Infant behaviour questionnaire – revised version: a psychometric study in a Portuguese sample

Raquel Costa & Bárbara Figueiredo

To cite this article: Raquel Costa & Bárbara Figueiredo (2018) Infant behaviour questionnaire – revised version: a psychometric study in a Portuguese sample, Journal of Reproductive and Infant Psychology, 36:2, 207-218, DOI: [10.1080/02646838.2018.1436752](https://doi.org/10.1080/02646838.2018.1436752)

To link to this article: <https://doi.org/10.1080/02646838.2018.1436752>

Published online: 22 Feb 2018.

Submit your article to this journal [↗](#)

Article views: 61

View related articles [↗](#)

View Crossmark data [↗](#)

Infant behaviour questionnaire – revised version: a psychometric study in a Portuguese sample

Raquel Costa^{a,b} and Bárbara Figueiredo^c

^aUniversidade Europeia, Lisboa, Portugal; ^bEPIUnit - Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal; ^cUniversity of Minho, Escola de Psicologia, Braga, Portugal

ABSTRACT

Background: Although the original characteristics of temperament tend to remain constant over the course of development, environmental circumstances may influence infants' reactions and behaviour. Parents' reports of infant temperament are rich informants of infant behaviours in different contexts.

Objective: This study aimed to examine the psychometric properties of the Portuguese version of the Infant Behaviour Questionnaire – Revised (IBQ-R) and test the adequacy of the original and other previously published structures to the Portuguese data.

Method: 330 mothers and 81 fathers of children aged between 3 and 12 months completed the Portuguese version of the IBQ-R.

Results: The confirmatory factorial analysis revealed a non-adequate model fit of the IBQ-R original structure to the Portuguese data; nonetheless, it did reveal an adequate model fit of a previous published IBQ-R structure. This structure, although only slightly different from the original one, seems to be more suitable for the Portuguese data.

Conclusion: This study provides data that indicates that the IBQ-R is a reliable questionnaire to evaluate infant temperament in the Portuguese culture.

ARTICLE HISTORY

Received 22 August 2016

Accepted 11 November 2017

KEYWORDS

Temperament; gender; infant behaviour questionnaire

Temperament has been defined as primarily biologically based individual differences in reactivity and self-regulation (Rothbart, 1981). Reactivity refers to characteristics of an individual's reactions to stimulus change that include behavioural, endocrine and nervous system responses, while self-regulation refers to processes that modulate reactivity (Rothbart, Derryberry, & Posner, 1994; Rothbart & Posner, 1985). Temperament can be observed as individual differences in patterns of emotionality, activity and attention (Goldsmith et al., 1987). After childbirth, an interactional process begins between temperament traits, other psychological dimensions and the environment, so that as a result of this process, changes or modifications may occur in one or more temperamental traits (Thomas & Chess, 1977). So, temperament is not absolutely stable over time, it mediates and shapes the influence of the environment on the individual's psychological structure. It also affects and is affected by the social context in a bidirectional influence (Chess & Thomas, 1986).

A variety of methods were developed to measure temperament. Parental ratings and observation in laboratory or naturalistic environments have been used to measure infant temperament over time. Both methods have advantages and disadvantages. Observation-based methods reduce bias by relying on standardised procedures and using trained observers who are not influenced by a relationship with the child, but these methods are confined to the children's behaviour in a novel situation and in response to elicited tasks in a limited period of time (Bates & Bayles, 1984). This can enhance in the children atypical behaviours that they do not usually express in response to everyday situations in familiar contexts. Moreover, it only catches infant behaviour at a given moment, and consequently it is confined to the infant state at that given moment (Strelau, 1983). On the other hand, parents can observe the infant in naturalistic environment across a variety of daily situations (Rothbart & Goldsmith, 1985). A debate on the validity of parental reports have emerged mainly because although some studies found significant association between parental temperament ratings and observational ratings (either naturalistic or laboratorial), others did not (Stifter, Willoughby, Towe-Goodman, & The Family Life Project Key Investigators, 2008). So a critique to this method is the doubtful accuracy of parents as reporters of their children's temperament. The lack of accuracy may be due to social desirability or other circumstances such as parental emotional characteristics, including anxiety and depression or parental expectancies about their child (Goldsmith & Hewitt, 2003; Mebert, 1991; Seifer, Sameroff, Barrett, & Krafchuk, 1994; Vaughn, Taraldson, Crichton, & Egeland, 1981). Several studies have emerged to understand parental characteristics associated with divergence between parental and observational methods. Early childhood experiences and depression (Forman et al., 2003; Leerkes & Crockenberg, 2003) are highly associated with low concordance between methods. Moreover, parents' personality, psychopathology, parental expectancies and stress are related to temperament ratings, indicating a possible alteration in the capacity of identifying and interpreting infant behaviours (Goldsmith & Hewitt, 2003; Mebert, 1991; Sameroff, Seifer, & Elias, 1982; Seifer et al., 1994; Vaughn et al., 1981).

Although parent report measures do contain some subjective parental components, some studies have demonstrated that they also contain a substantial objective component that enables the accurate assessment of children's individual characteristics (Wachs & Bates, 2001). Probably the problem of lack of agreement between parents and observers relies in that we are comparing the incomparable. Observers' and parents' measures are radically different so that observers are usually confined to an artificial situation in a limited period of time, whereas parents have access to the infant behaviour in an array of situations within familiar and unfamiliar environments to the child. Eventually, laboratory observations and parental reports serve different purposes. Perhaps parental reports of infant temperament may be more relevant to the understanding of child developmental outcomes, because parental perceptions of infant behaviour shape their own behaviour and interaction with the infant (Crockenberg & McCluskey, 1986). Considering the importance of parent–infant interaction for infant development, the way parents (mis)interpret infants' cues may be a relevant issue for understanding child development outcomes.

In order to study infant temperament dimensions, Rothbart (1981) developed a questionnaire to be completed by parents: the Infant Behaviour Questionnaire (IBQ), later revised by Gartstein and Rothbart (2003; Infant Behaviour Questionnaire – Revised; IBQ-R), to obtain a more differentiated measure. It assesses the following 14 dimensions of temperament: (1) activity level (15 items): movement of arms and legs, squirming and locomotor activity;

(2) distress to limitations (16 items): baby's fussing, crying or showing distress while (a) in a confining place or position, (b) involved in caretaking activities, (c) unable to perform a desired action; (3) approach (12 items): rapid approach, excitement and positive anticipation of pleasing activities; (4) fear (16 items): the baby's startle or distress to sudden changes in stimulation, novel physical objects or social stimuli, inhibited approach to novelty; (5) duration of orienting (12 items): the baby's attention to and/or interaction with a single object for extended periods of time; (6) smiling and laughter (10 items): the child's smile or laughter in general caretaking and playing situations; (7) vocal reactivity (12 items): amount of vocalisation exhibited by the baby in daily activities; (8) sadness (14 items): general low mood; lowered mood and activity specifically related to personal suffering, physical state, object loss, or inability to perform a desired action; (9) perceptual sensitivity (12 items): amount of detection of slight, low-intensity stimuli from the external environment; (10) high-intensity pleasure (11 items): amount of pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty and incongruity; (11) low-intensity pleasure (13 items): amount of pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty and incongruity; (12) cuddliness (17 items): the baby's expression of enjoyment and shaping of the body to express desire of being held by a caregiver; (13) soothability (18 items): baby's reduction of fussing, crying, or distress when the caretaker uses soothing techniques; (14) falling reactivity/rate of recovery from distress (13 items): rate of recovery from peak distress, excitement, or general arousal, ease of falling asleep. Three dimensions of temperament can be computed: (1) surgency/extraversion – computed from the mean scores of approach, vocal reactivity, high intensity pleasure, smiling and laughter, activity level and perceptual sensitivity; (2) negative affectivity – computed from the mean scores of distress to limitation, fear and falling reactivity; (3) orienting/regulation – computed from the mean scores of low-intensity pleasure, cuddliness, duration of orienting and soothability.

Temperament is influenced by the life contexts and cross-cultural comparisons are only possible using the same instrument; thus, the importance of the validation of this instrument for different cultures is unquestionable. Several studies on the adaptation of the IBQ-R to different cultures/countries were conducted through the study of the internal consistency (González Salinas, Hidalgo Montesinos, Carranza Carnicero, & Ato García, 2000), exploratory factor analysis (Ortega Samohano, Huitrón Vázquez, Osorio Guzmán, & Bazán Riverón, 2011), test–retest and inter-observer reliability (Lamb, Frodi, Hwang, & Frodi, 1983) and confirmatory factor analysis (Dragan, Kmita, & Fronczyk, 2011; Enlow, White, Hails, Cabrera, & Wright, 2016). The aim of this study is to examine the psychometric properties of the IBQ-R and the adequacy of previously published models of the IBQ-R structure to the Portuguese data, using confirmatory factorial analysis.

Method

Sample

The sample is composed of 330 mothers and 81 fathers of infants aged between 3 and 12 months (213 mothers and 39 fathers of a 3 to 6 months-old child; 52 mothers and 13 fathers of a 6 months+1 day to 9 months-old infant; and 65 mothers and 29 fathers of a 9 months+1 day to 12 months-old infant). The exclusion criteria are the inability to read or

write Portuguese and fetal congenital problems. The research objectives and procedures were explained to eligible parents. Information about the study objectives, methods and procedures was provided to parents. Participants were informed of their rights, namely to give up at any moment, confidentiality and use of data for research purposes only. An informed consent was provided, and after it was signed, data from the first assessment moment was collected. Table 1 provides information on sociodemographic and medical data.

Procedures

This research was conducted in a primary care centre and in several paediatricians' offices from two cities in Portugal, Porto and Braga. A total of 678 parents were approach to participate in the study; 48 refused to participate, 630 questionnaires were provided and parents were asked to complete them while waiting for the clinical consultation. Of the questionnaires, 526 (83.5%) were returned and 458 (72.7%) were fully completed. Of these, 47 concerned premature infants and were excluded from the statistical analysis. The Infant Behavior Questionnaire – Revised (IBQ-R, Gartstein & Rothbart, 2003) is a revised version of the IBQ (Rothbart, 1981) and was designed to measure temperament in infants aged between 3 and 12 months. The Portuguese version was obtained after translation and back-translation and was used in this study. The translation from English to Portuguese was conducted by a Portuguese native speaker psychologist and the back-translation was conducted by an English native speaker psychologist blind to the original version. The final translated version was compared to the original version and incongruences were analysed and discussed to reach a final version.

The IBQ-R is composed of 191 items scored on a seven-point Likert scale and assesses the following 14 dimensions of temperament. (1) Activity level: movement of arms and legs, squirming and locomotor activity. (2) Distress to limitations: baby's fussing, crying or showing distress. (3) Approach: rapid approach, excitement and positive anticipation of pleasurable activities. (4) Fear: the baby's startle or distress to sudden changes in stimulation, novel physical objects or social stimuli; inhibited approach to novelty. (5) Duration of orienting: the baby's attention to and/or interaction with a single object for extended periods of time. (6) Smiling and laughter: smiling or laughter from the child in general caretaking and play situations. (7) Vocal reactivity: amount of vocalisation exhibited by the baby in daily activities. (8) Sadness: general low mood; lowered mood and activity specifically related to personal suffering, physical state, object loss, or inability to perform a desired action. (9) Perceptual sensitivity: amount of detection of slight, low-intensity stimuli from the external environment. (10) High-intensity pleasure: amount of pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty and incongruity. (11) Low-intensity pleasure: amount of pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty, and incongruity. (12) Cuddliness: the baby's expression of enjoyment and moulding of the body to being held by a caregiver. (13) Soothability: baby's reduction of fussing, crying, or distress when the caretaker uses soothing techniques. (14) Falling reactivity/rate of recovery from distress: rate of recovery from peak distress, excitement, or general arousal; ease of falling asleep. These dimensions were originally grouped in three factors using exploratory factor analysis (Gartstein & Rothbart, 2003): (1) Surgency/Extraversion, consisting of the scales Approach, Vocal reactivity, High-intensity pleasure, Smiling and laughter, Activity level

Table 1. Sociodemographic and medical data.

Parental and gestational data		Mothers	Fathers	Neonatal data		
		(%)	(%)			(%)
Age	≥ 20 Λ ≤ 34 > 34	65.733.3	64.635.4	Type of delivery	VaginalCesarean/instrumental	38.561.5
Years of education	< 9 ≥ 9	13.586.5	14.285.8	Type of anaesthesia	NoneEpiduralGeneral	3.893.23.0
Marital status	MarriedCohabitingSingleDivorced	87.310.41.50.8	84.715.3	Gender	FemaleMale	57.942.1

and Perceptual sensitivity; (2) Negative Affectivity, consisting of the scales Sadness, Distress to limitations, Fear, and Falling reactivity/rate of recovery from distress; and (3) Orienting/Regulation, consisting of the scales Low-intensity pleasure, Cuddliness, Duration of orienting and Soothability (Gartstein & Rothbart, 2003).

Data analysis

Confirmatory factor analysis (CFA) was conducted in order to test the goodness-of-fit of previously published adequate models of the IBQ-R structure in different cultures (Dragan et al., 2011; Enlow et al., 2016; Gartstein & Rothbart, 2003) to the Portuguese data. Not all scales normality was verified, maximum likelihood estimation with robust standard errors was utilised when conducting the factor analyses. Several measures of fit were considered as recommended by Sharma, Mukherjee, Kumar, and Dillon (2005): Chi-square, the Tucker–Lewis Index (TLI), the Comparative Fit Index (CFI) and the root mean square error of approximation (RMSEA). For the TLI and CFI, values of 1 will be interpreted as indicating perfect fit, and the cut-off value of 0.90 will be considered as an adequate fit, as recommended by Bentler and Bonett (1980). For the RMSEA, values of 0.05 or less will be interpreted as indicating a close fit, values of 0.08 or less as indicating a reasonable fit and values over 0.1 as indicating unacceptable fit (Browne & Cudeck, 1993). This analysis was conducted using structural equation modelling (SEM) software: AMOS 17 (Arbuckle, 2008). The internal consistency of the factors in the model with an adequate fit was assessed through item-total correlation, average inter-item correlation and Cronbach's alpha coefficients. Nunnally and Bernstein (1994) suggest that good internal consistency is indicated when Cronbach's alpha is over .70 and mean inter-item correlations exceed .15. Field (2005) also argues that item-total correlation should be over .30.

Results

Confirmatory factorial analysis

The CFA conducted to study the goodness-of-fit of the original IBQ-R structure (Gartstein & Rothbart, 2003) to the Portuguese data indicates a non-adequate fit: $\chi^2(74) = 288.48$, $p < .001$; TLI = .74; CFI = .79; RMSEA = .136. The CFA conducted to study the goodness-of-fit of the original IBQ-R structure without cuddliness to the Portuguese data did not improve the model fit: $\chi^2(62) = 251.22$, $p < .001$; TLI = .74; CFI = .80; RMSEA = .139. The CFA conducted to study the goodness-of-fit of the IBQ-R structure for a Californian sample to the Portuguese data indicates non-adequate fit: $\chi^2(58) = 223.06$, $p < .001$; TLI = .76; CFI = .82; RMSEA = .135. The CFA statistics conducted to study the goodness-of-fit of the IBQ-R structure for a Russian sample to the Portuguese data indicate an improved model fit: $\chi^2(43) = 93.89$, $p < .001$; TLI = .90; CFI = .93; RMSEA = .087. The CFA statistics conducted to study the goodness-of-fit of the IBQ-R structure for a combined US sample to the Portuguese data indicate a reasonably adequate fit: $\chi^2(42) = 81.70$, $p < .001$; TLI = .92; CFI = .95; RMSEA = .078. The CFA statistics conducted to study the goodness-of-fit of the IBQ-R structure for the polish sample (Dragan et al., 2011, p. 546) to the Portuguese data indicate a non-adequate fit: $\chi^2(67) = 190.13$, $p < .001$; TLI = .84; CFI = .88; RMSEA = .108. Lastly, the CFA statistics

conducted to study the goodness-of-fit of the IBQ-R structure for a US sample (Enlow et al., 2016) to the Portuguese data indicate a non-adequate fit: $\chi^2(37) = 125.35$, $p < .001$; TLI = .83; CFI = .88; RMSEA = .123.

Psychometric properties

The IBQ-R structure in this adequate model for the Portuguese sample is the model also found to be suitable for a combined US sample. It excludes two scales – cuddliness and high-intensity pleasure – and is composed of (1) Surgency/Extraversion including Vocal reactivity, Perceptual sensitivity, Approach, Activity level, Smiling and laughter, with secondary loadings on Fear and Falling reactivity; (2) Negative Affectivity including Fear, Distress to limitation, Falling reactivity and Sadness with secondary loadings on Duration of orienting, Perceptual sensitivity, Approach, Activity level and Smiling and laughter; and (3) Orienting Regulation including Duration of orienting, Soothability and Low-intensity pleasure. Attending to the adequate fit of this model, the alpha of Cronbach for all IBQ-R dimensions was computed. Average item correlation, item-total correlation and alpha if item deleted are presented in Table 2.

Internal consistency of the IBQ-R scales

The alpha of Cronbach for all IBQ-R scales was computed and values over .70 were observed for all scales: Activity level ($\alpha = .78$), Distress to limitation ($\alpha = .73$), Fear ($\alpha = .93$), Duration of orienting ($\alpha = .88$), Smiling and laughter ($\alpha = .80$), Low-intensity pleasure ($\alpha = .87$), Soothability ($\alpha = .93$), Falling reactivity ($\alpha = .71$), Perceptual sensitivity ($\alpha = .87$), Sadness ($\alpha = .85$), Approach ($\alpha = .85$) and Vocal reactivity ($\alpha = .72$). Monomethod discriminant validity was examined through the correlation between scales. Nonetheless, Distress to limitation scale was not significantly correlated to three scales (Duration orienting, Smiling and laughter, and Low-intensity pleasure), while Activity level scale is not related to two scales (Soothability and Falling reactivity), the Falling reactivity was not correlated with two scales

Table 2. Internal consistency: item-total correlation and alpha if item deleted.

	ITC	α if itemdeleted
<i>Surgency/Extraversion ($\alpha = .80$)</i>		
Vocal reactivity	0.69	0.73
Perceptual sensitivity	0.58	0.77
Approach	0.64	0.74
Activity level	0.46	0.80
Smiling/laughter	0.58	0.76
<i>Negative Affectivity ($\alpha = .60$)</i>		
Fear	0.42	0.29
Distress to limitation	0.26	0.56
Falling reactivity	0.09	0.71
Sadness	0.48	0.31
<i>Orienting/Regulation ($\alpha = .68$)</i>		
Duration of orienting	0.40	0.51
Soothability	0.13	0.76
Low pleasure	0.44	0.38

(Activity level and Fear) and the Smiling and laughter scale was not correlated to one scale (Sadness) (see Table 3).

Discussion

The CFA revealed that the original structure of the IBQ-R is not adequate for the Portuguese data. Other previously published structures also showed a non-adequate fit to the Portuguese data (Dragan et al., 2011; Enlow et al., 2016). Nevertheless, the fit of the IBQ-R structure for a combined US sample, was adequate for the Portuguese sample. Overall, the structure of the IBQ-R in this model is quite similar to the original structure (Gartstein & Rothbart, 2003), strengthening the concept developed by the authors of the questionnaire. Comparing against the original IBQ-R structure, in this IBQ-R model: (1) the High-intensity pleasure scale in the Surgency/Extraversion dimension was excluded, with the other five scales remaining the same – Vocal reactivity, Perceptual sensitivity, Approach, Activity level, Smiling and laughter – additionally, there are secondary loadings on Fear and Falling reactivity; (2) the Negative Affectivity dimension remains with the four scales – Fear, Distress to limitation, Falling reactivity, sadness – with secondary loadings on Duration of orienting, Perceptual sensitivity, Approach, Activity level and Smiling and laughter; and (3) the Cuddliness in the Orienting Regulation dimension was excluded, with the other three scales remaining the same – Duration of orienting, Soothability and Low-intensity pleasure. Additionally, a covariance between Falling reactivity and both Approach and Soothability indicates higher association of these temperament domains in the Portuguese sample. In general, the dimensions in this structure include the exact same scales from the IBQ-R original structure with the exclusion of High-intensity pleasure and Cuddliness, which means that, in the Portuguese sample, these scales seem to be less reliable in the assessment of temperament. High-intensity pleasure refers to the ‘pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty, and incongruity’ and cuddliness refers to the ‘expression of enjoyment and moulding of the body to being held by a caregiver’ (Gartstein & Rothbart, 2003, p. 72) and there have been reports of differences between cultures in the way some of these temperament features are manifested in different cultures (Gartstein, Bogale, & Meehan, 2016). Very recently, Enlow et al. (2016) also found cuddliness to have low factor loadings and excluded this scale for a better fit model. On the other hand, the secondary loadings of Fear and Falling reactivity on the Surgency/Extraversion dimension as well as of Duration of orienting, Perceptual sensitivity, Approach, Activity level and Smiling and laughter on the Negative Affectivity dimension indicate a pronounced association between this temperament dimension in the Portuguese sample. These differences in Portuguese children’s temperament might be the result of a different approach to early care in Portugal. In fact, several studies have reported cultural differences in temperament (Gartstein, Slobodskaya, & Kinsht, 2003), which supports the association of parenting practices and the manifestation of temperament features.

In our study, the scales’ internal consistency ranged from .71 to .93; in other studies there is also some variability: in Finland from .74 to .93 (Gaias et al., 2012), in Italy from .73 to .92 (Montiroso, Cozzi, Putnam, Gartstein, & Borgatti, 2011), in Japan from .61 to .92 (Nakagawa & Sukigara, 2005), in Russia from .75 to .92 (Gartstein et al., 2003), in the UK from .75 to .91 (de Lauzon-Guillain et al., 2012) and in the USA from .77 to .89 (Gartstein & Rothbart, 2003). The scales with lower internal consistency, with Cronbach’s alphas under .80, were Activity level, Falling reactivity, Vocal reactivity and Distress to limitation, similar to what was found

Table 3. Descriptive statistics for the IBQ scales and Pearson correlation between IBQ scales.

	1	2	3	4	5	6	7	8	9	10	11	12	Mean(SD)	Min-Max.
1 Act	1												4.26(1.21)	1.80–7.67
2 DL	0.31**	1											4.20(6.77)	1.44–6.77
3 Fear	0.44**	0.29**	1										3.78(1.75)	1.00–8.00
4 DO	0.21	0.19*	0.38**	1									5.02(1.58)	1.75–8.00
5 SL	0.31**	0.10	0.33**	0.60**	1								5.41(1.19)	2.10–8.00
6 LP	0.15	0.10	0.29**	0.65**	0.58**	1							5.39(1.26)	2.23–8.00
7 Sooth	–0.08	0.19*	0.19*	0.25**	0.31**	0.38**	1						5.32(1.07)	2.00–8.00
8 Fall	–0.08	–0.17	0.18*	0.33**	0.41**	0.40**	0.50**	1					4.97(0.95)	2.46–6.88
9 PS	0.41**	0.30**	0.48**	0.49**	0.37**	0.40**	0.16	0.21*	1				5.13(1.61)	1.58–8.00
10 Sad	0.33**	0.50**	0.64**	0.31**	0.19*	0.33**	0.23**	–0.01	0.46**	1			4.36(1.36)	1.07–8.00
11 App	0.36**	0.29**	0.45**	0.47**	0.50**	0.53**	0.18*	0.27**	0.50**	0.38**	1		5.52(1.25)	1.58–8.00
12 VR	0.20	0.22**	0.43**	0.54**	0.70**	0.54**	0.35**	0.38**	0.43**	0.31**	0.63**	1	5.26(1.06)	2.18–8.00

*p < .05;
**p < .01;
Note: Act, Activity level; DL, Distress limitation; DO, Duration of orienting; SL, Smiling and laughter; LP, Low-intensity pleasure; Sooth, Soothability; Fall, Falling reactivity; PS, Perceptual sensitivity; Sad, Sadness; App, Approach; VR, Vocal reactivity.

in other studies (e.g. Dragan et al., 2011), while those with higher internal consistency with Cronbach's alphas above .90 were Soothability and Fear. In the original study of the IBQ-R questionnaire, Cronbach's alphas under .80 were also found for Falling reactivity in children aged 6–9 months; in Vocal reactivity in children aged 6–12 months, but not for Distress to limitation; Fear was also one of the scales with a higher Cronbach alpha in the original study (Gartstein & Rothbart, 2003) and in other studies, namely some with samples from different cultures (Dragan et al., 2011). In this study the Cronbach's alpha of the Surgency/Extraversion is over .70, but the Negative Affectivity is under the recommended .70 and the Orienting/Regulation dimensions is slightly under the recommended .70 (Field, 2005). Nonetheless, Kline (1999) claims that when dealing with psychological constructs, values below .70 can be expected. In other studies, Cronbach's alphas lower than .70 were also reported for the Orienting/Regulation dimension (de Lauzon-Guillain et al., 2012).

This study has several limitations including the sample size and lack of other infant developmental measures; a large sample size would have increased the power of the results and other infant developmental measures would allow examining concurrent and predictive validity. Nonetheless, it provides data that indicate that the IBQ-R is a reliable documentation of infant temperament in the Portuguese culture, as it was found to be in other cultures (Dragan et al., 2011; González Salinas et al., 2000; Lamb et al., 1983; Ortega Samohano et al., 2011). The identification of specific aspects of infant temperament is essential not only for research but also for clinical practice. For researchers, the discrimination of specific aspects of temperament is important because it enables examination of the association with infant development outcomes. For clinicians, it is important for the design of specific early intervention.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Fundação para a Ciência e a Tecnologia [grant number SFRH/BD/18249/2004 – Infant Developmental Trajectory].

ORCID

Raquel Costa <http://orcid.org/0000-0003-1614-540X>

Bárbara Figueiredo <http://orcid.org/0000-0002-8209-7445>

References

- Arbuckle, J. L. (2008). *AMOS 17 user's guide*. Chicago, IL: SPSS.
- Bates, J. E., & Bayles, K. (1984). Objective and subjective components of mothers' perceptions of their children from age 6 months to 3 years. *Merrill-Palmer Quarterly*, 30, 111–130.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606.
- Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park (CA): Sage Publications.
- Chess, S., & Thomas, A. (1986). *Temperament in clinical practice*. New York, NY: Guilford Press.

- Crockenberg, S. B., & McCluskey, K. (1986). Change in maternal behavior during the baby's first year of life. *Child Development*, 57, 746–753.
- Dragan, W. Ł., Kmita, G., & Fronczyk, K. (2011). Psychometric properties of the polish adaptation of the infant behavior questionnaire – revised (IBQ-R). *International Journal of Behavioral Development*, 35(6), 542–549. doi:10.1177/0165025411422181
- de Lauzon-Guillain, B., Wijndaele, K., Clark, M., Acerini, C. L., Hughes, I. A., Dunger, D. B., ... Ong, K. (2012). Breastfeeding and infant temperament at age three months. *PLoS ONE*, 7(1), e29326. doi:10.1371/journal.pone.0029326
- Enlow, M., White, M., Hails, K., Cabrera, I., & Wright, R. (2016). The infant behavior questionnaire-revised: Factor structure in a culturally and sociodemographically diverse sample in the United States. *Infant Behavior and Development*, 43, 24–35. doi:10.1016/j.infbeh.2016.04.001
- Field, A. (2005). *Discovering statistics using SPSS*. London: SAGE Publications.
- Forman, D. R., O'Hara, M. W., Larsen, K., Coy, K. C., Gorman, L. L., & Stuart, S. (2003). Infant emotionality: Observational methods and the validity of maternal reports. *Infancy*, 4(4), 541–565.
- Gaias, L. M., Räikkönen, K., Komsí, N., Gartstein, M. A., Fisher, P. A., & Putnam, S. P. (2012). Cross-cultural temperamental differences in infants, children, and adults in the United States of America and Finland. *Scandinavian Journal of Psychology*, 53(2), 119–128. doi:10.1111/j.1467-9450.2012.00937.x
- Gartstein, M., Bogale, W., & Meehan, C. (2016). Adaptation of the infant behavior questionnaire-revised for use in Ethiopia: Expanding cross-cultural investigation of temperament development. *Infant Behavior and Development*, 45, 51–63. doi:10.1016/j.infbeh.2016.09.004
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior & Development*, 26(1), 64–86.
- Gartstein, M. A., Slobodskaya, H. R., & Kinsht, I. A. (2003). Cross-cultural differences in the first year of life: United States of America (U. S.) and Russia. *International Journal of Behavioral Development*, 27(4), 316–328.
- Goldsmith, H. H., & Hewitt, E. C. (2003). Validity of parental report of temperament: Distinctions and needed research. *Infant Behavior & Development*, 26(1), 108–111.
- Goldsmith, H. H., Buss, A., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., & ... Hinde, R. B. (1987). Roundtable: What is temperament? four approaches. *Child Development*, 58, 505–529.
- González Salinas, C., Hidalgo Montesinos, M., Carranza Carnicero, J. A., & Ato García, M. (2000). Elaboración de una Adaptación a Población Española del Cuestionario Infant Behavior Questionnaire para la Medida del Temperamento en la Infancia. *Psicothema*, 12(4), 513–519.
- Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London, UK: Routledge.
- Lamb, M. E., Frodi, A. M., Hwang, C., & Frodi, M. E. (1983). Interobserver and test–retest reliability of Rothbart's Infant Behavior Questionnaire. *Scandinavian Journal of Psychology*, 24(1), 153–156. doi:10.1111/j.1467-9450.1983.tb00487.x
- Leerkes, E. M., & Crockenberg, S. C. (2003). The impact of maternal characteristics and sensitivity on the concordance between maternal reports and laboratory observations of infant negative emotionality. *Infancy*, 4(4), 517–539.
- Mebert, C. J. (1991). Dimensions of subjectivity in parents' ratings of infant temperament. *Child Development*, 62(2), 352–361.
- Montirosso, R., Cozzi, P., Putnam, S., Gartstein, M., & Borgatti, R. (2011). Studying crosscultural differences in temperament in the first year of life: United States and Italy. *International Journal of Behavioral Development*, 35(1), 27–37.
- Nakagawa, A., & Sukigara, M. (2005). How are cultural differences in the interpretation of infant behavior reflected in the Japanese revised infant behavior questionnaire? *Japanese Journal of Educational Psychology*, 53(4), 491–503.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New-York: McGraw-Hill.
- Ortega Samohano, E., Huitrón Vázquez, B., Osorio Guzmán, M., & Bazán Riverón, G. (2011). Validación en una muestra mexicana del Infant Behavior Questionnaire-R (IBQ-R): Estudio preliminar. *Apuntes De Psicología*, 29(1), 133–144.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, 52(2), 569–578.
- Rothbart, M. K., & Goldsmith, H. H. (1985). Three approaches to the study of infant temperament. *Developmental Review*, 5(3), 237–260.

- Rothbart, M. K., & Posner, M. I. (1985). Temperament and the development of self regulation. In L. C. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 93–123). New York, NY: Plenum.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences and the interface of biology and behavior* (pp. 83–116). Washington, DC: American Psychological Association.
- Sameroff, A. J., Seifer, R., & Elias, P. K. (1982). Sociocultural variability in infant temperament ratings. *Child Development*, 53, 564–578.
- Seifer, R., Sameroff, A. J., Barrett, L. C., & Krafchuk, E. (1994). Infant temperament measured by multiple observations and mother report. *Child Development*, 65(5), 1478–1490.
- Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, 58(7), 935–943. doi:10.1016/j.jbusres.2003.10.007
- Stifter, C. A., Willoughby, M. T., Towe-Goodman, N. R., & The Family Life Project Key Investigators (2008). Agree or agree to disagree? Assessing the convergence between parents and observers on infant temperament. *Infant and Child Development*, 17, 407–426.
- Strelau, J. (1983). *Temperament, personality, activity*. New York, NY: Academic Press.
- Thomas, A. & Chess, S. (1977). *Temperament and development*. New York, NY: Brunner/Mazel.
- Vaughn, B. E., Taraldson, B. J., Crichton, L., & Egeland, B. (1981). The assessment of infant temperament: A critique of the Carey infant temperament questionnaire. *Infant Behavior & Development*, 4, 1–17.
- Wachs, T. D., & Bates, J. E. (2001). Temperament. In G. Brenner & A. Fogel (Eds.), *Blackwell handbook of infant development: Handbooks of developmental psychology* (pp. 465–501). Malden MA: Blackwell.