Int. J. Prenatal and Perinatal Psychology and Medicine Vol. 6 (1994) No. 2, 221–231

Neonatal Temperament Questionnaire (NTQ): Psychometric Characteristic

J. Kožený and E. Šulcová

Laboratory of Psychometric Studies, Psychiatric Centre Prague, Prague, Czech Republic

Abstract

The structural dimensionality of the newly developed 16 item Neonatal Temperament Questionnaire was tested using exploratory, congeneric and confirmatory factor analyses based on evaluation by mothers of their two-week-old (N = 122) and four-week-old newborns (N = 121).

The full LISREL 7 model generally supported the claimed two-dimensional structure (self-regulation and reactivity) of the TPQ instrument, with acceptable adjusted goodness of fit index and root mean square residual estimates.

Zusammenfassung

Die strukturellen Dimensionen des neu entwickelten Temperament-Fragebogens im Neugeborenenalter wurden mittels explorativer, kongenerischer und konfirmatorischer Faktorenanalyse getestet. Grundlage dafür waren die Bewertungen von Müttern bei zwei Wochen alten (N = 122) und vier Wochen alten (N = 121) Neugeborenen.

Das volle LISREL 7 Modell unterstützt die behauptete zweidimensionale Struktur des TPQ-Instruments im Hinblick auf Selbst-Regulation und Reaktivität mit ausreichender Konstruktvalidität und innerer Konsistenz.

Correspondence to: Jiří Kožený, Laboratory of Psychometric Studies, Psychiatric Centre Prague, 181 03 Prague, Czech Republic

Introduction

Research interest into the problems of temperament at an early age began thirty years ago with the New York Longitudinal Study (NYLS – Thomas, Chess, Birth, Hertzig and Korn, 1963). The authors created the NYLS Parent Interview, which was comprised of nine dimensions with "difficult" temperament being the important derived quality. Temperament was defined as an individual's difference in behaviour style, with style referring to the "how" as opposite to the "what" or content behaviour (Thomas et al. 1968).

At the same time Haar at al. (1964) developed a specific rating instrument for newborns, the Nurses Scale for Rating Neonates (NSRN), with 16 items and a 4-point rating scale. Four dimensions: were distinguished: activity level, tenseness, cuddliness, and responsiveness at a biological level of behaviour.

A different approach is represented by the instrument EASY-I of Buss and Plomin (1975). They searched for the stable genetically encoded dimensions of temperament and identified these dimensions as emotionality, activity, sociability, and impulsiveness as a derived trait.

The psychobiological holistic theory of Rothbart and Derryberry (1981) approaches the problem from still others theoretical and instrumental points of view. Their 87 item parent-rated scale, the Infant Behaviour Questionnaire, discriminated six scales each with a 7-point rating scale: activity level, smiling and laughing, distress limitations, fear, soothability, and duration of orienting. They found two dimensions of temperament: reactivity based on constitutional differences and self-regulation. They defined reactivity as an individual reaction to changes in the environment, based on somatic, endocrine and autonomic nervous systems, and self-regulation as a process modulating reactivity, attentiveness and behavioural functions. They also called attention to the fact that temperament represents an "energetic" aspect of personality rather than a "conceptual" one and that maturation and social processes might influence the difference between children. This psychobiological concept was also supported by findings of Goldsmith and Campos findings (1982).

Temperament in a newborn is often measured by the complex method of the Neonatal Behaviour Assessment Scale (NBAS). Some items, especially those including style and tempo, are considered by Brazelton (1973) as characteristic of a newborn's temperament. Crockenberg and Smith (1982) tried to combine various items of NBAS into an irritability cluster.

Based on the analytical and critical studies of research on temperament at an early age (Hubert, Wachs et al. 1982, Bornstein, Gaugham and Homel 1986, Malatesta and Haviland 1986, Bates 1987), it follows that the problem of temperament measurement is far from being solved from both instrumental and conceptual viewpoints. Traditionally, temperament was usually defined in terms of component behaviour organised into dimensions such as duration, frequency and intensity (Campos et al., 1983; Derryberry and Rothbart, 1984). Most of the definitions and theories seem to have finally converged and agree that temperament includes individual behavioural differences in affective expressiveness, motor activity and stimulus sensitivity.

Our approach is based on the supposition, that temperament in a newborn is a natural phenomenon and manifests itself on the psychobiological level in both positive and negative forms of motor, emotional, perceptual and cognitive behaviours. These behaviours differ in their course, frequency and intensity of reactions and can be taken as innate precursors of adult temperament. The warning of Carey and McDewit (1978a, 1978b) against confusing normal temperament variation with behaviour reflecting physiological dysfunctions was taken into consideration. Therefore, the most frequently used Neonatal Perception Inventory (Broussard and Hartner, 1970) which is concerned with problems on a biological level and refrains from taking perception and cognition into account, differs considerably from our approach.

Material and Methods

Subjects and Procedure

The temperament of 122 physiological newborns (67 boys and 55 girls; Apgar score 1 min.: M = 9.1, SD = 0.87, range = 6–10; Apgar score 5 min: M = 9.8, SD = 0.48, range = 8-10; Apgar score 10 min.: M = 9.9, SD = 0.2, range = 8–10, delivery duration: 1st stage: $<$ 180 min. = 30, $>$ 360 min. = 20, 2nd stage $<$ 5 min. = 49, $>$ 20 min. = 7), birth weight: M = 3399.18 g, SD = 461.1 g, range = 2180–4500, rooming-in: full = 83, partial = 27) was rated by their mothers (age: M = 26.6, SD = 4.5, range = 18–41; education: elementary = 28, high school = 54, college = 40; primiparae = 64, primigravidae = 38, artificial abortion = 36; without obstetric medication = 40) at the age of 14 days and again at the age of 30 days at home using the NTQ device. During a short briefing the mothers were instructed to evaluate the modal behaviour of their children. One mother failed to provide the second assessment.

Instrument

Table 1 contains a listing of the 17 abbreviated NTQ items (Šulcová, 1994), which are rated by mothers on a 5-point scale. Each point is verbally defined and illustrated with examples. The highest score denotes behaviour expected of a normal physiological newborn.

The chosen items refer to behaviours of newborns that are manifest, highly loaded on temperament, occurring in everyday domestic care situations, and behaviour that mothers are familiar with and able to report upon.

Results and Discussion

Descriptive Statistics

The data from both measurements were negatively skewed (items skewness range $-0{,}85 \div 0{,}05$) with most items reporting little problems for their mothers in terms of the newborns behaviour symptoms and associated moods. This was expected, given that the sample comprised physiological newborns. The preliminary factor analysis indicated that the item 17 (physical activity) had very low

Kaiser-Meyer-Olkin measure of sampling adequacy (0.20) and did not belong to the psychometric family, therefore the item was not included in the following analysis. We suppose that we failed to define the item unambiguously enough.

The 16 × 16 matrices of product-moment correlations for the NTQ item responses from the first and the second measurement, indicated that in general, the magnitude of most coefficients was low to moderate, only about 52% – 56% being statistically insignificant and the range was –0,003 ÷ 0,52. The attenuation was almost certainly due to restriction in variance associated with the skewness of the data. Probably some of the NTQ items might even exhibit some degree of collinearity among normally distributed data.

Table 1. Neonatal Temperament Questionnaire; scale items.

1. Dominant mood
2. Intensity of positive emotional reactivity
3. Intensity of negative emotional reactivity
4. Adjustability to novel situations
5. Rapidity of reactions
6. Degree of alertness
7. Interest in physical surroundings
8. Attention to known persons
9. Behaviour endurance
10. Consolability with intervention
11. Behaviour to unknown persons
12. Sensitivity to unexpected, unknown, disturbing stimuli
13. Differential reactivity to inner stimuli
14. Mood after waking up
15. Irritability
16. Relaxation, well-being
17. Physical activity

Exploratory Factor Analysis

Using the SPSS/PC (SPSS, Inc., 1993), the intercorrelation matrices based on the NTQ items description of the two-week-old and the four-week-old newborns were subjected to an exploratory factor analysis (EFA) using an iterative maximum likelihood procedure with two factors extracted on the basis of the scree test (Cattell, 1966). The matrices were rotated to an oblique (direct oblimin) solution.

The Kaiser-Meyer-Olkin measure of sampling adequacy was .75 for both measurements suggesting that the item intercorrelation matrices, which were according to Bartlett test not identity matrices, might have a rather borderline quality for factor analysis.

The factor pattern based on the two-week-old newborns' data appears in the left side of Table 2. The two factor pattern approximating a simple structure accounts for 35.2%, of the item responses variance.

Table 2. The NTQ: oblique factor pattern solution (direct oblimin).

	two-week-old newborns (N = 122)					four-week-old newborns (N = 121)				
ITQ Items	I.	II.	h^2	M	SD	I.	II.	h^2	M	SD
1.	03	–64	42	4.08	68	02	70	49	4.06	66
2.	48	–03	24	3.79	91	63	–06	37	4.17	83
3.	56	58	40	3.55	88	39	–25	14	3.97	89
4.	19	–31	18	4.04	71	03	49	25	4.13	75
5.	68	07	43	3.59	77	64	06	43	3.75	72
6.	58	–01	34	4.22	89	60	14	45	4.33	83
7.	67	–15	55	3.40	79	69	15	58	3.58	77
8.	54	–12	35	4.02	87	65	10	48	4.21	80
9.	60	29	31	4.07	75	59	–27	30	4.19	77
10.	–06	–62	33	3.77	66	–06	69	45	3.79	59
11.	15	–54	37	3.76	89	37	39	40	4.01	88
12.	–12	–41	14	3.58	90	–32	61	32	3.65	91
13.	52	–21	40	3.56	93	48	27	40	3.86	86
14.	12	–44	24	3.54	90	14	30	14	3.63	78
15.	29	–47	41	4.34	70	–05	50	23	4.40	69
16.	–09	–72	48	4.10	69	02	71	52	4.01	70
eigenvalue:										
	3.85	1.7				84.22	1.74			
hyperplane count (±0.10):										
	2	3				5	2			

(Decimal points omitted.)

Factor I accounted for 24.1% of the variance associated with the rotated ML components. This factor loaded highly on the items representing rapidity of reactions (item 5), interest in physical surroundings (item 7), behaviour endurance (item 9), degree of alertness (item 6), attention to known persons (item 8), differential reactivity to inner stimuli (item 13) and intensity of positive emotional reactivity (item 2). This factor might be considered as expressing both reactivity to extraneous stimuli and mental activity.

Factor II involving 11.1% of variance loaded on relaxation and well-being (item 16), dominant mood (item 1), consolability with intervention (item 10), behaviour toward unknown persons (item 11), irritability (item 15), mood after waking up (item 14), sensitivity to unexpected, unknown, disturbing stimuli (item 12) and adjustability to novel situations (item 4). This factor corresponds to reactivity to inner stimuli, which shall be called self-regulation.

Item 3 (intensity of negative emotional reactivity) was equally loaded on both the first and the second factor. The factors being correlated (r = 0.38.4) had 14.7% of the variance in common.

The factor pattern based on the four-week-old newborns' data presented in the right side of Table 2 was very similar to that based on the two-week-old newborns' data with factors loading highly on the same items. The two factors were sufficient to account for 37.3% of the item responses variance.

Factor I (accounting for 26.4% of the variance) loaded in addition on item 3. Factor II (10.9% of variance) was not significantly loaded on item 3, and item 11 saturated both factors equally. As the interfactor correlation was .38 the factors shared about 14.0% of the variance.

The ±.10 hyperplane, after square root communalities adjustment, contained 5 ($p = 0.05$) and 7 ($p = 0.01$) zero-loadings in two-factor solutions based respectively on the two-week-old and four-week-old newborns' data, suggesting statistical significance of the obtained simple structures (Cattell, 1978).

The percentage of residuals with absolute value higher than 0.05 was 72% and 60% for the two-week-old newborns and the four-week-old newborns, respectively, indicating that a considerable amount of variance remained unexplained by the factor patterns.

Taking into account the rather small sample and the basic training of mothers in judging newborn expressive behaviour, the factor patterns invariance over the measurements and the determined hyperplanes are encouraging.

Congeneric Factor Analyses

In contrast to exploratory method of analysis where the results are generally taken as data-driven and findings frequently based on statistical artefact, the use of the LISREL (Jöreskog and Sörbon, 1989) congeneric and confirmatory methods enables testing of hypothesized models and the results are considered as conceptually-driven.

To avoid arbitrary solutions, which would serve only to conflate theory, the dimensions derived from EFA were taken as hypothesis and tested using congeneric factor analysis for the items in each of the separate NTQ subscales.

Since the responses to the NTQ items were measured on a five Likert-type ordinal scale, polychoric correlation coefficient matrices, which generally have less bias associated with them than do product-moment correlation coefficients, were used as an input for the LISREL, maximum-likelihood estimation procedure to calculate χ^2, the goodness of fit index (GFI), the adjusted goodness of fit index (AGFI), and root mean square residuals (RMR). As the χ^2 statistic and the GFI are directly influenced by sample size, the AGFI, which is adjusted for degree of freedom, and the RMR – an index of the degree to which the initial correlation matrix is not reproduced by the estimated factor model – were taken as the criteria of the adequate model fit. It is common consensus (Cuttance, 1987) that a model with an AGFI less than 0.8 is inadequate and most acceptable models would appear to have an AGFI index greater than 0.9.

The results of the congeneric factor analyses for each of the two NTQ subscales based on two-week-old and four-week-old newborns data are shown in Table 3.

The maximum likelihood GFI and AGFI estimates for subscale reactivity based on the two-week-old newborns' data were quite high but the estimates based on the four-week-old newborns' data were not entirely satisfactory indicating a rather imperfect fit of the congeneric model.

Table 3. Congeneric factor models for the NTQ subscales (ML standardized LISREL estimates λC).

a) two-week-old newborns				b) four-week-old newborns			
Reactivity scale							
NTQ items	Parameter value	Standard error	R^2	NTQ items	Parameter values	Standard error	R^2
5.	0.49	0.10	0.24	7.	0.80	0.08	0.64
7.	0.58	0.10	0.33	8.	0.62	0.09	0.38
9.	0.27	0.11	0.07	5.	0.60	0.09	0.36
6.	0.59	0.10	0.35	2.	0.47	0.09	0.22
8.	0.59	0.10	0.35	6.	0.71	0.09	0.50
14	0.55	0.10	0.30	9.	0.25	0.10	0.06
2.	0.35	0.10	0.12	13.	0.46	0.09	0.21
				3.	0.10	0.10	0.01
chi-square(df14) = 8.97, p = 0.83				chi-square(df20) = 65.7, p = 0.01			
coefficient of determination 0.72				coefficient of determination 0.83			
GFI = 0.979				GFI = 0.880			
AGFI = 0.958				AGFI = 0.784			
RMS = 0.04				RMS = 0.094			
Self-regulation scale							
16.	0.65	0.10	0.42	16.	0.69	0.09	0.48
1.	0.65	0.09	0.43	1.	0.65	0.09	0.43
10.	0.47	0.09	0.22	10.	0.59	0.10	0.35
3.	–0.17	0.10	0.03	12.	0.33	0.10	0.11
11.	0.43	0.10	0.18	15.	0.40	0.10	0.16
15.	0.52	0.10	0.27	4.	0.43	0.10	0.19
14.	0.47	0.10	0.22	11.	0.44	0.10	0.19
12.	0.18	0.10	0.03	14.	0.34	0.10	0.11
4.	0.32	0.10	0.10				
chi-square(df27) = 36.53, p = 0.104				chi-square(df20) = 23.3, p = 0.103			
coefficient of determination 0.74				coefficient of determination 0.76			
GFI = 0.932				GFI = 0.948			
AGFI = 0.887				AGFI = 0.907			
RMS = 0.069				RMS = 0.061			

GFI = goodness of fit index
AGFI = adjusted goodness of fit index
RMS = root mean square residual

The self-regulation scale maximum likelihood estimates based on both the two-week-old newborns' data and the four week-old newborns' data supported the model satisfactorily.

The total coefficients of determination were rather low indicating that the communality in all subscales had been accounted for moderately. Perusal of the standardized regression equations suggested that items 2, 3, 4, 9, and 12 (for

the two-week-old newborns) and items 3, 9, 12, and 14 (for the four-week-old newborns) contributed least toward the measurement of the subscales.

In general, the results of the congeneric factor analyses for the NTQ subscales support the construct validity of two separate dimensions among two-week-old and four-week-old newborns, although the reactivity dimension (four-week-old newborns' data) was less clearly defined in terms of fit to the congeneric model. From associated standardized regression equations it follows that a number of the NTQ items were not contributing adequately to the measurement of common variance. To test the overall NTQ model more efficiently these items were removed from the subsequent confirmatory factor analysis (CFA).

Confirmatory Factor Analyses

Using PRELIS (Jöreskog and Sörbom, 1993) the polychoric correlation matrices of the NTQ items 1, 5, 6,7, 8, 10, 11, 13, 14, 15, 16 and 1, 2, 4, 5, 6, 7, 8,10, 11, 13, 15, 16, based respectively on two-week-old newborns and four-week-old newborns, were utilised for the CFA. The results (Table 4) were supportive of the two-dimensional structure of the NTQ instrument among both samples as the GFI and AGFI were above the generally agreed upon low boundary for good fit. The χ^2 was statistically significant only for the four-week-old newborns' data. As the χ^2 test is sample size dependent and very sensitive to departures from multivariate normality of observed variables it has in most cases little meaning. It has been suggested (Jöreskog and Sörbom, 1989) that it should not be used for hypothesis testing but rather as a goodness-of-fit measure. The ration df/χ^2 should serve as a standard by which to judge whether χ^2 is large or small. In our case the highest ratio was 1.5, which is considered as indicating an acceptable fit (Carmines and McIver, 1981). The t-values (parameter estimates divided by their standard errors) were all significant and the maximum modification index provided by the LISREL was 0.23 and 0.41 respective for two-week-old and four-week-old newborns' data, which is regarded as quite acceptably low. The factors being set free correlated 0.51 (for two-week-old newborns) and 0.60 (for four-week-old newborns).

Internal Consistency

Reliability analysis yielded normalized Cronbach alpha coefficients 0.72 (two-week-old newborns) and 0.78 (four-week-old newborns) for the reactivity scale, and 0.71 (two-week-old and four week-old newborns) for the self-regulation scale. The values are within an acceptable range, especially in comparison to other findings in the field.

Standardisation

The description of the NTQ subscale is presented in Table 5. Dichotomising the NTQ scales at their means we created a two-dimensional space, which resembles to the ancient Galen-Kant-Wundt scheme of four temperaments and also to the well known Eysenkian space defined by extroversion and emotional balance. The placement of a newborn in a particular quadrant may indicate his/her

Table 4. Confirmatory factor analyses of the NTQ items (ML standardized LISREL estimates λC)

4a) two-week-old newborns			b) four-week-old newborns		
NTQ items	factor I	factor II	NTQ items	factor I	factor II
5.	0.48		7.	0.79	
7.	0.60		8.	0.63	
6.	0.54		5.	0.59	
8.	0.58		2.	0.46	
13.	0.57		6.	0.72	
16.		0.61	13.	0.49	
1.		0.69	16.		0.67
10.		0.46	1.		0.62
11.		0.39	10.		0.56
15.		0.55	15.		0.40
14.		0.51	4.		0.47
			11.		0.50
chi-square(df43) = 42.01, p = 0.514			chi-square(df53) = 79.4, p = 0.01		
coefficient of determination 0.905			coefficient of determination 0.938		
GFI = 0.938			GFI = 0.900		
AGFI = 0.905			AGFI = 0.853		
RMS = 0.059			RMS = 0.075		

Factor I – Reactivity
Factor II – Self-Regulation

modal temperament characteristics. Because the scales are not entirely orthogonal, these quadrants must be interpreted with caution. Table 6 shows absolute frequencies of the newborns falling into four quadrants, which correspond to the following temperamental types:

1) High self-regulation and reactivity – sanguine
2) Low self-regulation and reactivity – melancholic
3) High self-regulation and low reactivity – phlegmatic
4) Low self-regulation and high reactivity – choleric

Allocation of newborns into the NTQ scales space brings some support to our hypothesis (Šulcová, 1994), that there might be four types of temperamental precursors which are identifiable early after birth. Of course, further longitudinal research is called for to verify this interpretation.

Conclusion

Overall, despite having certain limitations psychometrically, it is evident from our findings that the NTQ has acceptable construct validity and internal consistency. The differences in structure between two-week-old and four-week-old newborns' data might account for developmental changes in newborns in the process of maturation, subtle emotional and reactivity differentiation, and ad-

Table 5. The NTQ scales central tendency and variability characteristic.

	M	SD	Skew	range actual	range theoretical	N
two-week-old newborns						
Reactivity	18.7	3.0	–0.46	10–25	5–25	122
Self-regulation	23.5	2.9	–0.35	15–24	5–30	
four-week-old newborns						
Reactivity	23.9	3.2	–0.55	15–29	6–30	121
Self-regulation	24.4	2.7	–0.64	13–30	6–30	

Table 6. Distribution of newborns in the NTQ scales space.

	two-week-old Sel-Regulation			four-week-old Self-Regulation		
Reactivity	Low	Hight	Total	Low	Hight	Total
Low	33	19	52	36	14	50
Hight	26	44	70	26	45	71
Total	59	63	122	62	59	121

Low – score lower than scale mean
Hight – score higher than scale mean

justment to physical and social environment. Of course, there might be confounding effects present, e.g. mothers' increasing experience and empathy in handling their newborns or the small sample size.

The structural dimensionality of the NTQ experimental version appears reasonably well supported, especially when compared to instruments currently available and further research might be devoted to refining the item content.

Acknowledgement. This research was supported by research grant IGA 0290-3 from the Czech Department of Health to Jiří Koženy.

References

Bates, J.E. (1980). The concept of difficult temperament. Merrill-Palmer Quarterly, **26** (4), 299–319

Bates, J.E. (1987). Temperament in Infancy. In: Osofsky, J. D. (ed.) *Handbook of Infant Development.* Wiley Intersciens, New York, pp. 1101–1149

Bornstein, M.H., Gaughran, J.M., Homel, P. (1986). Infant temperament: theory, tradition, critique, and new assessment. In: Izard, C.E., Read, P.B. (eds.) *Measuring emotions in infants and children.* Vol. II., Cambridge Univ. Press, pp. 172–199

Brazelton, T.B. (1973). *Neonatal Behavior Assessment Scale.* Philadelphia, Lippincott

Broussard, E.R., Hartner, M. (1970). Maternal perception of the neonate as related to development. *Child Psychiatry and Human Development* **1**, 16–25

Buss, A.H., Plomin, R. (1975). *A Temperament Theory of Personality Development.* Wiley, New York

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

Campos, J.J., Barrett, K.C., Lamb, K.E., Goldsmith, H.H., Stenberg, C. (1983): Socioemotional development. In: P.H. Mussen (Ed.) *Handbook of child psychology.* Vol. 2, 4th ed., New York: Wiley, pp. 783–915

Carey, W.B., McDevitt, S.C. (1978a). Revision of the Infant Temperament Questionnaire. *Pediatrics* **61**, 735–739

Carey, W.B., McDevitt, S.C. (1978b). Stability and change in individual temperamental diagnoses from infancy to early childhood. *Journal of the American Academy of Child Psychiatry* **17**, 331–337

Carmines, E.G., McIver, J.P. (1981). Analyzing models with unobserved variables: analysis of covariance structures. In: G. W. Bohrnstedt, E. F. Borgatta (Eds.) *Social measurement: Current issues.* Newbury Park, CA: Sage, pp. 65–115

Cattel, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research* **1**, 140–161

Cattel, R.B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* New York: Plenum Press

Crockenberg, S.B., Smith, P. (1982). Antecedents of man Academy of Child Psychiatry 17, 331-33ther-infant interaction and infant irritability in the first three months of life. Infant Behavior & Development 5, 105-119

Cuttance, P. (1987). Issues and problems in the application of structural equation models. In: Cuttance, P., and Ecob, R. (Eds.) *Stuctural modelling by example: Applications in educational, sociological, and behavioral research.* New York: Cambridge

Derryberry, D., Rothbart, M.K. (1984): Emotion, attention, and temperament. In: C.E. Izard, J. Kagen, R.B. Zajons (Eds.) *Emotions, cognition, and behavior.* Cambridge Univ. Press, pp. 132–166

Goldsmith, H.H., Campos, J.J. (1982). Toward a theory of infant temperament. In: Emde, R.N., Harmon, R. (eds.) *The development of attachment and affiliative systems.* New York: Plenum, pp. 161–193

Haar, E., Welkowitz, J., Blau, A., Cohen, J. (1964). Personality differentiation of neonates. *American Academy of Child Psychiatry* **3**, 330–342

Hubert, N.C., Wachs, T.D., Peters-Martin, P., Gandour, M.J. (1982). The Study of Early Temperament: Measurement and Conceptual Issues. *Child Development* **53**, 571–600

Jöreskog, K.G., Sörbom, D. (1989). LISREL 7: User's reference guide. In: *Scientific Software Inc.*, Mooresville

Jöreskog, K.G., Sörbom, D. (1993). PRELIS: A Preprocessor for LISREL. In: *Scientific Software Inc.*, Chicago

Malatesta, C.Z., Haviland, J.M. (1986). Measuring change in infant emotional expressivity: two approaches applied in longitudinal investigation. In: Izard, C.E., Read, P.B. (eds.) *Measuring emotions in infants and children.* Vol. II., Cambridge University Press

Rothbart, M.K., Derryberry, D. (1981). Development of individual differences in temperament. In: Lamb, M.E., Brown, A.L. (eds.) *Advances in developmental psychology 1*, Lawrence Erlbaum, Hillsdale, pp. 37-86

SPSS LISREL 7 and PRELIS. SPSS Inc., Software, Chicago, 1993

Šulcová, E. (1994): Vývoj dotazníku pro hodnocení temperamentu novorozenců, (Development of newborn temperamental questionnaire), *Čs. psychol.* (in press)

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

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