Early Human Development 89 (2013) 561–576

Contents lists available at SciVerse ScienceDirect

Early Human Development

journal homepage: www.elsevier.com/locate/earlhumdev

Psychometric properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public child daycare centers

Alberto Filgueiras [a,b], Pedro Pires [b,c], Silvia Maissonette [b], J. Landeira-Fernandez [a,b,d,*]

[a] Pontifícia Universidade Católica do Rio de Janeiro, Brazil
[b] Instituto Brasileiro de Neuropsicologia e Comportamento, Rio de Janeiro, Brazil
[c] Universidade Federal do Rio de Janeiro, Brazil
[d] Universidade Estácio de Sá, Brazil

ARTICLE INFO

Article history:
Received 7 March 2012
Received in revised form 3 November 2012
Accepted 12 February 2013

Keywords:
Child development assessment
Ages and Stages Questionnaire
Brazilian public child daycare centers
Classical Test Theory
Item Response Theory

ABSTRACT

Well-designed screening assessment instruments that can evaluate child development in public daycare centers represent an important resource to help improve the quality of these programs, as an early detection method for early developmental delay. The Ages and Stages Questionnaire, 3rd edition (ASQ-3), comprises a series of 21 questionnaires designed to screen developmental performance in the domains of communication, gross motor skills, fine motor skills, problem solving, and personal–social ability in children aged 2 to 66 months. The purpose of the present work was to translate and adapt all of the ASQ-3 questionnaires for use in Brazilian public child daycare centers and to explore their psychometric characteristics with both Classical Test Theory and Rating Scale analyses from the Rasch model family. A total of 18 Ages & Stages Questionnaires — Brazilian translation (ASQ-BR) questionnaires administered at intervals from 6 to 60 months of age were analyzed based on primary caregiver evaluations of 45,640 children distributed in 468 public daycare centers in the city of Rio de Janeiro. The results indicated that most of the ASQ-BR questionnaires had adequate internal consistency. Exploratory factor analyses yielded a one-factor solution for each domain of all of the ASQ-BR questionnaires. The only exception was the personal–social domain in some of the questionnaires. Item Response Theory based on Rating Scale analysis (infit and outfit mean squares statistics) indicated that only 44 of 540 items showed misfit problems. In summary, the ASQ-BR questionnaires are psychometrically sound developmental screening instruments that can be easily administered by primary caregivers.

1. Introduction

Developmental neuroscience research indicates that early life experience can have a major impact on cognitive, emotional, behavioral, and social development later in life [1,2]. Although most of the brain structures in early childhood are already present, they are still extremely immature and thus subjected to a wide range of environmental interactions. These early experiences represent the main underpinning of brain development that determines the strength and function of several neural circuits [3]. Accordingly, social deficits, generally associated with poverty and environmental degradation, might lead to a disruption of normal brain development in children [4].

These findings have important implications for political and economic decisions with regard to public investment in early childhood development programs [5,6]. The efficiency of these programs for low-income families, such as government-funded child daycare centers, is a dynamic process that depends, among other factors, on continuous evaluation. Assessing the development of children enrolled in public daycare centers represents one aspect of this evaluation process and might contribute to program enhancement and guide policy decisions [7]. The developmental assessment of child daycare centers might also help identify children who might need early intervention. For example, Gleason et al. [8] reported that approximately 10% of children between 1 month and 5 years of age had some kind of serious psychopathology. Moreover, less than 10% of these children were properly identified before they reach school age [9].

Most of the early developmental instruments, such as the Bayley Scales [10], require specialized training and are time-consuming and expensive, which might impose difficulty for a routine examination program in a daycare center. Conversely, some much simpler instruments are less expensive and time-consuming and generally designed to screen for developmental delays. Developmental screening consists of a brief process of evaluating large numbers of children to identify those who might be at high risk for developmental delay and for that reason need further evaluation [11].

The Ages and Stages Questionnaire (ASQ) is a screening instrument used for developmental assessment during the first 5 years of life [12]. The third edition of the ASQ (ASQ-3) comprises a series of 21

* Corresponding author at: Núcleo de Neuropsicologia Clínica e Experimental, Laboratório de Análise de Dados, Departamento de Psicologia, Pontifícia Universidade Católica do Rio de Janeiro, Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ 22453-900, Brazil. Tel.: +55 21 3527 2075; fax: +55 21 3527 1187.
E-mail address: landeira@puc-rio.br (J. Landeira-Fernandez).

0378-3782/$ – see front matter
http://dx.doi.org/10.1016/j.earlhumdev.2013.02.005

questionnaires for infants (2, 4, 6, and 8 months of age), toddlers (9, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, and 33 months of age), and preschoolers (36, 42, 48, 54, and 60 months of age) designed to screen developmental performance in the domains of communication, gross motor skills, fine motor skills, problem solving, and personal–social skills. Each domain has six items, and each item is scored as "yes" (10 points), "sometimes" (5 points), or "not yet" (0 points) [13].

Several reports indicate that the ASQ has well-established psychometric properties in a clinical context, such as test–retest reliability, internal consistency, criterion validity, sensitivity, and specificity [14–20]. Moreover, the ASQ has been employed as an instrument to assess the impact of Early Head-Start programs and public childcare systems, such as The Florida Infant Mental Health Pilot Program [21], and the five programs developed by the Early Promotion & Intervention Research Consortium [22]. Indeed, the ASQ is the screening instrument that has the most published research in the academic setting [23] and appears to be a reliable instrument to measure infant development in childcare centers [7,24].

The ASQ has been cross-culturally validated in other languages, such as Portuguese from Portugal [25], Spanish [26], French [27], Dutch [17], Norwegian [28,29], Danish [30], Chinese [31], Korean [32], and Hindi [33]. However, the ASQ has not yet been translated into Brazilian Portuguese. Therefore, the main purpose of the present work was to translate all of the ASQ-3 questionnaires into Brazilian Portuguese and explore their psychometric characteristics.

This study was also driven by the need to develop a reliable assessment instrument that might be used to help evaluate Brazilian public child daycare programs and allow the screening of children for possible developmental delays. This is an important issue because almost 18% of Brazilian children between 0 and 5 years old attend public daycare centers [34]. In Rio de Janeiro, 26.5% of all children within this age range specifically attend public childcare centers [34]. For that reason, the present work also sought to adapt the ASQ-3 questionnaires to these public institutions and evaluate whether childcare providers are able to administer these questionnaires.

2. Methods

2.1. Participants

Data from the present study were collected from children enrolled in all of the 468 public daycare centers in the city of Rio de Janeiro. Children in the age range of 4 to 60 months were distributed according to the 20 age intervals defined by the ASQ. The 2 month questionnaire was not used because public child daycare centers only accept children who are older than 4 months. The project was approved by the PUC-Rio Ethical Committee, Rio de Janeiro, Brazil.

2.2. The ASQ-BR

The Brazilian version of the ASQ-3 adapted for public child daycare centers (ASQ-BR) was based on the original ASQ-3 [13]. Initially, each of the 20 questionnaires of the ASQ-3 was translated into Brazilian Portuguese by three independent native Portuguese speakers with professional experience in English–Portuguese translation. Each translated item was then evaluated by a multidisciplinary panel of specialists with a high level of English fluency and different expertise in psychometrics and cross-cultural adaptation instruments, public child daycare systems, child development and education, economics, and public programs for low-income families. The multidisciplinary panel was also allowed to make changes to any of the translated items. The Spanish ASQ-3 [24] was employed to help solve difficulties that emerged during this phase.

Conceptual equivalence, cultural adaptation, and language idiosyncrasies were considered whenever necessary. Although efforts were made to maintain the exact meaning of each item, words generally used in the United States were modified to better fit the Brazilian context. For example, "inch" was replaced with "centimeter," "feet" was replaced with "meter," and "Cheerio" was replaced with "piece of biscuit." Additionally, Brazilian children do not usually learn their surname until they are 6 years old, so the expression "last name" was replaced with "name of the mother/father." Item content could also be adjusted to the context of public daycare centers. For example, "feed himself a cracker or a cookie" became "feed himself a fruit" because children who attend public childcare centers are encouraged to eat more fruits and vegetables than industrialized food.

At the end of this phase, a preliminary ASQ-BR version was then back-translated into English by a native American English speaker with high fluency in Brazilian Portuguese. The back-translated ASQ-BR and original ASQ-3 were examined by three native American speakers and the multidisciplinary panel. Only minimal differences were detected, and changes were made when necessary. Afterward, a pilot test was performed with 120 children (six per questionnaire) from different public child daycare centers. Caregivers were responsible for completing the questionnaire with minimal training so problems of understanding and item comprehension could be detected. Comments and suggestions from caregivers were evaluated by the multidisciplinary panel, and a few of the suggestions were incorporated into the final version of the ASQ-BR questionnaires.

2.3. Procedure

All of the directors of the 468 public child daycare centers were invited to participate in a 1-day, 8-hour meeting previously scheduled by the Education Secretary of the city of Rio de Janeiro. Each meeting had approximately 30 daycare directors, and the 20 ASQ-BR questionnaires were presented by a person previously trained in the ASQ-BR by a member of the multidisciplinary panel. All of the directors were responsible for taking the ASQ-BR questionnaires to their daycare centers so each child could be evaluated by the daycare caregivers. Brazilian public child daycare center typically have several classrooms, each containing approximately 30 children. The daycare director presented the ASQ-BR to the caregivers who were responsible for a classroom with children within the ASQ-BR age range. The caregivers, in turn, were responsible for administering the ASQ-BR to the children in their classrooms. Afterward, the caregivers or directors entered the child ASQ-BR data into a website using a computer located in the daycare center. Data collection occurred between October 12 and December 17, 2010.

2.4. Statistical analyses

Descriptive statistics were employed to characterize the study population and performance of all age intervals in each of the ASQ-BR domains. Cronbach's alpha [35] and item-total correlation coefficients were employed to evaluate the internal consistency of the six items in each domain of the ASQ-BR questionnaires. A Cronbach's alpha equal to or greater than 0.65 [36] and an item-total correlation equal to or greater than 0.3 [37] are considered satisfactory.

Exploratory factor analysis (EFA) was employed to evaluate the unidimensionality of each of the five domains across the different age intervals. Factors were extracted through principal axis factoring because this is the preferable method for factor extraction when employed in an exploratory manner [38]. Factor rotation was performed using the oblique method (Promax, $K = 4$) because of the likelihood of considerable conceptual correlation among the factors. Velicer's minimum average partial [39] and parallel [40] analyses were employed to determine the number of factors. Both of these procedures were performed using SPSS syntax developed by O'Connor [41].

The unidimensionality of each domain was expected to be found in the Confirmatory Factor Analysis (CFA) and EFA. An invariance assessment was conducted using Multiple-Group Factor Analysis (MGFA), which is part of the CFA family [42]. The MGFA was performed for each developmental domain in the ASQ-BR, with age intervals as groups. Three models were proposed to study invariance: configural invariance (to ensure that the groups have the same basic factor structure), metric invariance (in addition to the same pattern of fixed and freed parameters among groups, to analyze whether the loadings are equivalent), and error variance invariance (to test whether the groups present the same level of measurement error and the other two previous aspects of invariance) [42,43]. Chi-squared (χ^2) values, degrees of freedom (*df*), significance for χ^2 statistics (*p* value), the Root Mean Square Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), and Goodness-of-Fit Index (GFI) are reported. The best model should be the one with the most power of explanation with no statistical difference between the model and empirical data among groups ($p \geq 0.05$). The Akaike Information Criterion (AIC) was used to determine the best model for each domain of the ASQ-BR [43]. Finally, the model should present RMSEA and SRMR < 0.05 and GFI > 0.90 [42]. The MGFA was conducted using AMOS 18.0.

Rating Scale (RS) analyses from the Rasch modeling family were used to evaluate the psychometric characteristics of each ASQ-BR item [44,45]. The RS model tested the fit between the empirical data and theoretical model of an underlying trait. This analysis is considered part of the Rasch Measurement Model family [44–46]. When Rasch is used to analyze polytomous data, two methods are possible: the RS from Andrich [45,46] and the Partial Credit Model (PCM) from Masters [46,47]. The PCM is conducted when items present unique RS structures [46]. For example, if one has the same instrument with several dichotomous items, such as gender, and other polytomous items, such as ordinal attitude Likert categories, then the PCM represents the best statistical procedure, whereas the scale presents one factor.

Conversely, RS analysis is performed when items are unidimensional and share the same polytomous rating structure, such as the ASQ-BR [46]. Linacre's studies of these two Rasch family models indicate the importance of correctly using the PCM or RS analysis to understand the structure of the scale using ordinal items [46,48]. Based on this evidence, we used the RS analysis for the ASQ-BR items by considering the classical $m + 1$ rating structure [45]. To enable this analysis, we performed a data transformation by considering the traditional “0, 1, 2” scoring scheme, in which $0 =$ “not yet,” $1 =$ “sometimes,” and $2 =$ “yes.” We did not use the structure of the ASQ-BR (i.e., “0, 5, 10”) for the descriptive statistics and other measures for the ASQ-BR in the present study. This type of data transformation respects the RS and does not affect mean square infit and outfit statistics. The lower threshold is the same as the higher threshold (i.e., 1 for the traditional “0, 1, 2” structure and 5 for the ASQ-BR “0, 5, 10” scoring scheme). Consequently, thresholds remain stable throughout the scale.

The infit and outfit statistics are the most widely used misfit diagnostic RS parameters. Both statistics indicate how well items fit the underlying construct. The infit statistic is more sensitive to unexpected responses to well-targeted items in which performance is close to a respondent's level (inliers). The outfit statistic, in turn, is more sensitive to unexpected responses to items in which performance is far from the respondent's level (outliers). Infit and outfit statistics are expressed as mean square standardized residuals (MNSQ). Item fit values between 0.5 and 1.5 are considered adequate [48]. Infit or outfit values less than 0.5 indicate that the observed item response has little variation and thus overfits the model. In this case, the item is too predictive and thus superfluous for the questionnaire because the information carried by the item is redundant with the other items of the questionnaire. Infit and outfit values greater than 1.5 show that the observed response to an item has excessive variance and thus underfits the RS. This is an indication that the item was answered inconsistently [49]. Rating Scale analyses were performed using Winsteps 3.72.3 [48].

Finally, we used the Test Equating method [48,50] from the Rasch family analyses based on the RS model to link scales and build one Construct Map for each development scale in the ASQ-BR. This procedure allows comparisons of children's performance despite being different scales among age intervals. We followed the linking procedures suggested by Lee and Wright [51]. However, because the ASQ-BR is a polytomous scale, we used the One-step Equating method by respecting the RS model as recommended by Linacre [48].

The first step was to analyze the fit of the items when calibrated to a common scale (i.e., a unique scale formed by all items across the age intervals of the ASQ-BR). Four important types of information were extracted from this common scale: theta or logit (i.e., how many latent constructs the item reflects), logit standard error, and the infit and outfit statistics. As suggested by Linacre [48], we excluded items due to misfit, from 0.5 to 1.5. The following linking-equating procedure was conducted. First, we created five calibration matrices by considering which items were common among the scales to build the analysis table. The calibration matrices are presented in Fig. 1a–e. Second, the common items were used to build a common scale for each developmental domain. Third, from the common scale, we calibrated the entire set of items from the ASQ-BR (i.e., common and non-common items) to reveal the person's and items' theta measures and fit indices in each developmental domain. Fourth, the Construct Map analysis was performed using all of the calibrated items from the ASQ-BR.

The five common scales that emerged from the Equating, one for each developmental domain from the ASQ-BR, were used to produce the Construct Maps of the entire set of items across questionnaires, providing valuable information about the relationships between items, persons, and latent traits [48,51]. These items were the common and non-common items calibrated by the common-scale, thus the full content of the ASQ-BR for each developmental domain.

3. Results

3.1. Descriptive and reliability analysis

A total of 45,640 ASQ-BR questionnaires were included in the analyses. Data from the 4 month questionnaire were excluded because the sample size was too small to enable statistical inferences ($n = 10$). Data from the 9 and 10 month questionnaires were merged because the items on these two questionnaires are exactly the same. Therefore, a total of 18 ASQ-BR questionnaires were analyzed. Table 1 presents the number of questionnaires answered by each age interval and the percent distribution of boys and girls and their respective mean ages and standard deviations. High variation was found in the number of questionnaires answered by each age interval, ranging from 47 (60 months) to 8859 (42 months). Almost 40,000 of the answered ASQ-BR questionnaires (87.5%) were distributed across the 24 to 54 month age intervals.

The percentage of boys in the sample ranged from 25.0% to 43.6% (mean = 36.0%), and the percentage of girls ranged from 46.4% to 75% (mean = 64%). The lower percentage of boys compared with girls across the 18 age intervals was consistent. Finally, the mean ages of the boys and girls were within the estimate range across all of the intervals. This was expected because the ASQ-BR questionnaire was chosen according to the children's ages.

Table 2 depicts the score means, standard deviations, person's average theta (logit) calibrated with the common scale, and theta's standard error of the five domains of the ASQ-BR across the 18 age intervals. The scores and theta are presented for the entire sample and stratified according to the gender sample. Gender differences in mean ASQ-BR scores were compared using Student's *t*-test. Overall, the ASQ-BR scores of the girls tended to be higher than the boys in all of the domains, with the exception of the gross motor domain, in which boys tended to present higher scores than girls.

The theta was calibrated to the common scale for each developmental domain using the linking-equating procedure. The theta for

a

Calibration Matrix

Communication

Age Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
6	1	2	3	4	5	6																																						
8			1	2	4	3	5	6																																				
10					1	2		3	4	5	6																															
12								1	2	3	4	5	6																													
14											1		2	3	4	5	6																									
16															1	2	4	3	5	6																						
18																	2	1	4	3	5	6																				
20																			1	2	3	6	4	5																		
22																						6	1	2	3	4	5															
24																			2		1	5	4	3			6															
27																							2	1			4	3	5	6												
30																							1	2				3	4	5	6											
33																												1	2	3	4	5	6									
36																												1	2	3	4	5	6									
42																														1	2	3	4	5	6							
48																																		5	6	1	2	3	4			
54																																		4	2			1	3	5	6	
60																																		1			5			2	3	4	6

b

Calibration Matrix

Gross Motor

Age Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
6	1	2	3	4	5	6																																		
8			2	1	3	4	5	6																														
10					1		2	3	4	5	6																											
12									1	2	3	4	5	6																								
14												1	2	3	4	5	6																					
16														1	2	3	4	5	6																			
18																1	2	3	4	5	6																	
20																		2	1	3	4	5	6															
22																				3	1	2	4	5	6													
24																				1	2	4	3	5	6													
27																						2	1	3	4	5	6											
30																						1	2	4	3		5	6										
33																						1		3	2		4	5	6									
36																								2	1		3	4	5	6								
42																											1	2	3	4	5	6						
48																															1	2	3	4	5	6		
54																															4		2	1	3	5	6	
60																															2		1			3	4	5	6

c

Calibration Matrix

Fine Motor

Age Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
6	1	2	3	4	5	6																																						
8			1	2	3	4	5	6																																				
10						1	2	3	4	5	6																															
12									1	2	3	4	5	6																												
14												1	2	3	4	5	6																									
16													2	1		5	4	3	6																							
18													1			3	4	2	5	6																						
20																1	2		3	4	5	6																				
22																			4	1	2	3	5	6																		
24																			2	1		3	4	6	5																	
27																						1	2	5	4	3	6															
30																						1		3		2	4	5	6													
33																								2		1	3	4	5	6												
36																								2		1	4	3		5	6											
42																											2	1		3	4	5	6									
48																														2		1		3	4	5	6					
54																																		1	2		3	4	5	6		
60																																						1	2	3	4	5	6

d

Calibration Matrix

Problem Solving

Age Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
6	1	2	3	4	5	6																																							
8			2	1	4	3	5	6																																					
10					1		2	3	4	5	6																																		
12									1	2	3	4	5	6																															
14												1	2	3	4	5	6																												
16														1	2	3	4	5	6																										
18																1	2	4		3	5	6																							
20																		1			2		3	4	5	6																			
22																		1			4	5		6	2		3																		
24																					1	2					3	4	5	6															
27																											1	2	4	5	3	6													
30																													2	3	1	4	5	6											
33																													3	2	1	4	5	6											
36																													2	1		3	4		5	6									
42																																1	2		3	4	5	6							
48																																				1	2	5	3	4	6				
54																																					4	2		1	3	5	6		
60																																					1			2		3	5	4	6

e

Calibration Matrix

Personal-Social

Age Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
6	1	2	3	4	5	6																																				
8			1	2	4	3	5	6																																		
10					1		2	3	4	5	6																															
12									1	2	3	4	5	6																												
14												1	2	3	4	5	6																									
16														3	1	2	5	4	6																							
18														2			3	1	4	5	6																					
20																	2	1		3	4	5	6																			
22																				4	1	5	3	2	6																	
24																				1	2	4	3		5	6																
27																					1	3	2		4	5	6															
30																					1				3		4	2	5	6												
33																									2		3	1	4	5	6											
36																									2		4	1		3	5	6										
42																											2			1	3	4	5	6								
48																																	1	3	2	4	5	6				
54																																	4	1	5	2	3		6			
60																																	1		3				4	2	5	6

Fig. 1. a. Calibration matrix with all items from the communication domain of the ASQ-BR. Note: The items were numbered according to the order of the first questionnaire (6 months). The subsequent sets were then matched with the previous items if they were the same item in both scales. With regard to the order of appearance, the subsequent items were numbered according to this criterion. The numbers in white represent the order of the items from each of the questionnaires by age interval. The same procedure was conducted for the other domains. b. Calibration matrix with all items from the gross motor domain of the ASQ-BR. c. Calibration matrix with all items from the fine motor domain of the ASQ-BR. d. Calibration matrix with all items from the problem solving domain of the ASQ-BR. e. Calibration matrix with all items from the personal–social domain of the ASQ-BR.

each person was then calculated with the calibrated common scale with both common and non-common items. The global average of the theta and respective girls and boys averages were calculated. This enabled an overall analysis by considering all scales from the ASQ-BR because we now had comparable theta averages. We conducted an 18×2 two-way ANOVA using age interval (18) and gender (2) as factors. Each developmental domain was analyzed separately. To identify possible significant differences between factors, Fisher's Least Significant Difference (LSD) post hoc test was used. The p values from the LSD test that compared gender within each age interval are presented in Table 2. In the communication domain, we found a significant age interval × gender interaction ($F_{7,45,605} = 83.33$, $p < 0.01$). Significant main effects of age interval ($F_{1,45,605} = 171.52$, $p < 0.01$) and gender ($F_{1,45,605} = 231.56$, $p < 0.01$) were found. In the gross motor domain, we found a significant age interval × gender interaction ($F_{17,45,605} = 63.65$, $p < 0.01$). Significant main effects of age interval ($F_{1,45,605} = 60.33$, $p < 0.01$) and gender ($F_{1,45,605} = 100.64$, $p < 0.01$) were found. In the fine motor domain, we found a significant age interval × gender interaction ($F_{17,45,605} = 59.39$, $p < 0.01$). Significant main effects of age interval ($F_{1,45,605} = 57.82$, $p < 0.01$) and gender ($F_{1,45,605} = 118.76$, $p < 0.01$) were found. In the problem solving domain, no significant age interval × gender interaction was found ($F_{17,45,605} = 3.05$, $p = 0.12$). Significant main effects of age interval ($F_{1,45,605} = 73.18$, $p < 0.01$) and gender ($F_{1,45,605} = 246.79$, $p < 0.01$) were found. Finally, in the personal–social domain, the average theta presented a significant age interval × gender interaction ($F_{17,45,605} = 83.67$, $p < 0.01$). Significant main effects of age interval ($F_{1,45,605} = 62.53$, $p < 0.01$) and gender ($F_{1,45,605} = 278.90$, $p < 0.01$) were found.

Table 2 also presents the Cronbach's alpha for the five ASQ-BR domains across the 18 age intervals. The communication domain did not present any questionnaires with a Cronbach's alpha less than 0.65. The gross and fine motor skills domains had two of 18 questionnaires (11%) with Cronbach's alpha less than 0.65, whereas the problem-solving domain had three of 18 questionnaires (17%). The personal–

Table 1
Number of ASQ-BR questionnaires (*N*), percentage distributions of boys and girls, and their respective ages (mean and standard deviation [SD]) across the 18 age intervals.

Age interval (months)	*N*	Percentage		Mean age (SD)	
		Boys	Girls	Boys	Girls
6	55	43.6%	46.4%	6.35 (0.94)	6.43 (0.56)
8	104	25.0%	75.0%	8.22 (0.59)	7.97 (0.94)
10	236	36.3%	64.7%	9.96(0.77)	9.72 (0.69)
12	434	43.3%	56.7%	12.06 (0.77)	11.99 (0.93)
14	752	42.4%	57.6%	13.96 (1.14)	13.61 (1.11)
16	987	39.9%	60.1%	15.89 (0.82)	15.85 (1.73)
18	1103	41.8%	58.2%	17.91 (0.73)	17.87 (1.70)
20	1031	41.1%	58.9%	19.77 (0.80)	19.68 (0.79)
22	955	34.9%	65.1%	21.84 (0.81)	21.74 (0.98)
24	1454	39.6%	60.4%	24.18 (0.92)	24.10 (1.56)
27	2222	38.8%	61.2%	26.83 (1.34)	26.83 (1.51)
30	2814	35.5%	64.5%	29.95 (0.97)	29.78 (1.19)
33	3316	34.5%	65.5%	32.87 (0.95)	32.78 (1.21)
36	5291	34.1%	65.9%	36.60 (1.40)	36.49 (1.73)
42	8859	34.9%	65.1%	41.90 (1.78)	41.72 (2.78)
48	8528	35.3%	64.7%	47.74 (1.87)	47.63 (1.99)
54	7452	36.3%	63.7%	53.10 (3.56)	53.07 (2.97)
60	47	25.5%	74.5%	57.39 (0.95)	58.23 (1.13)
Total	45,640	36.0%	64.0%	–	–

social domain had 12 of 18 questionnaires (67%) with a Cronbach's alpha less than 0.65. Importantly, eight of these questionnaires (44%) presented a Cronbach's alpha less than 0.60.

An item-total correlation test was performed to evaluate whether the six items of each domain were consistent across the 18 ASQ-BR questionnaires. Overall, 45 of the 540 items from all of the questionnaires (8.3%) presented item-total correlations less than 0.30. The gross motor skills domain presented only one of 108 items (0.9%) with item-total correlations less than 0.3 (i.e., item 1 from the 42 month questionnaire). The communication domain had three of 108 items (2.8%) with item-total correlations less than 0.3 (i.e., item 2 from the 8 month questionnaire, item 1 from the 16 month questionnaire, and item 6 from the 36 month questionnaire). The fine motor skills domain presented 4 of 108 items (3.7%) with item-total correlations less than 0.3 (i.e., item 6 from the 22 month questionnaire, item 1 from the 24 month questionnaire, item 2 from the 27 month questionnaire, and item 6 from the 30 month questionnaire). The problem-solving domain had 10 of 108 items (9.3%) with item-total correlations less than 0.3 (i.e., item 5 from the 18 month questionnaire, item 4 from the 22 month questionnaire, item 1 from the 24 month questionnaire, item 2 from the 30 month questionnaire, item 3 from the 33 month questionnaire, item 2 from the 54 month questionnaire, items 2 and 3 from the 20 month questionnaire, and items 2 and 3 from the 36 month questionnaire). Finally, the personal–social domain had 27 of 108 items (25%) with item-total correlations less than 0.3 (i.e., item 1 from the 16 month questionnaire, item 5 from the 18 month questionnaire, items 1 and 6 from the 20 month questionnaire, items 1 and 3 from the 24 month questionnaire, items 4 and 6 from the 36 month questionnaire, items 1 and 2 from the 42 month questionnaire, items 1, 2, and 3 from the 10 month questionnaire, items 2, 3, and 4 from the 22 month questionnaire, items 1, 4, and 5 from the 42 month questionnaire, items 1, 2, 3, and 4 from the 27 month questionnaire, and items 1, 2, 4, and 5 from the 54 month questionnaire).

3.2. Exploratory factor analysis

Velicer's minimum average partial and parallel analyses indicated that the six items of each domain of the ASQ-BR 18 age intervals presented a single factor solution. The only exception was the personal–social domain that presented a two-factor solution for the 10, 54, and 60 month questionnaires. The first factor of the 10 month personal–social domain was responsible for 34.5% of the variance with an eigenvalue of 2.07. This factor incorporated items 3, 4, 5, and 6. The second factor explained 17.4% of the variance with an eigenvalue of 1.04. This factor was composed of items 1 and 2. The 54 month personal–social domain also presented a two-factor solution. The first factor explained 31.8.4% of the variance with an eigenvalue of 1.91. Items 1, 2, 5, and 6 were loaded in this factor. The second factor was responsible for 16.9% of the variance with an eigenvalue of 1.02. This factor was composed of items 3 and 4. Finally, the 60 month personal–social domain presented a two-factor solution. The first factor was responsible for 48.2% of the variance with an eigenvalue of 2.89. This factor included items 2, 3, 5, and 6. The second factor explained 19.4% of the variance with an eigenvalue of 1.17. This factor contained items 1 and 4.

3.3. Multiple-Group Confirmatory Factor Analysis (MGCFA)

MGCFA was performed using the six items from each development domain in a unidimensional model. The groups were the 18 age intervals divided by the questionnaires of the ASQ-BR. We considered the following hierarchy for the model, from less to more rigorous: (a) configural invariance, (b) metric invariance, and (c) error invariance variance. However, any of these models were accepted if the results met the proposed criteria. We report data only from the best model, respecting the delimited hierarchy.

The communication domain presented metric invariance as the more rigorous accepted model: $\chi^2 = 538.72$, $p < 0.01$, $df = 108$, $\chi^2/df = 4.99$, RMSEA = 0.04, SRMR = 0.05, and GFI = 0.93. The gross motor domain depicted error invariance variance as the best model: $\chi^2 = 276.03$, $p < 0.01$, $df = 126$, $\chi^2/df = 2.19$, RMSEA = 0.04, SRMR = 0.02, and GFI = 0.95. The fine motor domain showed metric invariance with the best results: $\chi^2 = 455.48$, $p < 0.01$, $df = 130$, $\chi^2/df = 3.50$, RMSEA = 0.06, SRMR = 0.08, and GFI = 0.91. The problem solving domain presented configural invariance as the best model: $\chi^2 = 581.85$, $p < 0.01$, $df = 142$, $\chi^2/df = 4.09$, RMSEA = 0.05, SRMR = 0.05, and GFI = 0.92. Finally, the personal–social domain depicted configural invariance as the best model: $\chi^2 = 1493.15$, $p < 0.01$, $df = 216$, $\chi^2/df = 6.91$, RMSEA = 0.08, SRMR = 0.09, and GFI = 0.80. This domain did not meet any of the established criteria. However, Hair et al. [52] stated that a good model needs not only adequate fit indices but also fine parsimony indices to ensure its acceptance. An appropriate model is not only the better fitted but also the more parsimonious model. These authors recommended referring to the Tucker–Lewis Index (TLI) and Normed Fit Index (NFI) as parsimony measures. The NFI and TLI need to be as close as possible to 1.0 to characterize a more parsimonious model. Regarding this index, all of the ASQ-BR scales presented good parsimony fit. The communication domain had TLI = 0.96 and NFI = 0.98 in metric invariance, which substantiated the fit of the invariance model. The gross motor domain had TLI = 0.94 and NFI = 0.92 for error invariance. The fine motor domain had TLI = 0.97 and NFI = 0.96 in metric invariance. The problem solving domain had TLI = 0.90 and NFI = 0.91 in configural invariance. The personal–social domain had TLI = 0.89 and NFI = 0.88 in the configural invariance model, which is very close to 0.90 (i.e., the value recommended in the literature [52]). Based on the parsimony criteria, all of the scales from the ASQ-BR substantiated their respective invariance models. Regarding the personal–social domain, we considered that we did not have sufficient empirical evidence to reject configural invariance as an adequate model for all of the personal–social domain scales when considering fit and parsimony indices.

3.4. Rating Scale fit statistics

Infit and outfit MNSQ statistics were calculated for each item of the five domains across the 18 age intervals. These two RS parameters are summarized in Table 3. Overall, 44 of 540 items (8.1%) of the ASQ-BR questionnaires showed misfit problems. Infit MNSQ problems were found in only four of the 44 misfit items (9.1%; one overfitting and three underfitting the model). Outfit MNSQ problems were found in 40 of the 44 misfit items (90.9%; eight overfitting and 32 underfitting the model).

This analysis found nine ASQ-BR domains with only one misfit item. Items that presented outfit greater than 1.5 were item 6 from the 8 month communication domain, item 3 from the 10 month personal–social domain, item 6 from the 20 month personal–social domain, item 2 from the 20 month communication domain, item 3 from the 24 month communication domain, item 1 from 24 month from the fine motor skills domain, item 1 from the 24 month problem-solving domain, item 3 from the 24 month personal–social domain, and item 2 from the 30 month problem-solving domain. Finally, item 2 from the 30 month problem-solving domain also presented infit statistics greater than 1.5.

Ten domains across the 18 ASQ-BR questionnaires presented two items with misfit. Items 1 and 2 from the 10 month gross motor skills domain showed outfit greater than 1.5. Items 1 and 5 from the 16 month communication domain had outfit greater than 1.5. Items 2 and 6 from the 16 month gross motor skills domain had outfit greater than 1.5. Items 1 and 5 from the 18 month problem-solving domain had outfit greater than 1.5. Items 2 and 3 from the 20 month problem-solving domain had outfit greater than 1.5. Items 2 and 3 from the 22 month personal–social domain had outfit greater than 1.5. Items 2 and 3 from the 27 month personal–social domain had outfit greater than 1.5. Items 2 and 4 from the 54 month problem-solving domain had outfit greater than 1.5.

Finally, only four domains of the ASQ-BR questionnaire had three or more misfit items. Two of them presented three misfit items. Item 4 from the 14 month gross motor skills domain had infit less than 0.5 and outfit greater than 1.5. Item 3 from the 14 month gross motor skills domain had outfit less than 0.5. Item 2 from the 14 month gross motor skills domain had outfit greater than 1.5. Items 1, 2, and 5 from the 18 month communication domain had outfit greater than 1.5. The 6 month communication domain had four misfit items. Items 1 and 4 had outfit less than 0.5, and items 3 and 6 had outfit greater than 1.5. The 60 month gross motor skills domain had five misfit items. Item 3 had both infit and outfit statistics less than 0.5. Item 6 had both infit and outfit greater than 1.5. Items 1, 2, and 4 had outfit less than 0.5.

3.5. Test Equating and common scales of the ASQ-BR

Logit, standard error, and infit and outfit statistics for the common scales of the ASQ-BR using the entire set of items (i.e., both common and non-common items calibrated to the common scale) are presented in Table 4.

The number of common scale items for each domain is the following: communication, 43 items; gross motor, 39 items; fine motor, 43 items; problem solving, 45 items; and personal–social, 42 items. No items presented misfit problems when calibrated together in a common scale. Fig. 2A–E depicts the five construct maps with the entire set of items for the ASQ-BR scales.

4. Discussion

Currently, no brief and low-cost early developmental screening instruments with psychometric strength have been adapted for Brazilian public daycare centers. Accordingly, the present study represents the first evaluation of the psychometric properties of the Brazilian–Portuguese translation and adaptation of the ASQ-3 for public daycare centers. The results confirmed that the ASQ-BR questionnaires consist of a set of easy and brief instruments that can be completed in a brief period of time. Public childcare providers were able to understand and properly administer each of the 18 ASQ-BR questionnaires with minimal training. These results are consistent with previous reports that indicated that the ASQ can be properly administered by caregivers as a child developmental instrument in daycare centers [7,53]. For example, Allen [7] reported that caregivers from daycare centers located in Florida were also able to successfully administer the ASQ. Indeed, having caregivers rather than parents complete a developmental screening instrument is more convenient for child development monitoring purposes in a daycare center [7]. Childcare staff knows well the habits and abilities of the children and are trained in a wide range of childhood developmental subjects that might help the screening evaluation process. Moreover, caregivers have several opportunities to observe the dynamics of child development on a daily basis when the child is present in the daycare center.

The present study also indicated that most of the 18 ASQ-BR questionnaires presented internal reliability estimates within expected values, including Cronbach's alphas and item-total correlations. These results are consistent with other studies performed with the original English version of the ASQ-3 [13] and other versions of these questionnaires translated into different languages [17,24–33]. The only exception to these general psychometric characteristics was the personal–social domain of some of the ASQ-BR questionnaires. For example, the personal–social domains from eight ASQ-BR questionnaires (10, 22, 24, 30, 42, 48, 54, and 60 months of age) presented Cronbach's alphas less than 0.6. Moreover, the 27 and 54 month personal–social domain presented four of six items with item-total correlations less than 0.3. Other studies also detected difficulty with the ASQ personal–social domain. For example, Tsai et al. [31] reported that the personal–social domain presented the lowest Cronbach's alphas in the Taiwanese version of the 36 month ASQ questionnaire. This study also found that Taiwanese parents and preschool teachers evaluated the same child differently in the personal–social domains, suggesting inconsistencies within this domain.

Exploratory factor analysis examined the dimensionality of each of the ASQ-BR domains. The results indicated that all five of the ASQ-BR domains across the 18 age intervals are unidimensional, with the possible exception of the personal–social domain from the 10, 54, and 60 month questionnaires, which presented a two-factor solution. Because of the low number of items that loaded in each of these factors, envisioning possible labels for these factors was impossible.

To further explore the factorial characteristics of the ASQ-BR, we conducted a MGCFA as recommended by Little and Slegers [42]. The results confirmed the information provided by the EFA. The invariance of the scales was more related to the same basic unidimensional structure. This means that the organization of the item sets based on the factor loadings for each item varies from one questionnaire to another. The only conclusion that the MGCFA allows us to make is that the ASQ-BR scales are unidimensional, but we cannot extend this statement to the items' factor loadings or levels of measurement errors.

Importantly, the impact of the different types of ASQ-BR invariance models tested in the present study should be emphasized [42,43]. The results indicated that all of the ASQ-BR domains presented configural invariance, meaning that the questionnaires measure the same single dimension across different age intervals. Given the adequate configural invariance of all of the ASQ-BR domains, we also tested both metric (factor loading) and error (level of measurement error) invariance. These analyses showed that the communication, gross motor, and fine motor domains depicted adequate metric invariance, suggesting that the factor structure and item factor loadings of these domains are similar across all of the age intervals. Finally, the gross motor domain also presented satisfactory error invariance, indicating that the items of this domain measure the same latent trait with an equivalent level of error measurement.

The ASQ personal–social domain involves items related to child independence and social behavior, such as “toileting skills” or “playing cooperatively with peers.” However, other items that are present in this domain, such as the child's ability to “feed himself with a spoon” or “dress and undress himself, including buttoning and zipping front zipper,” might be related to motor skills. Items in the ASQ personal–social domain might also be associated with cognitive skills, such as questions related to gender or remembering the family name.

Table 2
Mean and standard deviation (SD) of ASQ-BR raw scores for all of the subjects divided by gender for each of the five ASQ-BR domains across the 18 age intervals. Respective p values from Student's t-test comparisons between the two groups are presented, with significant differences between groups marked with an asterisk. The theta measure and standard error (SE) are presented, calculated using the One-step Equating method from the Rating Scale model of the ASQ-BR for all the subjects divided by gender for each development domain across age intervals. The p values from Fisher's Least Significant Difference (LSD) post hoc test following an 18 × 2 two-way ANOVA are presented to compare genders within each age interval. The reliability index, Cronbach's alpha (α), is also presented.

Age intervals (months)	Communication								α	Gross motor							
	Mean (SD)				Theta (SE)					Mean (SD)				Theta (SE)			
	Global	Boys	Girls	t-test	Global	Boys	Girls	Fisher's LSD		Global	Boys	Girls	t-test	Global	Boys	Girls	Fisher's LSD
6	37.1 (13.6)	34.5 (9.6)	40.5 (16.2)	$p = 0.12$	0.88 (0.93)	0.78 (0.22)	0.95 (0.29)	$p = 0.13$	0.70	33.3 (15.9)	39.1 (8.4)	32.1 (14.6)	$p = 0.55$	1.34 (0.29)	1.53 (0.31)	1.18 (0.27)	$p = 0.81$
8	37.4 (14.1)	38.0 (7.8)	35.6 (7.3)	$p = 0.13$	0.96 (0.16)	1.39 (0.34)	0.82 (0.17)	$p = 0.14$	0.67	41.1 (17.9)	40.9 (10.8)	41.3 (10.3)	$p = 0.98$	0.94 (0.14)	0.98 (0.28)	0.93 (0.16)	$p = 0.88$
10	26.7 (17.2)	27.9 (18.9)	25.8 (14.8)	$p = 0.45$	1.27 (0.12)	1.45 (0.18)	1.16 (0.17)	$p = 0.25$	0.76	38.5 (16.9)	37.5 (15.7)	40.3 (16.7)	$p = 0.39$	1.27 (0.16)	1.12 (0.28)	1.36 (0.19)	$p = 0.48$
12	36.4 (16.6)	33.9 (16.8)	38.2 (16.2)*	$p < 0.01$	1.56 (0.07)	1.32 (0.11)	1.73 (0.09)*	$p < 0.01$	0.75	43.3 (17.9)	43.1 (18.2)	43.3 (17.7)	$p = 0.89$	1.54 (0.10)	1.51 (0.16)	1.55 (0.13)	$p = 0.86$
14	31.3 (16.7)	28.8 (16.3)	32.1 (16.6)*	$p < 0.01$	1.14 (0.06)	1.07 (0.09)	1.30 (0.08)*	$p < 0.01$	0.75	47.5 (17.9)	45.9 (19.6)	48.7 (16.5)*	$p < 0.05$	1.47 (0.07)	1.30 (0.12)	1.59 (0.09)	$p = 0.06$
16	29.9 (14.6)	29.6 (13.8)	30.2 (15.1)	$p = 0.54$	1.03 (0.06)	1.02 (0.09)	1.06 (0.08)	$p = 0.49$	0.75	52.2 (13.7)	53.4 (13.3)	51.4 (13.9)*	$p < 0.03$	1.63 (0.06)	1.80 (0.08)	1.52 (0.07)*	$p < 0.02$
18	33.5 (16.3)	32.3 (16.1)	34.5 (16.6)*	$p < 0.03$	1.47 (0.06)	1.33 (0.09)	1.58 (0.08)*	$p < 0.04$	0.73	55.8 (7.9)	56.7 (7.1)	55.2 (8.5)*	$p < 0.01$	1.43 (0.03)	1.55 (0.04)	1.34 (0.04)*	$p < 0.01$
20	33.1 (19.2)	29.8 (18.6)	35.3 (18.9)*	$p < 0.01$	0.54 (0.06)	0.19 (0.08)	0.85 (0.06)*	$p < 0.01$	0.77	52.3 (11.5)	53.0 (10.6)	51.9 (12.1)	$p = 0.12$	1.10 (0.04)	1.16 (0.06)	1.06 (0.05)	$p = 0.20$
22	35.4 (18.4)	31.5 (18.3)	36.9 (18.2)*	$p < 0.01$	0.78 (0.06)	0.56 (0.07)	0.89 (0.06)*	$p < 0.01$	0.82	48.2 (13.0)	47.2 (13.2)	48.8 (13.0)	$p = 0.06$	1.19 (0.03)	1.15 (0.03)	1.22 (0.02)*	$p < 0.02$
24	44.1 (17.5)	40.9 (18.2)	46.1 (16.9)*	$p < 0.01$	1.44 (0.05)	1.11 (0.08)	1.65 (0.06)*	$p < 0.01$	0.81	52.1 (10.6)	52.6 (10.6)	51.9 (10.7)	$p = 0.19$	1.32 (0.03)	1.39 (0.03)	1.27 (0.04)*	$p < 0.01$
27	48.1 (14.3)	46.4 (14.7)	49.2 (14.1)*	$p < 0.01$	1.67 (0.06)	1.49 (0.05)	1.83 (0.06)*	$p < 0.01$	0.84	50.9 (11.9)	52.0 (10.9)	50.3 (12.5)*	$p < 0.01$	1.24 (0.02)	1.28 (0.03)	1.20 (0.02)*	$p < 0.01$
30	50.3 (13.0)	49.6 (13.5)	50.8 (12.8)*	$p < 0.03$	1.41 (0.03)	1.32 (0.05)	1.56 (0.03)*	$p < 0.02$	0.77	53.5 (10.1)	55.0 (8.5)	52.7 (10.9)*	$p < 0.01$	1.09 (0.03)	1.19 (0.02)	1.05 (0.04)*	$p < 0.01$
33	47.8 (14.3)	46.6 (15.2)	48.6 (13.7)*	$p < 0.01$	1.87 (0.04)	1.74 (0.03)	1.95 (0.04)*	$p < 0.01$	0.75	52.2 (11.2)	52.6 (11.3)	52.0 (11.3)	$p = 0.15$	1.18 (0.04)	1.23 (0.03)	1.13 (0.05)*	$p < 0.03$
36	46.8 (12.4)	45.9 (13.0)	47.3 (12.1)*	$p < 0.01$	1.74 (0.06)	1.71 (0.08)	1.79 (0.04)*	$p < 0.01$	0.74	53.8 (10.4)	54.5 (9.9)	53.4 (10.8)*	$p < 0.01$	1.46 (0.06)	1.54 (0.03)	1.41 (0.06)*	$p < 0.01$
42	46.9 (12.8)	46.2 (13.3)	47.4 (12.5)*	$p < 0.01$	1.82 (0.02)	1.79 (0.01)	1.84 (0.03)*	$p < 0.01$	0.66	54.8 (8.7)	55.2 (8.8)	54.7 (8.8)*	$p < 0.01$	1.27 (0.04)	1.34 (0.02)	1.23 (0.03)*	$p < 0.01$
48	48.2 (13.9)	46.5 (14.9)	49.3 (13.4)*	$p < 0.01$	1.85 (0.02)	1.82 (0.02)	1.87 (0.02)*	$p < 0.01$	0.66	53.3 (10.4)	53.2 (10.5)	53.5 (10.4)	$p = 0.15$	1.31 (0.02)	1.30 (0.03)	1.32 (0.02)	$p = 0.23$
54	52.4 (11.7)	51.1 (12.9)	53.2 (11.0)*	$p < 0.01$	1.53 (0.03)	1.31 (0.04)	1.72 (0.02)*	$p < 0.01$	0.78	54.3 (9.8)	54.0 (10.3)	54.6 (9.5)*	$p < 0.01$	1.36 (0.03)	1.41 (0.03)	1.33 (0.02)*	$p < 0.01$
60	47.9 (15.2)	50.0 (8.7)	47.8 (15.7)	$p = 0.82$	1.52 (0.18)	1.63 (0.22)	1.46 (0.29)	$p = 0.78$	0.78	48.5 (17.0)	57.0 (3.2)	47.7 (17.3)	$p = 0.23$	1.19 (0.20)	1.36 (0.22)	1.09 (0.29)	$p = 0.81$

Table 2 (*continued*)

Age intervals (months)	α	Fine motor								α	Problem solving							
		Mean (SD)				Theta (SE)					Mean (SD)				Theta (SE)			
		Global	Boys	Girls	*t*-test	Global	Boys	Girls	Fisher's LSD		Global	Boys	Girls	*t*-test	Global	Boys	Girls	Fisher's LSD
6	0.71	36.8(18.7)	33.5(8.3)	40.7(12.4)	$p = 0.47$	1.57(0.21)	1.22(0.44)	1.60(0.21)	$p = 0.38$	0.78	39.5(17.5)	42.5(2.3)	38.9(16.9)	$p = 0.81$	1.31(0.10)	1.36(0.12)	1.26(0.16)	$p = 0.54$
8	0.79	41.1(17.5)	37.5(12.4)	42.8(10.5)	$p = 0.64$	1.34(0.10)	1.30(0.12)	1.39(0.14)	$p = 0.28$	0.79	40.2(16.5)	43.4(10.4)	38.3(17.5)	$p = 0.55$	1.46(0.14)	1.52(0.16)	1.33(0.17)	$p = 0.38$
10	0.79	37.6(17.2)	33.6(14.5)	40.9(18.9)	$p = 0.23$	1.29(0.14)	1.20(0.19)	1.41(0.22)	$p = 0.30$	0.78	35.5(16.6)	36.7(16.8)	32.4(19.8)	$p = 0.14$	1.07(0.12)	1.18(0.14)	1.04(0.08)	$p = 0.19$
12	0.85	38.1(17.5)	37.8(17.8)	38.3(17.3)	$p = 0.74$	1.14(0.09)	1.08(0.12)	1.18(0.13)	$p = 0.47$	0.77	35.7(17.6)	33.8(18.5)	37.2(16.8)*	$p < 0.05$	1.18(0.11)	1.09(0.12)	1.22(0.16)	$p = 0.28$
14	0.85	35.8(17.4)	34.8(17.9)	36.6(17.1)	$p = 0.16$	1.42(0.08)	1.34(0.10)	1.46(0.12)	$p = 0.86$	0.77	34.1(15.5)	32.2(18.8)	35.5(17.6)*	$p < 0.01$	1.56(0.07)	1.50(0.09)	1.61(0.11)	$p = 0.86$
16	0.89	43.2(16.7)	44.3(16.1)	42.6(17.3)	$p = 0.12$	1.25(0.06)	1.29(0.06)	1.21(0.08)	$p = 0.20$	0.75	36.7(18.2)	37.0(17.8)	36.5(18.5)	$p = 0.67$	1.03(0.06)	1.05(0.07)	1.00(0.08)	$p = 0.33$
18	0.85	44.7(14.6)	46.1(13.7)	43.7(15.3)*	$p < 0.01$	1.31(0.04)	1.41(0.05)	1.19(0.06)	$p < 0.01$	0.79	37.3(15.3)	38.6(15.1)	36.4(15.5)*	$p < 0.02$	1.67(0.06)	1.74(0.07)	1.60(0.07)*	$p < 0.01$
20	0.64	41.2(14.8)	40.7(14.4)	41.6(15.1)	$p = 0.38$	1.17(0.03)	1.11(0.06)	1.22(0.04)	$p = 0.28$	0.72	38.2(13.9)	37.9(13.3)	38.5(14.3)	$p = 0.48$	1.87(0.04)	1.84(0.05)	1.90(0.05)*	$p = 0.14$
22	0.74	40.4(13.6)	39.6(13.7)	40.9(13.6)	$p = 0.19$	1.22(0.04)	1.20(0.04)	1.23(0.05)	$p = 0.13$	0.67	39.3(13.9)	37.5(14.3)	40.3(13.7)*	$p < 0.01$	1.85(0.02)	1.78(0.03)	1.91(0.04)*	$p < 0.01$
24	0.71	42.8(12.6)	42.9(12.2)	42.9(13.3)	$p = 0.94$	1.45(0.03)	1.44(0.04)	1.45(0.04)	$p = 0.14$	0.64	41.1(13.4)	40.4(13.4)	41.7(13.3)	$p = 0.08$	1.31(0.03)	1.22(0.04)	1.36(0.03)*	$p < 0.03$
27	0.67	34.1(15.2)	33.8(14.4)	34.4(15.8)	$p = 0.39$	1.02(0.04)	0.89(0.05)	1.10(0.05)	$p < 0.01$	0.62	47.8(12.5)	47.6(12.0)	47.9(12.8)	$p = 0.52$	1.14(0.03)	1.14(0.03)	1.15(0.04)	$p = 0.38$
30	0.71	34.2(17.7)	33.2(17.2)	34.9(17.9)*	$p < 0.02$	0.93(0.03)	0.86(0.04)	1.01(0.03)	$p < 0.01$	0.70	45.0(14.8)	44.6(14.6)	45.3(14.9)	$p = 0.25$	1.45(0.04)	1.43(0.05)	1.48(0.05)	$p = 0.14$
33	0.67	36.4(18.7)	35.0(18.5)	37.3(18.7)*	$p < 0.01$	0.95(0.04)	0.90(0.05)	1.03(0.06)	$p < 0.01$	0.78	46.8(14.1)	45.9(14.5)	47.3(13.8)*	$p < 0.01$	1.32(0.02)	1.28(0.03)	1.35(0.03)*	$p < 0.01$
36	0.69	42.2(17.9)	41.5(18.1)	42.7(17.7)*	$p < 0.02$	1.28(0.03)	1.19(0.03)	1.34(0.06)	$p < 0.01$	0.79	48.4(13.4)	47.9(13.6)	48.7(13.3)*	$p < 0.04$	1.01(0.05)	0.94(0.06)	1.08(0.07)*	$p < 0.02$
42	0.71	41.3(15.5)	40.6(15.5)	41.7(15.4)*	$p < 0.01$	1.41(0.03)	1.37(0.04)	1.45(0.04)	$p < 0.01$	0.79	48.7(13.2)	47.6(13.7)	49.3(12.8)*	$p < 0.01$	1.78(0.02)	1.71(0.03)	1.86(0.03)*	$p < 0.01$
48	0.64	39.0(16.4)	36.3(16.9)	40.5(16.0)*	$p < 0.01$	1.19(0.03)	1.15(0.03)	1.25(0.04)	$p < 0.01$	0.71	44.1(14.5)	41.5(15.4)	45.4(13.9)*	$p < 0.01$	1.43(0.04)	1.35(0.05)	1.48(0.05)*	$p < 0.01$
54	0.70	44.4(14.5)	54.0(10.3)	54.6(9.5)*	$p < 0.01$	1.14(0.02)	1.07(0.04)	1.18(0.03)	$p < 0.01$	0.72	40.6(14.5)	41.8(15.4)	46.0(13.8)*	$p < 0.01$	1.36(0.03)	1.27(0.05)	1.39(0.04)*	$p < 0.01$
60	0.71	46.8(15.4)	28.3(5.7)	48.1(15.1)*	$p < 0.03$	1.35(0.34)	1.02(0.33)	1.57(0.38)	$p = 0.73$	0.72	46.7(15.3)	36.7(10.4)	47.4(15.4)	$p = 0.24$	1.24(0.24)	1.10(0.28)	1.33(0.30)	$p = 0.45$

(*continued on next page*)

Table 2 (*continued*)

Age intervals (months)	α	Personal–social								α
		Mean (SD)				Theta (SE)				
		Global	Boys	Girls	*t*-test	Global	Boys	Girls	Fisher's LSD	
6	0.79	33.8(16.7)	38.9(11.9)	30.7(18.4)	$p = 0.46$	0.85(0.22)	1.03(0.28)	0.73(0.24)	$p = 0.64$	0.70
8	0.78	45.6(14.3)	45.9(15.5)	45.2(13.8)	$p = 0.24$	1.45(0.14)	1.48(0.17)	1.43(0.18)	$p = 0.16$	0.67
10	0.72	35.8(13.8)	39.1(16.7)	32.7(16.5)	$p = 0.34$	1.14(0.16)	1.18(0.18)	1.11(0.18)	$p = 0.19$	0.58
12	0.77	33.8(17.1)	31.6(18.5)	35.5(15.8)*	$p < 0.03$	1.25(0.21)	1.00(0.23)	1.49(0.22)*	$p < 0.03$	0.76
14	0.77	30.6(16.5)	29.4(16.7)	32.7(16.6)*	$p < 0.01$	1.14(0.06)	1.08(0.06)	1.22(0.07)*	$p < 0.01$	0.76
16	0.80	34.0(15.5)	32.7(15.2)	35.0(15.7)*	$p < 0.03$	1.47(0.06)	1.43(0.07)	1.51(0.09)	$p = 0.38$	0.71
18	0.79	41.6(14.5)	40.3(15.2)	42.5(13.9)*	$p < 0.02$	0.78(0.06)	0.69(0.08)	0.91(0.08)	$p = 0.86$	0.64
20	0.72	36.2(12.9)	33.9(13.1)	37.9(12.5)*	$p < 0.01$	1.10(0.04)	1.01(0.05)	1.17(0.05)*	$p < 0.01$	0.64
22	0.63	38.1(12.6)	35.1(13.2)	39.7(12.0)*	$p < 0.01$	1.09(0.03)	1.03(0.06)	1.11(0.02)*	$p < 0.01$	0.57
24	0.65	35.7(13.1)	33.0(13.1)	37.4(12.6)*	$p < 0.01$	1.46(0.05)	1.36(0.07)	1.54(0.08)*	$p < 0.01$	0.59
27	0.63	34.3(11.4)	32.1(11.5)	35.8(11.2)*	$p < 0.01$	1.63(0.06)	1.53(0.07)	1.68(0.06)*	$p < 0.01$	0.63
30	0.62	46.2(11.7)	46.0(11.2)	46.4(12.0)	$p = 0.39$	1.19(0.03)	1.17(0.04)	1.20(0.04)*	$p < 0.03$	0.53
33	0.69	46.8(12.9)	45.0(13.6)	47.8(12.4)*	$p < 0.01$	1.09(0.03)	1.01(0.04)	1.13(0.03)*	$p < 0.01$	0.61
36	0.67	49.5(10.7)	48.1(11.5)	50.2(10.2)*	$p < 0.01$	1.65(0.03)	1.50(0.04)	1.72(0.04)*	$p < 0.01$	0.66
42	0.67	48.0(10.4)	47.0(11.1)	48.6(9.9)*	$p < 0.01$	1.83(0.04)	1.73(0.03)	1.94(0.05)*	$p < 0.01$	0.57
48	0.67	47.6(11.2)	46.6(11.5)	48.1(10.9)*	$p < 0.01$	1.04(0.02)	1.00(0.03)	1.09(0.03)*	$p < 0.01$	0.52
54	0.68	49.3(10.5)	38.0(14.9)	42.1(14.1)*	$p < 0.01$	1.19(0.03)	1.12(0.04)	1.20(0.05)*	$p < 0.01$	0.53
60	0.70	51.6(11.3)	50.0(8.7)	51.7(11.6)	$p = 0.81$	0.68(0.10)	0.50(0.12)	0.74(0.08)*	$p < 0.05$	0.52

Table 3
Minimum (Min) and maximum (Max) mean square infit and outfit statistics and respective mean and standard deviation (SD) for each of the five ASQ-BR domains across the 18 age intervals.

Age intervals (months)	Statistics	Communication			Gross motor			Fine motor			Problem solving			Personal–social		
		Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)
6	Infit	0.55	1.15	0.88 (0.22)	0.93	1.22	1.02 (0.11)	0.81	1.14	0.98 (0.12)	0.74	1.40	1.04 (0.21)	0.67	1.24	0.98 (0.18)
	Out fit	0.43	9.23	2.47 (3.14)[d]	0.51	1.26	0.86 (0.28)	0.62	1.48	1.05 (0.27)	0.65	1.43	1.07 (0.36)	0.55	1.49	1.04 (0.28)
8	Infit	0.62	1.27	0.96 (0.21)	0.73	1.25	1.03 (0.18)	0.69	1.35	1.00 (0.22)	0.89	1.18	1.01 (0.07)	0.88	1.22	0.99 (0.13)
	Out fit	0.46	1.56	1.15 (0.35)[a]	0.51	1.23	0.92 (0.23)	0.63	1.45	1.00 (0.32)	0.83	1.20	1.00 (0.13)	0.77	1.37	1.03 (0.20)
10	Infit	0.77	1.44	1.01 (0.22)	0.64	1.50	1.01 (0.33)	0.74	1.16	1.02 (0.14)	0.75	1.17	0.98 (0.15)	0.70	1.37	0.99 (0.18)
	Out fit	0.59	1.50	0.97 (0.31)	0.61	1.98	1.32 (0.57)[b]	0.59	1.50	1.03 (0.31)	0.70	1.48	1.10 (0.28)	0.61	2.52	1.48 (1.09)[a]
12	Infit	0.93	1.08	1.01 (0.05)	0.83	1.13	0.97 (0.11)	0.82	1.14	1.00 (0.11)	0.74	1.35	1.01 (0.20)	0.83	1.26	1.00 (0.14)
	Out fit	0.85	1.35	1.02 (0.17)	0.76	1.49	1.07 (0.28)	0.76	1.24	1.01 (0.18)	0.65	1.50	1.05 (0.33)	0.67	1.30	0.97 (0.20)
14	Infit	0.78	1.16	0.99 (0.13)	0.50	1.62	0.87 (0.36)[a]	0.73	1.19	1.00 (0.17)	0.70	1.22	0.97 (0.17)	0.87	1.14	1.00 (0.09)
	Out fit	0.74	1.49	1.05 (0.26)	0.34	5.57	1.53 (1.82)[c]	0.63	1.34	1.02 (0.26)	0.68	1.50	1.15 (0.36)	0.83	1.30	1.02 (0.14)
16	Infit	0.78	1.19	0.96 (0.13)	0.68	1.14	0.89 (0.16)	0.79	1.34	1.01 (0.19)	0.91	1.18	1.03 (0.10)	0.79	1.22	1.01 (0.15)
	Out fit	0.77	2.70	1.51 (0.58)[b]	0.65	1.57	1.16 (0.36)[b]	0.69	1.49	1.02 (0.29)	0.88	1.12	0.97 (0.09)	0.72	1.49	1.01 (0.25)
18	Infit	0.83	0.99	0.94 (0.06)	0.83	1.09	0.97 (0.11)	0.72	1.25	1.02 (0.20)	0.77	1.24	1.01 (0.20)	0.86	1.20	0.99 (0.13)
	Out fit	0.81	2.26	1.48 (0.60)[c]	0.67	1.26	0.99 (0.18)	0.59	1.48	1.01 (0.34)	0.65	2.24	1.24 (0.55)[b]	0.76	1.30	0.98 (0.19)
20	Infit	0.77	1.11	0.99 (0.13)	0.89	1.14	1.01 (0.09)	0.82	1.18	1.00 (0.13)	0.83	1.22	1.00 (0.12)	0.89	1.17	1.00 (0.09)
	Out fit	0.76	1.58	1.10 (0.24)[a]	0.77	1.12	0.95 (0.14)	0.72	1.21	0.99 (0.20)	0.77	1.81	1.20 (0.38)[b]	0.79	3.53	1.39 (0.97)[a]
22	Infit	0.77	1.20	0.99 (0.13)	0.82	1.11	1.01 (0.11)	0.78	1.10	0.98 (0.14)	0.68	1.25	1.03 (0.19)	0.84	1.11	1.00 (0.11)
	Out fit	0.70	1.32	1.17 (0.49)	0.74	1.20	1.00 (0.14)	0.81	1.13	1.02 (0.20)	0.85	1.19	1.00 (0.11)	0.67	4.79	1.70 (1.44)[b]
24	Infit	0.76	1.31	1.00 (0.19)	0.91	1.14	1.02 (0.10)	0.85	1.20	0.99 (0.12)	0.84	1.15	0.98 (0.10)	0.84	1.48	1.02 (0.22)
	Out fit	0.75	2.56	1.23 (0.63)[a]	0.74	1.28	1.03 (0.18)	0.85	1.55	1.12 (0.26)[a]	0.85	1.70	1.19 (0.29)[a]	0.88	3.25	1.42 (0.84)[a]
27	Infit	0.77	1.36	1.00 (0.19)	0.78	1.36	0.98 (0.19)	0.68	1.30	1.03 (0.31)	0.81	1.18	0.97 (0.09)	0.84	1.15	0.99 (0.11)
	Out fit	0.69	1.42	1.03 (0.26)	0.73	1.50	1.06 (0.27)	0.79	1.28	1.11 (0.19)	0.76	1.13	1.00 (0.12)	0.81	2.50	1.38 (0.59)[b]
30	Infit	0.83	1.27	1.00 (0.15)	0.86	1.08	0.99 (0.09)	0.83	1.39	1.09 (0.24)	0.69	2.13	1.16 (0.47)[a]	0.85	1.08	0.99 (0.07)
	Out fit	0.77	1.49	1.09 (0.29)	0.80	1.13	0.99 (0.11)	0.80	1.27	1.11 (0.29)	0.87	1.59	1.06 (0.17)[a]	0.76	1.21	1.06 (0.14)
33	Infit	0.88	1.10	0.99 (0.12)	0.85	1.10	1.00 (0.08)	0.73	1.12	1.08 (0.21)	0.89	1.26	1.01 (0.22)	0.91	1.11	1.00 (0.08)
	Out fit	0.81	1.17	0.96 (0.14)	0.79	1.17	1.00 (0.11)	0.89	1.14	1.01 (0.08)	0.92	1.15	1.07 (0.07)	0.81	1.15	0.96 (0.15)
36	Infit	0.79	1.10	0.97 (0.11)	0.77	1.22	1.00 (0.16)	0.75	1.30	1.00 (0.22)	0.78	1.31	0.99 (0.20)	0.92	1.16	1.04 (0.08)
	Out fit	0.89	1.32	1.06 (0.28)	0.66	1.40	1.01 (0.24)	0.67	1.36	0.98 (0.26)	0.68	1.38	1.00 (0.25)	0.86	1.06	0.94 (0.07)
42	Infit	0.88	1.13	0.99 (0.06)	0.91	1.18	1.00 (0.08)	0.81	1.18	1.00 (0.18)	0.83	1.14	1.00 (0.12)	0.86	1.13	0.97 (0.09)
	Out fit	0.90	1.38	1.07 (0.18)	0.84	1.26	0.99 (0.13)	0.72	1.25	1.03 (0.27)	0.71	1.13	0.97 (0.16)	0.84	1.23	1.04 (0.15)
48	Infit	0.85	1.32	1.03 (0.15)	0.87	1.26	1.03 (0.14)	0.68	1.46	0.98 (0.34)	0.92	1.09	1.01 (0.07)	0.89	1.08	0.98 (0.08)
	Out fit	0.82	1.31	1.00 (0.17)	0.82	1.20	1.00 (0.14)	0.65	1.35	0.99 (0.37)	0.86	1.19	1.00 (0.10)	0.82	1.26	1.02 (0.15)
54	Infit	0.79	1.46	1.04 (0.22)	0.90	1.18	1.00 (0.10)	0.54	1.23	0.87 (0.39)	0.83	1.21	1.00 (0.14)	0.88	1.10	0.99 (0.07)
	Out fit	0.72	1.49	0.97 (0.25)	0.86	1.14	0.96 (0.10)	0.56	1.19	0.89 (0.28)	0.73	1.83	1.15 (0.42)[b]	0.72	1.23	1.03 (0.18)
60	Infit	0.70	1.34	1.01 (0.22)	0.45	2.77	1.02 (0.81)[b]	0.65	1.18	0.96 (0.18)	0.64	1.53	1.04 (0.32)[a]	0.78	1.50	0.98 (0.26)
	Out fit	0.61	1.37	1.07 (0.17)	0.44	4.45	1.20 (1.47)[e]	0.62	1.47	1.07 (0.30)	0.32	1.48	0.92 (0.32)[b]	0.27	2.09	1.06 (0.56)[b]

[a] 1 item with MsSq fit statistics outside limit.
[b] 2 item with MsSq fit statistics outside limit.
[c] 3 items with MnSq fit statistics outside limit.
[d] 4 items with MsSq statistics outside limit.
[e] 5 items with MnSq fit statistics outside limit.

To deal with the complexity of personal–social abilities, Squires et al. [54] developed the ASQ Social–Emotional (ASQ-SE). This is a screening instrument for possible developmental delays in social–emotional functioning for children 6 to 60 months of age. Accordingly, the ASQ-SE evaluates different aspects of the personal–social domain, such as self-regulation, compliance, communication, adaptive functioning, autonomy, affect, and interaction with people.

Rasch family analyses allow the evaluation of ASQ-BR psychometric characteristics that cannot be analyzed using Classic Test Theory techniques, such as how well an item fits into a unidimensional logistic model for the probability of responses. The fact that more than 90% of the 540 items had infit and outfit MNSQ within the acceptable range of 0.5 to 1.5 provided additional psychometric support for the ASQ-BR, indicating that the majority of the ASQ-BR items targeted the intended construct.

Infit and outfit statistics evaluate different aspects of an item based on differences in how they are calculated. The infit statistic places less importance on extreme responses and is weighted according to items that have better alignment to the person's level of ability. The outfit statistic is not weighted and is sensitive to extreme off-targeted responses. For that reason, infit statistics are more useful for assessing the utility of an item because they are not affected by outliers [55]. Conversely, outfit statistics are less critical for item evaluation because they are focused on sensitivity in instances when a large difference exists between an item's difficulty and the person's ability.

Accordingly, ASQ-BR items fitted the model particularly well with regard to infit statistics; only four of the 540 items (0.7%) presented fitting problems. Most of the misfit items presented outfit statistic values greater than 1.5 (32 of 44; 72.3%), suggesting that they might have problems related to culturally inappropriate content or ambiguous wording and thus lead to item misunderstanding. Therefore, these items might need to be further reviewed by our panel of specialists to consider alternative adaptations or language evaluation.

To understand the relationship among items in the entire set for each development domain from the ASQ-BR, we linked the scales of the different age intervals in a common scale using the One-step Equating method [50,51]. This analysis allowed us to infer that items from the ASQ-BR have good properties and are well distributed through the latent trait. This statement is based on the fact that no items presented misfit problems once they were calibrated in the common scale. The items with misfit depicted in the RS analyses probably showed problems in one specific age interval, but when calibrated with other scales, the misfit was diluted. Our hypothesis is that the calibration in Equating makes room for different fits of the entire item sets. A misfitted item in one specific age interval could be fit to items from other age intervals. For example, item 3 from the 22-month-old personal–social questionnaire presents an outfit issue (>1.5). It asks about the "toileting skills" of the child but is the only item that addresses this matter in the scale. However, the same item appears in the 20, 24, and 27 month questionnaires. In these other scales, the outfit problem is only shown in the 27-month scale. The calibration

Table 4
Item numbers in the common scale (Item), considering all items (i.e., both common and non-common), when calibrated to the equated scale, logits, standard errors, and infit and outfit statistics for each development domain when the items were calibrated together using the One-step Equating method for linking the scales of the ASQ-BR.

Common scales of the ASQ-BR																								
Communication					Gross motor					Fine motor					Problem solving					Personal–social				
Item	Item difficulty in logits	Stand. error	Infit	Outfit	Item	Item difficulty in logits	Stand. error	Infit	Outfit	Item	Item difficulty in logits	Stand. error	Infit	Outfit	Item	Item difficulty in logits	Stand. error	Infit	Outfit	Item	Item difficulty in logits	Stand. error	Infit	Outfit
16	1.24	.03	1.05	1.04	6	.78	.09	1.04	1.04	34	.86	.01	.93	.92	21	1.39	.02	1.14	1.16	23	2.38	.03	1.30	1.44
20	1.22	.02	1.03	1.01	5	.70	.09	1.05	1.05	29	.69	.01	.97	.96	42	1.00	.01	.90	.90	33	1.26	.01	1.08	1.08
33	.91	.01	1.23	1.23	26	.63	.03	1.04	1.06	36	.59	.01	.95	.95	17	.96	.02	1.01	1.01	15	1.24	.03	1.05	1.03
22	.88	.02	1.11	1.11	11	.51	.05	1.01	.97	28	.54	.01	.93	.90	43	.77	.01	.96	.96	27	.85	.01	.95	.95
11	.88	.03	1.09	1.09	15	.42	.03	1.03	1.02	12	.53	.05	1.03	1.03	40	.44	.01	.90	.88	26	.72	.02	1.01	1.01
25	.87	.04	1.05	1.05	4	.35	.09	1.00	.96	22	.53	.02	1.08	1.09	10	.43	.05	.98	.98	2	.63	.15	.99	.99
6	.83	.07	.92	.93	8	.32	.08	1.03	1.03	35	.51	.01	.96	.94	14	.43	.02	1.03	1.03	16	.57	.03	.97	.97
19	.72	.02	1.03	1.03	10	.32	.05	.86	.81	30	.48	.01	.96	.91	25	.41	.03	.99	.99	18	.53	.02	1.04	1.04
7	.49	.11	1.02	1.04	24	.32	.01	1.02	.96	38	.29	.01	1.00	1.02	15	.40	.03	1.02	1.01	13	.45	.03	.94	.93
26	.48	.04	.97	.96	28	.30	.01	1.08	1.08	27	.28	.02	1.10	1.10	39	.25	.01	.87	.87	1	.44	.15	.95	.94
13	.33	.03	.97	.97	31	.26	.01	.87	.96	17	.28	.02	1.03	1.03	19	.25	.04	1.08	1.07	35	.42	.01	1.01	.98
27	.28	.02	1.05	1.04	33	.25	.01	.96	.90	24	.26	.01	1.07	1.09	11	.23	.05	.98	.97	12	.39	.03	.94	.93
39	.27	.01	.82	.83	29	.24	.01	1.08	1.08	32	.17	.01	1.06	1.09	37	.18	.01	1.04	1.03	11	.36	.05	.81	.81
5	.19	.06	1.05	1.07	14	.20	.03	1.10	1.00	31	.16	.01	1.13	1.19	8	.16	.07	.96	.95	10	.30	.05	.95	.95
17	.18	.02	1.00	1.00	38	.16	.18	.82	.80	3	.14	.09	1.00	.99	32	.16	.01	1.14	1.11	32	.27	.01	.82	.91
14	.13	.04	.96	.95	37	.12	.02	1.07	1.16	13	.13	.03	1.04	1.04	6	.12	.09	1.02	1.02	19	.19	.03	1.04	1.04
12	.09	.06	.99	.98	35	.08	.01	.92	.80	7	.13	.11	1.04	1.02	34	.10	.02	.99	.98	22	.07	.02	1.01	1.00
10	.04	.05	1.02	1.03	9	.07	.06	.97	.89	10	.11	.05	.99	.99	22	.07	.02	1.04	1.04	4	.06	.09	1.04	1.03
35	.01	.01	.88	.90	19	.07	.03	1.06	1.07	40	.11	.01	.92	.94	13	.07	.03	.92	.91	7	.02	.07	1.11	1.11
42	−.01	.17	1.05	1.03	2	.05	.17	1.11	1.06	33	.10	.01	1.04	1.04	41	.07	.01	.96	.90	20	−.01	.02	.89	.86
43	−.01	.17	.86	.85	25	.03	.02	1.03	.94	37	.00	.01	.97	.96	36	.07	.01	1.02	1.01	9	−.06	.05	.95	.92
21	−.01	.02	1.01	1.03	30	.03	.02	1.10	1.14	8	−.01	.04	1.06	1.07	44	.04	.17	.94	.91	17	−.09	.02	1.02	.97
8	−.02	.05	1.05	1.07	13	.02	.04	.97	.84	15	−.06	.03	.99	.96	7	−.03	.07	.98	.97	14	−.09	.02	.99	.97
34	−.04	.01	.94	.92	36	−.04	.02	.88	.73	11	−.09	.03	1.04	1.04	33	−.04	.01	1.04	1.03	37	−.12	.01	1.02	.99
31	−.04	.01	1.03	1.05	3	−.18	.08	1.06	1.15	25	−.10	.04	1.03	1.02	30	−.11	.01	.97	.96	25	−.14	.01	.96	.97
29	−.05	.01	1.09	1.03	27	−.18	.02	1.06	1.10	16	−.10	.03	1.04	1.01	26	−.13	.04	1.01	1.00	6	−.14	.10	1.02	1.01
23	−.08	.02	1.07	1.04	22	−.22	.02	.91	1.00	39	−.10	.01	.95	.98	2	−.13	.16	1.03	1.02	21	−.21	.02	1.04	.96
32	−.12	.01	1.00	.97	12	−.24	.05	.96	.80	42	−.11	.17	.92	.92	5	−.14	.07	.94	.93	29	−.35	.02	1.01	.98
9	−.12	.05	1.00	1.01	34	−.26	.02	.89	.67	20	−.12	.02	1.00	1.03	12	−.14	.04	.99	.96	28	−.37	.02	.86	.90
2	−.13	.16	.80	.79	7	−.28	.10	1.07	.95	41	−.15	.02	1.01	1.03	3	−.18	.10	.91	.89	5	−.39	.08	1.00	.97
30	−.37	.01	1.02	.96	16	−.35	.04	1.10	.89	6	−.21	.07	1.07	1.09	20	−.19	.04	1.03	1.01	31	−.39	.01	1.18	1.10
38	−.46	.01	.93	.82	18	−.35	.04	1.00	.98	2	−.24	.17	.93	.89	24	−.21	.03	.99	.98	39	−.42	.02	.94	1.00
37	−.48	.02	.86	.79	23	−.40	.02	1.11	1.02	14	−.27	.05	.98	.97	9	−.22	.05	1.00	.99	30	−.43	.01	1.14	1.06
1	−.53	.19	1.09	1.08	21	−.40	.03	1.08	1.08	9	−.32	.05	1.01	1.00	18	−.24	.02	1.05	1.05	8	−.56	.09	1.13	1.11
41	−.54	.02	1.07	1.07	17	−.45	.04	1.13	.89	18	−.36	.03	1.04	1.03	45	−.31	.19	.92	.86	38	−.56	.02	.89	.85
18	−.58	.04	1.00	.99	1	−.53	.22	1.11	1.00	4	−.38	.11	1.03	.97	29	−.33	.01	1.03	1.08	42	−.64	.22	.81	.77
36	−.64	.02	.97	.83	20	−.60	.04	1.05	1.08	43	−.40	.19	.83	.80	38	−.34	.01	.92	.96	24	−.73	.06	1.08	1.08
15	−.65	.04	1.01	.96	39	−.69	.27	1.11	.96	23	−.42	.02	1.05	1.06	35	−.35	.01	.98	.97	3	−.83	.14	.88	.78
28	−.89	.02	.95	.82	32	−1.05	.03	1.15	1.00	19	−.48	.02	1.04	1.01	31	−.46	.02	1.07	1.09	34	−.99	.02	1.05	1.02
3	−.97	.15	.92	.96	–	–	–	–	–	5	−.52	.08	1.05	1.01	16	−.46	.03	1.03	.98	36	−1.00	.02	1.13	.86
24	−.97	.02	1.12	1.00	–	–	–	–	–	21	−.73	.03	1.01	1.11	1	−.47	.19	1.03	1.00	41	−1.27	.33	1.03	.86
40	−.98	.03	1.15	1.05	–	–	–	–	–	1	−.79	.22	1.02	.86	27	−.50	.02	.98	1.00	40	−1.39	.35	1.17	1.20
4	−1.35	.19	.85	.84	–	–	–	–	–	26	−.90	.06	1.03	1.03	23	−.60	.05	1.07	1.07	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	28	−1.17	.04	.98	.94	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	4	−1.26	.19	1.02	.72	–	–	–	–	–

allowed the item to fit in the common scale when the item was compared with a larger set of items. This procedure enabled us to build a Construct Map of the five ASQ-BR developmental domains.

The Construct Map is a tool that is used to overlook the difficulty of the distribution of items and persons in a scale [48]. The item difficulties of all of the ASQ-BR's domains were normally distributed and

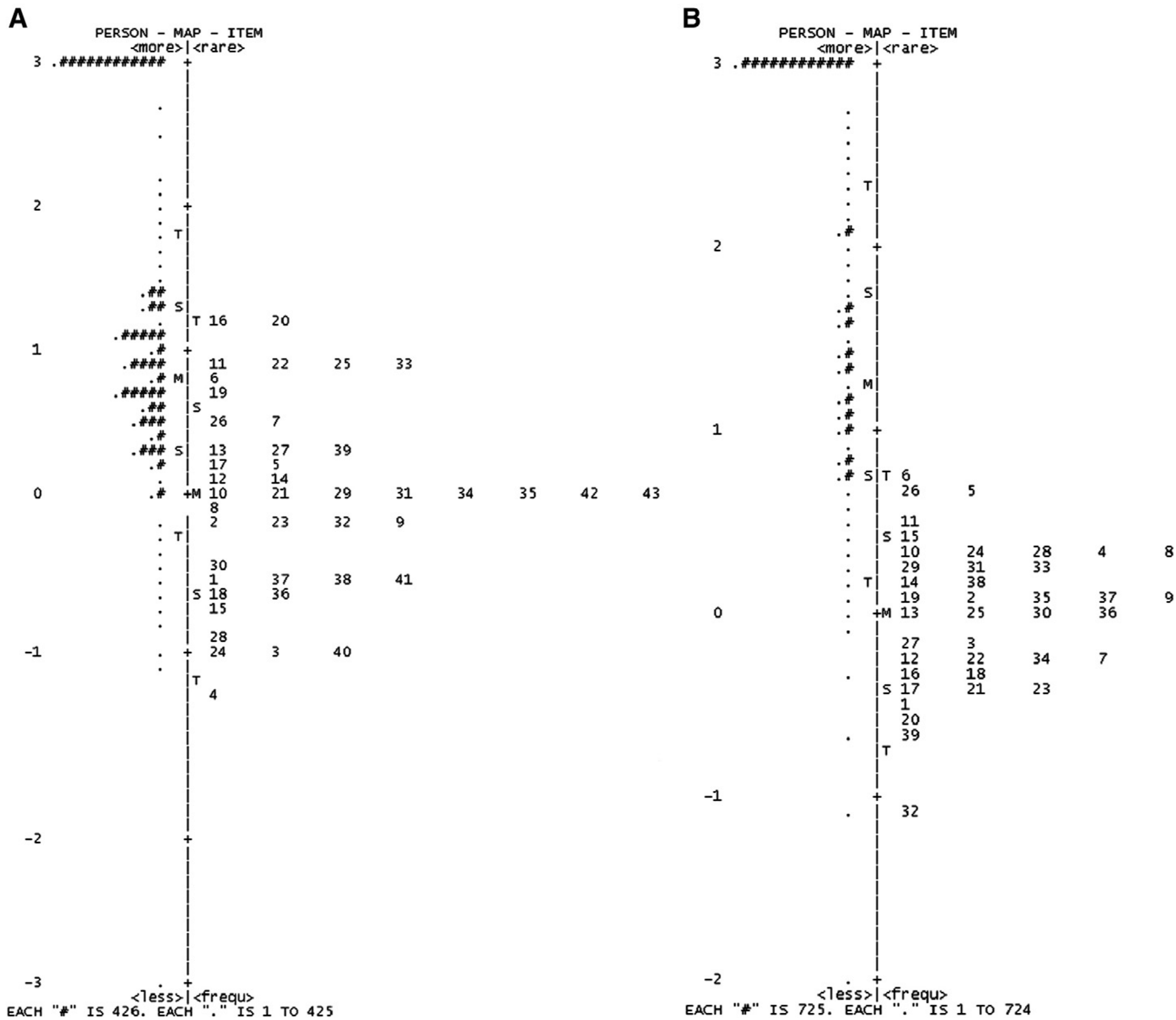

Fig. 2. A. Construct Map for the communication common scale of the ASQ-BR. Note: Example of items: logit +1.2, item 16, "Does your baby say four or more words in addition to 'Mama' and 'Dada'?"; logit +0.8, item 6, "If you copy the sounds your baby makes, does your baby repeat the same sounds back to you?"; logit +0.5, item 26, "Does your child say 15 or more words in addition to 'Mama' and 'Dada'?"; logit 0, item 42, "Does your child use comparison words, such as 'heavier,' 'stronger,' or 'shorter'?"; logit −0.5, item 38, "Does your child tell you at least two things about common objects? For example, if you say to your child, 'Tell me about the ball,' does she say something like, 'It's round. I throw it. It's big'?"; logit −1.0, item 3, "If you call your baby when you are out of sight, does she look in the direction of the voice?"; logit −1.3, item 4, "When a loud noise occurs, does your baby turn to see where the sound came from?" B. Construct Map for the gross motor common scale of the ASQ-BR. Note: Example of items: logit +0.8, item 6, "Does your baby get into a crawling position by getting up on her hands and knees?"; logit +0.4, item 15, "Does your baby climb onto the furniture or other large objects, such as large climbing blocks?"; logit 0, item 13, "When you hold one hand just to balance her, does your baby take several steps forward?"; logit −0.5, item 17, "Does your baby move around by walking, rather than by crawling on his hands and knees?"; logit −1.1, item 32, "Does your child climb the rungs of a ladder of a playground slide and slide down without help?" C. Construct Map for the fine motor common scale of the ASQ-BR. Note: Example of items: logit +0.5, item 12, "Without resting her arm or hand on the table, does your baby pick up a crumb or Cheerio with the *tips* of her thumb and a finger?"; logit +0.3, item 17, "Does your baby stack three small blocks or toys on top of each other by herself?"; logit 0, item 37, "Does your child color mostly within the lines in a coloring book or within the lines of a 5-centimeter circle that you draw?"; logit −0.5, item 19, "Does your child turn the pages of a book by himself?"; logit −0.9, item 26, "After your child watches you draw a line from the top of the paper to the bottom with a pencil, crayon, or pen, ask him to make a line like yours. Do not let your child trace your line. Does your child copy you by drawing a single line in a vertical direction?" D. Construct Map for the problem solving common scale of the ASQ-BR. Note: Example of items: logit +1.4, item 21, "After watching you draw a line from the top of the paper to the bottom with a crayon (or pencil or pen), does your child copy you by drawing a single line on the paper in any direction?"; logit +0.8, item 43, "Does your child know the names of numbers? 3–1–2"; logit 0, item 7, "Does your baby pick up two small toys, one in each hand, and hold onto them for about 1 minute?"; logit −0.5, item 1, "When a toy is in front of your baby, does she reach for it with both hands?"; logit −1.2, item 28, "Does your child put things away where they belong? For example, does he know his toys belong on the toy shelf, his blanket on his bed, and dishes in the kitchen?" E. Construct Map for the personal–social common scale of the ASQ-BR. Note: Example of items: logit +2.4, item 23, "Does your child eat with a fork?"; logit +0.9, item 27, "Does your child put on a coat, jacket, or shirt by himself?"; logit 0, item 20, "Does your child drink from a cup or glass, putting it down again with little spilling?"; logit −0.6, item 8, "Does your baby feed himself a cookie or fruit?"; logit −1.4, item 40, "Does your child wash her hands and face using soap and water and dry off with a towel without help?"

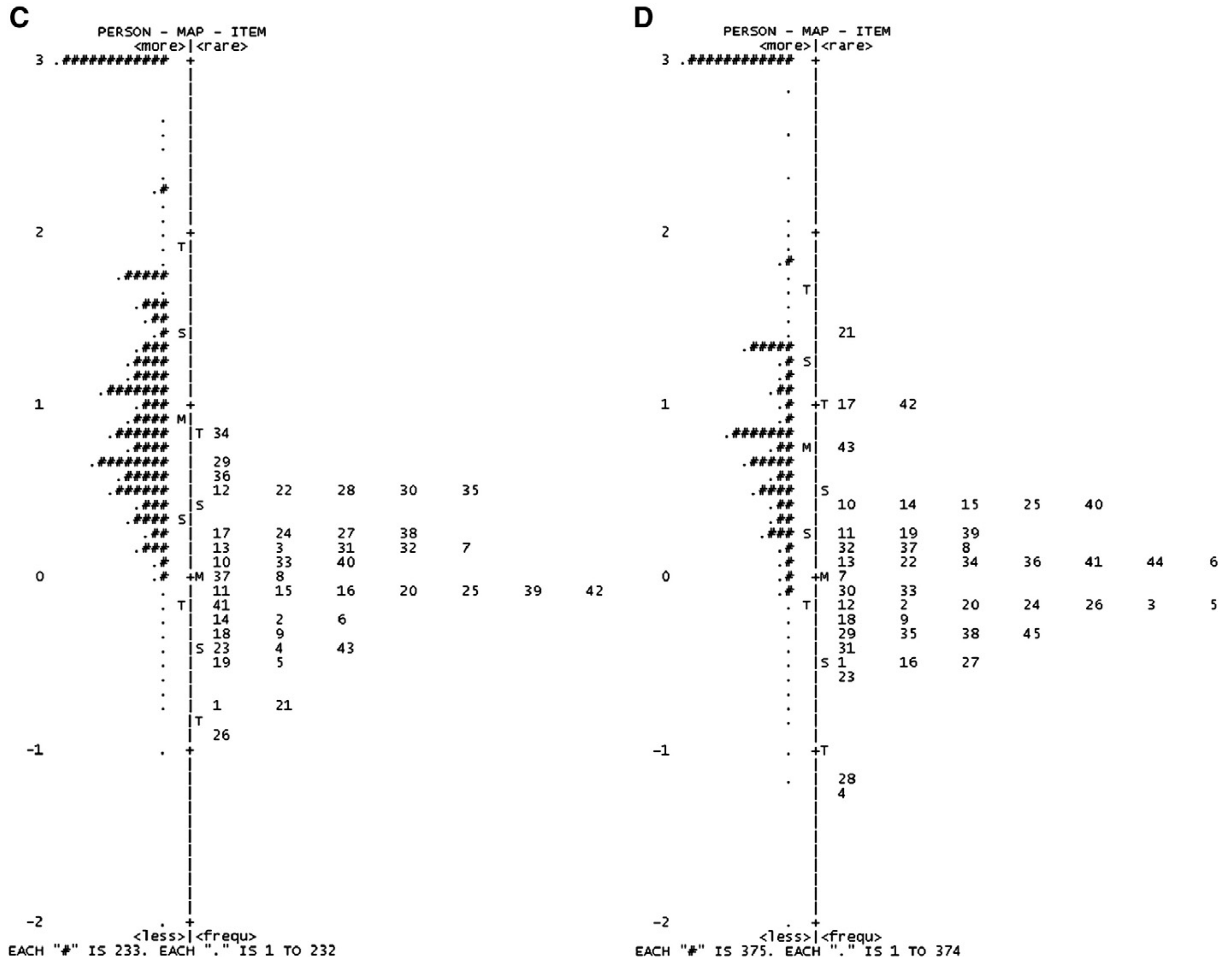

Fig. 2 (*continued*).

well spread. The only exception was item 23 from the personal–social domain. This item was clearly the most difficult item in the ASQ-BR and corresponds to an already discussed problem (i.e., "feed himself with a fork"). In Rio de Janeiro's childcare centers, children are not allowed to use forks for safety reasons. The item evaluation was likely compromised once teachers and professionals involved in caring for the children did not feel sufficiently secure to test the children. Thus, "never" was the most marked answer.

The distributions of persons' abilities were skewed. The ASQ-BR, such as the original instrument, the ASQ-3 [13], is essentially a screening instrument. It is expected that children whose items are marked by the caregivers already present complete or nearly complete development of the measured trait. The original proposition of the ASQ-3 [13] was to determine whether the child has adequate development for her age and not to separate persons by their abilities.

The Construct Map provided a glimpse into the persons' distributions for a screening test, such as ASQ-BR. A high concentration of "Yes" responses was expected. Thus, a total score of 60 (i.e., the maximum score) only indicates that the child is within a normal developmental range. This child does not necessarily have the same development as a child with the same score. It only means that both children are developing adequately. The importance of this fact is that once you understand the ASQ-BR as a screening tool, you use it more carefully. Conclusions about any evaluated person should be made using several instruments, not simply one [56]. The purpose of this study was to analyze whether the ASQ-BR is an adequate instrument in the educational context. The Construct Map allowed us to understand that persons tend to concentrate at the top of the scale; thus, it should be used only for screening purposes and not to separate children.

The present work also addressed the issue of gender differences in early development. Overall, ASQ-BR scores in girls tended to be higher than in boys in all of the domains, with the exception of the gross motor skills domain, in which boys tended to present higher scores than girls. These results are consistent with other studies that found the same gender differences using the Norwegian [29] and Dutch [17] versions of the ASQ. Different developmental assessment instruments have also indicated that preschool girls from different cultural regions have better developmental scores outside gross motor skills [57–60], including Brazilian preschool children [61]. The fact that the ASQ-BR was able to detect gender differences in developmental milestones at early ages confers psychometric value to these questionnaires.

Once the domain scales were linked-equated in a common scale, calibrating and studying items across both age intervals and persons were possible. The procedure revealed an average theta for boys and girls. We compared differences across age intervals. As expected, the difference between genders remained. The interactions revealed by the ANOVA were interesting. The results of the ANOVA corroborated

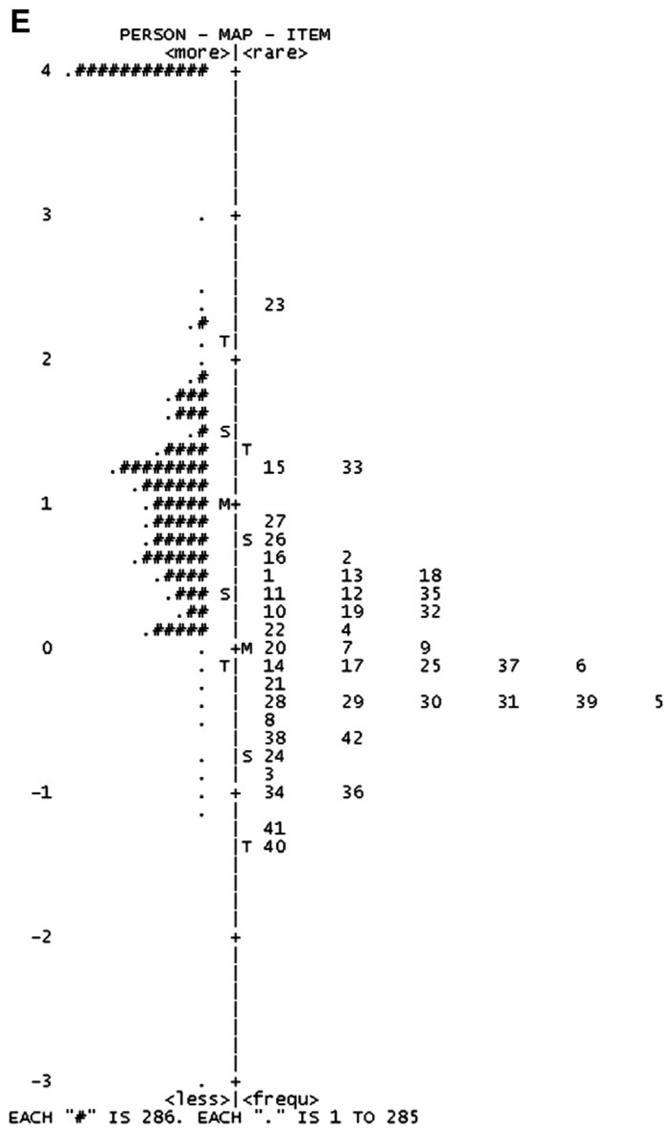

Fig. 2 (*continued*).

the *t*-test results and indicated that the gender differences were age-dependent. Differences in the gross motor domain, for example, did not appear until 14 months of age and tended to disappear around 4 years of age. This likely occurred because gender differences in gross motor skills are more related to movement skills than physical fitness; therefore, the children's experiences during development impact gross motor skills [58,59,62]. According to our empirical findings with regard to the ASQ-BR results, fine motor skills and problem solving skills also have different pathways during development. These differences are well-known in the literature and likely are related to the way that different genders explore their world through experiences with parents, family, school, and other children [60,63]. With regard to the problem solving domain, no interaction was found in the ANOVA. This may provide evidence of parallel development of the skill [63] (i.e., the ability continuously grows in both genders in two parallel paths). The statistical difference remained the same across age intervals but was not influenced by gender. With regard to gender, differences in personal–social skills emerge in Brazilian children around 20 months of age. The development of social skills in boys and girls is different and probably go through paths with few intersections since preschool [64]. The ANOVA results in the present study are important to understand how we deal with the development of infants, toddlers, and preschoolers with regard to gender differences.

Ii is important to note that the present study had several limitations that can be explored in future research. For example, test–retest analysis would be important to evaluate the consistency of ASQ-BR scores over time. The convergent validity of the ASQ-BR could also have been investigated if the study used another early developmental screening instrument, such as the Denver Developmental Screening Test, or the Bayley Infant Neurodevelopment Screener which have been validated for the Brazilian population [61,65]. Moreover, the sensitivity and specificity of each of the ASQ-BR questionnaires could be evaluated with the presence of a clinical sample of children with developmental disabilities. Finally, although the present study employed an exceptionally large number of subjects, they were drawn from child daycare centers with low-income families who live in Rio de Janeiro. This is an important issue because the geographically narrow region, which probably restrained the cultural diversity of our sample, might limit the generalizability of the present findings.

Acknowledging these limitations, we may conclude that the ASQ-BR is a brief and low-cost method that can be employed cost-effectively to gather information from a large number of children enrolled in public daycare centers during a short period of time. Cultural and linguistic adaptation allowed caregivers to complete the questionnaires without major difficulty. Classical Test Theory and RS analysis from the Rasch model family indicated that the ASQ-BR has suitable psychometric properties for assessing early development. ASQ-BR results might help improve the quality of public child daycare centers and help the decision-making process with regard to political and economic governmental investments. ASQ-BR results might also help the child screening process for possible developmental delays as early as possible so that children can be referred to trained examiners for further evaluation and different forms of interventions can be planned in advance.

Conflict of interest

None.

Acknowledgments

The present work was only possible because of the invaluable support of many people. We are deeply grateful to Ricardo Paes e Barros, Mirela Carvalho, Rosana Mendonça, Marcia Gil, Sheila Najberg, Daniel Santos, Eduardo Pádua, and Jane Squires.

References

[1] Hannon P. Developmental neuroscience: implications for early childhood intervention and education. Curr Pediatr 2003;13:58–63.
[2] Knudsen E. Sensitive periods in the development of the brain and behavior. J Cogn Neurosci 2004;16:1412–25.
[3] Pascual-Leone A, Amedi A, Fregni F, Merabet L. The plastic human brain cortex. Annu Rev Neurosci 2005;28:377–401.
[4] Shonkoff JP. Protecting brains, not simply stimulating minds. Science 2011;333:982–3.
[5] Doyle O, Harmon CP, Heckman JJ, Tremblay RE. Investing in early human development: timing and economic efficiency. Econ Hum Biol 2009;7:1–6.
[6] Heckman JJ. Skill formation and the economics of investing in disadvantaged children. Science 2006;312:1900–2.
[7] Allen SF. Assessing the development of young children in child care: a survey of formal assessment practices in one state. Early Childhood Educ J 2007;34:455–65.
[8] Gleason MM, Zeanah CH, Dickstein S. Recognizing young children in need of mental health assessment: development and preliminary validity of the Early Childhood Screening Assessment. Infant Mental Health J 2010;31:335–57.
[9] Maternal Child Health Bureau. The national survey of children with special health care needs: chartbook 2005–2006. http://www.uconnucedd.org/actearlyct/PDFs/National.pdf. [accessed August 4, 2011].
[10] Bayley N. Bayley Scales of infant development. New York: Psychological Corporation; 1969.
[11] Rydz D, Shevell MI, Majnemer A, Oskoui M. Developmental screening. J Child Neurol 2005;20:4–21.
[12] Bricker DD, Squires J. Ages and Stages Questionnaires: a parent-completed, child-monitoring system. 2nd ed. Baltimore: Paul H. Brookes Publishing; 1999.
[13] Squires J, Bricker DD, Twonbly E, Potter L. Ages & Stages Questionnaires: a parent-completed child monitoring system. 3rd ed. Baltimore: Paul H. Brookes Publishing; 2009.

[14] Handal AJ, Lozoff B, Breilh J, Harlow SD. Effect of community of residence on neurobehavioral development in infants and young children in a flower-growing region of Ecuador. Environ Health Perspect 2007;115:128–33.
[15] Hix-Small H, Marks K, Squires J, Nickel R. Impact of implementing developmental screening at 12 and 24 months in a pediatric practice. Pediatrics 2007;120:381–9.
[16] Hutchison BL, Stewart AW, Mitchell EA. Characteristics, head shape measurements and developmental delay in 287 consecutive infants attending a plagiocephaly clinic. Acta Paediatr 2009;98:1494–9.
[17] Kerstjens JM, Bos AF, ten Vergert EMJ, de Meer G, Butcher PR, Reijneveld SA. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. Early Hum Dev 2009;85:443–7.
[18] Limbos MM, Joyce DP. Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. J Dev Behav Pediatr 2011;32: 499–511.
[19] Msall ME. Developing preschool surveillance tools for adaptive functioning: lessons for neuro-oncology. Eur J Paediatr Neurol 2010;14:368–79.
[20] O'Leary C, Zubrick SR, Taylor CL, Dixon G, Bower C. Prenatal alcohol exposure and language delay in 2-year-old children: the importance of dose and timing on risk. Pediatrics 2009;123:547–54.
[21] Osofsky JD, Kronenberg M, Hammer JH, Lederman JC, Katz L, Adams S, et al. The development and evaluation of the intervention model for the Florida Infant Mental Health Pilot Program. Infant Mental Health J 2007;28:259–80.
[22] Beeber LS, Chazan-Cohen R, Squires J, Harden BJ, Boris NW, Scott-Heller S, et al. The Early Promotion and Intervention Research Consortium (E-PIRC): five approaches to improving infant/toddler mental health in Early Head Start. Infant Mental Health J 2007;28:130–50.
[23] Macy M. The evidence behind developmental screening instruments. Infant Young Child 2012;25:19–61.
[24] Thomas SA, Cotton W, Pan X, Ratliff-Schaub K. Comparison of systematic developmental surveillance with standardized developmental screening in primary care. ClinPediatr (Phila) 2012;51(2):154–9.
[25] Graça PRM, Teixeira MSC, Lopes SCG, Serra AMPH, Campos ARS. Momento da avaliação na intervenção precoce: o envolvimento da família: estudo das qualidades psicométricas do ASQ −2 dos 30 aos 60 meses. Rev Bras Educ Esp 2010;16:177–96.
[26] Ages & Stages Questionnaires, Third Edition (ASQ-3): user's guide. Spanish version. San Antonio, TX: Paul H. Brookes Publishing; 2009.
[27] Bonin M, Robaey P, Vandaele S, Bastin GL, Lacroix V. Questionnaire sur les étapes de développement: Évaluation de l'enfant par les parents. 2nd ed. San Antonio, TX: Paul H. Brookes Publishing; 2000.
[28] Janson H, Squires J. Parent-completed developmental screening in a Norwegian population sample: a comparison with US normative data. Acta Paediatr 2004;93:1525–9.
[29] Richter J, Janson H. A validation study of the Norwegian version of the Ages and Stages Questionnaires. Acta Paediatr 2007;96:748–52.
[30] Plomgaard AM, Hansen BM, Greisen G. Measuring developmental deficit in children born at gestational age less than 26 weeks using a parent-completed developmental questionnaire. Acta Paediatr 2006;95:1488–94.
[31] Tsai HLA, McClellan MM, Pratt C, Squires J. Adaptation of the 36-month Ages and Stages Questionnaire in Taiwan: results from a preliminary study. J Early Intervent 2006;28:213–25.
[32] Heo KH, Squires J, Yovanoff P. Cross-cultural adaptation of a pre-school screening instrument: comparison of Korean and US populations. J Intellect Disabil Res 2008;52:195–206.
[33] Juneja M, Mohanty M, Jain R, Ramji S. Ages and stages questionnaire as a screening tool for developmental delay in Indian children. Indian Pediatrics 2012;49(6): 457–61.
[34] Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional por Amostra de Domicílios. http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2006/pnad_sintese_2006.pdf; 2006. [accessed February 28, 2012].
[35] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334.
[36] Dukes KA. Cronbach's alpha. In: Armitage P, Colton T, editors. Encyclopedia of biostatistics. 2nd ed. Chichester: Wiley; 2005 [http://www.mrw.interscience.wiley.com/emrw/9780470011812/eob/article/b2a13018/current/html; accessed February 28, 2012].
[37] Cohen RJ, Swerdlik ME, Smith DK. Psychological testing and assessment: an introduction to tests and measurement. Mountain View, CA: Mayfield Publishing; 1992.
[38] Kline P. An easy guide to factor analysis. London: Routledge; 1994.
[39] Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976;41:321–7.
[40] Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. Org Res Methods 2004;7:191–205.
[41] O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behav Res Methods Instrum Comput 2000;32:396–402.
[42] Little TD, Slegers DW. Factor analysis: multiple groups. In: Everitt BS, Howell DC, editors. Encyclopedia of statistics in behavioural science. Chichester: John Wiley & Sons; 2005. p. 617–23.
[43] Milfont TL, Fisher R. Testing measurement invariance across groups: applications in cross-cultural research. Int J Psychol Res 2010;3(1):111–21.
[44] Rasch G. On general laws and the meaning of measurement in psychology. In: Neyman J, editor. Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. Berkeley: University of California Press; 1961. p. 321–44.
[45] Andrich D. A rating scale formulation for ordered response categories. Psychometrika 1978;43:561–73.
[46] Linacre JM. Comparing "Partial Credit Models" (PCM) and "Rating Scale Models" (RSM). Rasch Meas Trans 2000;14(3):768.
[47] Masters GN. A Rasch model for partial credit scoring. Psychometrika 1982;47: 149–74.
[48] Linacre JM. A user's guide to Winsteps: Rasch-model computer programs. www.winsteps.com; 2010. [accessed February 28, 2012].
[49] Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum; 2001.
[50] Wright BD. Equitable Test Equating. Rasch Measurement Transactions 1993;7(2): 298–9.
[51] Lee OK, Wright BD. Mathematics and Reading Test Equating. Annual meeting of the American Educational Research Association: San Francisco; 1992.
[52] Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate data analysis. 5th ed. New York: Prentice Hall; 2007.
[53] Branson D. The effects of using the ASQ on childcare providers' developmental assessment abilities. Reno: University of Nevada; 2009.
[54] Squires J, Bricker DD, Twombly E. Ages and Stages Questionnaires, social–emotional. Baltimore: Paul H. Brookes Publishing; 2003.
[55] Wright BD, Stone MH. Best test design. Chicago: Mesa Press; 1979.
[56] Camara WJ, Nathan JS, Puente AE. Psychological test usage: implications in professional psychology. Prof Psychol Res Pract 2000;31(2):141–54.
[57] Hyde JS, Linn MC. Gender differences in verbal ability: a meta-analysis. Psychol Bull 1988;104:53–69.
[58] Lung FW, Chiang TL, Lin SJ, Feng JY, Chen PF, Shu BC. Gender differences of children's developmental trajectory from 6 to 60 months in the Taiwan Birth Cohort Pilot Study. Res Dev Disabil 2011;32:100–6.
[59] Thomas JR, French KE. Gender differences across age in motor performance: a meta-analysis. Psychol Bull 1985;98:260–82.
[60] Walker S, Irving K, Berthelsen D. Gender influences on preschool children's social problem-solving strategies. J Genet Psychol 2002;163:197–209.
[61] Drachler ML, Marshall T, de Carvalho-Leite JC. A continuous-scale measure of child development for population-based epidemiological surveys: a preliminary study using Item Response Theory for the Denver Test. Paediatr Perinat Epidemiol 2007;21:138–53.
[62] Crombie G. Gender differences: implications for social skills assessment and training. J Clin Child Psychol 1988;17(2):116–20.
[63] Halpern DF. Sex differences in intelligence: implications for education. Am Psychol 2011;52(10):1091–102.
[64] Ikeda T, Aoyagi O. Relationship between gender difference in motor performance and age, movement skills and physical fitness among 3- to 6-year-old Japanese children based on effect size calculated by meta-analysis. School Health 2009;5: 9–23.
[65] Guedes DZ, Primi R, Kopelman BI. BINS validation — Bayley neurodevelopmental screener in Brazilian preterm children under risk conditions. Infant Behav Dev 2011;34(1):126–35.