JOURNAL OF PERSONALITY ASSESSMENT, *82*(1), 71–79
Copyright © 2004, Lawrence Erlbaum Associates, Inc.

ARTICLES

Temperament in Adults—Reliability, Stability, and Factor Structure of the EAS Temperament Survey

Ane Nærde, Espen Røysamb, and Kristian Tambs
Division of Epidemiology
Norwegian Institute of Public Health
Oslo, Norway

We assessed the factor structure and psychometric properties of the Emotionality, Activity, and Sociability (EAS) Temperament Survey (Buss & Plomin, 1984) for adults using a longitudinal sample of adult women. The stability estimates of the EAS instrument were assessed over a period of 3 years. The results indicated an acceptable fit for the basic theoretical EAS model, implying that the scale is functioning satisfactory. However, the results also suggest that the measure could be improved. Across time, latent stability factors explained within-scale covariances. Both latent stability factors and time-specific factors accounted for cross-sectional covariances between subscales. Additional research is warranted to guide the further development of the EAS model.

There is a general agreement that *temperament* refers to the stylistic component of an individual's mental or behavioral repertoire (Buss & Plomin, 1984). Temperament theories differ regarding the definition of temperament, the number and structure of temperament traits, and the importance of specific factors determining individual differences in temperament. Although much research has focused on temperament in relation to infant behavior, there are fewer studies on temperament in adults.

The theory of temperament on which the Emotionality, Activity, and Sociability (EAS) Temperament Survey (Buss & Plomin, 1984) is based regards temperament as a subclass of personality traits characterized by appearance during the 1st year of life, persistence later in life, and a high contribution of heredity (Buss & Plomin, 1984). The three personality traits that meet these criteria are emotionality, activity, and sociability, from which are derived the acronym EAS (Buss & Plomin, 1984). *Emotionality* is defined as primordial distress, which is assumed to differentiate into fear and anger during the first 6 months of life. *Activity* is defined as the sheer expenditure of physical energy, and *sociability* is defined as a preference for being with others rather than being alone. Emotionality, activity, and sociability are found in various forms in almost every model of temperament (Rutter, 1989). (For a more detailed description of the components of the three temperaments as well as the motivational aspects of each, see Buss, 1989; Buss & Plomin, 1984, 1986). The EAS theory was introduced by Buss and Plomin in 1975 and originally included four temperaments—emotionality, activity, sociability, and impulsivity (EASI). The items relating to impulsivity were eventually omitted because the evidence for its inheritance was mixed, and also, this factor did not reliably emerge from factor analytic studies (Buss & Plomin, 1975; Rowe & Plomin, 1977). Two instruments were originally developed by the use of exploratory factor analysis: a temperament survey for children (parental ratings) and a self-report inventory for adults. The instruments have undergone several revisions (EASI I, EASI II, EASI III). The final adult EAS version (Buss & Plomin, 1984) is a 20-item scale measuring each of the following traits: activity, sociability, and the three subdimensions anger, distress, and fearfulness, being components of emotionality.

A considerable amount of research has been conducted with the adult EAS scale and its variants (Angleitner & Ostendorf, 1994; Digman, 1989, 1990; John, 1989; McCrae & Costa, 1985; Plomin, Pedersen, McClearn, Nesselroade, & Bergman, 1988; Ruch, Angleitner, & Strelau, 1991; Windle, 1989a, 1989b). Published studies on the scales' psychometric properties are surprisingly scarce apart from those in which the scales were initially derived. We have only been able to identify one study reporting on the psychometric properties of the adult EASI–III scale

(Braithwaite, Duncan-Jones, Bosly-Craft, & Goodchild, 1984). Thus, for the final EAS version of the adult temperament survey, there is no available psychometric information other than the original statistics reported by Buss and Plomin (1984). Moreover, common to all the EASI/EAS versions is that the data used by Buss and Plomin (1975, 1984) to derive the scales are based on nonrandom and nonrepresentative samples consisting of college students. Also, previous stability estimates of the EAS scale over time are based on test–retest correlations only, and there is a lack of studies applying more refined methods such as structural equation programs.

The EAS theory represents one of several theoretical approaches to temperament. We necessarily need to limit our coverage of alternative models. A previous roundtable discussion of four different approaches to the concept of temperament (including the EAS model) illustrates some links and divergences among different theories (Goldsmith et al., 1987).

The purpose of this study is twofold. First, we assessed the factor structure and psychometric properties of the EAS Temperament Survey for adults. Second, we assessed the stability estimates of the EAS instrument over a period of 3 years. More specifically, the purpose of the study was to (a) test the fit of the factor structure presented by Buss and Plomin (1984), including a five-factor and a three-factor solution; (b) estimate the correlations between factors when controlling for measurement errors; (c) investigate whether test–retest correlations are due to occasion-specific latent stability factors or due to transmission of effects from one time point to another; and (d) explore how interfactor correlations observed at any point in time are due to covariances between the stable or the time-specific parts of the factors.

MATERIALS AND METHOD

Participants

The data were derived from a longitudinal epidemiological survey of mothers and their preschool children to investigate early childhood behavior problems (Mathiesen & Sanson, 2000). Participants were recruited from a nationwide, mandatory child public health program that includes about 95% of all families with preschool children in Norway. Most often the mothers, and in a few cases, the fathers or both parents, accompany their child at the health examinations. This study utilizes only data from the mothers. The participants completed a questionnaire when their children were 1.5 years, 2.5 years, and 4 years old. The details concerning the participation rate at each time point have been presented previously (Nærde, Tambs, Mathiesen, Dalgard, & Samuelsen, 2000). The participation rate was 87% at Time 1 (T1; $n = 939$; 1,081 eligible parents), 74% at T2 ($n = 804$; 1,087 eligible parents) and 72% at T3 ($n = 759$; 1,059 eligible parents). Altogether, 63% of the families participated at all three occasions ($n = 682$).

Demographic information for the nonparticipants was obtained from records maintained at the child health clinic (mothers' age, number of children in the family, marital status, mothers' education, and employment status). Attrition analyses showed no significant differences between the participants and the nonparticipants regarding these variables. The age of the participating mothers ranged from 19 to 46 years, with a mean age of 30 years ($SD = 4.7$). Forty-nine percent of the families had only 1 child, 37% had 2, and 15% had 3 to 10 children. Most mothers (91%) were living with a partner. For a more detailed account of the demographic characteristics of the sample, see Mathiesen, Tambs, and Dalgard (1999). The demographic profile of the sample did not vary significantly across time points. Comparisons on the demographic variables between the participants with complete data ($n = 682$) and the participants who dropped out at either T2 or T3 ($n = 262$) revealed two significant differences between the groups (Nærde, Tambs, & Mathiesen, 2001). The dropouts were significantly younger than the participants (28 years vs. 30 years), $t(942) = 3.24$, $p = .001$. Also, significantly fewer of the dropouts were employed compared to the participants (56% vs. 67%; $p = .002$). Finally, there were no significant differences between the participants with complete versus incomplete data regarding the EAS measures.

Method

Temperament was assessed by the EAS Temperament Survey for adults (Buss & Plomin, 1984). This is the newest version of the instrument and has 4 items corresponding to each of the five subscales. Each of the items was rated on a Likert scale ranging from 1 (*not characteristic or typical of yourself*) to 5 (*very characteristic or typical of yourself*). The scores from the questions belonging to each of the five scales were added and each scale score was divided by 4 (number of items per scale). The 20 EAS items are listed in Table 1. The standard EAS measure was translated from English to Norwegian by a British psychologist who had spent several years working at a Norwegian research center. To assure that the meaning of the original items were retained, a blind back translation from Norwegian to English was completed by a Canadian researcher working at the same center. The back-translated version had slight differences in meaning for three of the EAS questions. These differences were discussed among the two translators and five Norwegian researchers until consensus was reached regarding Norwegian equivalents for the original items.

Statistical Analyses

Structural equation modeling (SEM) was used to evaluate various aspects of the EAS. In general, the application of SEM involves several advantages as compared to more traditional multivariate methods. First, SEM enables the simultaneous test of a number of interrelations between variables.

Second, SEM analyses provide fit measures for an entire model, in addition to regular path coefficients and multiple *R*s. These fit measures reflect how well the proposed model recaptures the variance–covariance structure of the data under analysis. Thus, a theoretically based model can be tested against the data. Third, random measurement errors that are typically comprised in ordinary sum-score indexes can be eliminated by modeling latent factors (Bollen, 1989; Hoyle, 1995; Jöreskog & Sörbom, 1993; Loehlin, 1998). Due to skewness in the data, the analyses were based on polychoric correlations and asymptotic covariance matrices computed with Prelis (Jöreskog & Sörbom, 1993).

In these analyses, SEM was first used to perform confirmatory factor analyses (CFA) in which the theoretical five-dimensional EAS model (activity, sociability, anger, distress, and fearfulness) was tested against data at each of the three time points. In the next step, we tested for the presence of higher order factors corresponding with the three-dimensional EAS model (emotionality, activity, and sociability) and tested whether this model fitted the data better than the five-dimensional model. We adopted a strategy of analysis involving first the testing of the basic theoretical model at each of the time points. The basic model comprised a simple structure, allowing for no cross-loadings and no correlated errors. Further, to identify covariances not accounted for by this model and investigate to what extent the model could be improved, we tested alternative models by allowing cross-loadings based on the modification indexes provided.

In the next set of analyses, a longitudinal model was analyzed using SEM and the EQS computer program (Bentler, 1995).[1] This model investigated the nature of the cross-time, same-trait covariances as well as the same-time, cross-trait covariances within the EAS. These analyses were based on the variance–covariance matrices of the sum-score indexes of each of the five temperament dimensions.[2] The strategy of analysis involved testing alternative hypotheses by relaxing sets of constraints in a stepwise fashion. We elaborate on the specific constraints in the Results section.

The overall fit of the models was assessed with the goodness of fit index (GFI), comparative fit index (CFI) and the root mean square error of approximation (RMSEA). There is not full consensus about the selection of a standard for determining "adequate" fit. Thus, we chose to use standards commonly accepted today. An adequate model should have a GFI or a CFI of at least 0.90 (Hoyle, 1995). Moreover, a RMSEA index lower than 0.05 indicates a very good fit, whereas one up to 0.08 indicates a reasonably good fit (Browne & Cudeck, 1993; Jöreskog & Sörbom, 1993). Chi-square statistics were used to compare alternative models and to develop modifications to the basic models (Bentler, 1995; Bollen, 1989; Hoyle, 1995; Jöreskog & Sörbom, 1993).

RESULTS

Descriptive Statistics for the 20 Items and the Five Subscales of the EAS

Descriptive statistics for the 20 items of the EAS scale at T1 (means and standard deviations) and polychoric test–retest correlations are summarized in Table 1. The single item test–retest correlations pooled across the 20 items were .62 for T1 to T2, .63 for T2 to T3, and .58 for T1 to T3.[3]

Descriptive statistics for the five subscales of the EAS (means and standard deviations) and Cronbach alpha reliability estimates for the subscales at T1, T2, and T3 are presented in Table 2. The means for each of the subscales did not change significantly over time. The Cronbach alpha reliability ranged from 0.53 to 0.75. The skewness for the five subscales at the three time points ranged from –0.76 to 0.51, whereas the kurtosis ranged from –0.22 to 0.58.

The test–retest correlations between the EAS subscales at T1, T2, and T3 and pooled Pearson intercorrelations (at T1, T2, and T3) between the five dimensions are listed in Table 3. The test–retest correlations ranged from .61 to .72 and did not change significantly over time. The intercorrelations between the temperament dimensions ranged from –.12 to .61.

Factor Structure of the EAS

The basic five-dimensional model yielded an acceptable fit. Based on the T1 data, $\chi^2(160, N = 921) = 856.70$; GFI = 0.96; CFI = 0.94; and RMSEA = 0.069. A similar fit was obtained for the T2 data, $\chi^2(160, N = 784) = 879.85$; GFI = 0.95; CFI = 0.94; RMSEA = 0.076. Likewise, the fit at T3 was $\chi^2(160, N = 736) = 840.66$; GFI = 0.95; CFI = 0.94; and RMSEA = 0.076.

To obtain factor loadings based on all three time points, a three-group model was tested that included one group for each point of time and in which all parameters were con-

[1]The reason for switching from LISREL to EQS was that inasmuch as the first set of analyses was based on categorical and skewed data, the use of polychoric correlations, asymptotic covariance matrices, and weighted least squares estimation would be more appropriate than applying ordinary covariance matrices and maximum likelihood estimation. Because Lisrel and Prelis are better developed to handle these situations, we used Lisrel for the first set of analyses and then returned to our main SEM program, namely EQS, for the remaining analyses involving sum-score indexes as observed variables.

[2]Regarding the choice of the specific type of association matrix used, polychoric correlations were used in the first set of analyses in which the observed variables were categorical and skewed and the analyses were cross-sectional. Further, in the remaining analyses based on sum-score indexes as observed variables, we used variance–covariance matrices, which in general comprise more information than ordinary correlation matrices and typically are preferable in longitudinal analyses.

[3]By *pooled* we mean the average of the *z* transforms of the correlation values transformed back to the corresponding correlation value.

TABLE 1
Descriptive Statistics (T1 Only) and Polychoric Test–Retest Correlations for the 20 Items of the Emotionality, Activity, Sociability Temperament Survey

Item	*M*	*SD*	*Test–Retest Correlation*[a]
1. I like to be with people (Soc)	4.09	0.80	.71
2. I usually seem to be in a hurry (Act)	3.39	0.79	.63
3. I am easily frightened (E–f)	2.16	1.10	.67
4. I frequently get distressed (E–d)	2.77	0.91	.63
5. When displeased, I let people know it right away (E–a)	3.56	0.91	.66
6. I am something of a loner (rev., Soc)	1.90	1.08	.70
7. I like to keep busy all the time (Act)	3.16	0.85	.61
8. I am known as hot blooded and quick-tempered (E–a)	3.01	1.13	.77
9. I often feel frustrated (E–d)	2.67	0.99	.60
10. My life is fast paced (Act)	2.99	1.01	.62
11. Everyday events make me troubled and fretful (E–d)	2.43	1.05	.53
12. I often feel insecure (E–f)	2.70	0.99	.66
13. There are many things that annoy me (E–a)	2.99	0.90	.56
14. When I get scared, I panic (E–f)	1.80	1.05	.55
15. I prefer working with others rather than alone (Soc)	3.47	0.94	.62
16. I get emotionally upset easily (E–d)	2.89	0.99	.58
17. I often feel as if I'm bursting with energy (Act)	3.24	0.84	.56
18. It takes a lot to make me mad (rev., E–a)	3.12	0.88	.67
19. I have fewer fears than most people my age (rev., E–f)	2.95	1.01	.39
20. I find people more stimulating than anything else (Soc)	3.26	0.81	.66

Note. T = time; Soc = sociability; Act = Activity; E–f = Emotionality–fearfulness; E–d = Emotionality–distress; E–a = Emotionality–anger; rev. = reversed item.
[a]Pooled values based on estimates from T1 to T2 and T2 to T3.

TABLE 2
Descriptive Statistics and Chronbach Alpha for the Five Emotionality, Activity, Sociability, Temperament Survey at T1, T2, and T3

	Mean			*SD*			α		
Dimension	*T1*[a]	T2b	*T3*[c]	*T1*	*T2*	*T3*	*T1*	*T2*	*T3*
Distress	2.36	2.33	2.32	0.73	0.72	0.69	.73	.75	.74
Fearfulness	2.25	2.22	2.19	0.63	0.60	0.57	.57	.55	.56
Anger	2.99	2.94	2.92	0.70	0.69	0.66	.57	.62	.58
Activity	3.01	3.10	3.16	0.69	0.69	0.70	.65	.67	.72
Sociability	3.74	3.76	3.73	0.58	0.60	0.62	.53	.61	.65

Note. T = time.
[a]$N = 921$. [b]$N = 784$. [c]$N = 736$.

strained to be equal across groups. Due to the lack of independence between observations, the fit statistics for such a model are not valid; however, the parameter estimates represent the common structure across time. Figure 1 shows factor loadings and correlations based on the three-group model. As can be seen, factor loadings ranged from 0.17 to 0.72. With the exception of two items (i.e., Item 19 on fearfulness and Item 5 on anger), all loadings were 0.40 or higher. Although the general picture involves evidence of an acceptable fit and an acceptable level of factor loadings, we also found indications of less than optimal aspects of the scale.

The factor correlations ranged from .11 to .96 for the five-dimensional model. In particular, note the high correlation between distress and fearfulness, suggesting that when random measurement error is eliminated, these two factors are practically representing the same phenomenon. We tested this notion specifically by constraining the factor correlation to unity. At T1 and T2 this model did not fit worse than the original model in which the parameter was freely estimated, $\Delta\chi^2(1, N = 921) = 2.32$, *ns* (T1); $\Delta\chi^2(1, N = 784) = 1.19$, *ns* (T2). At T3, there was a slight but significant reduction in fit, $\Delta\chi^2(1, N = 736) = 6.29$, $p < .05$. In total, these analyses support the notion of the Distress and Fearfulness scales in the EAS as measures of virtually the same underlying dimension.

In the next set of analyses, a higher order, three-dimensional EAS model was tested against the data at each of the three time points. This model also yielded acceptable fit. Based on the T1 data, $\chi^2(164, N = 921) = 887.42$; GFI = 0.96; CFI = 0.93; and RMSEA = 0.069. At T2, a similar fit was obtained, $\chi^2(164, N = 784) = 906.17$; GFI = 0.95; CFI = 0.93; and RMSEA = 0.076. Likewise, the fit at T3 was $\chi^2(164, N = 736) = 867.58$; GFI = 0.95; CFI = 0.94; and RMSEA = 0.076. The three-dimensional model did, however, fit significantly worse than the five-dimensional model at all time points, $\Delta\chi^2(4, N = 921) = 30.72$, $p < .001$ (T1); $\Delta\chi^2(4, N = 784) = 26.32$, $p < .001$ (T2); and $\Delta\chi^2(4, N = 736) = 26.92$, $p < .001$ (T3). For this reason, we did not perform any further analyses with the three-dimensional model.

The next step in the confirmatory analyses was to modify the five-dimensional model by relaxing the constraints of no cross-loadings. This model allowed for the estimation of the parameters that would yield the highest contribution to decrease in chi-square. In a step-wise manner, parameters were freed one at a time until no significant improvements could be found. No restrictions were applied regarding which factor loaded on which item, and thereby, all potential cross-loadings could be estimated. These analyses were performed on the T1 sample. A total of 16 significant cross-loadings were identified, and the model yielded good fit, $\chi^2(144, N = 921) = 400.82$; GFI = 0.98; CFI = 0.98; RMSEA = 0.044. However, when such modifications are performed ad hoc, there is a risk of capitalizing on chance characteristics. To avoid the time-specific characteristics and to identify only the cross-loadings that were present over time, this modified model was then tested on the T2 data. Four of the 16 cross-loadings from T1 were no longer significant and were hence omitted. Thereafter, this model was tested on the T3 data, and another 4 cross-loadings were found to be insignificant. After testing

TABLE 3
Test–Retest Correlations for the Five Emotionality, Activity, Sociability, Temperament Survey (for T1, T2, and T3) and Pooled Intercorrelations (for T1, T2, and T3) Between the Dimensions

				Correlation Matrix							
Dimension	*T1 to T2*	*T2 to T3*	*T1 to T3*	*Fearfulness*		*Anger*		*Activity*		*Sociability*	
Distress	.66	.66	.62	.61	.52	.30	.37	.07	−.08	−.12	−.04
Fearfulness	.62	.68	.62			.10	.17	−.11	−.02	−.12	−.06
Anger	.73	.76	.71					.28	.05	−.01	−.06
Activity	.67	.68	.65							.28	.21
Sociability	.66	.71	.65								

Note. Intercorrelations from Buss and Plomin (1984) are underlined ($n = 220$). All correlations were significant at the .01 level (two-tailed). T = time.

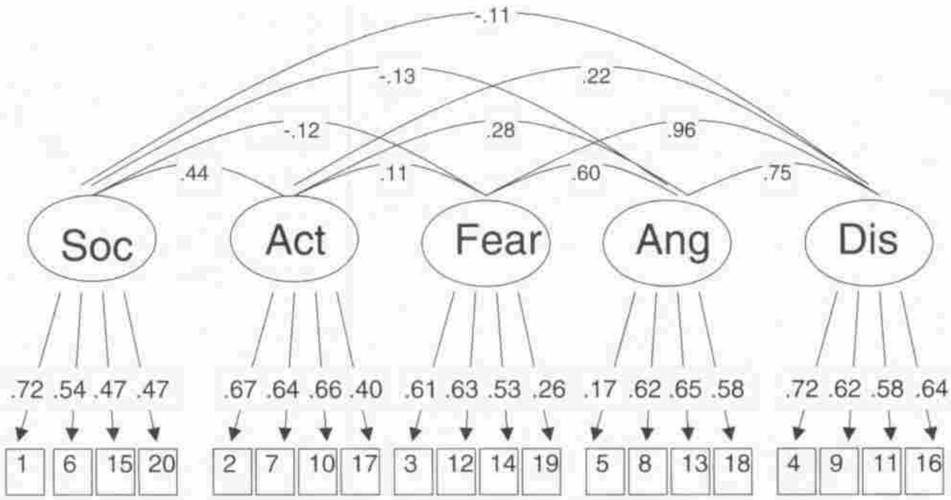

FIGURE 1 Confirmatory factor analysis of the five-dimensional Emotionality, Activity, and Sociability Temperament Survey; standardized coefficients. Soc = sociability; Act = activity; Fear = fearfulness; Ang = anger; and Dis = distress.

across time, a total of 8 cross-loadings were found to be significant at all three time points, whereas the remaining cross-loadings appeared to be due to chance characteristics specific to single time points. The final five-dimensional model with the remaining 8 cross-loadings yielded good fit. Based on the T1 data, $\chi^2(152, N = 921) = 496.15$; GFI = 0.98; CFI = 0.97; and RMSEA = 0.050. At T2, a similar fit was obtained, $\chi^2(152, N = 784) = 448.16$; GFI = 0.98; CFI = 0.98; and RMSEA = 0.050. Likewise, the fit at T3 was $\chi^2(152, N = 736) = 455.47$; GFI = 0.97; CFI = 0.97; and RMSEA = 0.052.

The factor loadings of the final three-group model ranged from 0.20 to 0.85, with a total of eight cross-loadings. Particularly noteworthy is Item 13 loading only 0.28 on the original Anger factor but loading 0.53 on the Fearfulness factor. Item 2 loaded 0.51 on the original activity dimension as well as 0.31 on the sociability dimension, whereas Item 5 loaded 0.44 on the original anger dimension and –0.35 on the fear dimension. Note also that three items in the Emotionality factor (4, 18, and 19) all load negatively on the Activity factor. For example, a high score on Item 19 appears to indicate high distress and low activity rather than a pure representation of level of distress only.

Longitudinal Analyses

The first model to be tested comprised a latent stability factor for each temperament. For example, a latent activity factor was modeled that yielded equally strong influence on the activity indexes at each time point and which accounted for all the covariance between the three activity measures. The five latent stability factors were allowed to correlate, but the residual variances for each measure were constrained to be uncorrelated. Figure 2 shows the tested model, with the solid lines indicating the free parameters in the first model.[4] This model yielded acceptable yet not very high fit measures,

[4]Correlations between latents are not shown in the figure due to visual complexity but are reported in Table 4. Also note that the figure depicts standardized coefficients. In the modeling analyses, the factor loadings (the three time points on the latent stability factor) were constrained to be equal in terms of unstandardized values. As expected, when standardizing, the loadings varied slightly due to differences in variances. Moreover, whereas a three-indicator factor model without constraints would only be just identifiable and thus unable to provide a true test of fit, by constraining loadings to be equal, overidentified models are tested.

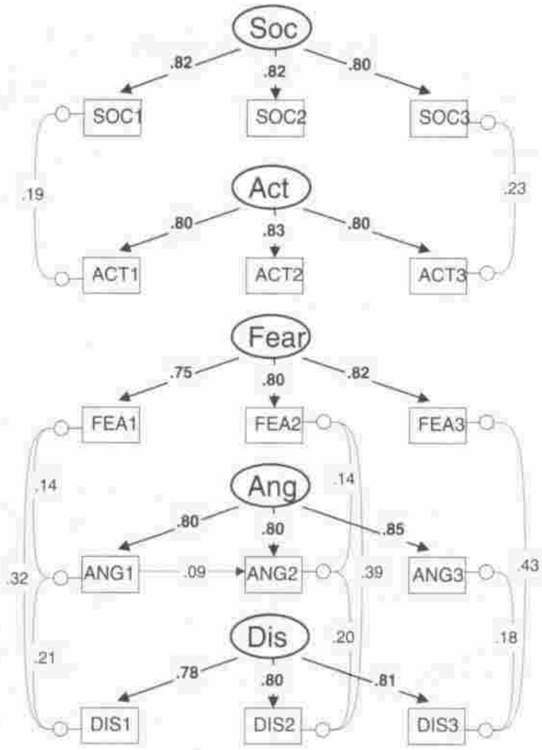

FIGURE 2 Longitudinal model of the five-dimensional Emotionality, Activity, Sociability Temperament Survey with standardized parameters. Solid lines indicate basic model; dotted lines show added parameters contributing to increased fit. $\chi^2(79, N = 682) = 137.53$; GFI = 0.97; CFI = 0.99; and RMSEA = 0.033. Rectangular boxes represent observed sum-score indexes at Time 1 (T1), T2, and T3; ellipses represent latent stability factors. Soc = sociability; Act = activity; Fear = fearfulness; Ang = anger; Dis = distress.

$\chi^2(90, N = 682) = 466.11$; GFI = 0.92; CFI = 0.93; and RMSEA = 0.078.

The first model involved the rather strict assumption of no direct transmission of effects over time. An alternative model would include direct effects across time within each temperament. We applied the Lagrange Multiplier Test (LM; Bentler, 1995) to identify potential direct effects. This test evaluates the effect of adding free parameters to a restricted model (i.e., reducing the restrictions of the model) and the LM statistics can be interpreted as an approximate decrease in model goodness-of-fit chi-square resulting from freeing previously fixed parameters and from eliminating equality restrictions (Bentler, 1995). Among the 10 possible direct effects, only one path yielded a significantly lower chi-square; that is the effect of anger at T1 on anger at T2 (depicted as dotted arrow in Figure 2). The fit was $\chi^2(89, N = 682) = 451.30$; GFI = 0.92; CFI = 0.94; and RMSEA = 0.077. Except for this single freed parameter, all the cross-time covariances between the same temperaments appeared to be accounted for by latent stability factors.

The next step involved investigating cross-trait residual correlations within each time point. Whereas the first model assumed that all cross-trait correlations could be accounted for by correlations between the latent stability factors, we also wanted to identify potential time-specific correlations across traits. Again, we used the LM Test to identify which parameters, if freed, would contribute to significantly better fit. A total of 10 residual correlations were identified indicating time-specific correlations between Sociability and Activity, and between the three Emotionality dimensions (depicted as dotted curves in Figure 2). Estimation of these parameters strongly contributed to improved fit, and this final model accounted for the observed data structure to a high extent, $\chi^2(79, N = 682) = 137.53$; GFI = 0.97; CFI = 0.99; and RMSEA = 0.033. See Figure 2 for parameter estimates. Thus, it appeared that from one time point to another, there was a clear codevelopment among the measures, for example, a relative change in fearfulness from T1 to T2 (indicated by the residual time-specific variance at T2) occurs in parallel with a corresponding change in distress.

The correlations between the latent factors are not included in Figure 2 but are presented in Table 4. The values vary considerably; of particular note is that the correlation between the stable fearfulness factor and the stable distress factor is .73. Using regular path analytic tracing rules, we decomposed the total fearfulness–distress correlation into amounts due to stable and time-specific correlations. Averaging across the three time points, this calculation estimated that 77% of the observed correlations between the fearfulness and distress indexes are attributable to the correlation between the stable fearfulness and distress factors, and the remaining 23% of the correlations are accounted for by time-specific correlations.

In summary, the results show that latent stable factors account for the covariances within each temperament measure across time. Furthermore, the covariances between the different temperament indexes were both due to correlations between the latent stability factors and correlations between the time-specific variances not accounted for by the latent factors.

DISCUSSION

The purpose of this study was to examine the psychometric properties of the newest version of Buss and Plomin's (1984)

TABLE 4
Estimated Correlations Between the Latent Stable Factors of the Emotionality, Activity, Sociability, Temperament Survey

Factor	*1*	*2*	*3*	*4*	*5*
1. Sociability					
2. Activity	.35*				
3. Fearfulness	–.15*	–.19*			
4. Anger	.00	.37*	.11*		
5. Distress	–.17*	.04	.73*	.33*	

*$p < .05$.

EAS Temperament Survey for Adults. The lack of available psychometric information for the different EAS versions is surprising given their widespread use and popularity. Although it was pointed out almost 20 years ago that one of the weaknesses with the EAS measure was that the relative newness of the instrument resulted in a paucity of information about its psychometric properties (Braithwaite et al., 1984), this picture is not very different today. To date, there have been no published CFAs of the EAS scale. Our study provides some of this missing information by taking advantage of the implicit strength of more advanced methods such as SEM as well as contributing with longitudinal data.

Comparisons between the results from our study and the results from other studies employing the EAS Temperament Survey for Adults are complicated for several reasons. First, most other studies have used earlier versions of the EAS (e.g., Angleitner & Ostendorf, 1994; Braithwaite et al., 1984; Ruch et al., 1991; Strelau, 1991; Windle, 1989b). Second, the original studies by Buss and Plomin (1984) are based on small samples that except for the 2-week test–retest data provided data from one point in time only.

The means and standard deviations for the five EAS subscales in our study based on indexes of the summed scores are in general in accordance with the ones presented by Buss and Plomin (1984). Our 1-year test–retest correlations are generally lower (on average, 0.15) than those based on a 2-week period (which ranged from 0.75 to 0.85) as reported by Buss and Plomin (1984). This result was expected considering the different time lags in the two studies. An earlier psychometric study of the EAS Temperament Survey for children (Mathiesen & Tambs, 1999) based on the same longitudinal survey as this study reported high stability estimates (mean value corrected for measurement error was 0.81 for children from 18 to 30 months of age, 0.79 for the period 30 to 50 months, and 0.68 for the period 18 to 50 months).

There is hardly any substantial difference between the intercorrelations reported by Buss and Plomin (1984) and ours except for the estimates for the anger and the activity dimension. Assuming that our pooled correlations have the same standard errors as each of the observed correlations (which is a conservative approximation), calculations show a highly significant difference between the two studies for the correlation between anger and activity. It is likely that these correlations between the dimensions are underestimated because they were assessed directly between indexes somewhat contaminated by measurement error. Applying methods that correct for attenuations of correlations due to measurement errors provide more correct estimates, and our results from the SEM analyses do imply higher intercorrelations. Several of the earlier EAS studies are inconsistent in that they apply varimax rotation while at the same time reporting fairly high correlations between the different dimensions (Angleitner & Ostendorf, 1994; Braithwaite et al., 1984; Buss & Plomin, 1984; Ruch et al., 1991). Our results imply that all the dimensions in the temperament of emotionality (distress, fearfulness, and anger) are highly intercorrelated. In particular, we found an almost perfect correlation between fearfulness and distress. Thus, serious doubts can be raised as to whether these two scales are measuring empirically distinct constructs.

Due to the low number of items, the estimates of internal consistency for the five EAS scales are only moderate. The highest alpha coefficients across the three time points were found for the Distress scale (an average of .74), whereas the lowest occur for the Fearfulness scale (an average of .56). Buss and Plomin (1984) presented no corresponding information on the reliability of the EAS temperament scales. The reliability estimates reported by Strelau (1991) on the total EAS scale is in accordance with our results. Moreover, studies on earlier versions of the EAS (EASI–II and EASI–III) found moderate to low reliability coefficients (Angleitner & Ostendorf, 1994; Braithwaite et al., 1984; Ruch et al., 1991). It is interesting to note that the dimensions with the weakest alpha reliability coefficients constitute two of the three subdimensions representing the temperament emotionality (fearfulness and anger). Our test–retest correlations, which can be interpreted as reliability estimates, ranging from .62 to .76, suggest that the alpha values slightly underestimate the reliability.

Generally, the main finding from the CFAs suggests an acceptable fit for the basic theoretical five-dimensional EAS model and the results presented herein suggest that the scale is functioning satisfactorily. However, although the model fit is acceptable, it is nevertheless far from perfect and could be improved. The higher order, three-dimensional model yields acceptable but significantly poorer fit. Particularly noteworthy is the unexpected result for the factor loadings of the final three-group, five-dimensional model in which Item 13 ("There are many things that annoy me") loads much higher on the fearfulness dimension than on the original anger dimension. We can only speculate about the reason for this result. Of course it might be related to the actual translation of the measure. The word *annoy* was translated into a Norwegian word, which is synonymous with *irritated* in English. Although we do think that the word *irritated* has a clear connotation of anger, it might well be that the discrepancy in meaning between *annoyed* and *irritated* is a contributing factor to our finding. There might also be other explanations for this result such as cultural differences regarding the expression of emotions. Through SEM analyses, we are able to identify more reliably the poorly functioning items than has been done in earlier studies. Accordingly, if the results from this study can be replicated, a modification of the scale might be advisable.

The longitudinal analyses based on EQS are performed to investigate the nature of stability and change of the EAS scale. Although earlier studies have presented test–retest scores, we investigated these scores further. First we explored, for each of the five temperaments, whether there was a stable factor explaining the correlations across time. Our results do indeed imply that the test–retest correlations are due to a stable factor affecting the measurements at each point in time. Next, we tested whether there were time-specific correlations between the different temperaments. The results suggest several such correlations, particularly for the emotionality dimension. More specifically, the relative change in fearfulness coevolves with the relative change in distress and anger. However, path analyses show that the larger part of covariation between the emotionality scores was due to stable factors, and only approximately one fourth was due to time-specific factors.

A few caveats are warranted regarding these findings. First, the sample includes women only. A balanced number of male and female participants would have permitted assessment of gender differences and similarities for the constructs measured. Nevertheless, it has been claimed that the factor structure of the EAS is similar for women and men (Plomin et al., 1988), and Buss and Plomin (1984) reported only small differences between women and men regarding the results from factor analysis. No strong gender differences were found in the earlier mentioned study on the EAS questionnaire for children (Mathiesen & Tambs, 1999). Nonetheless, future studies would clearly benefit from including both sexes. Also, our sample comprised women between the ages of 19 and 46, with a mean age of 30 years (SD = 4.7). It would be advantageous to include samples with a broader age range.

This study has provided information on the psychometric properties of the EAS Temperament Survey for adults. In general, our results confirm the structure and stability of the EAS. They also indicate that there is scope for improvements. Additional research is warranted to guide the further development of the EAS model.

ACKNOWLEDGMENTS

This research was supported by grants from the Norwegian Council for Mental Health and by the Norwegian Research Council. We thank Kristin Schjelderup Mathiesen for generously giving us access to the data and Jennifer Harris for offering advice about the language. We are also grateful to the mothers who participated in the study and the public health nurses who assisted in extensive work related to data collection.

REFERENCES

Angleitner, A., & Ostendorf, F. (1994). Temperament and the big five factors of personality. In C. F. Halverson, G. A. Kohnstam, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 69–90). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

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

Braithwaite, V., Duncan-Jones, P., Bosly-Craft, R., & Goodchild, M. (1984). A psychometric investiagtion of the usefulness of the EASI–III temperament survey in the Australian general population. *Australian Journal of Psychology, 36*, 85–95.

Browne, M. W., & Cudeck, R. (1993). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 10*, 83–93.

Buss, A. (1989). Temperament 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 Associates, Inc.

Buss, A. H., & Plomin, R. (1986). The EAS approach to temperament. In R. Plomin, & J. Dunn (Eds.), *The study of temperament: Changes, continuities and challenges* (pp. 67–79). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440.

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., et al. (1987). Roundtable: What is temperament? Four approaches. *Child Development, 58*, 505–529.

Hoyle, R. H. (1995). *Structural equation modeling. Concepts, issues, and applications*. Thousand Oaks, CA: Sage.

John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 261–271). New York: Springer.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 user's reference guide*. Chicago: SSI.

Loehlin, J. C. (1998). *Latent variable models. An introduction to factor, path, and structural analysis*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mathiesen, K. S., & Sanson, A. (2000). Dimensions of early childhood behavior problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology, 28*, 15–31.

Mathiesen, K. S., & Tambs, K. (1999). The EAS Temperament Questionnaire—Factor structure, age trends, reliability, and stability in a Norwegian sample. *Journal of Child Psychology and Psychiatry and Allied Sciences, 40*, 431–439.

Mathiesen, K. S., Tambs, K., & Dalgard, O. S. (1999). The influence of social class, strain and social support on symptoms of anxiety and depression in mothers of toddlers. *Social Psychiatry and Psychiatric Epidemiology, 34*, 61–72.

McCrae, R. R., & Costa, P. T. (1985). Openness to experience. In R. Hogan & W. H. Jones (Eds.), *Perspectives in personality* (Vol. 1, pp. 145–172). Greenwich, CT: JAI.

Nærde, A., Tambs, K., & Mathiesen, K. S. (2001). Child related strain and maternal mental health: A longitudinal study. *Acta Psychiatrica Scandinavica, 105*, 301–309.

Nærde, A., Tambs, K., Mathiesen, K. S., Dalgard, O. S., & Samuelsen, S. O. (2000). Symptoms of anxiety and depression among mothers of preshool children: Effects of chronic strain related to children and child care-taking. *Journal of Affective Disorders, 58*, 181–199.

Plomin, R., Pedersen, N. L., McClearn, G. E., Nesselroade, J. R., & Bergeman, C. S. (1988). EAS temperament during the last half of the life span: Twins reared apart and twins reared together. *Psychology and Aging, 3*, 43–50.

Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment, 41*, 150–156.

Ruch, W., Angleitner, A., & Strelau, J. (1991). The Strelau temperament inventory–Revised (STI–R): Validity studies. *European Journal of Personality, 5*, 287–308.

Rutter, M. (1989). Temperament: Conceptual issues and clinical implications. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 463–482). Chichester, England: Wiley.

Strelau, J. (1991). Individual differences in temperament: An international perspective. In J. G. Adair, D. Belanger, & K. L. Dion (Eds.), *Advances in psychological science, Vol. 1: Social, personal and cultural aspects* (pp. 33–50). Hove, England: Psychology Press.

Windle, M (1989a). Predicitng temperament–mental health relationships: A covariance structure latent variable analysis. *Journal of Research in Personality, 23,* 118–144.

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

Ane Nærde
Division of Epidemiology
Norwegian Institute of Public Health
P.O. Box 4404 Nydalen
N0–0403 Oslo, Norway
E-mail: ane.nerde@fhi.no

Received September 25, 2002
Revised May 28, 2003

Copyright of Journal of Personality Assessment is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.