Infant Behavior & Development 34 (2011) 126–135

Contents lists available at ScienceDirect

Infant Behavior and Development

BINS validation – Bayley neurodevelopmental screener in Brazilian preterm children under risk conditions

Deborah Z. Guedes[a,*], Ricardo Primi[b], Benjamin I. Kopelman[a,*]

[a] Department of Pediatrics – Universidade Federal de São Paulo (UNIFESP), Brazil
[b] Department of Psychology – Universidade São Francisco (USF), Brazil

ARTICLE INFO

Article history:
Received 17 March 2009
Received in revised form 15 January 2010
Accepted 18 November 2010

Keywords:
Psychological assessment
Screening
Validation
Child Development
Preterm infant
At-risk child

ABSTRACT

Psychometric researches increase in Brazil. Bayley Infant Neurodevelopment Screener – BINS (Aylward, 1995) is a low cost, fast instrument. In 10' it classifies children under developmental risk degrees. This research purpose was investigating BINS psychometric properties. 61 low-income Brazilian preterm, were divided in groups: 31 children (12 months) and 30 children (24 months), both sex, birth weight <2000 g. Socio-demographic-psychological profile was previously registered. Neurologists examined them through Amiel-Tison and Gosselin (2001) and physicians with Denver-DDST-II (Frankenburg, Dodds, Archer, & Bresnick, 1990). Psychologist assessed children at chronological age, with Bayley Scales–BSID-II (Bayley, 1993) and BINS (12 m) and BINS (24 m). Results demonstrated homogeneous characteristics sample. Reliability indexes were over requested standards. Validity evidences based on external variables were positive moderated and BINS (24 m)/BSID-II (mental) presented high correlation. Validity evidences based on content were attested by expertise. High sensitivity was found. So, BINS can be considered an instrument of adequate psychometric properties, able to screen children under risk, according to Psychological Association requests.

Currently, Developmental Psychology Brazilian university groups are researching instruments administered to obtain effective and quality measurements. They must follow the Standards for Educational and Psychological Testing, the American Psychological Association (APA, 1985), the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME) fulfilling criteria for Reliability, Validity, Sensitivity and Specificity (Noronha, Vendramini, Souza, Franco, & Filizatti, 2003; Pasquali, 2001; Urbina, 2007).

It is essential that every test while submitted to a fully documented validity process, go through rigorous expertise translation and back-translation when adapting measures for different languages and cultures (Beaton, Bombardier, Guillemin, & Ferraz, 2000; Pasquali, 2000; Pedromônico, 2003).

Early detection of developmental delays helps prevent developmental, behavioral, and emotional risks or delays and disabilities, before they become lifelong problems. Therefore detection and intervention gives every child a chance to succeed (Rydz, Shevell, Majnemer, & Oskoui, 2005).

Interest in children assessment tools, named “Baby Tests (Anastasi, 1967; Bee, 2003; Nunes, Sisdelli, & Fernandes, 1995) started in the 1920s in USA (Fig. 1) and in the 1980s and 1990s scientific researches increased, specially regarding child

* Corresponding authors at: UNIFESP – Universidade Federal de São Paulo, Neonatal Pediatric Department, R Diogo de Faria, 764, São Paulo, SP 04037-002, Brazil. Tel.: +55 34 9994 1997.
E-mail address: dzguedes@yahoo.com (D.Z. Guedes).

0163-6383/
doi:10.1016/j.infbeh.2010.11.001

PUBLICATION	INSTRUMENT	AUTHOR	KIND	ITEMS	AGE	TIME
1925	**Developmental Scales**	Gesell	Development assessment	144	4 weeks - 42 months	–
1930	**An Instrument For Assessing Infant Psychological Development**	Uzgiris, Hunt	Development assessment	126	0 - 24 months	–
1933	**The California First-year Mental Scale**	Bayley	Development assessment	-	2 - 30 months	1 hour - 1:30 hour
1966	**A developmental Screening inventory for infants**	Knobloch et al	Screening	-	0 – 30 months	–
1967	**Denver Development Screening Test (DDST)**	Frankenburg, Dodds	Screening	125 items	0 – 6 yo	20'/30'
1969	**Bayley Scales of Infant Development (BSID)**	Bayley	Development assessment	Mental 163 Motor 81 + Behavior	0 a 42 months	1:00 - 1:30 hour
1981	**L'échelle De Développement Psychomoteur De La Première Enfance**	Brunet-Lézine	Development assessment	10 items by age	1 month – 5 yo	30' – 1:00 hour
1984	**Vineland Adaptive Behavior Scales**	Sparrow, Balla, Cicchetti	Screening	577 items to individual parent/caregiver interview	0 – 18 yo	20' a 90'
1988	**Battelle Developmental Inventory Screening**	Newborg et al	Screening	100	0 - 7/11 yo	30'
1992	**Alberta infant motor scale (AIMS)**	Piper et al	Screening	58	0 - 18 months	20'
1993	**Bayley Scales of Infant Development (BSID-II)**	Bayley (APA)	Development assessment	Mental 178 Motor 111 Behavior 30	0 - 42 months	1 hour - 1:30 hour
1995	**Bayley Infant Neurodevelopmental Screening – BINS**	**Aylward**	**Screening**	**11 to 13 items to each age**	**3 – 24 months**	**15'**
1997	**Escala de Desenvolvim. Comportamento da Criança: O 1° Ano de Vida**	Batista, Vilanova, Vieira,	Development assessment	64	0 - 12 months	30' – 1:00 hour
2005	**Bayley Scales of Infant and Toddler Development -Screening (Bayley – III)**	Bayley (APA)	Screening	Nearly 100 items	0 - 42 months	15' to 25'

Fig. 1. Development assessment instruments for diagnostic and screening. Refs. Batista, Vilanova, and Vieira (1997), Bayley (1933, 1969, 2005), Brunet and Lèzine (1981), Frankenburg & Dodds (1967), Gesell (1925), Knobloch, Pasamanick, and Sherard (1966), Newborg, Stock, Wnek, Guidubaldi, and Svinicki (1988), Piper, Pinnell, Darrah, Maguire, and Byrne (1992), Sparrow, Balla, and Cicchetti (1984) and Uzgiris and Hunt (1966).

developmental screening, answering American government needs of detecting children with developmental and behavioral delays, due to the Public Law 99-457 "The education of the Handicapped Act Amendts" and the health program "EPSDT – Early and Periodic Screening Diagnosis and Treatment".

Tests for newborns, infants or young children present peculiarities, which are not present when assessing adults. They privilege developmental aspects of behavior, through multiple items evaluating milestones in each stage of child development (Duarte & Bordin, 2000; Theuer & Mendonza, 2003). They also establish quantitative scores for the evaluated aspects and compare groups through them. The test can determine the best comprehension of the item by the child and insure secure investigation desired by the professional, in favor of assessment efficiency (Alchieri & Cruz, 2003).

Screening and assessment of children are a continuous process, but independent procedures support the final child evaluation (Pedromônico, 2003). Developmental assessment gives a deep comprehension of the presence and extension of problems, identifying specific abilities and determining appropriate strategies for intervention, while screening tracks children who appear to be in a situation of potential problems, out of the standards of reference. Screening tests capture initial glimpses of children in need of accurate and fast intervention and flags those who need further assessment (Bellis & Burke, 1996; Urbina, 2007). It is highly recommended that children screened at risk, be referred to intervention, an appropriate source for follow up or specific assessment to improve their developmental path (Aylward, 2005).

Children from underdevelopment countries face multiple challenges, fighting infant mortality daily and this can favor serious health consequences in later life. They are subjected to extremely low birth weight, premature birth, no adequate stimulation, chaotic social situation, environmental violence, economic poverty, lack of family bonding and support, etc. Researches demonstrate association between child neurodevelopmental delay and degree of environmental and biological adversities (Grantham McGregor, Powel, Stewart, & Schofield, 1982).

Due to the vast amount of adversities in Brazilian children lives and the lack of valid, adapted instruments, we came across Bayley Infant Neurodevelopmental Screening – BINS (Aylward, 1995), designed to screen infants at risk for developmental delays or neurodevelopmental problems (Fig. 1).

It assesses cognitive processes (e.g., memory and problem solving), receptive functions (e.g., visual and auditory input), expressive functions (e.g., oral and motors skills) and basic neurological functions/intactness (e.g., muscle tone and head control).

It is an instrument quick to administer by a trained professional. In about 10 or 15 min, the sum of scores can possibly point out neurological impairments or place the infant (3–24 months) in low, moderate, or high risk category for developmental delay (Hess, Papas, & Black, 2004; Ergenekon et al., 2007).

BINS test–retest reliability ranged from .71 to .81, and internal consistency was reported to be moderate to strong. Coefficient alphas ranged from .73 to .85 across age. The inter-rater reliability was also established and ranged from .79 to .96. Construct validity was related to indices of severity of the sample medical problems. Criterion validity in a high risk infant population was established by comparing BINS score to those obtained in the BSID-II. Sensitivity and specificity were calculated to be 64% and 87% respectively (Aylward, 1995; Bess & Humes, 1998; Naar-King, Ellis, & Frey, 2004).

The objective of this research was to determine BINS psychometric properties (Validity, Reliability, Sensitivity and Specificity) and examine it as a screening technique for Brazilian at-risk children.

1. Methods

1.1. Participants

BINS was administered to 61 preterm infants, from low-income families and all were Brazilian unified health system users. They were divided in 2 groups: 31 children up to 12 months (12 m) and 30 children up to 24 months (24 m), both sexes, bellow 2000 g at birth weight. There were no restrictions to race/ethnicity categories, however syndromic, impaired children, twins or infants with any sensory disorders were taken out of this sample. Children assessed at 12 months were not re-assessed at 24, in this research.

1.2. Procedures

Children came to the follow up program with caretakers, for their monthly routine appointment with the doctor, and were invited to take part in this research. Almost 95% of the caregivers gave consent. After that, an interview about the family socio-demographic profile was performed. A pediatric neurologist applied a neurodevelopmental examination on the children (Amiel-Tison & Gosselin, 2001) and a pediatrician applied the Denver Developmental Screening Test II – DDST-II (Frankenburg, Dodds, Archer, & Bresnick, 1990) at the pediatric routine visit. The psychologist assessed children through Bayley Scales of Infant Development, 2nd ed. – BSID-II (Bayley, 1993), considered a golden standard instrument and Bayley Infant Neurodevelopmental Screener – BINS (Aylward, 1995) for the 12 and the 24 month old children.

BINS translation, adaptation and validation to the Portuguese language (Brazilian version) were carried out in the city of São Paulo, which hosts a huge cultural diversity of inhabitants, originated from many regions of the country. We followed these steps: translation, back-translation, correction and semantic adaptation, content validation by professional experts (judges) and a final critical assessment in 10 children from the target population in a tryout experiment version. These children did not take part on the final sample of this research (Beaton et al., 2000; Pasquali, 2000).

2. Results

Sociodemographic aspects (Table 1) and children birth risk conditions (Table 2) presented homogeneous characteristics in the 2 groups. Every child from 0 to 6 years old, who presents developmental delay, physical impairment or emotional-mental disorder are eligible for the Early Intervention Program in Brazil. From the entire group (61 children) who were under screening, 54 children were eligible for the early intervention program: 30 infants at 12 months and 24 children at 24 months. Children were referred to specialists (developmental pediatrician, pediatric neurologist, psychiatrist, optometrist, speech pathologists or psychologist) for further evaluation, diagnostic testing and for continuation of early intervention in the Brazilian unified health system.

Table 1
Sociodemographic characteristics.

	12 months	24 months
Age (month)	12.37	23.88
Gender		
Male	45.16%	50%
Female	54.83%	50%
Race/ethnic groups		
White	63%	65%
Black	31%	30%
Others	6%	5%
Family monthly income (minimum wage)	2.20	1.98
Maternal education (years of schooling)	7.52	7.51

Table 2
Children birth risk conditions.

	12 month (31 children)	24 month (30 children)
Mean birthg weight (g)	1 461 g	1 433 g
Birth condition	14 SGA	5 SGA
	17 AGA	25 AGA
Mean hospitalization time (days)	36.25	38.16
Medical complications		
Respiratory distress syndr.	25.8%	26.6%
Intraventricular hemorrage	25.8%	16.6%
Apnea	12.9%	6.6%

2.1. Reliability

To guarantee BINS internal consistency and precision, two procedures, Cronbach's Alpha (Cronbach, 1951) and Rasch Model were performed (Rasch, 1960). The precision index at 12 months was 0.64 (Cronbach's Alpha 0.67). At 24 months the precision index was 0.72 (Cronbach's Alpha 0.76). The internal consistency of BINS fulfills the criteria requested by the Brazilian Psychological Association for instruments Reliability (index over 0.60) and guarantees its reproductibility and reliability.

Regarding the Rasch model, the babies (24 m) presented better performance of the requested items (0.23 above the average, centered on zero). Data pointed that there were no difficulties to answer items in both groups, nor were there any disagreements among items (Table 3).

High children performance at 24 months can be credited to the "catching up" phenomena. Scientific developmental literature says that child normal evolution process places them in a balanced position able to be compared to other children of the same age, accomplishing tasks and keeping standards as their pairs (Chaudari, Kulkarni, Pajnigar, Pandit, & Deshmukh, 1991; Isotani, Pedromônico, Perissinoto, & Kopelman, 2002; Oliveira, Lima, & Goncalves, 2003).

2.2. Validity

For the *Validity Evidence Based on Internal Structure*, the items location were identified through the Rasch model maps, calibrating simultaneously BINS and BSID-II items. The higher (and more positive) the left side scores are, the more difficult is to answer the item in relation to the other items of the test. The "x" represents children. On the right side, there are BINS and BSID-II items. The map position indicates the performance difficult level, on the test.

Item BINS-3 (12 m) demonstrated to be very difficult and no child performed it. While items were located between the −2 and 2 position, individuals centered between 1/−1. (Fig. 2) Item BINS-13 (24 m) seems to be the easiest. Items are located bellow the center measurement "0", regarding their difficulty and they are concentrated as the difficulties −2 and 0 (Fig. 3).

Table 3
Children's ability to answer items at BINS (12 months) and BINS (24 months).

	BINS (12m)			BINS (24m)		
	Measure	Infit	Outfit	Measure	Infit	Outfit
Mean	**0.17**	0.97	1.09	**0.23**	0.98	1.02
S Deviation	1.46	0.33	1.04	1.40	0.20	0.64
Maximum	2.94	1.67	5.82	2.74	1.42	3.96
Minimun	−3.01	0.50	0.19	−2.82	0.67	0.25

Measure tells about the children's ability to answering the item;.
Infit and outfit are adjustment items. The ideal value is ≤1.2; ≥1.5 (moderate desadjustment); ≥2.0 (high desadjustment).
Maximum/minimum are indication of ability's level, from the lower to the higher ability.

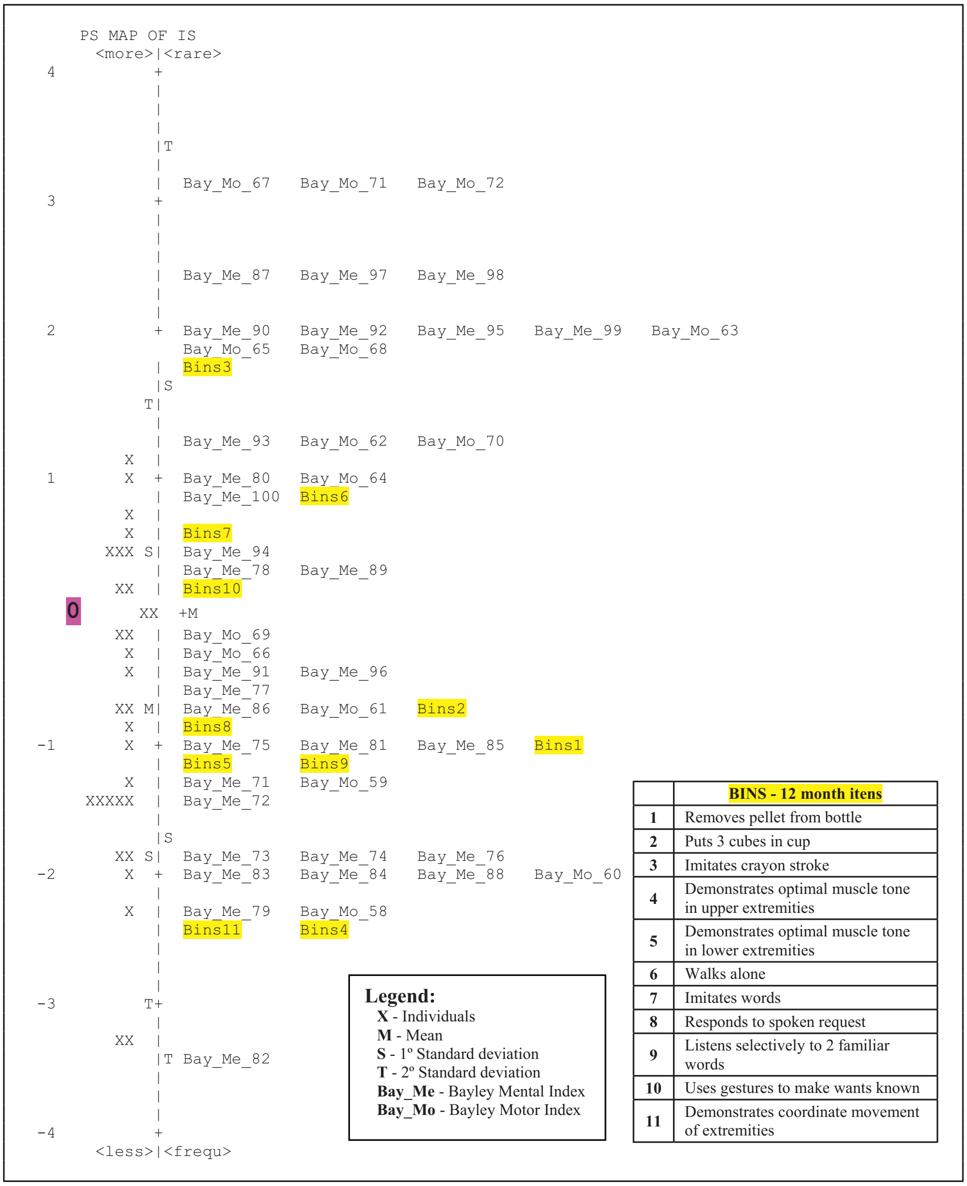

Fig. 2. BINS (12 months) X BSID-II map of items.

Winstep program provides the classical "item-total" correlation, verifying the discriminative power of each item (Linacre & Wright, 1991). At both ages 12 and 24 months, values are adequate.

BINS items (Table 4) showed the mean square (*Mnsq*) presented in the *infit* format (level of difficulty of the item) and *outfit* format (score variation). If the relation response/expected model is according to the expected, *Mnsq* must be $\leq$1.2. (Linacre & Wright, 1991; Linacre, 2002) 1.4 it is an erratic score item, in which individuals with high ability used to receive low score or high score for low abilities. (Fisher, 1993; Bond & Fox, 2001) Values under 0.7 are predictable very consistent items and they have a short score variation, putting in jeopardy the test validation, because predictable scores are not clear about the individual performance. Erratic score items present a higher impact than the predictable items.

Also in Table 4, BINS-3 (12 m) *Imitates crayon stroke* and BINS-13 (24 m) *Absence of drooling and motor overflow* were over 1.4, indicating misadjusted items to the expected model, due to its easiness/difficulty. This means they are not adjusted to the test, due to item announcement difference in the original and in the final version form or they are not appropriated to Brazilian children at this age. The essence of the test may be altered if cultural adaptation is not done.

The *Content Validity Evidences* were met in the translation, back-translation and comparison of the translated instrument with its original form, in the evaluation of the semantic equivalence and in the final version best chosen by the expertise. BINS 10 (12 m) – *Use gestures to make wants known* was considered different from the original test. In Portuguese, it turned out

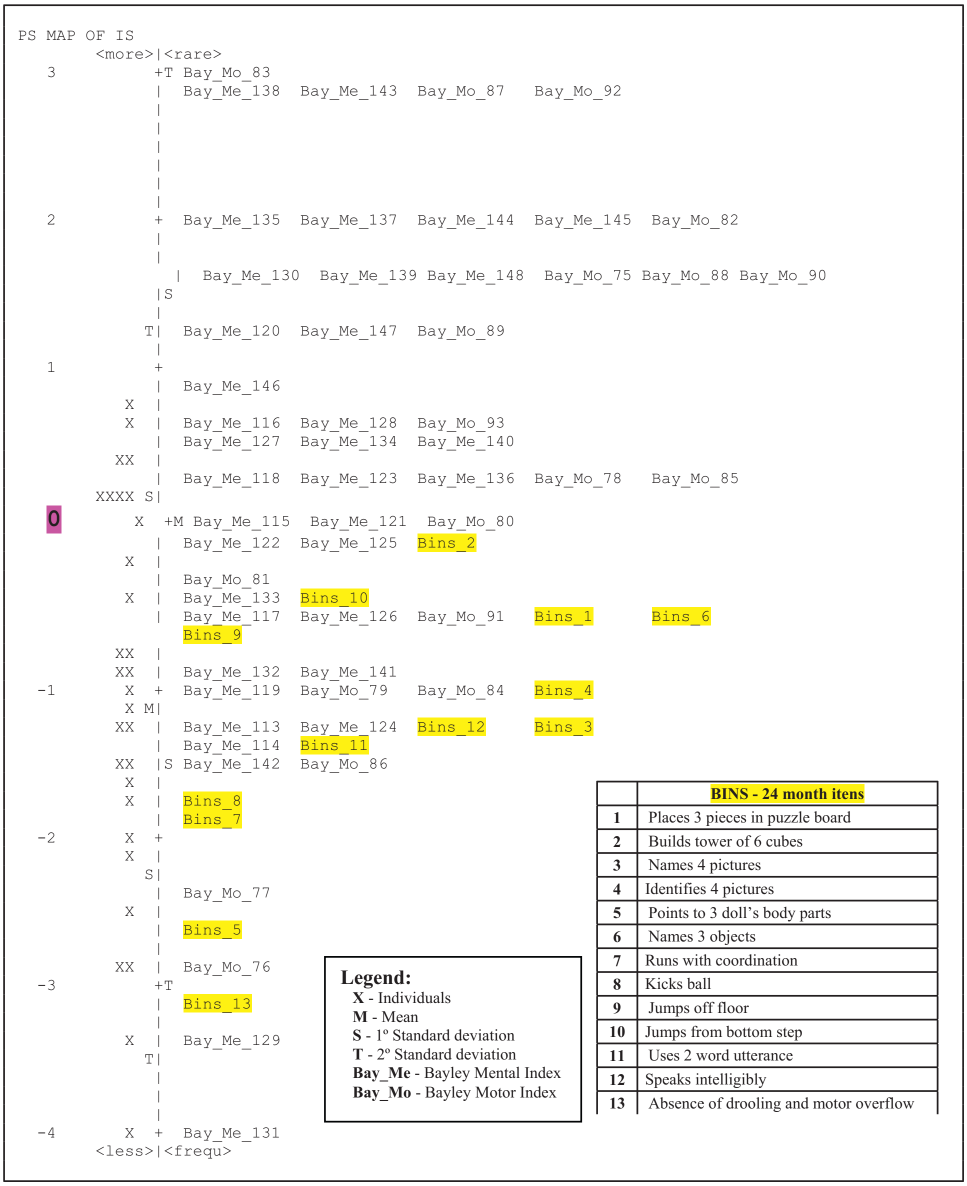

Fig. 3. BINS (24 months) X BSID-II map of items.

to be similar to "Make gestures to communicate yourself". It is understood that a gesture, antecessor of verbal competence, is a highly complex expression of wish, which assumes communication. If a child uses gesture and somebody understands what the intention of the baby is, the child was able communicate by eliciting its wish (Guedes, Pedromonico, Goulart, & Gomes, 2005).

The *Evidences based in Convergent and Discriminant Validation* were made through Pearson's correlation coefficient, between BINS (12 m) and BINS (24 m) and external variables: Denver Developmental Screening Test II – DDST-II (Frankenburg et al., 1990) and Bayley Scales of Infant Development – BSID II (Bayley, 1993) in chronological form, in the mental and motor indexes. Neurological assessment (Amiel-Tison & Gosselin, 2001) was correlated to BINS, by evidences based in the relation with external variables and the test-criteria.

Correlation between BINS (12 m), BINS (24 m) and neurological examination was positive and moderate, however Amiel-Tison and Gosselin (2001) neurological aspects are not as complex as in BINS, which focus spreads to mental, motor and neurological aspects. If child examination does not have neurological and development interaction, indexes of agreement tend to be lower (Aylward, 2005).

Moderate positive meaningful correlations were identified between BINS (12 m) and variables, suggesting they go in the same direction (Table 5). Moderate positive meaningful correlation between BINS (24 m) and Neurological assessment,

Table 4
(Mis) adjusted items at BINS (12 months) and BINS (24 months).

BINS (12 m)					BINS (24 m)				
Item	Measure	Infit	Outfit	Ptmea corr.	Item	Measure	Infit	Outfit	Ptmea corr.
4	2.05	0.95	1.13	0.51	**2**	1.22	1.20	1.32	0.34
11	2.05	0.94	0.74	0.56	**10**	0.84	1.27	1.20	0.34
5	0.65	0.81	0.70	0.64	**1**	0.65	1.14	0.99	0.44
9	0.65	0.88	0.85	0.59	**6**	0.65	1.12	1.12	0.43
1	0.45	1.25	1.37	0.37	**9**	0.65	0.73	0.61	0.67
8	0.27	0.88	0.73	0.60	**4**	0.29	0.79	0.84	0.63
2	0.09	1.00	1.29	0.48	**3**	0.10	0.94	0.97	0.55
10	−0.83	0.97	0.83	0.48	**12**	0.10	0.79	0.65	0.66
7	−1.23	0.91	0.79	0.48	**11**	−9.08	1.15	0.99	0.46
6	−1.46	0.90	0.68	0.48	**8**	−0.47	0.92	0.96	0.57
3	−2.68	**1.44**	2.88	**0.03**	**7**	−0.67	0.75	0.60	0.68
					5	−1.36	0.90	0.83	0.56
					13	−1.93	**1.40**	2.21	**0.23**

Measure tells about the difficulty of the item.
Infit and outfit are adjustment items. The ideal value is ≤1.2; ≥1.5 (moderate misadjustment); ≥2.0 (high misadjustment).
Ptmea corr. tells us the association between the item and the entirely test. The ideal value is ≥0.30.

DDST-II and BSID II (motor index) was also identified. BINS (24 m) and BSID-II (mental index) present high meaningful correlation (Table 5). Data suggest similar children classification tendency in both, indicating the presence of the same construct, even tough they are instruments for different purposes. BINS tracks children with developmental delays, while BSID-II diagnoses/classifies global or specific delays.

A high degree of agreement between the screening instruments BINS and Denver was expected. Nevertheless, Denver only measures mental and motor functions of the child and there are severe critics about its sensitivity, being that it sometimes underestimates or superstimates children (Applebaum, 1978; Meisels, 1989; Glascoe et al., 1992).

No meaningful differences were found between the 2 groups using Student's *t*-test.

2.3. Sensitivity and specificity

The diagnostic performance or the accuracy of a test to discriminate diseased cases from normal cases is evaluated using Receiver Operating Characteristic (ROC) curve analysis (Metz, 1978; Zweig & Campbell, 1993). A ROC curve analysis was undertaken to determine *sensitivity* and evaluate BINS *specificity*. The scores were compared with results of the golden standard instrument, BSID-II. Index scoring was dichotomized as: inadequate performance (<85) or adequate performance (≥85).

Statistics presented high sensitivity (0.824) between BINS (12 m) and BSID-II (mental). Concerning similarities between BINS (12 m) and BSID-II (motor), the 0.646 index is considered reasonable, according to Brazilian Psychological instrument measurement requirements. Higher sensitivity index was found at 24 months: 0.64 for BINS (24 m) and BSID-II (mental) and 0.785 for BINS e BSID-II (motor). These indexes show BINS high sensitivity in detecting children for development risk. The ROC curve drawing, more pronounced at northwest (Fig. 4) demonstrated BINS sensitivity and specificity, in the discrimination of different groups. Sensitivity was even higher than specificity due to BINS being a screening test.

The adoption of the cut score ≤5.5 for both ages (12 and 24 months) can efficiently classify the child at risk for adequate development. Competence in test administration and its accurate cut score is necessary to reduce costs in public management. If one intends to check for risk of delays in public hospitals, it is important to have a high sensitivity instrument. It is

Table 5
Correlation among BINS (12 m)/BINS (24 m) and external variables.

		BINS (12 m)	BINS (24 m)
Amiel-Tison	Pearson	0.36*	**0.35**
	Sig.	0.04	0.05
Denver (DDST-II)	Pearson	**0.62****	**0.59****
	Sig.	0.00	0.00
Bayley Me (BSID-II Me)	Pearson	**0.62****	**0.62****
	Sig.	0.00	0.00
Bayley Mo (BSID-II Mo)	Pearson	**0.36***	**0.62****
	Sig.	0.04	0.00
BINS (12 m)	Pearson	**1**	–
	Sig.		
BINS (24 m)	Pearson	–	**1**
	Sig.		

BINS (12 m): *Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).
BINS (24 m): *Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

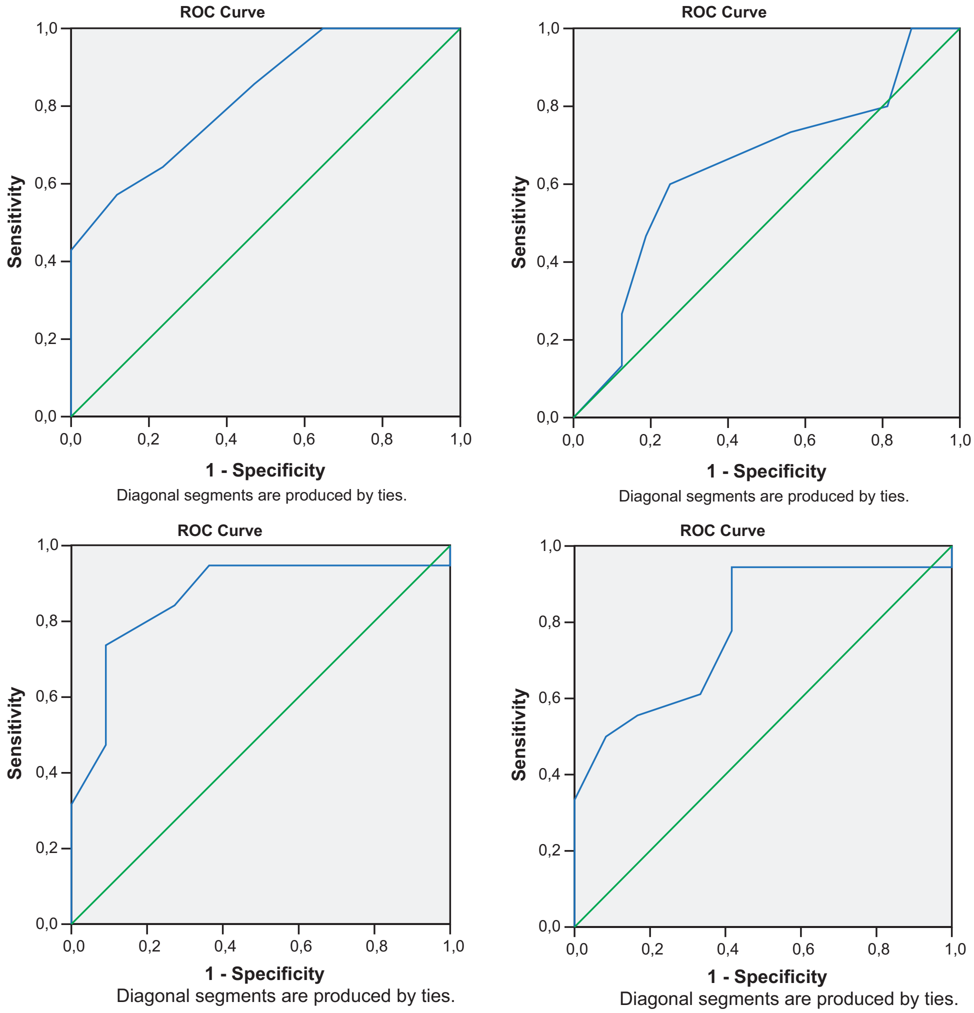

Fig. 4. ROC Curve – BINS X BSID-II (Mental and Motor index) at12 and 24 months.

better to have false positive children identified, who could overcome problems and change their developmental perspective in the future, than taking the risk of skipping them.

3. Discussion

The risk concept definition (Horowitz, 1992) comprehends that children face biological, environmental and psychiatric risks, when development does not happen in the way it was expected to in a specific population (Hutz, Koller, Bandeira, & Forster, 1995; Hutz & Silva, 2002).

Preterm babies are vulnerable children at risk, due to their cumulative risk factors exposure (medical, environmental, social, economic, etc.), which can harm their growing and developmental process (Barratt, Roach, & Leavitt, 1996; Hack & Costello, 2007; Halpern, Giugliani, Victora, Barros, & Horta, 2000; Leonard, Piecuch, & Cooper, 2001; Meisels & Wasik, 1990; Sameroff, 1986, 1990; Stjernqvist & Svenningsen, 1999) and also may place them at risk for motor developmental delays and compromised cognitive skills (Leonard et al., 2001) or lead them to high mental retard index, emotional problems, behavioral disturbances and disabilities (Brooks Gunn, 1990; Grantham McGregor et al., 1982; McCornick, 1989; Vohr, 1991).

A screening test should capture children, who need closer attention. However, any screening test can make classification errors, both in underidentifying or overidentifying children (Colombo, 1993; Fagan & Singer, 1983; Leonard et al., 2001; McCall, Hogarty, & Hurlburt, 1972). Qualitative transitory aspects of organic changes in development mark childhood. Due to simultaneous discontinuities in child development and instability of individual differences, specific in each child stage, some pathologies need to be reviewed through other proposals, combining subjectivity and neuroplasticity. In this way, scores of baby tests can also be diffused.

Therefore, developmental scales or screening tests cannot be a long-term predictive instrument. They can evaluate child cognitive and motor aspects, at the very moment. The professional must not elicit any linear determinist anticipation of development. Test measurement only gives indication of aspects to be observed or a hint for follow-up conduction (Leonard et al., 2001). It is essential to keep an individualized perception for every each unique child.

In this study, with BINS, items distribution at 12 months were balanced in the 4 areas assessed, and at 24 month they privileged expressive functions, due to fluency and grammar introjections of the language expected at this stage. Some item accomplishments are extremely important and if a child overcomes them, there is a small chance of not being at risk.

For BINS (12 m), the BINS-2 items: *Puts 3 cubes in cup* (Cognitive function), BINS-6 *Walks alone* (Expressive function – Gross Motor), BINS-7 *Imitates words* (Expressive function – verbal) and *BINS-8 Respond to spoken requests* (Receptive function – verbal) are examples of important item accomplishments. This last item especially indicates that a child is capable of understanding a request; processing it and reacting to it, which implicates in elaborated neural integrated connection.

For BINS (24m), the items: BINS 1 - *Puts 3 pieces in the puzzle board* (Cognitive function), BINS-2 *Builds tower of 6 cubes* (Expressive function), BINS-3 *Names 4 pictures* (Expressive function – verbal) and the BINS-6 item, *Names 3 objects* (Expressive function – verbal) are the most discriminative ones.

BINS is a test connected to development assessment, and some people understand it as an intelligence scale. The misconceptions are due to the fact that BINS presents an examination designed for babies, with items requesting non-verbal tasks, based on sensory-motor activities. However, BINS does not tell us about any intellectual aspects of the child. Besides that, the Flynn effect guarantees that cognitive abilities are always changing and increasing from time to time in every generation (Flynn, 1999, 2008) making it hard work to define intelligence constructs for babies.

Health professionals who recognize potential adversities in infant development, who undergo early screening with BINS, can improve babies long-term functional and structural development, due to strategies of early delay prevention, interfere in multiple risks intensity, reduce their intensification and promote acquisition of social protective factors (mothering, parent schooling, community partnership) in the children and their families (Masten & Garmezy, 1985; Msall, 2004; Rutter, 1985, 1987, 1993; Sameroff, 1986, 1990; Yunes, 2003).

More psychometric researches including screening tests reliability and validity could positively contribute to this study, as well as exchanging processes with other fields such as neurology, psychiatry, physiotherapy, and speech pathology.

Therefore, the BINS screener adoption by pediatric professionals justifies its practice by being a low cost and fast instrument, with high reliability and validity degree, generating long-term benefit for changing the health status of at risk children. Obtaining early and quickly BINS scores and fast intervention for the child, reduces the epidemiological impact of psychiatric and neurological disorders in childhood and also mainly cuts short health hospitalization costs.

Acknowledgements

This study was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), a Brazilian Federal Agency for Support and Evaluation of Graduate Education.

References

Alchieri, J. C., & Cruz, R. M. (2003). *Avaliação psicológica: Conceito, métodos e instrumentos*. São Paulo: Casa do Psicólogo.
Amiel-Tison, C., & Gosselin, J. (2001). *Neurological development from birth to six years: Guide for examination and evaluation*. Baltimore: Johns Hopkins University Press.
Anastasi, A. (1967). *Testes psicológicos* (2a ed.). São Paulo: Heder.
American Psychological Association (APA). (1985). *American Psychological Association standards for educational and psychological testing*. Washington: American Psychological Association Inc.
Applebaum, A. (1978). Validity of the revised Denver developmental screening test for referred and non referred samples. *Psychological Report*, *43*, 227–233.
Aylward, G. P. (2005). The conundrum of prediction. *Pediatrics*, *116*(2), 491–492.
Aylward, G. P. (1995). *The Bayley infant neurodevelopmental screener (BINS)*. San Antonio: The Psychological Corporation.
Barratt, M. S., Roach, M. A., & Leavitt, L. A. (1996). The impact of low-risk prematurity on maternal behavior and toddler outcomes. *International Journal of Behavioral Development*, *19*(3), 581–602.
Batista, P. E., Vilanova, L. C. P., & Vieira, R. M. (1997). *O desenvolvimento do comportamento da criança no primeiro ano de vida: Padronização de uma escala para a avaliação e o acompanhamento*. São Paulo: Casa do Psicólogo.
Bayley, N. (1933). *The California first year mental scale*. *Syllabus Series* University California.
Bayley, N. (1969). *Bayley scales of infant development*. New York: Psychological Corporation.
Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). San Antonio: Psychological Corporation.
Bayley, N. (2005). *Bayley scales of infant and Toddler development* (3rd ed.). New York: Psychological Corporation.
Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191.
Bee, H. (2003). *A Criança em Desenvolvimento* (9a ed.). Porto Alegre: Artmed.
Bellis, T. J., & Burke, J. R. (1996). Screening. In T. J. Bellis (Ed.), *Assessment and management of central auditory processing disorders in the educational settings* (pp. 91–112). San Diego, CA: Singular.
Bess, F. H., & Humes, L. E. (1998). *Fundamentos de audiologia*. Porto Alegre: Artes Médicas.
Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
Brooks Gunn, J. (1990). Enhancing the development of young children. *Current Opinion in Pediatrics*, *2*, 873–877.
Brunet, O., & Lèzine, I. (1981). *Desenvolvimento psicológico da primeira infância*. Porto Alegre: Artes Médicas.
Chaudari, S., Kulkarni, S., Pajnigar, N. A., Pandit, N. A., & Deshmukh, S. (1991). A longitudinal follow-up of development of preterm infants. *Indian Pediatrics*, *28*, 873–880.
Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. Newbury Park, CA: Sage.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.
Duarte, C. S., & Bordin, I. A. S. (2000). Instrumentos de avaliação. *Jornal Brasileiro de Psiquiatria*, *22*(2), 55–58.
Ergenekon, E., Unal, S., Gücüyener, K., Soysal, S. E., Koç, E., Okumus, N., et al. (2007). Hypernatremic dehydration in the newborn period and long-term follow up. *Pediatrics International*, *49*(1), 19–23.

Fagan, J. F., & Singer, L. T. (1983). Infant recognition memory as a measure of intelligence. *Advanced Infancy Research*, *2*, 31–78.

Fisher, A. G. (1993). The assessment of IADL motor skills: An application of many Rasch analyses. *The American Journal of Occupational Therapy*, *47*(4), 319–329.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *The American Psychologist*, *54*, 5–20.

Flynn, J. R. (2008). Aumentos de QI. *Scientific American Viver Mente e Cérebro*, *15*(184), 44–49.

Frankenburg, K. W., Dodds, J., Archer, P., & Bresnick, B. (1990). *Denver II technical manual*. Denver, CO: Denver Developmental Materials Inc.

Frankenburg, W. K., & Dodds, J. B. (1967). The Denver developmental screening test. *The Journal of Pediatrics*, *71*(2), 181–191.

Gesell, A. L. (1925). *The mental growth of the preschool child*. New York: MacMillan.

Glascoe, F. P., Byrne, K. E., Ashford, L. G., Johnson, K. L., Chang, B., & Strickland, B. (1992). Accuracy of the Denver-II in developmental screening. *Pediatrics*, *89*, 1221–1224.

Grantham McGregor, S. M., Powel, C., Stewart, M., & Schofield, W. N. (1982). Longitudinal study of growth and development of young Jamaican children recovering from severe protein-energy malnutrition. *Developmental Medicine and Child Neurology*, *24*, 321–331.

Guedes, D. Z., Pedromonico, M. R. M., Goulart, A. L., & Gomes, M. (2005). The pointing gestures in 12 month old infants born preterm. [abstract], *Pediatric Research*, *57*(6), 923 [Presented at XLII Annual meeting of the Latin American Society for Pediatric Research (SLAIP); 2004; Lima, Peru].

Hack, M., & Costello, D. W. (2007). Decrease in frequency of cerebral palsy in preterm infants. *Lancet*, *369*(9555), 7–8.

Halpern, R., Giugliani, E. R. J., Victora, C. G., Barros, F. C., & Horta, B. L. (2000). Fatores de risco para suspeita de atraso no desenvolvimento neuropsicomotor aos 12 meses de vida. *The Journal of Pediatrics*, *76*(6), 421–428.

Hess, C. R., Papas, M. A., & Black, M. M. (2004). Use of the Bayley infant neurological screener with environmental risk groups. *Journal of Pediatric Psychology*, *29*, 321–330.

Horowitz, F. D. (1992). The concept of risk: A reevaluation. In S. L. Friedman, & M. D. Sigman (Eds.), *The psychological developmental of birthweight children* (pp. 61–88). Norwood, NJ: Ablex.

Hutz, C. S., & Silva, D. F. M. (2002). Avaliação psicológica com crianças e adolescentes em situação de risco. *Avaliação Psicológica*, *1*, 73–79.

Hutz, C. S., Koller, S. H., Bandeira, D. R., & Forster, L. M. (1995). *Methodological and ethical issues in research with street children* [Presented at the meeting of the Society for Research in Child Development; Indianapolis, USA].

Isotani, S. M., Pedromônico, M. R. M., Perissinoto, J., & Kopelman, B. I. (2002). O desenvolvimento de crianças nascidas pré-termo no terceiro ano de vida. *Folha Médica*, *121*(2), 85–92.

Knobloch, H., Pasamanick, P. H., & Sherard, E. S. (1966). A developmental screening inventory for infants. *Pediatrics*, *38*, 1095–1108.

Leonard, C. H., Piecuch, R. E., & Cooper, B. A. (2001). Use of the Bayley infant neurodevelopment screener with low birth weight infants. *Journal of Pediatric Psychology*, *26*(1), 33–40.

Linacre, J. M., & Wright, B. D. (1991). *WINSTEPS – Rasch-Model [computer program]*. Chicago: MESA Press.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 887.

Masten, A. S., & Garmezy, N. (1985). Risk, vulnerability and protective factors in developmental psychopathology. In B. B. Lahey, & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (8th ed., Vol. 243, pp. 1–52). New York: Plenum Press.

McCall, R. B., Hogarty, P. S., & Hurlburt, N. (1972). Transitions in infant sensorimotor development and the prediction of childhood IQ. *The American Psychologist*, *27*(8), 728–748.

McCornick, M. C. (1989). Long-term follow-up of infants discharged from neonatal intensive care units. *The Journal of the American Medical Association*, *261*(12), 1767–1772.

Meisels, J. S. (1989). Can developmental screening testes identify children who are developmentally at risk? *Pediatrics*, *83*(4), 578–585.

Meisels, J. S., & Wasik, B. A. (1990). Who should be served? Identifying children in need of early intervention. In S. J. Meisels, & J. Shonkoff (Eds.), *Handbook of early intervention* (pp. 605–632). Cambridge: Cambridge University Press.

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*, 283–298.

Msall, M. E. (2004). Developmental vulnerability and resilience in extremely preterm infants. *The Journal of the American Medical Association*, *292*, 2399–2401.

Naar-King, S., Ellis, D. A., & Frey, M. A. (2004). *Assessing children's well-being: A handbook of measures*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Newborg, J., Stock, J., Wnek, L., Guidubaldi, J., & Svinicki, J. (1988). *Battelle developmental inventory screening test with recalibrated technical data and norms*. Allen, TX: DLM-Teaching Resources.

Noronha, A. P. P., Vendramini, C. M. M., Souza, R. L., Franco, M. O., & Filizatti, R. (2003). Propriedades psicométricas apresentadas em manuais de testes de inteligência. *Psicologia em Estudos*, *8*(1), 93–99.

Nunes, L. R. D. P., Sisdelli, R. O., & Fernandes, R. L. C. (1995). O valor dos testes de bebês e suas implicações para Psicologia do Desenvolvimento. *Revista Brasileira de Educação Especial*, *5*, 107–125.

Oliveira, L. N., Lima, M. C. M. P., & Goncalves, V. M. G. (2003). Acompanhamento de lactentes com baixo peso ao nascimento: Aquisição de linguagem. *Arquivos de Neuro-Psiquiatria*, *61*(3B), 802–807.

Pasquali, L. (2000). Princípios de elaboração de escalas psicológicas. In C. Gorenstein, L. H. S. G. Andrade, & A. W. Zuardi (Eds.), *Escalas de avaliação clínica em psiquiatria e psicofarmacologia* (pp. 15–21). São Paulo, Brazil: Lemos.

Pasquali L. (Ed.). (2001). *Técnicas de exame psicológico – TEP (V.1)*. São Paulo: Casa do Psicólogo.

Pedromônico, M. R. M. (2003). *Problemas de desenvolvimento da criança: Prevenção e intervenção* [Presented in II Encontro de Estudos do Desenvolvimento Humano em Condições Especiais; São Paulo, Brazil].

Piper, M. C., Pinnell, L. E., Darrah, J., Maguire, T., & Byrne, P. J. (1992). Construction and validation of the Alberta Infant Motor Scale (AIMS). *Canadian Journal of Public Health. Revue Canadienne De Sante Publique*, *83*(2), 46–50.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedogogische Institut.

Rutter, M. (1985). Resilience in the face of adversity: Protective factors and resistance to psychiatric disorder. *The British Journal of Psychiatry*, *147*, 598–611.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Orthopsychiatric Association*, *57*, 316–331.

Rutter, M. (1993). Resilience: Some conceptual considerations. *The Journal of Adolescent Health*, *14*, 626–631.

Rydz, D., Shevell, M. I., Majnemer, A., & Oskoui, M. (2005). Topical review: Developmental screening. *Journal of Child Neurology*, *20*, 4–21.

Sameroff, A. J. (1986). Environmental context of child development. *The Journal of Pediatrics*, *109*(1), 192–200.

Sameroff, A. J. (1990). Neo-environmental perspectives on developmental theory. In R. M. Hodopp, J. A. Burack, & E. Zigler (Eds.), *Issues in the developmental approach to mental retardation* (pp. 93–111). New York: Cambridge.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland adaptive behavior scales* (interview ed.). Circle Pines, MN: American Guidance Service.

Stjernqvist, K., & Svenningsen, N. W. (1999). Ten-year follow-up of children born before 29 gestational weeks: Health, cognitive development, behaviour and school achievement. *Acta Paediatrica*, *88*(5), 557–562.

Theuer, R. V., & Mendonza, C. E. F. (2003). Avaliação da inteligência na primeira infância. *Psico-USF*, *8*(1), 21–32.

Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.

USA Public Law 99-457. (1991). *The Education of the Handicapped Act Amendts (EHA) reauthorized*. Washington, USA: Federal Register.

Uzgiris, I. C., & Hunt, J. Mc. V. (1966). *An instrument for assessing infant psychological development*. Urbana, IL: University of Illinois Press.

Vohr, R. B. (1991). Preterm cognitive development: Biologic and environmental influences. *Infants and Young Children*, *3*, 20–29.

Yunes, M. A. M. (2003). Psicologia positiva e resiliência: O foco no indivíduo e na família. *Psicologia em Estudos*, *8*(esp.), 75–84.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*, 561–577.