYON, M., & PLOMIN, R. (1981). The measurement of temperament using parental ratings. *Journal of Child Psychology and Psychiatry*, 22, 47-54.
- MARTIN, R. (1988). *The Temperament Assessment Battery for Children*. Brandon, VT: Clinical Psychology Publishing Co.
- MCCLLOWRY, S. G., GIANGRANDE, S. K., TOMMASINI, N. R., CLINTON, W., FOREMAN, N. S., LYNCH, K., & FERKETICH, S. (1994). The effects of child temperament, maternal characteristics, and family circumstances on the maladjustment of school-age children. *Research in Nursing and Health*, 17, 25-35.
- MCCLLOWRY, 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.
- PRESLEY, R., & MARTIN, R. P. (1994). Toward a structure of preschool temperament: Factor structure of the Temperament Assessment Battery for Children. *Journal of Personality*, 62, 415-448.
- PRIOR, M. R., SANSON, A. V., & OBERKLAI, F. (1989). The Australian Temperament Project. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood*. Chichester, England: Wiley.

- QUE SOFTWARE. (1992). *RightWriter* (Version 5). Carmel, IN: Prentice Hall Computer Publishing.
- REICH, W., & EARLS, F. (1987). Rules for making psychiatric diagnoses in children on the basis of multiple sources of information: Preliminary strategies. *Journal of Abnormal Child Psychology*, *15*, 601-616.
- ROTHBART, M. K., & MAURO, J. A. (1990). Questionnaire measures of infant temperament. In J. W. Fagen & J. Columbo (Eds.), *Individual differences in infancy: Reliability, stability, and prediction*. Hillsdale, NJ: Erlbaum.
- ROWE, D. C., & PLOMIN, R. (1977). Temperament in early childhood. *Journal of Personality Assessment*, *41*, 150-156.
- SANSON, A., OBERKLAID, F., PEDLOW, R., & PRIOR, M. (1991). Risk indicators: Assessment of infancy predictors of pre-school behavioural maladjustment. *Journal of Child Psychology and Psychiatry*, *32*, 609-626.
- SANSON, A., SMART, D., PRIOR, M., OBERKLAID, F., & PEDLOW, R. (1994). The structure of temperament from age 3 to 7 years: Age, sex, and sociodemographic influences. *Merrill-Palmer Quarterly*, *40*, 233-252.
- SLABACH, E. H., MORROW, J., & WACHS, T. D. (1991). Questionnaire measurement of infant and child temperament. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament*. New York: Plenum Press.
- THOMAS, A., & CHESS, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- WINDLE, M., & LERNER, R. M. (1986). Reassessing the dimensions of temperament individuality across the life span: The revised Dimensions of Temperament Survey (DOTS-R). *Journal of Adolescent Research*, *1*, 213-230.
- ZELLER, R. A., & CARMINES, E. D. (1980). *Measurement in the social sciences*. London: Cambridge University Press.