A CONFIRMATORY FACTOR ANALYSIS OF THE STRUCTURE OF TEMPERAMENT QUESTIONNAIRE

DAVID I. BISHOP
Luther College

MATTHEW J. HERTENSTEIN
DePauw University

This study examined the factor structure of scores on the English-language version of the Structure of Temperament Questionnaire. Scores from 300 college students were subjected to maximum-likelihood confirmatory factor analyses (CFA). A first-order model consisting of eight correlated factors and a second-order model consisting of two superordinate factors and eight first-order factors were tested. The results of the analyses indicated that the first-order correlated model was the best fit for the data. The English-language version of the Structure of Temperament Questionnaire can be used to measure eight correlated yet theoretically unique dimensions of temperament.

Keywords: *temperament; personality; confirmatory factor analysis; score validity*

The distinction between the constructs of personality and temperament has been, and continues to be, disputed in differential psychology. Whereas some models treat temperament and personality as essentially synonymous, others emphasize and formalize this distinction. According to the Rusalov (1989) model of temperament, temperament (the "formal" psychobiological properties of the nervous system) and personality (the sociopsychological manifestation at the behavioral level) are distinct phenomena. Temperament is a generalized and qualitatively new system of stable properties that emerges when basic biological properties are restructured and reorganized by various kinds of activity. According to Rusalov, four fundamental dimensions characterize variation in temperament: ergonicity (level of energy,

Educational and Psychological Measurement, Vol. 64 No. 6, December 2004 1-11
DOI: 10.1177/0013164404264843

activity, or motivation), plasticity (degree of openness or flexibility), tempo (rapidity of behavioral execution), and emotional sensitivity (emotional upset or reactivity). These dimensions of individuality are manifest somewhat independently in the world of objects (things) and in the social world (people); hence, there is an object-related and social-related aspect to each dimension.

The Russian-language Structure of Temperament Questionnaire (STQ-R) was developed to operationalize the eight dimensions of temperament specified in the structure of temperament model. The English-language version of the Structure of Temperament Questionnaire (STQ-E) was subsequently developed by Rusalov (1989) as an analogue to the original STQ-R. Preliminary attempts to assess the factorial validity of scores on the STQ-E employed exploratory factor analytic (EFA) methods at the scale level. Following the strategy employed by Rusalov with the STQ-R, two studies subjected STQ-E scale scores to a series of factor analyses that extracted two through nine (eight temperament plus one lie scale) factors. Stough, Brebner, and Cooper (1991) administered the STQ-E to a sample of Australian university students and found a factor structure quite similar to the factor structure in the original Russian sample. For example, the two-factor solution showed the expected clustering of the object-related and social-related aspects of temperament. Bishop, Jacks, and Tandy (1993), using a sample of American college students, also found a factor structure similar to the original factor structure.

Although such scale-level factor analyses help to elucidate the interrelationships between the scales (for example, the factor solutions should reveal the object-related and social-related clustering of the scales) and provide a means to compare the factor structure of the instrument across translations, scale-level analysis cannot explicitly confirm the behavior of individual items with regard to their respective scales. Consequently, Dumenci (1996) administered the STQ-E to a sample of American college students and conducted an item-level exploratory factor analysis. He found that the obtained nine-factor solution was inconsistent with the hypothesized nine dimensions of the model. Instead, a three-factor solution (subsequently named Object-Related Activity, Emotionality, and Communicative) emerged in this analysis. Dumenci concluded that the STQ-E should be revised to represent these three factors. It should be noted, however, that a standard factor analysis of dichotomous data (where items are scored yes-no) may produce biased results. Both least squares and maximum likelihood estimation techniques require item distributions that are continuous and normal, assumptions that may be violated with dichotomous response categories (Bandalos, 2002). To achieve unbiased estimates from dichotomous data, researchers may employ programs that utilize special analytic solutions or base their analysis on item parcels (sums of two or more items assessing the same construct). Item parcels have scores that fall on greater-than-dichotomous scales and thus are more likely to satisfy the assumptions of interval scaling and normality. Because Dumenci did not employ either of these options, the inferences drawn from his item-based analysis may be questioned.

The purpose of the present study was to use confirmatory factor analysis (CFA) to assess the factor structure of the STQ-E with a sample of college students. To avoid the problems associated with the factor analysis of dichotomous items, parcels were formed by summing three items from a common scale. The resulting parcel scores fall on a greater-than-dichotomous scale and reduce the number of parameters to be estimated.

Two models were tested. Model 1 was a first-order model consisting of eight dimensions: object-related ergonicity (Er), social-related ergonicity (SEr), object-related plasticity (Pl), social-related plasticity (SPl), object-related tempo (Tp), social-related tempo (STp), object-related emotionality (Em), and social-related emotionality (SEm). Previous research by Rusalov (1989) suggested that several of these dimensions are correlated. Model 2 was a second-order model consisting of two superordinate dimensions of temperament: object-related temperament (O) and social-related temperament (S). In this model, the four object-related dimensions of temperament and the four social-related dimensions of temperament were treated as first-order factors. Using theory and previous research, both the second-order and first-order models seemed plausible. Because there was not a correlation between every social-related aspect of temperament or between every object-related aspect of temperament on the STQ-R (Rusalov, 1989), it could be argued that superordinate factors may not need to be invoked to explain the intercorrelations between the first-order factors. Consequently, we predicted that the first-order model would provide the best fit for the data.

Method

Participants

The STQ was administered to 300 undergraduates enrolled at a private liberal arts college. The sample consisted of 123 males (mean age in years = 19.15, *SD* = 0.99) and 177 females (mean age in years = 18.94, *SD* = 1.90). Although information on ethnic origin was not collected from the participants, the student body of this college is predominately White. All participants received extra credit in a general psychology course.

Instrument

The STQ is a 105-item, self-report instrument designed to measure eight theoretically derived aspects of temperament. Four scales measure object-related individual differences: Er, Pl, Tp, and Em. Four scales measure

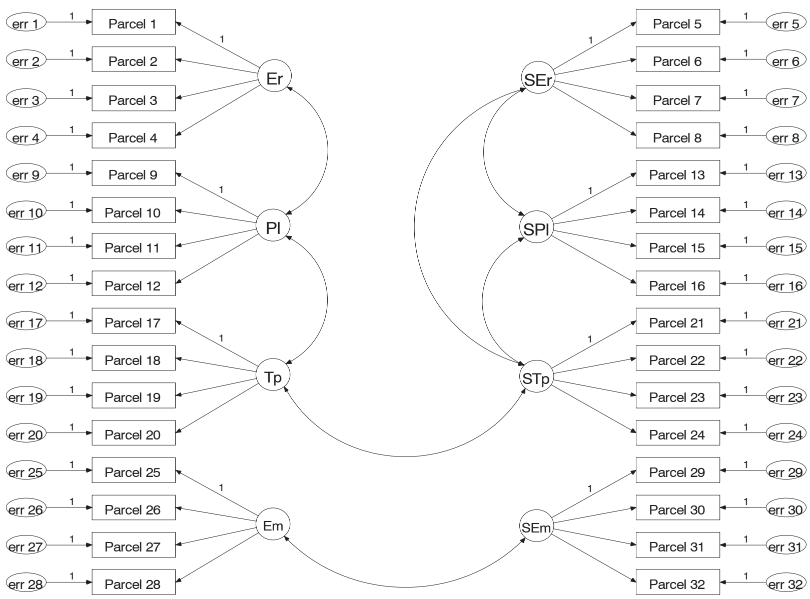

Figure 1. Path diagram for the correlated first-order model.
Note. Er = object-related ergonicity; SEr = social-related ergonicity; Pl = object-related plasticity; SPl = social-related plasticity; Tp = object-related tempo; STp = social-related tempo; Em = object-related emotionality; SEm = and social-related emotionality.

socially related individual differences: SEr, SPl, STp, and SEm. Each scale consists of 12 dichotomous (yes-no) items. A Lie Scale consisting of 9 dichotomous items is used as a measure of social desirability.

Parcel Formation

The 12 items composing each scale were randomly assigned to 4 parcels of 3 items each. Thirty-two parcels were thus created and served as the observed variables in this analysis.

Specification of Models

Two models were evaluated in this study. The first-order model consisted of the eight temperament scales. Utilizing the STQ-R scale intercorrelations above .30 reported by Rusalov (1989), several of the scales were allowed to correlate. Because the measurement scale for each unobserved variable was indeterminate, one factor loading for each unobserved variable was arbitrarily set to one (see Figure 1). The second-order model consisted of two

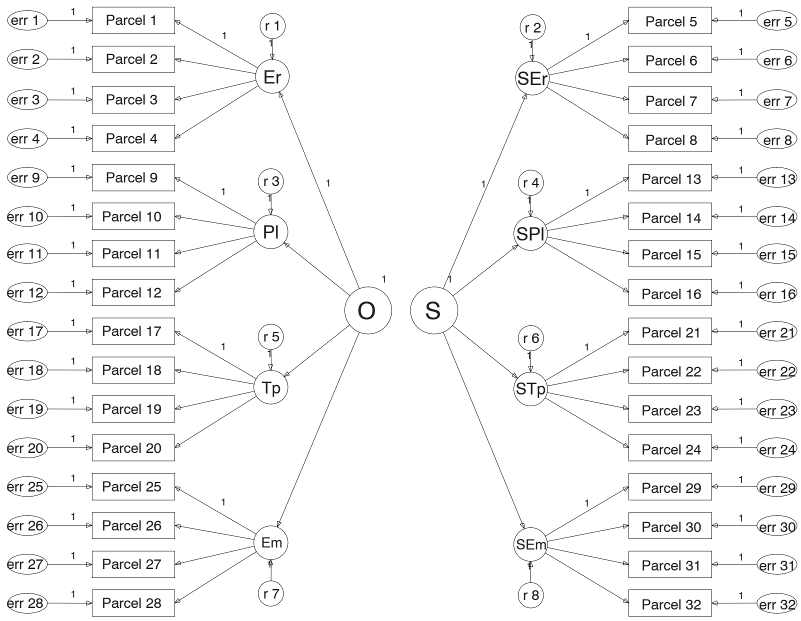

Figure 2. Path diagram for the second-order model.
Note. Er = object-related ergonicity; SEr = social-related ergonicity; Pl = object-related plasticity; SPl = social-related plasticity; Tp = object-related tempo; STp = social-related tempo; Em = object-related emotionality; SEm = and social-related emotionality.

second-order and eight first-order temperament factors. In this model, the object-related and the social-related aspects of each temperament factor served as first-order factors. Once again, one factor loading for each unobserved variable was arbitrarily set to one (see Figure 2).

Results

Descriptive Statistics

The item parcels were examined for their distributional characteristics. Parcel means, standard errors, skewness, and kurtosis statistics are presented in Table 1. An inspection of the table suggests that the item parcels generally satisfied assumptions of normality, with some tendency toward platykurtic distributions for a few parcels.

Table 1
Descriptive Statistics for Parcels

Parcel	*M*	*SE*	Skewness	Kurtosis
1	1.66	0.05	–0.07	–0.85
2	2.07	0.05	–0.53	–0.27
3	1.88	0.05	–0.38	–0.47
4	1.98	0.04	–0.40	0.29
5	2.42	0.05	–1.19	0.64
6	2.24	0.05	–0.92	–0.13
7	1.66	0.06	–0.25	–0.86
8	2.34	0.05	–1.11	0.22
9	2.53	0.04	–1.41	1.32
10	1.88	0.06	–0.38	–0.85
11	1.81	0.06	–0.35	–1.01
12	2.16	0.05	–0.82	0.00
13	1.60	0.06	–0.01	–1.22
14	1.49	0.06	–0.03	–1.02
15	1.53	0.06	–0.11	–0.90
16	1.23	0.06	0.33	–1.09
17	2.05	0.06	–0.65	–0.66
18	1.85	0.06	–0.35	–1.03
19	2.25	0.05	–0.88	0.10
20	2.10	0.05	–0.62	–0.48
21	2.14	0.06	–0.78	–0.60
22	1.81	0.06	–0.37	–1.12
23	1.88	0.06	–0.45	–0.77
24	2.21	0.04	–0.70	0.08
25	1.29	0.06	0.19	–1.13
26	1.27	0.06	0.31	–1.19
27	1.73	0.04	0.08	–0.56
28	1.33	0.06	0.27	–1.19
29	2.40	0.04	–0.85	–0.03
30	1.45	0.05	0.21	–0.59
31	1.42	0.05	0.07	–0.70
32	2.21	0.04	–0.65	–0.10

CFA

The two models were subjected to a maximum-likelihood CFA using AMOS 3.62 (Arbuckle, 1997). Table 2 presents the fit statistics for the two models. Several fit indices were examined to evaluate the overall fit of each model. The chi-square goodness-of-fit statistic was statistically significant for both models, suggesting that neither fit the data. However, the chi-square statistic is sensitive to sample size, so it is rarely used as a sole index of model fit. An adjunct discrepancy-based fit index is the ratio of chi-square to degrees of freedom (χ^2/df). Carmines and McIver (1981) suggested that ratios in the range of 2 to 3 are indicative of an acceptable fit between the

Table 2
Fit Indices for the Two Models

Model	χ^2	χ^2/df	NFI	TLI	CFI	RMSEA
First-order	1,207.416	2.642	.939	.955	.961	.074
Second-order	1,431.954	3.127	.928	.942	.950	.084

Note. NFI = normed fit index; TLI = Tucker-Lewis Index; CFI = comparative fit index; RMSEA = root mean square error of approximation.

hypothesized model and the sample data. Using this standard for fit, the first-order model demonstrated acceptable fit.

Three incremental indices of fit were examined next: the normed fit index (NFI), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). Incremental indices reflect the improvement in fit gained by a given factor model relative to the most restrictive (null or independence) model. All three incremental indices are scaled from 0 (no fit) to 1 (perfect fit). Hu and Bentler (1999) advised that values close to .95 are indicative of good fit. By this standard, the first-order model had two coefficients that suggested a good fit. Finally, the root mean square error of approximation (RMSEA) is a population discrepancy function that compensates for the effects of model complexity. The closer the RMSEA coefficient is to 0, the better the fit of the model. According to Browne and Cudek (1993), a RMSEA value of .05 or less indicates a close fit of the model in relation to the degrees of freedom, whereas a value of .08 or less indicates a reasonable error of approximation. Once again, the first-order model had an RMSEA coefficient within the acceptable range. Taken together, these five indices suggested that the hypothesized first-order model was a reasonable fit for the data in this study. The improvement in fit between the first-order and second-order model was statistically significant ($\Delta\chi^2 = 224.54$, $\Delta df = 1$, $p < .001$).

The factor intercorrelations for the eight first-order factors are presented in Table 3. With one exception, the moderate correlations between factors suggest adequate discriminant validity between these related aspects of temperament. In other words, as specified by theory, the scales measure related but still separate aspects of temperament. The moderate correlations also explain the reasonable (but not better) fit of the second-order model. In contrast, the strong correlation between Em and SEm suggests a lack of discriminant validity. However, Rusalov (1989) argued that the correlation between Em and SEm reflects the common underlying mechanisms of emotional sensitivity that are independent of the sphere of human activity. Hence, from a theoretical point of view, this correlation is not problematic.

The first-order model factor pattern and structure coefficients are presented in Table 4. All pattern coefficients for the 32 parcels were statistically significant ($p < .01$). Thompson (1997) and Graham, Gutherie, and Thompson (2003) argued that both pattern and structure coefficients should

Table 3
Factor Intercorrlations for the First-Order Model

Factor	Er	SEr	Pl	SPl	Tp	STp	Em	SEm
Er	—	—	.52*	—	—	—	—	—
SEr		—	—	.63*	—	.63*	—	—
Pl			—	—	.35*	—	—	—
SPl				—	—	.69*	—	—
Tp					—	.22*	—	—
STp						—	—	—
Em							—	.90*
SEm								—

Note. Er = object-related ergonicity; SEr = social-related ergonicity; Pl = object-related plasticity; SPl = social-related plasticity; Tp = object-related tempo; STp = social-related tempo; Em = object-related emotionality; SEm = and social-related emotionality.
$^{*}p < .01$.

be examined in CFA when the model in question has correlated factors. A failure to consult the structure coefficients could lead to interpretation errors. An inspection of the table reveals that the structure coefficients for the parcels with pattern coefficients fixed to zero were generally smaller than the structure coefficients that were free to vary. This suggests an appropriate model specification and no need to free additional parameters, yielding an interpretation consistent with the pattern coefficients.

Discussion

A CFA assessed the factor structure of scores on the STQ-E in a sample of young adult college students. Using item parcels, the confirmatory analysis indicated that the subscales of the STQ-E are best represented as eight separate but interrelated (correlated) dimensions of temperament. Contrary to the item-based exploratory factor analysis performed by Dumenci (1996), the current CFA structure of the STQ-E conformed to both theory and prior research. Although the factor structure of scores on the STQ-E can be deemed adequate with this adult population, the homogeneity of the sample used in this study certainly places limits on the generalizability of these findings. Further confirmatory research with other adult age groups and populations is needed. In addition, research is needed to validate scores from the STQ-E against multiple external criteria. Given that the model suggests that temperament may be manifest somewhat differently in the social world versus the world of objects, future research could focus on existing groups where social versus object differences in temperament are likely to be manifest.

Table 4
Factor Pattern and Structure Coefficients for the First-Order Model

	Er		SEr		Pl		SPl		Tp		STp		Em		SEm	
Parcel	p	r_s	p	r_s	p	r_s	p	r_s	p	r_s	p	r_s	p	r_s	p	r_s
1	.68	.68	.00	.00	.00	.35	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
2	.46	.46	.00	.00	.00	.24	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
3	.71	.71	.00	.00	.00	.37	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
4	.63	.63	.00	.00	.00	.32	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
5	.00	.00	.69	.69	.00	.00	.00	.43	.00	.00	.00	.44	.00	.00	.00	.00
6	.00	.00	.74	.74	.00	.00	.00	.47	.00	.00	.00	.47	.00	.00	.00	.00
7	.00	.00	.71	.71	.00	.00	.00	.45	.00	.00	.00	.45	.00	.00	.00	.00
8	.00	.00	.52	.52	.00	.00	.00	.33	.00	.00	.00	.33	.00	.00	.00	.00
9	.00	.26	.00	.00	.50	.50	.00	.00	.00	.18	.00	.00	.00	.00	.00	.00
10	.00	.37	.00	.00	.72	.72	.00	.00	.00	.25	.00	.00	.00	.00	.00	.00
11	.00	.40	.00	.00	.77	.77	.00	.00	.00	.27	.00	.00	.00	.00	.00	.00
12	.00	.35	.00	.00	.69	.69	.00	.00	.00	.24	.00	.00	.00	.00	.00	.00
13	.00	.00	.00	.31	.00	.00	.50	.50	.00	.00	.00	.34	.00	.00	.00	.00
14	.00	.00	.00	.43	.00	.00	.68	.68	.00	.00	.00	.47	.00	.00	.00	.00
15	.00	.00	.00	.44	.00	.00	.70	.70	.00	.00	.00	.48	.00	.00	.00	.00
16	.00	.00	.00	.37	.00	.00	.58	.58	.00	.00	.00	.40	.00	.00	.00	.00
17	.00	.00	.00	.00	.00	.24	.00	.00	.67	.67	.00	.15	.00	.00	.00	.00
18	.00	.00	.00	.00	.00	.20	.00	.00	.58	.58	.00	.13	.00	.00	.00	.00
19	.00	.00	.00	.00	.00	.23	.00	.00	.64	.64	.00	.14	.00	.00	.00	.00
20	.00	.00	.00	.00	.00	.27	.00	.00	.75	.75	.00	.17	.00	.00	.00	.00
21	.00	.00	.00	.29	.00	.00	.00	.31	.00	.10	.45	.45	.00	.00	.00	.00

(continued)

Table 4 *(Continued)*

	Er		SEr		Pl		SPl		Tp		STp		Em		SEm	
Parcel	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s
22	.00	.00	.00	.52	.00	.00	.00	.56	.00	.18	.82	.82	.00	.00	.00	.00
23	.00	.00	.00	.37	.00	.00	.00	.41	.00	.13	.59	.59	.00	.00	.00	.00
24	.00	.00	.00	.40	.00	.00	.00	.43	.00	.14	.63	.63	.00	.00	.00	.00
25	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.67	.67	.00	.60
26	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.77	.77	.00	.69
27	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.64	.64	.00	.58
28	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.73	.73	.00	.66
29	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.37	.41	.41
30	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.45	.50	.50
31	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.58	.64	.65
32	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.45	.49	.49

Note. Er = object-related ergonicity; SEr = social-related ergonicity; Pl = object-related plasticity; SPl = social-related plasticity; Tp = object-related tempo; STp = social-related tempo; Em = object-related emotionality; SEm = and social-related emotionality.

References

Arbuckle, J. L. (1997). *Amos user's guide version 3.6*. Chicago: SmallWaters Corporation.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, *9*(1), 78-102.

Bishop, D., Jacks, H., & Tandy, S. B. (1993). The Structure of Temperament Questionnaire (STQ): Results from a U.S. sample. *Personality and Individual Differences*, *14*, 485-487.

Browne, M. W., & Cudek, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models.* Newbury Park, CA: Sage.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables. In G. W. Bohrnstedt, & E. F. Borgatta (Eds.) *Social measurement: Current issues*. Beverly Hills, CA: Sage.

Dumenci, L. (1996). Factorial validity of scores on the Structure of Temperament Questionnaire. *Educational and Psychological Measurement*, *56*, 487-493.

Graham, J. M., Gutherie, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling*, *10*, 142-153.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Rusalov, V. M. (1989). Object-related and communicative aspects of human temperament: A new questionnaire of the structure of temperament. *Personality and Individual Differences*, *10*, 817-827.

Stough, C., Brebner, J., & Cooper, C. (1991). The Rusalov Structure of Temperament Questionnaire (STQ): Results from an Australian sample. *Personality and Individual Differences*, *12*, 1355-1357.

Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement*, *57*, 5-19.