Archives of
Rehabilitation

April 2022. Vol 23. Num 1

Research Paper

Standardization of the Bayley Scales of Infant and Toddler Development for Persian Children

Farin Soleimani[1], *Nadia Azari[1], Roshanak Vameghi[2], Seyyed Hamed Barekati[3], Hamidreza Lornejad[3], Adis Kraskian[4]

1. Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
2. Department of Clinical Education, Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
3. Ministry of Health and Medical Education, Tehran, Iran.
4. Department of psychology, Islamic Azad University, Branch of Karaj, Karaj, Iran.

Use your device to scan and read the article online

Citation Soleimani F, Azari N, Vameghi R, Barekati S H, Lornejad H, Kraskian A. [Standardization of the Bayley Scales of Infant and Toddler Development for Persian Children (Persian)]. Archives of Rehabilitation. 2022; 23(1):8-31. https://doi.org/10.32598/RJ.23.1.42.4

https://doi.org/10.32598/RJ.23.1.42.4

CC BY NC

ABSTRACT

Objective The first years of life are particularly important because vital developments occur in all development domains including motor, cognition, communication, and social-emotional. Early detection of developmental delays is essential in planning for early intervention. This study aims to standardize the score of bayley scales of infants and toddlers development-third edition (Bayley-III) for Iranian samples aged 1-42 months.

Materials & Methods Participants included 1700 children divided into 17 age groups of 100 children. The normative information was based on a national sample representative of Iran's population for infants aged 1-42 months based on the 2011 national population and housing census. For each age group, the total raw scores of each Bayley-III subscale (cognition, receptive communication, expressive communication, fine motor, and gross motor) were converted to scaled scores with a Mean±SD of 10±3. The composite scores (cognition, language and motor domains) were obtained by summing up the scaled scores. The composite scores were scaled to a metric with Mean±SD of 100±15 ranged 40-160. Growth scores (ranged 200-800, with a Mean±SD 500±100), percentile ranks (ranged from 1-99, with 50 as the mean and median), and developmental age equivalent were also determined. Confidence intervals for all five subscales were also determined. To compare the development level of the US children (norm samples) and Iranian children, their mean raw scores in five subscales were compared for finding the difference in scores.

Results The Bayley-III was performed on 1744 children aged 0-42 months, including 908 (52.1%) girls and 836 (47.9%) boys. In comparing the development level of Iranian and US children, it was found that the mean scores were significantly different in 28 age groups ($P<0.05$). In seven age groups, the US children's scores were higher (age groups <6 months), and in 21 age groups, the scores of Iranian children were higher (age groups >6 months).

Conclusion For early detection and intervention of children with developmental delays, a test with a normalized score should be used in Iran. Using the US norm-based scores in Iranian children leads to inaccuracy in the early detection of children with developmental delays.

Keywords Child, Developmental delay, Standardization, Bayley Scales of Infants and Toddlers Development

Received: 24 Mar 2021
Accepted: 06 Oct 2021
Available Online: 01 Apr 2022

Corresponding Author:
Nadia Azari, PhD.
Address: Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
Tel: +98 (21) 71732846
E-Mail: azarinadia@gmail.com

Extended Abstract

Introduction

The childhood years are the most vital years of life which make the foundation of future life when humans experience developments in Cognition, motor, language, emotional, and social domains [1, 2]. More than 200 million children in developing countries, i.e, more than one-third of children worldwide, do not experience proper development. As a result, their countries lose approximately 20% of their manpower [3]. The critical period for growth and development is during the early years of childhood (first 5 years) [4]. To ensure optimal development, a growth assessment in early childhood is needed; therefore, developmental screening for timely interventions in early childhood can reduce the incidence of developmental delay [5-9]. About 6-9% of children under 6 years of age suffer from developmental delay or disorders and need early intervention [10-14]. An study in the United States (US) indicated that 1 in 6 children with developmental disorders have speech and language disorders, mental retardation, cerebral palsy, and autism [11]. Most children who need early intervention are not diagnosed before school age [9, 15]. Children whose developmental delay is undiagnosed and, consequently, receive no intervention until school age are at risk for academic failure, behavioral problems, and socio-emotional disorders [16].

The World Health Organization has emphasizes that the provision of better care in early childhood improves the child's school performance [17-19]. To definitively diagnose children with developmental disorders, valid and reliable tools should be used. In developing countries, the tools developed by developed countries are often used. These tools include items that may be adaptable to other cultures. Hence, the tools may fail to assess children from different cultures [20]. These tools should be translated and localized to be adaptable to other cultures [21]. On the other hand, the use of norm scores based on the original version of the tools can lead to errors in estimating the developmental level of children in other countries; therefore, determining the norm scores for different cultures is necessary [22].

The standardization of developmental tests for children is limited in Iran; the ages & stages questionnaire (ASQ) [24] and the Denver Developmental Screening Test [25] have been translated and their validity and reliability have been examined, but their norm scores have not been determined. The bayley scales of infant and toddler development, third edition (Bayley-III) assesses the child's developmental status in five developmental domains including cognition, language (receptive, expressive communication), motor (gross and fine), social-emotional, and adaptive behavior in early childhood, i.e, from day 15 to 42 months. It is used as the gold standard for grading other developmental assessment tools [23], and is applied in many countries. In a study, it was found reliance on US norm-based scores of Bayley-III led to the misclassification of developmental delays in Malawian children [26]. A study conducted in Taiwan suggested that higher Bayley-III scores should be considered in Taiwanese children as a sign of developmental delay [27]. Godamunne used cognition and motor subscales of Bayley-III for Sri Lankan children and compared their performance with US children based on the subscales. The results showed that the scores of 12-month-old children were higher in the cognition domain and lower in the motor domain compared to the US children, while the scores of 24-month-old children were lower in the cognition domain compared to the US children [28]. In another study which was conducted on preterm infants, it was warned that we should be cautious in interpreting Bayley-III scores for detecting developmental delays in this group and the norm scores of this test should be considered higher than US norm-based scores [29].

Since a comprehensive and appropriate test is not available for health care providers in Iran to assess the development status of children in various fields and be used as a gold standard, and given that the use of scores based on the norms of a country may lead to the wrong estimation of disorders in other countries, this study aims to standardize the Bayley-III for Iranian children. The mean raw scores in the two main and Persian versions were examined to compare the development level of US and Iranian children.

Materials and Methods

Study design and participants

This is a cross-sectional study. Sampling was conducted using a stratified sampling method. First, Iran was divided into eight geographical zones (center, north, west, south, southeast, southwest, northeast, northwest, and center) and 10 cities including Tehran, Isfahan, Mashhad, Rasht, Tabriz, Hamadan, Ahvaz , Shiraz, Qom, and Zahedan were selected. Two experienced childcare experts from each city were selected for training and testing. The sampling phase lasted for 3-6 months, depending on the population rate of different cities. The selected samples were 1700 children aged 16 days to 42 months 15 days including 17 age groups of 100 samples from A to Q. The age range was categorized into 1-6 months groups with a one-month interval, 6-12 months group with a two-months interval, 12-30 months group with a 3-month interval, and 30-42 months group with a 6-month interval. This sample size was calculated based on the 2011 national population and housing census related to the population aged

Table 1. Number of norm samples from the selected cities in Iran

No.	City	Population*	Norm samples
			No. (%)
1	Hamedan	45500	50(2.9)
2	Rasht	51239	57(3.3)
3	Esfahan	147738	161(9.2)
4	Qom	104387	114(6.5)
5	Mashhad	268278	295(16.9)
6	Shiraz	117987	130(7.5)
7	Zahedan	90047	98(5.6)
8	Ahvaz	135231	148(8.5)
9	Tabriz	126061	141(8.1)
10	Tehran	476210	550(31.5)
Total		1562678	1744(100)

* Children aged 0-4 years according to the 2011 national population and housing census.

Archives of
Rehabilitation

<4 years for each city [30]which are presented in Table 1. The samples were classified based on demographic factors including age, gender, and geographical area. Inclusion criteria were no significant medical complications and not being under treatment for any cognition, physical, or behavioral problems. Exclusion criteria were attendance in rehabilitation interventions, being unable to speak in Persian, and having sensory disorders. After receiving written consent from the parents of children, they completed a demographic form (surveying child's age and gender, parents' education, gestational age at birth, birth weight, birth defects). In case of a developmental delay, the child was introduced to the Integrated Center for Child Development for diagnostic measures and health promotion training.

Data collection

The Bayley-III is an assessment tool that evaluates developmental level of children in 5 domains including cognition (91 items), receptive communication (49 items), expressive communication (48 items), fine motor (66 items), and gross motor (72 items). Norm data of this test were achieved from 1700 samples in the US version [23]. The assessment was conducted objectively by the examiner. The norm scores were determined for 48 age groups to increase the accuracy. The test reports four groups of norm scores: Scaled scores, composite scores, percentile ranks, and growth scores. Confidence intervals are provided for these scores, and developmental age equivalents are available for the relevant scales. The Persian version of Bayley-III was used in this study. The Cronbach's alpha coefficients for all domains are higher than 0.76 and the Pearson correlation coefficient in the test-retest method and inter-rater method for different domains was 0.987 (P<0.001) and 0.991 (P<0.001), respectively [31].

Data analysis

At first, 48 age groups were determined based on the available samples. To determine norm scores from the main version; i.e, the conversion of raw scores into scaled, composite, and growth scores, the percentile ranks and developmental age equivalents were used at 48 age groups from 15 days to 42 months. The scaled scores obtained from the raw scores were converted linearly into scores in the range of 1-19 with a Mean±SD of 10±3. The composite scores were derived from the sums scores of the subscales and converted into a score of 40-160 with a Mean±SD of 100±15. The percentile scores were converted into a score of 1-99 with a mean of 50. The sums percentile ranks (range from 1 to 99, with 50 as the mean and median), and 90% and 95% confidence intervals were determined for cognition, language, and motor subscales. The growth scores were obtained by converting raw scores to scores in the range of 200-800 with a mean of 500 and a SD of 100. Accordingly, the growth curves were plotted to show the child's growth over time for cognition,

receptive and expressive communications and fine and gross motors at 5, 10, 25, 50, 75, 90, and 95 percentiles. To determine the developmental age equivalent, the raw score which was equivalent to the score of 10 for each age group, was calculated. To compare the developmental level of the US and Iranian children, their mean raw scores were examined in five domains at 48 age groups, and the difference in scores was statistically evaluated.

Results

In this study, participants were 1744 children aged 15 days to 42 months; 908 girls (52.1%) and 836 boys (47.9%). Their norm scores based on the Bayley-III were determined in 48 age groups. Based on the parents' educational level, 64.5%, 34.2%, and 1.3% of fathers, and 58%, 41%, and 1% of mothers had diploma and lower, bachelor's degree, master's degree and higher, respectively. Table 2 presents the mean

Table 2. Mean raw scores of age groups in the norm samples

	Age group	No. (%)	Mean±SD				
			Cognition	Receptive Communication	Expressive Communication	Fine Motor	Gross Motor
A	16 days – 1 month and 15 days	104(6)	5.56±3.17	3.61±1.69	1.61±0.7	3.61±1.85	3.94±1.43
B	1 month and 16 days – 2 months and 15 days	100(5.7)	8.70±2.99	4.40±1.14	3.75±1.25	5.53±1.68	8.45±3.20
C	2 months and 16 days – 3 months and 15 days	103(5.9)	11.50±2.88	6.62±2.82	5.54±3.07	6.67±2.43	11.33±3.85
D	3 months and 16 days – 4 months and 15 days	99(5.7)	14.74±4.03	7.05±1.18	5.34±1.17	11.46±2.52	16.45±2.75
E	4 months and 16 days – 5 months and 15 days	100(5.7)	14.74±4.03	7.05±1.18	5.34±1.17	11.46±2.52	16.45±2.75
F	5 months and 16 days – 6 months and 15 days	61(3.5)	26.15±3.34	9.65±1.92	7.12±2.23	19.62±2.61	24.62±3.50
G	6 months and 16 days – 8 months and 30 days	89(5.1)	30.10±2.01	11.03±1.80	9.25±2.62	21.69±1.89	29.69±4.41
H	9 months and 0 days – 10 months and 30 days	84(4.8)	33.97±3.10	12.06±1.81	9.84±2.57	25.87±2.23	34.68±3.34
I	11 months and 0 days – 13 months and 15 days	97(5.6)	40.71±3.79	14.97±2.56	14.16±3.13	28.06±2.52	42.63±4.37
J	13 months and 16 days – 16 months and 15 days	117(6.7)	48.91±3.91	19.49±2.90	19.54±3.38	32.34±2.45	50.71±3.64
K	16 months and 16 days – 19 months and 15 days	111(6.4)	52.47±4.20	20.87±2.97	21.28±4.27	34.53±2.24	52.19±3.46
L	19 months and 16 days – 22 months and 15 days	103(5.9)	58.50±4.06	25.20±2.93	26.85±3.27	37.85±3.35	55.40±2.16
M	22 months and 16 days – 25 months and 15 days	104(6)	62.50±6.07	30.22±3.86	32.78±6.50	40.11±3.01	57.28±4.06
N	25 months and 16 days – 28 months and 15 days	108(6.2)	65.82±5.57	31.71±5.19	31.76±7.40	41.67±4.41	58.81±3.63
O	28 months and 16 days – 32 months and 30 days	151(8.7)	70.39±5.54	35.26±4.21	38.52±6.19	44.82±4.16	61.27±3.37
P	33 months and 0 days – 38 months and 30 days	146(8.4)	74.72±4.93	38.24±4.42	41.41±4.76	50.33±4.03	63±4.45
Q	39 months and 0 days – 42 months and 15 days	67(3.8)	76.83±4.69	42.39±4.41	43±7.01	53.88±4.12	66.50±2.66

Table 3. Comparison of the mean scores of Bayley-III between the US and Iranian children

Age Group	Subscale	Sample	N	Mean±SD	P
1	Cognition	US	30	6±3	P<0.01
		Iran	18	4.22±1.35	
	Receptive communication	US	30	5.00±1.50	P<0.05
		Iran	18	4.11±1.32	
4	Receptive communication	US	30	5.50±1.62	P<0.05
		Iran	12	4.42±1.17	
5	Expressive communication	US	30	4±1.75	P<0.05
		Iran	30	5.20±1.67	
6	Receptive communication	US	30	6±1.62	P<0.05
		Iran	21	5±1.41	
14	Gross motor	US	30	20±3.75	P<0.01
		Iran	16	17.06±2.74	
15	Gross motor	US	30	21±3.87	P<0.5
		Iran	19	18.32±3.33	
16	Receptive communication	US	30	9.50±1.87	P<0.01
		Iran	88	10.81±1.65	
19	Gross movements	US	30	35±3.75	P<0.05
		Iran	58	32.90±4.27	
23	Receptive communication	US	30	15±2.75	P<0.01
		Iran	64	17.23±2.86	
24	Cognition	US	30	44.50±4.87	P<0.05
		Iran	50	46.86±3.81	
	Receptive communication	US	30	16±3.25	P<0.01
		Iran	50	18.10±2.69	
25	Receptive communication	US	30	17±3.25	P<0.05
		Iran	49	18.57±2.97	
	Fine motor	US	30	31±3	P<0.01
		Iran	49	32.73±2.33	
	Gross motor	US	30	47±4	P<0.01
		Iran	49	49.98±3.69	

Age Group	Subscale	Sample	N	Mean±SD	P
26	Receptive communication	US	30	18±3.50	P<0.05
		Iran	62	19.61±3.29	
	Gross motor	US	30	48.50±4.12	P<0.05
		Iran	62	50.52±4.24	
29	Receptive communication	US	30	21.50±4.12	P<0.05
		Iran	37	23.43±3.01	
34	Receptive communication	US	30	26.50±5.25	P<0.05
		Iran	19	29.63±3.15	
40	Receptive communication	US	30	31.50±6.37	P<0.05
		Iran	17	34.35±2.85	
42	Expressive communication	US	30	36.50±6.12	P<0.05
		Iran	16	40.44±4.77	
	Fine motor	US	30	45±4.87	P<0.05
		Iran	16	48.31±3.86	
43	Expressive communication	US	30	37±6.62	P<0.05
		Iran	23	40.74±5.07	
	Fine motor	US	30	45.50±5	P<0.05
		Iran	23	47.91±3.01	
44	Gross motor	US	30	62.50±4	P<0.05
		Iran	24	64.38±2.76	
47	Fine motor	US	30	49.50±5.87	P<0.05
		Iran	74	51.95±3.57	
48	Receptive communication	US	30	38±6.62	P<0.01
		Iran	71	41.62±4.29	
	Gross motor	US	30	65.50±3.75	P<0.05
		Iran	71	67.18±2.40	

Archives of Rehabilitation

raw scores and SD for the Bayley-III subscales in the norm sample for 17 age groups.

The norm scores in 48 age groups based on composite score, percentile rank, and developmental age equivalent for cognition, receptive and expressive communications, gross motor, and fine motor were also obtained. Results in Table 3 showed a significant difference between the US and Iranian children. The mean scores at 28 age groups (2 in cognition, 12 in receptive communication, 3 in receptive communication, 4 in fine motor, and 7 in gross motor) were different between the two groups.

Discussion and Conclusion

In this study, the Bayley-III score was normalized in a sample of 1744 children aged 1-42 months in Iran. Norm scores

(scaled, composite, and growth scores) were calculated, and the percentile ranks and developmental age equivalent of Iranian children were determined. The difference in mean norm scores in 28 age groups was different between the US and Iranian samples (P<0.05). In 7 age groups, the mean norm scores were higher than in the US children (age range <6 months) while in 21 age groups, the mean norm scores were higher than in the Iranian children (age range >6 months). In this study, only children with normal development were included. Since the US samples are representative of the general population, about 10% of the samples were devoted to children with special needs (21 children with trisomy, cerebral palsy, autism spectrum disorder, prematurity, specific speech disorders, maternal alcohol consumption during pregnancy, asphyxia at birth, and low birth weight) [23]. Previous studies have indicated that the inclusion of children with special needs in a standard sample can negatively affect the ability to distinguish children with natural development and developmental delay [32]. For this reason, this technqiue was not used in our study. The selection of normal children in Iran may have caused an increase in the mean scores, a difference between the US and Iranian samples, or a difference in biological, environmental, and cultural characteristics. The same differences between the two norms have been identified in the studies conducted in other countries [26-29, 33, 34].

In Iran, only the ASQ score has been normalized [24]. The ASQ has communication, fine motor, gross motor, problem-solving and personal-social domains for the ages of 4-60 months, which is completed by parents. Examining the difference between the mean scores of the Iranian and US children reported that the scores of Iranian children in the areas of fine motor, problem-solving and personal-social, especially at the ages of 8 and 12 months, were higher than those of the US children, while in the gross motor domain, especially at 36 months, their scores were lower than the US children (P<0.05) [24]. In our study, the mean scores of all Bayley-III subscales for the age group >6 months were higher in the Iranian children compared to the US children. In the Netherlands, the Bayley-III norm scores were obtained at 48 age groups for cognition, motor, and language scales. The mean scores of Dutch and US children were different. It was reported that the use of US norm-based scores in the Netherlands leads to the lack of early detection of developmental delay in children [35].

Cromwell et al. examined the validity of Bayley-III in Mali. They performed the test on 167 healthy Malawian children. The results indicated that the mean scores of Malawian children under the age of 6 months were higher than those of US peers in all subscales; however, the US children showed higher scores at the ages above 6 months [26]. Moore evaluated 185 children aged 29-41 months simultaneously with the Bayley-II and Bayley-III tools. Their results indicated that the development level of children's cognition and communication based on the Bayley-III score was higher than that based on the Bayley-II score. They suggested that it should be careful in using the Bayley-III scores to detect developmental delays in children [29].

Similarly, Johnson et al. conducted a study on 158 preterm childs to assess how to classify neurodevelopmental delay using Bayley-III compared to the Bayley-II. The scores obtained by the Bayley-III were higher than those by the Bayley-II [33]. In Taiwan, Yu et al. used the Bayley-II and Bayley-III simultaneously in 47 norm infants and 167 preterm infants at the ages of 6, 12, 18, and 24 months. They stated that the scores obtained from the Bayley-III determined the development level of Taiwanese children higher compared to the scores of Bayley-II [27]. Anderson et al. evaluated the Bayley-III's ability to diagnose developmental delay in 2-year-old children born prematurely (aged <28 weeks) or underweight (weight <1,000 grams) in Victoria, Australia. The mean of all subscale scores in the patient group was significantly lower than in the control group. However, their mean scores were consistent with the mean scores of the US samples; the mean scores in the control group were higher than in the US samples. They concluded that the US children's Bayley III scores underestimated the developmental delay in Australia [34]. Godamunne et al. used the cognition and motor domains of the Bayley-III in 150 Sri Lankan norm children in 6-, 12-, and 24-month-old age groups and compared their performance with the US children. The results showed that the cognition score of 12-month-old children was higher and their motor scores were lower compared to the US children, and the cognition score of 24-month-old children was lower than that of the US children [28].

The Bayley-III is apropriate for assessment of child development, because of its good psychometric properties and quantitative scoring system. The standardization of its Persian version not only allows early detection of developmental delay, but also eliminates the lack of a gold standard for comparing the results of other developmental tests in Iran.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation Sciences. Informed consent was obtained from the parents of the children (Code: IR.USWR.REC.1393.14357).

Funding

Financial support for this study was obtained from the Iranian Ministry of Health and Medical Education.

Authors' contributions

Conceptualization: Farin Soleimani, Nadia Azari, RoshanakVameghi;Methodology: Farin Soleimani, NadiaAzari, Adis Kraskian; Investigation: Farin Soleimani,Nadia Azari, Seyed Hamed barekati, Hamidreza Lornejad;Writing-original draft: Farin Soleimani, Nadia Azari;Writing-review & editing: All author.

Conflict of interest

According to the authors, there is no conflict of interest.

Acknowledgments

The authors would like to thank the University of Social Welfare and Rehabilitation Sciences and the Shahid Beheshti University of Medical Sciences, Iran, Tehran, Mashhad, Rasht, Tabriz, Hamedan, Ahvaz, Shiraz, Isfahan, Qom, and Zahedan for their cooperation and the parents of children who participated in the study

This Page Intentionally Left Blank

مقاله پژوهشی

هنجاریابی مقیاس تکاملی شیرخواران و نوپایان بیلی برای کودکان ایرانی

فرین سلیمانی[۱]، *نادیا آذری[۱]، روشنک وامقی[۲]، سیدحامد برکاتی[۳]، حمیدرضا لرنژاد[۳]، آدیس کراسکیان[۴]

۱. مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم توانبخشی و سلامت اجتماعی، تهران، ایران.
۲. گروه آموزش بالینی، مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم توانبخشی و سلامت اجتماعی، تهران، ایران.
۳. ستاد مرکزی وزارت بهداشت، درمان و آموزش پزشکی، تهران، ایران.
۴. گروه روانشناسی، دانشکده روانشناسی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران.

Use your device to scan and read the article online

Citation Soleimani F, Azari N, Vameghi R, Barekati S H, Lornejad H, Kraskian A. [Standardization of the Bayley Scales of Infant and Toddler Development for Persian Children (Persian)]. Archives of Rehabilitation. 2022; 23(1):8-31. https://doi.org/10.32598/RJ.23.1.42.4

https://doi.org/10.32598/RJ.23.1.42.4

CC BY NC

چکیده

هدف تکامل شامل حیطه‌های حرکتی، شناختی، ارتباطی و عاطفی‌اجتماعی است. تکامل دوران ابتدای کودکی به دلیل رشد سریع مغز اهمیت فراوانی دارد. تشخیص زودرس تأخیر تکاملی برای آزمون بهنگام امری ضروری محسوب می‌شود. هدف از این مطالعه، هنجاریابی نسخه فارسی مقیاس تکاملی بیلی درکودکان ۱ تا ۴۲ماهه ایرانی بود.

روش بررسی آزمودنی‌ها (۱۷۰۰ نمونه) با توجه به ۱۷ گروه سنی و با در نظر گرفتن ۱۰۰ نمونه در هر گروه، بر اساس آمار و سرشماری مرکز آمار ایران برای جمعیت کودکان صفر تا ۴۲ ماه انتخاب شدند. برای هر گروه سنی، کل نمرات خام ۵ مقیاس شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت، به نمرات ترازشده با میانگین ۱۰ و انحراف معیار ۳ تبدیل شد. نمرات مشتق از مجموع نمرات ترازشده مقیاس‌های شناختی، زبانی و حرکتی به‌دست آمد و به نمراتی با میانگین ۱۰۰ و انحراف معیار ۱۵ با دامنه ۴۰ تا ۱۶۰ تبدیل شد. نمرات رشدی (با دامنه ۲۰۰ تا ۸۰۰، میانگین ۵۰۰ و انحراف معیار ۱۰۰)، رتبه‌های درصدی (با دامنه ۱ تا ۹۹ و میانگین و میانه ۵۰) و معادل سن تکاملی محاسبه شد. فواصل اطمینان برای هر ۵ مقیاس محاسبه و منحنی‌های مربوطه ترسیم شد. به منظور مقایسه سطح تکامل کودکان آمریکایی (نسخه اصلی آزمون) و ایرانی و بررسی تفاوت نمرات دو جامعه، میانگین نمرات خام در مقیاس‌های پنج‌گانه از نظر معناداری اختلاف نمرات ارزیابی شد.

یافته‌ها آزمون بر روی ۱۷۴۴ کودک صفر تا ۴۲ماهه اجرا و در ۴۸ گروه سنی هنجاریابی شد. تعداد دختران ۹۰۸ (۵۲/۱ درصد) و تعداد پسران ۸۳۶ (۴۷/۹ درصد) نفر بود. میانگین نمرات ۲۸ گروه با توجه به سن و مقیاس بررسی بین دو جامعه متفاوت بود (۰۵/ P<۰)، در این بررسی، در ۷ گروه، میانگین نمرات نمونه آمریکایی بالاتر از نمونه ایرانی بود (طیف سنی کمتر از ۶ ماه) و در ۲۱ گروه، میانگین نمرات در نمونه ایرانی بالاتر از آمریکایی بود (طیف سنی بالاتر از ۶ ماه).

نتیجه‌گیری برای تشخیص و آزمون بهنگام، تأخیر تکاملی باید از نمرات هنجار آزمون در ایران استفاده کرد. استفاده از نمرات هنجار نسخه اصلی آزمون در کودکان ایرانی به اشتباه در تشخیص کودکان با تأخیر تکاملی منجر می‌شود.

کلیدواژه‌ها آزمون بیلی، هنجاریابی، تأخیر تکامل، کودک

تاریخ دریافت: ۰۴ فروردین ۱۴۰۰
تاریخ پذیرش: ۱۴ مهر ۱۴۰۰
تاریخ انتشار: ۱۲ فروردین ۱۴۰۱

* نویسنده مسئول:
دکتر نادیا آذری
نشانی: تهران، دانشگاه علوم توانبخشی و سلامت اجتماعی، مرکز تحقیقات توانبخشی اعصاب اطفال.
تلفن: ۷۱۷۳۲۸۴۶ (۲۱) ۹۸+
رایانامه: azarinadia@gmail.com

مقدمه

سال‌های ابتدایی کودکی مهم‌ترین سال‌های حیات است. تجربیات این سال‌ها زیربنای زندگی آینده هر انسانی خواهد بود [۱، ۲]. تکامل شامل حیطه‌های مختلفی از قبیل شناختی، حرکتی، زبانی و عاطفی‌اجتماعی است [۱، ۲]. بیش از ۲۰۰ میلیون کودک در کشورهای در حال توسعه، یعنی بیش از یک‌سوم کل کودکان به ظرفیت تکاملی خود نمی‌رسند. درنتیجه کشورهای آنان حدود ۲۰ درصد از نیروی مولد این افراد در بزرگسالی را از دست می‌دهد [۳].

«دوره بحرانی» در رشد و تکامل مغز در سال‌های ابتدایی کودکی (از زمان حاملگی تا ۵ سال اول زندگی) قرار دارد [۴]. مهم‌ترین فرصت برای اطمینان از کسب تکامل مطلوب، ارزیابی تکامل در ابتدای کودکی است. بنابراین غربالگری تکاملی برای آزمایشات زودهنگام در سال‌های ابتدای کودکی می‌تواند بروز تأخیر تکاملی کودکان را کاهش دهد [۵]. آزمایشات زودهنگام در کودکان آسیب‌دیده منجر به ارتقای سلامت، رشد شناختی و اجتماعی می‌شود [۶-۹].

حدود ۶ تا ۹ درصد کودکان کمتر از ۶ سال دچار تأخیر یا اختلالات تکاملی هستند و نیاز به آزمایش زودرس دارند [۱۰-۱۳]. تأخیر تکاملی از شایع‌ترین مشکلات طب کودکان در دنیا است و روند آن رو به افزایش است [۱۴، ۲]. یک مطالعه گسترده در آمریکا نشان داد از هر ۶ کودک یک نفر دچار اختلالات تکاملی (شامل اختلالات خفیفی چون مشکلات گفتار و زبان یا مشکلات جدی‌تری مانند ناتوانی‌های هوشی، فلج مغزی و اوتیسم[1]) است [۱۱]. اغلب کودکانی که به آزمایش زودرس نیاز دارند، قبل از سنین مدرسه تشخیص داده نمی‌شوند [۱۵، ۹]. کودکانی که تأخیر تکاملی آن‌ها تشخیص داده نشده است و درنتیجه تا سنین مدرسه بدون آزمایش می‌مانند، در خطر شکست تحصیلی، مشکلات رفتاری و اختلالات اجتماعی عاطفی هستند [۱۶].

سازمان بهداشت جهانی[2] بر این موضوع تأکید دارد که مراقبت بهتر در سال‌های اول کودکی، عملکرد کودک را در دبستان بهبود می‌بخشد [۱۷] و با توجه به اینکه آزمایشات زودرس می‌تواند باعث کاهش عواقب تأخیر تکاملی شود، این سازمان تشخیص زودرس این کودکان را در اولویت‌های خود قرار داده است [۱۸، ۱۹]. بدین ترتیب باید ابزارهای ارزیابی تکاملی معتبر و پایا در دسترس مراقبین بهداشتی قرار گیرند تا کودکانی که تکامل آن‌ها در حد انتظار نیست را تشخیص و به مراکز آزمایشات زودرس ارجاع دهند [۱۸].

به منظور تشخیص قطعی کودکان دچار تأخیر تکامل می‌بایست از ابزارهای با دقت و حساسیت بالا (روا و پایا) و استانداردشده استفاده کرد. در کشورهای در حال توسعه در مطالعات بالینی در مورد تکامل کودک، اغلب از آزمون‌های ساخته‌شده در کشورهای پیشرفته استفاده می‌شود. این آزمون‌ها شامل فعالیت‌ها و مواردی است که شاید برای سایر فرهنگ‌ها بیگانه باشد. بنابراین در ارزیابی کودکان در شرایط فرهنگی به غیر از آنچه برای آن طراحی شده‌اند، ممکن است با شکست مواجه شوند [۲۰].

البته این آزمون‌ها می‌توانند برای استفاده در کشورهای دیگر مناسب‌سازی شوند که شامل ترجمه، انطباق[3] و گاه تغییر[4] سؤالات است که بیان‌کننده خصوصیات فرهنگی آن منطقه است، در غیر این صورت ممکن است باعث سوءتعبیر نتایج شوند [۲۱]. حتی با این تغییرات، استفاده از نمرات هنجار[5] آزمون اصلی می‌تواند به اشتباه در برآورد سطح تکاملی کودک منجر شود. بنابراین تعیین نمرات هنجار در هر جامعه مورد تأکید است. به همین دلیل مطالعات هنجاریابی برای یک جمعیت در استفاده از ابزارهای ترجمه و انطباق داده‌شده باید صورت گیرد [۲۲].

هنجاریابی آزمون‌های تکاملی در ایران محدود هستند و تنها پرسش‌نامه سنین و مراحل[6] در ایران استاندارد شده است. پرسش‌نامه سنین و مراحل یک آزمون غربالگر تکاملی است و ۶۰ سؤال از والد در مورد تکامل کودک دارد و در صورت عدم کسب امتیاز کافی و موارد مشکوک به تأخیر تکاملی نیاز به آزمونی با استاندارد دقیق‌تر دارد.

آزمون تشخیصی مقیاس‌های تکاملی شیرخواران و نوپایان «بیلی ویرایش سوم»[7] برای سنجش تکامل بعد از غربالگری در ایران انتخاب شد. این آزمون دارای دقت و روان‌سنجی بالا (اعتبار و پایایی)، عینی‌بودن و انواع نمرات استاندارد است و وضعیت تکامل کودک را در تمام حیطه‌های تکاملی شامل شناختی، زبانی (ارتباط درکی و بیانی) و حرکتی (ظریف و درشت) در سنین ابتدای کودکی، یعنی از پانزده روزگی تا ۴۲ ماهگی ارزیابی می‌کند و به‌عنوان ابزار استاندارد طلایی برای درجه‌بندی سایر ابزارهای ارزیابی و غربالگری تکامل استفاده می‌شود [۲۳، ۲۴].

یکی از مقیاس‌های تکاملی دیگر که ابعاد مختلف تکامل را می‌سنجد، آزمون دنور[8] است که در ایران روایی و پایایی آن مشخص شده، اما نمرات هنجار آن تعیین نشده است [۲۵].

1. Autism Spectrum Disorder (ASD)
2. World Health Organization (WHO)

3. Adaptation
4. Modification
5. norm
6. Ages & Stages Questionnaire
7. Bayley Scales of Infant and Toddler Development, Third
8. Denver Developmental Screening Test

با توجه به اینکه نمرات هنجار هر آزمون در جمعیت‌های مختلف، متفاوت است، نیاز به نمرات هنجار در هر جمعیت تأکید شده است. در کشور مالی، روایی آزمون بیلی ویرایش سوم بررسی و مشخص شد که تفسیر آزمون با استناد به نمرات هنجار آمریکایی در کودکان مالاوی منجر به اشتباه در دسته‌بندی تأخیر تکاملی، به‌ویژه در حیطه‌های شناختی و ارتباطی می‌شود [۲۶].

مطالعه‌ای در تایوان نیز مشخص کرد نمرات به‌دست‌آمده از آزمون بیلی ویرایش سوم سطح تکامل کودکان تایوانی را بالاتر از بیلی ویرایش دوم ارزیابی می‌کند و محققین پیشنهاد کردند نمرات هنجار آزمون بیلی ویرایش سوم را باید در کودکان تایوانی بالاتر درنظر گرفت تا تأخیر تکاملی به موقع تشخیص داده شود [۲۷]. گودامون، حیطه‌های شناختی حرکتی آزمون بیلی ویرایش سوم را بر روی کودکان سریلانکایی در گروه‌های سنی ۶، ۱۲ و ۲۴ ماهه انجام داد و عملکرد این کودکان را با کودکان هم‌سن آمریکایی در حیطه‌های ذکرشده مقایسه کرد. نتایج نشان داد نمرات کودکان ۱۲ ماهه سریلانکایی در حیطه شناختی بالاتر، در حیطه حرکتی پایین‌تر از متوسط کودکان آمریکایی و در ۲۴ ماهگی نمرات شناختی نسبت به کودکان آمریکایی پایین‌تر بود [۲۸]. در مطالعه دیگری بر روی نوزادان نارس مشخص شد در تفسیر نمرات آزمون بیلی ویرایش سوم برای تشخیص تأخیر تکاملی در این گروه باید محتاط بود و نمرات هنجار در این آزمون برای تشخیص تأخیر تکامل باید بالاتر از نمرات آمریکایی درنظر گرفته شود [۲۹]. درنتیجه هنجاریابی آزمون به منظور تعیین نمرات هنجار برای هر جمعیتی مورد تأکید است.

با توجه به اینکه آزمون جامع و مناسبی که توانایی تشخیص و ارزیابی سطح تکامل کودکان را در ابتدای کودکی و به‌طور هم‌زمان در حیطه‌های مختلف تکاملی داشته باشد و به‌عنوان استاندارد طلایی برای ارزیابی آزمون‌های غربالگری استفاده شود، در اختیار مسئولین و مراقبین بهداشت کشور قرار نداشت و این موضوع که هنجاریابی و تعیین نمرات هنجار در هر جامعه مورد تأکید است و شاید استفاده از نمرات هنجار آزمون اصلی در جوامع دیگر باعث اشتباه در برآورد سطح تکامل و خطا در تشخیص زودرس تأخیر تکاملی شود، این مطالعه برای هنجاریابی نسخه فارسی آزمون تکاملی آزمون بیلی ویرایش سوم در جهت تشخیص تأخیر تکاملی کودکان ایرانی طراحی شد. به منظور بررسی امکان استفاده از نمرات هنجار آمریکایی (نسخه اصلی آزمون) در ایران و مقایسه سطح تکامل کودکان دو نمونه، میانگین نمرات خام در دو گروه مستقل در حیطه‌های پنج‌گانه در ۴۸ گروه سنی از نظر معنادار بودن اختلاف نمرات ارزیابی شد.

روش بررسی

این مطالعه به روش مقطعی و با هدف هنجاریابی کشوری نسخه فارسی آزمون بیلی ویرایش سوم برای کودکان ایرانی صفر تا ۴۲ماهه اجرا شد. روش نمونه‌گیری پژوهش حاضر به‌صورت نمونه‌گیری طبقه‌ای تا تکمیل حجم نمونه بود.

ابتدا کشور به ۸ پهنه جغرافیایی شامل (مرکز، شمال، غرب، جنوب، جنوب‌شرق، جنوب‌غرب، شمال‌شرق، شمال‌غرب و تهران) تقسیم شد، سپس به منظور اجرای طرح، ۱۲ دانشگاه علوم‌پزشکی در ۱۰ شهر، شامل دانشگاه‌های علوم‌پزشکی در شهرهای تهران (۳ دانشگاه تهران، شهید بهشتی و ایران)، اصفهان، مشهد، رشت، تبریز، همدان، اهواز، شیراز، قم و زاهدان انتخاب شدند. از هر شهر نیز ۲ مرکز بهداشتی درمانی بر اساس امکان دسترسی مراکز دیگر بهداشتی به آن و سهولت دسترسی مراجعین انتخاب شد. از هر شهر ۲ نفر کارشناس مرکز بهداشت با حداقل مدرک کارشناسی و حداقل ۵ سال سابقه کار با کودکان پس از مصاحبه برای انجام آزمون انتخاب شدند و در کل ۲۸ کارشناس آموزش دیدند.

مربیان به آزمونگران در یک کارگاه یک روزه تئوری «مبانی آزمون تشخیصی تکاملی بیلی» و سه روزه عملی کشوری در تهران آموزش دادند. آزمونگران سپس به منظور کسب مهارت‌های لازم به شهر خود بازگشتند و اجرای آزمون را تمرین کردند. این فاز بسته به زمان مورد نیاز برای تمرین در شهرهای مختلف حدود ۲ تا ۴ ماه طول کشید. پس از کسب آمادگی جهت اجرای آزمون، آزمونگران با ارسال ۱۰ فیلم اجرای آزمون در گروه‌های سنی متفاوت، توسط مربیان کشوری ارزیابی شدند.

کار اجرای پایلوت آزمون بر روی ۱۲۸۵ آزمودنی به مدت ۴ تا ۸ ماه در شهرهای منتخب انجام شد. پس از بررسی‌هایی شامل مهارت آزمونگر در اجرای آزمون، مهارت در ارتباط با کودک و والدین، بررسی واحد پذیرش مرکز نسبت به رعایت پراکندگی نوبت‌دهی از تمام مناطق شهری، شرایط انجام و محل آزمون در مراکز توسط کارشناسان کشوری، کار نمونه‌گیری بر روی تعداد مشخص‌شده در هر شهر آغاز شد. این فاز بسته به جمعیت مورد نیاز در شهرهای مختلف بین ۳ تا ۶ ماه طول کشید.

نمونه شامل ۱۷۰۰ کودک ۱۶ روزه تا ۴۳ ماه و ۱۵ روز شامل ۱۷ گروه سنی و هر گروه شامل ۱۰۰ نمونه بود. گروه‌های سنی در سنین ۱ تا ۶ ماه به فواصل ۱ ماه، در سنین ۶ تا ۱۲ ماه به فواصل هر ۲ ماه، در سنین ۱۲ تا ۳۰ ماه به فواصل ۳ ماه و در سنین ۳۰ تا ۴۲ ماه به فواصل ۶ ماه دسته‌بندی شدند. این حجم از نمونه بر اساس آمار و سرشماری نفوس و مسکن سال ۱۳۹۰ به نسبت جمعیت زیر ۴ سال هر شهر محاسبه شد [۳۰].

در جدول شماره ۱، تعداد کودکان زیر ۴ سال در شهرهای منتخب بر اساس سرشماری شبکه ملی آمار و تعداد شرکت‌کنندگان در نمونه نهایی نمایش داده شده است. نمونه‌ها بر اساس عوامل جمعیت‌شناختی شامل سن، جنس و منطقه جغرافیایی طبقه‌بندی شدند و از مراکز بهداشتی درمانی، شامل کودکان با تکامل طبیعی و معیارهای ورود مورد نظر جمع‌آوری شدند. تکامل طبیعی به‌صورت «هر کودکی که

بدون عوارض پزشکی قابل توجه متولد شده باشد، تاریخچه عوارض پزشکی نداشته باشد و در حال حاضر تشخیص یا درمانی به علت مشکل شناختی، جسمی یا رفتاری دریافت نکند»، تعریف شد. معیارهای خروج شامل دریافت آزمایشات توان‌بخشی، عدم توانایی درک یا صحبت کردن به زبان فارسی و اختلالات حسی چون مشکلات شنوایی یا بینایی بود.

سپس با دریافت رضایت‌نامه کتبی از والدین و پر کردن پرسش‌نامه جمعیت‌شناختی (شامل سن و جنس کودک، سطح تحصیلات والدین، سن حاملگی هنگام تولد، وزن تولد، مشکلات هنگام تولد و تاریخچه پزشکی)، آزمون اجرا و در صورت مشخص‌شدن تأخیر تکامل کودک برای اقدامات تشخیصی و آموزش ارتقای سلامت به مرکز جامع تکامل کودکان معرفی شد.

ابزار و نحوه جمع‌آوری داده‌ها

آزمون بیلی ویرایش سوم یک ابزار ارزیابی فردی است که عملکرد تکاملی را در ۵ حیطه شناختی (شامل ۹۱ گزینه)، ارتباط درکی (۴۹ گزینه)، ارتباط بیانی (۴۸ گزینه)، حرکات ظریف (۶۶ گزینه) و حرکات درشت (۷۲ گزینه) ارزیابی می‌کند. اطلاعات هنجار این آزمون در آمریکا از نمونه طبقه‌بندی شده با حجم ۱۷۰۰ نمونه به دست آمده است [۲۳].

ارزیابی هر ۵ حیطه به شکل عینی[9] و توسط آزمونگر انجام می‌شود. انجام این آزمون در کودکان زیر یک‌سال حدود ۵۰ دقیقه و در کودکان بزرگ‌تر حدود ۹۰ دقیقه طول می‌کشد. طیف سنی بررسی‌شده در این آزمون، کودکان صفر تا ۴۲ماهه هستند که جهت اجرا در ۱۷ گروه سنی A تا Q دسته‌بندی شده است و برای افزایش دقت، نمرات هنجار در ۴۸ گروه سنی تعیین می‌شود. این آزمون، ۴ گروه از نمرات هنجار را به شرح زیر به‌دست می‌دهد: نمرات تراز‌شده[10]، نمرات مرکب[11]، رتبه‌های صدکی[12] و نمرات رشدی[13]. به علاوه، فواصل اطمینان[14] برای نمرات این مقیاس‌ها فراهم بوده است و معادل‌های سن تکاملی[15] نیز برای مقیاس‌های مربوطه در دسترس هستند.

ابزار این پژوهش، نسخه فارسی آزمون بیلی ویرایش سوم بود که در ایران توسط آذری و همکاران، روایی و اعتبار آن به دست آمده است [۳۱]. در روان‌سنجی نسخه فارسی آزمون، میانگین ضرایب آلفای کرونباخ برای تمام حیطه‌ها بالاتر از ۰/۷۶ و ضریب همبستگی پیرسون در بازآزمایی و اعتبار ارزیاب‌ها در حیطه‌های مختلف، بین ۰/۹۸۷ (P<۰/۰۰۱) و ۰/۹۹۱ (P<۰/۰۰۱) بود.

روایی صوری و محتوایی با استفاده از نظرات خبرگان تأیید شده بود. روایی سازه هم توسط روش‌های تحلیل عاملی و مقایسه میانگین نمرات تأیید شده بود [۳۱].

روش تجزیه‌وتحلیل داده‌ها

توصیف داده‌ها، شامل جداول توزیع تعداد و درصد گروه آزمایش بر حسب متغیرهای مورد بررسی و تعیین شاخص‌های گرایش مرکزی و پراکندگی داده‌ها به تفکیک زیرگروه‌های ۴۸گانه سنی و حیطه‌های آزمون بود. برای هنجاریابی، ابتدا طبیعی بودن توزیع داده‌ها بررسی شد. در صورتی که داده‌های پژوهش دارای توزیع نرمال بودند و در مواردی که سطح معناداری آزمون بزرگ‌تر از ۰/۰۵ بود، فرض صفر مبنی بر عدم تفاوت بین توزیع داده‌ها و توزیع طبیعی تأیید شد، اما در هر توزیع طبیعی لزوماً میانگین در نمره‌های تراز‌شده برابر ۱۰۰ و در رتبه درصدی برابر ۵۰ قرار نمی‌گرفت. بنابراین جداول نمرات مرکب، رتبه‌های صدکی (درصدی) و فاصله اطمینان برای نمره‌های تراز‌شده به دو صورت مشخص شد:

در ابتدا توزیع در گروه‌های سنی ۴۸گانه بر اساس نمونه موجود تدوین شد که با توجه به حجم نمونه و انتخاب آزمودنی‌ها، معرف واقعی جامعه از نظر توزیع این متغیرها و نشان‌دهنده دقت بیشتر در تبدیل نمرات تراز‌شده به نمره مرکب بود. سپس صرفاً با تبدیل یک نمره استاندارد با میانگین ۱۰ و انحراف معیار ۳ به ۱ نمره استاندارد دیگر با میانگین ۱۰۰ و انحراف معیار ۱۵ یا تبدیل آن به توزیع درصدی با میانگین ۵۰ تدوین شد.

برای تعیین نمرات هنجار از روش آزمون اصلی، یعنی تبدیل نمرات خام به نمرات تراز، مرکب و رشدی، رتبه‌های صدکی و معادل سن تکاملی در گروه‌های سنی ۴۸گانه از ۱۵روزگی تا ۴۲ ماهگی استفاده شد. نمرات تراز‌شده از نمرات خام مقیاس‌ها به‌دست آمد که با تبدیل خطی به نمراتی در محدوده ۱۹-۱ با میانگین ۱۰ و انحراف معیار ۳ شد. نمرات مرکب از مجموع نمرات تراز‌شده مقیاس‌ها مشتق شدند. نمرات مرکب در سیستم متریکی با میانگین ۱۰۰ و انحراف معیار ۱۵ و محدوده ۴۰ تا ۱۶۰ تنظیم شدند. نمرات صدکی در محدوده ۱ تا ۹۹ با میانگین و میانه ۵۰ تبدیل شدند.

رتبه‌های صدکی (با توزیع درصدی با میانگین ۵۰) و فاصله اطمینان ۹۰ و ۹۵ درصدی برای مقیاس‌های شناختی، زبانی و حرکتی مشخص شد. نمرات رشدی از تبدیل نمرات خام به نمراتی در محدوده ۲۰۰ تا ۸۰۰ با میانگین ۵۰۰ و انحراف معیار ۱۰۰ به‌دست آمد که بر اساس آن‌ها منحنی‌های رشدی برای نشان‌دادن رشد کودک در طول زمان برای مقیاس‌های شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت در صدک‌های ۵، ۱۰، ۲۵، ۵۰، ۷۵، ۹۰ و ۹۵ ترسیم شد.

9. Objective
10. Scaled Score
11. Composite Score
12. Percentile Rank
13. Growth Score
14. Confidence Interval
15. Developmental Age Equivalent

معادل‌های سن تکامل از نمرات خام به‌دست‌آمده و نشان‌دهنده سن متوسطی است که در آن سن نمره خام کودک به‌شکل معمول در هنجارهای سنی هر مقیاس دیده می‌شود. برای تعیین معادل سن تکاملی، جداولی ترسیم شد که در آن نمره خامی که معادل نمره تراز ۱۰ برای هر گروه سنی در هر مقیاس بود، مشخص شد.

به منظور بررسی امکان استفاده از نمرات هنجار نسخه اصلی در کودکان ایرانی و مقایسه سطح تکامل کودکان آمریکایی و ایرانی، میانگین نمرات خام آن‌ها از طریق آزمون میانگین دو گروه مستقل در حیطه‌های پنج‌گانه در ۴۸ گروه سنی بررسی و از نظر معنادار بودن، اختلاف نمرات ارزیابی شد (P<۰/۰۵).

یافته‌ها

در این مطالعه به منظور هنجاریابی آزمون بیلی ویرایش سوم در ایران، آزمون بر روی ۱۷۴۴ کودک صفر تا ۴۲ماهه اجرا و نمرات هنجار در ۴۸ گروه سنی تعیین شد (برای تعیین نمرات هنجار حیطه زبانی تنها از اطلاعات ۱۶۰۳ نمونه آزمودنی که فارسی‌زبان بودند، استفاده شد). تعداد دختران ۹۰۸ (۵۲/۱ درصد) و تعداد پسران ۸۳۶ (۴۷/۹ درصد) نفر بود. در این نمونه، میزان تحصیلات والدین (شامل دیپلم و کمتر از آن، هرگونه تحصیلات دانشگاهی تا سطح کارشناسی، کارشناسی ارشد و بالاتر) در پدران به ترتیب ۶۴/۵، ۳۴/۲ و ۱/۳ درصد و در مادران ۵۸، ۴۱ و ۱ درصد بود.

در جدول شماره ۲، متوسط نمرات خام و انحراف معیار در مقیاس‌های شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت در نمونه هنجارشده بر حسب زیرگروه‌های سنی هفده‌گانه آزمون بیلی ویرایش سوم ارائه شده است.

نمرات هنجار در گروه‌های سنی ۴۸گانه (بر اساس نمرات ترازشده در طیف ۱ تا ۱۹ با میانگین ۱۰ و انحراف معیار ۳)، نمرات رشدی، معادل سن تکاملی برای مقیاس‌های شناختی، ارتباط درکی و بیانی، حرکات درشت و ظریف از دیگر یافته‌های این مطالعه است که در جداول شماره ۳، ۴، ۵، ۶ و ۷ بخشی از آن‌ها به‌صورت نمونه ارائه شده است.

در جدول شماره ۳، تعداد، میانگین، انحراف معیار، کمینه و بیشینه در مقیاس‌های شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت در گروه سنی ۴۸ (۳۹ ماه و ۱۶ روز و نیز ۴۲ ماه و ۱۵ روز) به‌عنوان نمونه آمده است.

در جدول شماره ۴، تبدیل نمره خام به نمره تراز در مقیاس‌های شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت در گروه سنی ۲۳ (۱۲ ماه و ۱۶ روز و نیز ۱۳ ماه و ۱۵ روز) آزمون بیلی ویرایش سوم به‌صورت نمونه ارائه شده است.

در جدول شماره ۵، تبدیل نمره تراز به مرکب، رتبه‌های صدکی و فاصله اطمینان بخشی از نمره‌های ترازشده مقیاس شناختی به‌صورت نمونه آمده است. در جدول شماره ۶، بخشی از تبدیل نمره خام به نمره رشدی در آزمون بیلی ویرایش سوم به‌صورت نمونه آمده است. در جدول شماره ۷، معادل سن تکاملی برای

جدول ۱. تعداد شرکت‌کنندگان در نمونه هنجار بر حسب شهرها منتخب

ردیف	شهر	*جمعیت	تعداد (درصد)
			شرکت‌کنندگان در نمونه هنجارشده
۱	همدان	۴۵۵۰۰	۵۰(۲/۹)
۲	رشت	۵۱۲۳۹	۵۷(۳/۳)
۳	اصفهان	۱۴۷۷۳۸	۱۶۱(۹/۲)
۴	قم	۱۰۴۳۸۷	۱۱۴(۶/۵)
۵	مشهد	۲۶۸۲۷۸	۲۹۵(۱۶/۹)
۶	شیراز	۱۱۷۹۸۷	۱۳۰(۷/۵)
۷	زاهدان	۹۰۰۴۷	۹۸(۵/۶)
۸	اهواز	۱۳۵۲۳۱	۱۴۸(۸/۵)
۹	تبریز	۱۲۶۰۶۱	۱۴۱(۸/۱)
۱۰	تهران	۴۷۶۲۱۰	۵۵۰(۳۱/۵)
	جمع	۱۵۶۲۶۷۸	۱۷۴۴(۱۰۰)

*کودکان صفر تا ۴ سال بر اساس سرشماری شبکه ملی آمار سال ۱۳۹۰

توانبخشی

جدول ۲. متوسط نمرات خام (انحراف معیار) گروه‌های سنی در نمونه هنجار

	گروه سنی	تعداد (درصد)	متوسط نمرات خام(انحراف معیار)				
			شناختی	درکی	بیانی	ظریف	درشت
A	۱۶ روز–۱ ماه و ۱۵ روز	۱۰۴(۶)	۵/۵۶(۳/۱۷)	۳/۶۱(۱/۶۹)	۱/۶۱(۰/۷)	۳/۶۱(۱/۸۵)	۳/۹۴(۱/۴۳)
B	۱ ماه و ۱۶ روز–۲ ماه و ۱۵ روز	۱۰۰(۵/۷)	۸/۷۰(۲/۹۹)	۴/۴۰(۱/۱۴)	۳/۷۵(۱/۲۵)	۵/۵۳(۱/۶۸)	۸/۴۵(۳/۲۰)
C	۲ ماه و ۱۶ روز–۳ ماه و ۱۵ روز	۱۰۳(۵/۹)	۱۱/۵۰(۲/۸۸)	۶/۶۲(۲/۸۲)	۵/۵۴(۳/۰۷)	۶/۶۷(۲/۴۳)	۱۱/۳۳(۳/۸۵)
D	۳ ماه و ۱۶ روز–۴ ماه و ۱۵ روز	۹۹(۵/۷)	۱۴/۷۴(۴/۰۳)	۷/۰۵(۱/۱۸)	۵/۳۴(۱/۱۷)	۱۱/۴۶(۲/۵۲)	۱۶/۴۵(۲/۷۵)
E	۴ ماه و ۱۶ روز–۵ ماه و ۱۵ روز	۱۰۰(۵/۷)	۲۳/۴۰(۲/۹۵)	۸/۸۰(۲/۴۴)	۶/۷۰(۱/۷۰)	۱۷/۸۰(۲/۵۷)	۲۰/۲۰(۳/۰۱)
F	۵ ماه و ۱۶ روز–۶ ماه و ۱۵ روز	۶۱(۳/۵)	۲۶/۱۵(۳/۳۴)	۹/۶۵(۱/۹۲)	۷/۱۲(۲/۲۳)	۱۹/۶۲(۲/۶۱)	۲۴/۶۲(۳/۵۰)
G	۶ ماه و ۱۶ روز–۸ ماه و ۳۰ روز	۸۹(۵/۱)	۳۰/۱۰(۲/۰۱)	۱۱/۰۳(۱/۸۰)	۹/۲۵(۲/۶۲)	۲۱/۶۹(۱/۸۹)	۲۹/۶۹(۴/۴۱)
H	۹ ماه و صفر روز–۱۰ ماه و ۳۰ روز	۸۴(۴/۸)	۳۳/۹۷(۳/۱۰)	۱۲/۰۶(۱/۸۱)	۹/۸۴(۲/۵۷)	۲۵/۸۷(۲/۲۳)	۳۴/۶۸ (۳/۳۴)
I	۱۱ ماه و صفر روز–۱۳ ماه و ۱۵ روز	۹۷(۵/۶)	۴۰/۷۱(۳/۷۹)	۱۴/۹۷(۲/۵۶)	۱۴/۱۶(۳/۱۳)	۲۸/۰۶(۲/۵۲)	۴۲/۶۳(۴/۳۷)
J	۱۳ ماه و ۱۶ روز–۱۶ ماه و ۱۵ روز	۱۱۷(۶/۷)	۴۸/۹۱(۳/۹۱)	۱۹/۴۹(۲/۹۰)	۱۹/۵۴(۳/۳۸)	۳۲/۳۴(۲/۴۵)	۵۰/۷۱(۳/۶۴)
K	۱۶ ماه و ۱۶ روز–۱۹ ماه و ۱۵ روز	۱۱۱(۶/۴)	۵۲/۴۷(۴/۲۰)	۲۰/۸۷(۲/۹۷)	۲۱/۲۸(۴/۲۷)	۳۴/۵۳(۲/۲۴)	۵۲/۱۹(۳/۴۶)
L	۱۹ ماه و ۱۶ روز–۲۲ ماه و ۱۵ روز	۱۰۳(۵/۹)	۵۸/۵۰(۴/۰۶)	۲۵/۲۰(۲/۹۳)	۲۶/۸۵(۳/۲۷)	۳۷/۸۵(۳/۳۵)	۵۵/۴۰(۲/۱۶)
M	۲۲ ماه و ۱۶ روز–۲۵ ماه و ۱۵ روز	۱۰۴(۶/۰)	۶۲/۵۰(۶/۰۷)	۳۰/۲۲(۳/۸۶)	۳۲/۷۸(۶/۵۰)	۴۰/۱۱(۳/۰۱)	۵۷/۲۸(۴/۰۶)
N	۲۵ ماه و ۱۶ روز–۲۸ ماه و ۱۵ روز	۱۰۸(۶/۲)	۶۵/۸۲(۵/۵۷)	۳۱/۷۱(۵/۱۹)	۳۱/۷۶(۷/۴۰)	۴۱/۶۷(۴/۴۱)	۵۸/۸۱(۳/۶۳)
O	۲۸ ماه و ۱۶ روز–۳۲ ماه و ۳۰ روز	۱۵۱(۸/۷)	۷۰/۳۹(۵/۵۴)	۳۵/۲۶(۴/۲۱)	۳۸/۵۲(۶/۱۹)	۴۴/۸۲(۴/۱۶)	۶۱/۲۷(۳/۳۷)
P	۳۳ ماه و صفر روز–۳۸ ماه و ۳۰ روز	۱۴۶(۸/۴)	۷۴/۷۲(۴/۹۳)	۳۸/۲۴(۴/۴۲)	۴۱/۴۱(۴/۷۶)	۵۰/۲۳(۴/۰۳)	۶۳(۴/۴۵)
Q	۳۹ ماه و صفر روز–۴۲ ماه و ۱۵ روز	۶۷(۳/۸)	۷۶/۸۳(۴/۶۹)	۴۲/۳۹(۴/۴۱)	۴۳(۷/۰۱)	۵۳/۸۸(۴/۱۲)	۶۶/۵۰(۲/۶۶)

توانبخشی

نمرات خام در مقیاس‌ها در چند نمونه گروه سنی ارائه شده است.

در جدول شماره ۸، مقایسه میانگین نمرات دارای تفاوت معنادار در ۲ نمونه کودکان آمریکایی و ایرانی آمده است. طبق جدول شماره ۸، میانگین نمرات در ۲۸ گروه سنی/زیرمقیاس بین ۲ نمونه متفاوت بود. تفاوت در ۲ گروه سنی در مقیاس شناختی، ۱۲ گروه در مقیاس درکی، ۳ گروه در مقیاس بیانی، ۴ گروه در مقیاس حرکات ظریف و ۷ گروه در مقیاس حرکات درشت مشخص شد. در این ۲۸ گروه سنی/مقیاس در ۷ گروه میانگین نمرات نمونه آمریکایی بالاتر از ایرانی بود که در طیف

جدول ۳. ویژگی‌های نمونه هنجار در گروه سنی ۴۸ (۳۹ ماه و ۱۶ روز – ۴۲ ماه و ۱۵ روز) در مقیاس‌های آزمون بیلی ویرایش سوم

مقیاس	تعداد	میانگین±انحراف معیار	کمینه	بیشینه	انحراف معیار			میانگین	انحراف معیار		
					+۳	+۲	+۱		-۱	-۲	-۳
شناختی	۵۶	۷۵/۹۸±۴/۹۳	۶۳	۸۶	۹۰/۷۷	۸۵/۸۴	۸۰/۹۱	۷۵/۹۸	۷۱/۰۵	۶۶/۱۲	۶۱/۱۹
ارتباط درکی	۵۲	۳۹/۷۷±۳/۹۳	۳۱	۴۸	۴۹/۰۰	۴۷/۶۳	۴۳/۷۰	۳۹/۷۷	۳۵/۸۴	۳۱/۹۱	۲۷/۹۸
ارتباط بیانی	۵۲	۴۲/۸۲±۳/۴۰	۳۵	۴۸	——	۴۸/۰۰	۴۶/۲۱	۴۲/۸۱	۳۹/۴۱	۳۶/۰۱	۳۲/۶۱
حرکات ظریف	۵۶	۵۴/۵۲±۴/۳۳	۴۱	۶۵	۶۶/۰۰	۶۳/۱۸	۵۸/۸۵	۵۴/۵۲	۵۰/۱۹	۴۵/۸۶	۴۱/۵۳
حرکات درشت	۵۶	۶۶/۷۵±۲/۹۹	۵۸	۷۲	——	۷۲/۰۰	۶۹/۷۴	۶۶/۷۵	۶۳/۷۶	۶۰/۷۷	۵۷/۷۸

توانبخشی

جدول ۴. تبدیل نمرات خام به نمرات تراز در گروه سنی ۲۳ (۱۲ ماه و ۱۶ روز۱۳– ماه و ۱۵ روز) در مقیاس‌های آزمون بیلی ویرایش سوم

نمره تراز	میانگین (انحراف معیار)	شناختی	ارتباط درکی	ارتباط بیانی	حرکات ظریف	حرکات درشت
۱	-3SD	۰–۳۱	۰–۸	۰–۵	۰–۲۱	۰–۲۶
۴	-2SD	۳۵	۱۰	۸	۲۴	۳۱–۳۲
۷	-1SD	۳۹–۴۰	۱۳	۱۲	۲۷	۳۷–۳۸
۱۰	M	۴۳–۴۴	۱۵	۱۵	۲۹	۴۲–۴۳
۱۳	+1SD	۴۷–۴۸	۱۸	۱۸–۱۹	۳۲	۴۸–۴۹
۱۶	+2SD	۵۲	۲۰	۲۲	۳۴	۵۴–۵۵
۱۹	+3SD	۵۶–۹۱	۲۳–۴۹	۲۵–۴۸	۳۷–۶۶	۵۹–۷۲

توانبخشی

سنی ۱۵ روز تا ۶ ماهگی قرار داشت و شامل گروه سنی یک در زیرمقیاس شناختی، ۱، ۴ و ۶ در زیرمقیاس ارتباط درکی و ۱۴، ۱۵ و ۱۹ در زیرمقیاس حرکات درشت بودند.

در ۲۱ گروه، میانگین نمرات در نمونه ایرانی بالاتر از نمونه آمریکایی در طیف سنی ۶ تا ۴۲ ماه به‌شکل؛ گروه ۲۴ در زیرمقیاس شناختی، گروه ۱۶، ۲۳، ۲۴، ۲۵، ۲۶، ۲۹، ۳۴، ۴۰ و ۴۸ در زیرمقیاس ارتباط درکی و گروه ۵، ۴۲ و ۴۳ در زیرمقیاس ارتباط بیانی، گروه ۲۵، ۴۲، ۴۳ و ۴۷ در زیرمقیاس حرکات ظریف و گروه ۲۵، ۲۶، ۴۴ و ۴۸ در زیرمقیاس حرکات درشت بود.

بحث

در این پژوهش نسخه فارسی «آزمون تشخیصی مقیاس‌های تکاملی بیلی ویرایش سوم» در یک نمونه ملی ۱۷۴۴ نفری از کودکان صفر تا ۴۲ماهه در ایران هنجار شد. نمرات هنجار (تراز، مرکب و رشدی) محاسبه و رتبه صدکی و معادل سن تکاملی و همچنین منحنی‌های رشدی کودکان ایرانی مشخص شد.

میانگین نمرات ۲۸ گروه سنی/مقیاسی بین دو جامعه مورد بررسی متفاوت بود (P<۰/۰۵) که ۷ مورد میانگین نمرات در نمونه آمریکایی بالاتر از نمونه ایرانی بود (طیف سنی ۱۵ روزه تا ۶ ماهه) و در ۲۱ مورد میانگین نمرات در نمونه ایرانی بالاتر از نمونه آمریکایی بود (طیف سنی بالاتر از ۶ ماه). برای مثال، در گروه سنی یک در مقیاس شناختی با تفاوت میانگین نمرات خام ۱/۷۸ (P<۰/۰۱) و در ارتباط درکی ۰/۸۹ (P<۰/۰۵) نمرات نمونه آمریکایی بالاتر از ایرانی بود و در گروه سنی ۲۵ در مقیاس حرکات درشت با تفاوت میانگین نمرات خام ۲/۹۸ (P<۰/۰۱) و در حرکات ظریف ۱/۷۳ (P<۰/۰۱) نمرات نمونه ایرانی بالاتر از آمریکایی بود.

در این مطالعه برای نمونه‌گیری فقط از کودکان با تکامل طبیعی استفاده شد. در نسخه هنجارشده بیلی ویرایش سوم به‌دلیل اینکه نمونه کاملاً نماینده جمعیت عمومی باشد، حدود ۱۰ درصد از

جدول ۵. تبدیل نمره تراز به نمره مرکب، رتبه‌های صدکی و تعیین فاصله اطمینان در مقیاس شناختی آزمون بیلی ویرایش سوم

نمره تراز	میانگین (انحراف معیار)	نمره مرکب	صدک	فاصله اطمینان	
				۹۰ درصد	۹۵ درصد
۱	-3SD	۵۵	۰/۱	۵۲–۶۶	۵۱–۶۷
۴	-2SD	۷۰	۲	۶۶–۷۹	۶۵–۸۱
۷	-1SD	۸۵	۱۶	۸۰–۹۳	۷۸–۹۴
۱۰	M	۱۰۰	۵۰	۹۳–۱۰۷	۹۲–۱۰۸
۱۳	+1SD	۱۱۵	۸۴	۱۰۷–۱۲۰	۱۰۶–۱۲۲
۱۶	+2SD	۱۳۰	۹۸	۱۲۱–۱۳۴	۱۲۹–۱۳۵
۱۹	+3SD	۱۴۵	۹۹/۹	۱۳۴–۱۴۸	۱۳۳–۱۴۹

توانبخشی

جدول ۶. تبدیل نمره خام به نمره رشدی در مقیاس‌های آزمون بیلی ویرایش سوم

نمره رشدی					نمره خام
حرکات درشت	حرکات ظریف	ارتباط بیانی	ارتباط درکی	شناختی	
۳۱۴	۳۱۷	۳۶۱	۳۳۷	۳۱۵	۰
۳۱۸	۲۳۲	۳۶۸	۳۴۵	۳۱۹	۱
۳۲۳	۳۳۰	۳۷۵	۳۵۴	۳۲۳	۲
۵۳۳	۶۱۴	۶۹۷	۷۴۰	۵۱۱	۴۶
۵۳۷	۶۲۰	۷۰۴	۷۴۹	۵۱۶	۴۷
۵۴۲	۶۲۷	۷۱۲	۷۵۸	۵۲۰	۴۸

توانبخشی

نمونه به کودکان با نیازهای ویژه (شامل تریزومی ۲۱، فلج مغزی، اختلالات طیف اوتیسم، نارسی، اختلالات خاص گفتاری، مصرف الکل مادر در دوران بارداری، اختناق یا آسفیکسی[16] هنگام تولد، وزن پایین نسبت به سن تولد) اختصاص یافته بود [۲۳].

تحقیقات قبلی مشخص کرد وارد کردن کودکان با نیازهای ویژه در یک نمونه استاندارد می‌تواند بر توانایی افتراقی بین تکامل طبیعی و تأخیر تکامل تأثیر منفی بگذارد و باعث پایین آوردن میانگین نمرات شده و توانایی ابزار را برای تشخیص کودکان دچار تأخیر تکاملی خفیف محدود کند [۳۲]. به همین دلیل در هنجاریابی در ایران از این روش استفاده نشد. شاید انتخاب کودکان طبیعی در ایران سبب بالا رفتن میانگین نمرات و تفاوت میانگین نمرات بین نمونه آمریکایی و ایرانی شده است یا تفاوت در ویژگی‌های بیولوژیکی، محیطی و فرهنگی موجب این تفاوت باشد که در مطالعات قبلی در کشورهای دیگر نیز همین تفاوت مشاهده شده است [۳۴-۲۶].

مطالعات هنجاریابی در آزمون‌های تکاملی کودکان در سطح کشوری محدود هستند و فقط پرسش‌نامه سنین و مراحل در ایران هنجار شده است [۲۳]. پرسش‌نامه سنین و مراحل یک پرسش‌نامه غربالگر است. دارای حیطه‌های ارتباطی، حرکات ظریف و درشت، حل مسئله و شخصی اجتماعی در سنین ۴ تا ۶۰ ماهگی است که با سؤالات از والدین تکمیل می‌شود.

در این مطالعه همچنین مشخص شد میانگین نمرات کودکان ایرانی در حیطه‌های حرکات ظریف، حل مسئله و شخصی اجتماعی، به‌ویژه در سنین ۸ و ۱۲ ماهگی بالاتر و در حیطه حرکتی، به‌ویژه در ۳۶ ماهگی کمتر از نمونه آمریکایی است (P<۰/۰۵) [۲۴]. در مطالعه ما نیز میانگین نمرات در تمام مقیاس‌های گروه سنی بالای ۶ ماه در کودکان ایرانی نسبت به نمونه آمریکایی بالاتر بود.

در بیشتر مطالعات انجام‌شده با آزمون بیلی ویرایش سوم، تعداد نمونه‌ها کم است و در بعضی سنین و مقیاس‌ها انجام شده و با میانگین نمرات آمریکایی مقایسه شده است. تنها در کشور هلند همانند این مطالعه نمرات هنجار در ۴۸ گروه سنی و در مقیاس‌های شناختی، حرکتی و زبانی به‌دست آمد [۳۵]. نتایج این مطالعه نیز نشان داد میانگین نمرات کودکان هلندی و آمریکایی تفاوت دارد و استفاده از نمرات آمریکایی در کودکان هلندی منجر به عدم تشخیص زودهنگام تأخیر

16. Asphyxie

جدول ۷. معادل سن تکاملی نمرات خام مقیاس‌های آزمون بیلی ویرایش سوم

حرکات درشت	حرکات ظریف	ارتباط بیانی	ارتباط درکی	شناختی	سن رشدی(ماه و روز)	گروه سنی
۰-۴	۰-۳	۰-۲	۰-۳	۰-۳	۲۰ روز	۱
۱۶	۱۳	۷	۸	۱۸-۲۰	۴ ماه و ۱۰ روز	۱۲
۴۱-۴۲	۲۸	۱۳-۱۴	۱۵	۴۰-۴۱	۱۲ ماه	۲۲
۵۵	۳۶-۳۷	۲۷	۲۴-۲۵	۵۷	۲۲ ماه	۳۲
۶۲	۴۶	۳۸	۳۴	۷۰	۳۲ ماه	۴۲
۶۷	۵۳-۵۵	۳۹	۴۰	۷۶	۴۲ ماه	۴۸

توانبخشی

جدول ۸. مقایسه میانگین نمرات دارای اختلاف معنادار در دو نمونه کودکان آمریکایی و ایرانی در آزمون بیلی ویرایش سوم

گروه سنی	مقیاس	نمونه	تعداد	میانگین±انحراف معیار	P
۱	شناختی	ایالات متحده	۳۰	۶/۰۰±۳/۰۰	P<۰/۰۱
		ایران	۱۸	۴/۲۲±۱/۳۵	
	ارتباط درکی	ایالات متحده	۳۰	۵/۰۰±۱/۵۰	P<۰/۰۵
		ایران	۱۸	۴/۱۱±۱/۳۲	
۴	ارتباط درکی	ایالات متحده	۳۰	۵/۵۰±۱/۶۲	P<۰/۰۵
		ایران	۱۲	۴/۴۲±۱/۱۷	
۵	ارتباط بیانی	ایالات متحده	۳۰	۴/۰۰±۱/۷۵	P<۰/۰۵
		ایران	۳۰	۵/۲۰±۱/۶۷	
۶	ارتباط درکی	ایالات متحده	۳۰	۶/۰۰±۱/۶۲	P<۰/۰۵
		ایران	۲۱	۵/۰۰±۱/۴۱	
۱۴	حرکات درشت	ایالات متحده	۳۰	۲۰/۰۰±۳/۷۵	P<۰/۰۱
		ایران	۱۶	۱۷/۰۶±۲/۷۴	
۱۵	حرکات درشت	ایالات متحده	۳۰	۲۱/۰۰±۳/۸۷	P<۰/۵
		ایران	۱۹	۱۸/۳۲±۳/۳۳	
۱۶	ارتباط درکی	ایالات متحده	۳۰	۹/۵۰±۱/۸۷	P<۰/۰۱
		ایران	۸۸	۱۰/۸۱±۱/۶۵	
۱۹	حرکات درشت	ایالات متحده	۳۰	۳۵/۰۰±۳/۷۵	P<۰/۰۵
		ایران	۵۸	۳۲/۹۰±۴/۲۷	
۲۳	ارتباط درکی	ایالات متحده	۳۰	۱۵/۰۰±۲/۷۵	P<۰/۰۱
		ایران	۶۴	۱۷/۲۳±۲/۸۶	
۲۴	شناختی	ایالات متحده	۳۰	۴۴/۵۰±۴/۸۷	P<۰/۰۵
		ایران	۵۰	۴۶/۸۶±۳/۸۱	
	ارتباط درکی	ایالات متحده	۳۰	۱۶/۰۰±۳/۲۵	P<۰/۰۱
		ایران	۵۰	۱۸/۱۰±۲/۶۹	
۲۵	ارتباط درکی	ایالات متحده	۳۰	۱۷/۰۰±۳/۲۵	P<۰/۰۵
		ایران	۴۹	۱۸/۵۷±۲/۹۷	
	حرکات ظریف	ایالات متحده	۳۰	۳۱/۰۰±۳/۰۰	P<۰/۰۱
		ایران	۴۹	۳۲/۷۳±۲/۳۳	
	حرکات درشت	ایالات متحده	۳۰	۴۷/۰۰±۴/۰۰	P<۰/۰۱
		ایران	۴۹	۴۹/۹۸±۳/۶۹	

گروه سنی	مقیاس	نمونه	تعداد	میانگین±انحراف معیار	P
۲۶	ارتباط درکی	ایالات متحده	۳۰	۱۸/۰۰±۳/۵۰	p<۰/۰۵
		ایران	۶۲	۱۹/۶۱±۳/۲۹	
	حرکات درشت	ایالات متحده	۳۰	۴۸/۵۰±۴/۱۲	p<۰/۰۵
		ایران	۶۲	۵۰/۵۲±۴/۲۴	
۲۹	ارتباط درکی	ایالات متحده	۳۰	۲۱/۵۰±۴/۱۲	p<۰/۰۵
		ایران	۳۷	۲۳/۴۳±۳/۰۱	
۳۴	ارتباط درکی	ایالات متحده	۳۰	۲۶/۵۰±۵/۲۵	p<۰/۰۵
		ایران	۱۹	۲۹/۶۳±۳/۱۵	
۴۰	ارتباط درکی	ایالات متحده	۳۰	۳۱/۵۰±۶/۳۷	p<۰/۰۵
		ایران	۱۷	۳۴/۳۵±۲/۸۵	
۴۲	ارتباط بیانی	ایالات متحده	۳۰	۳۶/۵۰±۶/۱۲	p<۰/۰۵
		ایران	۱۶	۴۰/۴۴±۴/۷۷	
	حرکات ظریف	ایالات متحده	۳۰	۴۵/۰۰±۴/۸۷	p<۰/۰۵
		ایران	۱۶	۴۸/۳۱±۳/۸۶	
۴۳	ارتباط بیانی	ایالات متحده	۳۰	۳۷/۰۰±۶/۶۲	p<۰/۰۵
		ایران	۲۳	۴۰/۷۴±۵/۰۷	
	حرکات ظریف	ایالات متحده	۳۰	۴۵/۵۰±۵/۰۰	p<۰/۰۵
		ایران	۲۳	۴۷/۹۱±۳/۰۱	
۴۴	حرکات درشت	ایالات متحده	۳۰	۶۲/۵۰±۴/۰۰	p<۰/۰۵
		ایران	۲۴	۶۴/۳۸±۲/۷۶	
۴۷	حرکات ظریف	ایالات متحده	۳۰	۴۹/۵۰±۵/۸۷	p<۰/۰۵
		ایران	۷۴	۵۱/۹۵±۳/۵۷	
۴۸	ارتباط درکی	ایالات متحده	۳۰	۳۸/۰۰±۶/۶۲	p<۰/۰۱
		ایران	۷۱	۴۱/۶۲±۴/۲۹	
	حرکات درشت	ایالات متحده	۳۰	۶۵/۵۰±۳/۷۵	p<۰/۰۵
		ایران	۷۱	۶۷/۱۸±۲/۴۰	

تکاملی می‌شود و درنتیجه محققین به هنجاریابی آزمون در هر جمعیت تأکید کردند.

کِرومول و همکاران طی مطالعه‌ای روایی آزمون بیلی ویرایش سوم را در کشور مالی بررسی کردند. آن‌ها این آزمون را بر روی ۱۶۷ کودک سالم مالاوی انجام دادند و ضمن استخراج نمرات هنجار، از نمره z[17] جهت دسته‌بندی تأخیر تکاملی استفاده کردند. نتایج به‌دست‌آمده حاکی از این بود که میانگین نمرات کودکان مالاوی در سنین کمتر از ۶ ماه در تمام مقیاس‌ها از میانگین نمرات هم‌سالان آمریکایی خود بالاتر بود، اما بعد از شش‌ماهگی کودکان آمریکایی میانگین نمرات بالاتری نشان می‌دادند.

17. Score

مطابق نتیجه‌گیری محققان، نمرات هنجار آزمون در جمعیت‌های مختلف، متفاوت است. تفسیر آزمون با استناد به نمرات هنجار آمریکایی در کودکان مالاوی منجر به اشتباه در دسته‌بندی تأخیر تکاملی به‌ویژه در حیطه‌های شناختی و ارتباطی می‌شود [۲۶].

مور طی پژوهشی به منظور مقایسه نمرات شاخص رشد شناختی[18] مربوط به آزمون بیلی ویرایش دوم و نمرات شناختی و ارتباطی ازمون بیلی ویرایش سوم و بررسی توافق بین دو آزمون در تشخیص تأخیر تکاملی و معرفی الگویی که طی آن نمرات دو آزمون قابل تبدیل به یکدیگر باشند، ۱۸۵ کودک ۲۹ تا ۴۱ماهه را با مقیاس شاخص رشد شناختی آزمون بیلی ویرایش دوم و مقیاس شناختی و ارتباطی ازمون بیلی ویرایش سوم به‌صورت هم‌زمان ارزیابی کرد. این پژوهشگر نمره ترکیبی متوسط حیطه‌های شناختی و ارتباطی آزمون بیلی ویرایش سوم تحت عنوان نمره ترکیبی بیلی ویرایش سوم[19] را برای مقایسه با نمره شاخص رشد شناختی آزمون بیلی ویرایش دوم استفاده کرد.

نتایج نشان داد نمرات حیطه‌های شناختی، ارتباطی و ترکیبی، به ترتیب ۳، ۱۰ و ۷ امتیاز بالاتر از نمره شاخص رشد شناختی، هستند، اما ارتباط آن‌ها به‌صورت خطی نبود. نمرات ترکیبی بیلی ویرایش سوم نسبت به شاخص رشد شناختی بالاتر بود و سطح تکامل حیطه‌های شناختی و ارتباطی کودک توسط آزمون بیلی ویرایش سوم بالاتر از آزمون بیلی ویرایش دوم ارزیابی شده و در صورتی که از نمره ترکیبی بیلی ویرایش سوم کمتر از ۷۰ برای تشخیص تأخیر تکاملی استفاده می‌شد، نمره ترکیبی بیلی ویرایش سوم در قیاس با نمره شاخص رشد شناختی، تأخیر تکاملی را کمتر از واقع (با حساسیت ۵۸ درصد و ویژگی ۱۰۰ درصد) برآورد می‌کرد، اما در صورتی که نمره ترکیبی بیلی ویرایش سوم کمتر از ۸۰ استفاده می‌شد، حساسیت به ۸۹ درصد و ویژگی به ۹۹ درصد می‌رسید و ارزش پیشگویی یا پیش‌بینی[20] افزایش می‌یافت.

این محققان نتیجه‌گیری کردند که در تفسیر نمرات آزمون بیلی ویرایش سوم برای تشخیص تأخیر تکاملی باید محتاط بود و توصیه کردند نمره ترکیبی بیلی ویرایش سوم کمتر از ۸۰ در نظر گرفته شود که با شاخص رشد شناختی کمتر از ۷۰ قرابت بیشتری دارد [۲۹].

مشابه همین روش، جانسون و همکاران، تحقیقی در مورد ضریب توافق تقسیم‌بندی تأخیر تکاملی با استفاده از دو آزمون بیلی ویرایش دوم و بیلی ویرایش سوم انجام دادند. آن‌ها شاخص رشد شناختی بیلی ویرایش دوم[21] و مقیاس شناختی و گفتاری بیلی ویرایش سوم را بر روی ۱۵۸ کودک نارس کمتر از ۲۷ هفته انجام دادند. نمرات به‌دست‌آمده در بیلی ویرایش سوم بالاتر از نمرات شاخص رشد شناختی بیلی ویرایش دوم بود.

محققان اعلام کردند نمرات جداگانه شناختی گفتاری کمتر از ۸۵ در بیلی ویرایش سوم با شاخص رشد شناختی کمتر ۷۰ در بیلی ویرایش دوم، توافق ۹۹ درصد داشته است و در حدود ۱/۱ درصد کمتر تأخیر تکامل را برآورد می‌سازد، اما نمرات ترکیبی شناختی و گفتاری زیر ۸۰، ضریب توافق ۹۸ درصد داشته است و به همان میزان تأخیر تکامل را تشخیص می‌دهد؛ بنابراین آن‌ها نتیجه گرفتند نمرات جداگانه شناختی وگفتاری زیر ۸۵ و نمرات ترکیبی شناختی و گفتاری کمتر از ۸۰ بهترین شاخص برای تشخیص تأخیر تکاملی متوسط تا شدید است و با شاخص رشد شناختی کمتر از ۷۰ در بیلی ویرایش دوم توافق دارد [۳۳]؛ بنابراین بالاتر بودن میانگین در نمونه ایرانی نیز می‌تواند به‌دلیل تفاوت نمونه‌گیری در نمونه ایرانی با نمونه آمریکایی یا به‌دلیل تفاوت در ویژگی‌های محیطی و ژنتیکی دو جمعیت باشد.

یو وای در تایوان، آزمون بیلی ویرایش دوم و سوم را به‌صورت هم‌زمان بر روی ۴۷ نوزاد ترم و ۱۶۷ نوزاد نارس در ۶ و ۱۲ و ۱۸ و ۲۴ ماهگی اجرا کرد. نتایج همبستگی سؤالات آزمون بیلی ویرایش دوم و سوم در حیطه‌های شناختی حرکتی خوب تا عالی و در حیطه گفتاری ضعیف تا عالی نشان داد. مطابق نتایج این مطالعه کرد نمرات به‌دست‌آمده از بیلی ویرایش سوم سطح تکامل کودکان تایوانی را بالاتر از آزمون بیلی ویرایش دوم ارزیابی می‌کند و درنتیجه نمرات بالاتری از آزمون بیلی ویرایش سوم مشخص‌کننده تأخیر تکاملی در این جمعیت است [۲۷].

اندرسون ارزیابی توانایی بیلی ویرایش سوم در تشخیص تأخیر تکاملی کودکان دو ساله‌ای که نارس (کمتر از ۲۸ هفته) یا کم وزن (کمتر از هزار گرم) بودند و در شهر ویکتوریای استرالیا متولد شده بودند را بررسی کرد. میانگین تمام نمرات در گروه آزمایش به شکل قابل‌توجهی پایین‌تر از گروه کنترل بود، اما میانگین نمرات گروه آزمایش منطبق با میانگین نمرات هنجارشده بود و در مقابل میانگین نمرات در گروه کنترل بالاتر از نمونه هنجارشده قرار داشت. میانگین نمرات بین گروه آزمایش و کنترل متفاوت بود ($P<0/01$). بنابراین نتیجه گرفت نمرات هنجار نمونه آمریکایی بیلی ویرایش سوم تأخیر تکاملی را کمتر از حد واقعی در استرالیا برآورد می‌سازد [۳۴].

گودامون و حیطه‌های شناختی حرکتی آزمون بیلی ویرایش سوم را بر روی ۱۵۰ کودک ترم[22] در گروه‌های سنی ۶، ۱۲ و ۲۴ ماهه سریلانکایی انجام داد و عملکرد این کودکان را با کودکان هم‌سن آمریکایی در حیطه‌های یادشده مقایسه کرد. نتایج نشان داد نمرات کودکان ۱۲ماهه سریلانکایی در حیطه شناختی بالاتر

18. Mental Developmental Index (MDI)
19. CB-III Score
20. Predictive Validity
21. Bayley II MDI

و در حیطه حرکتی پایین‌تر از نمرات نمونه آمریکایی و در ۲۴ ماهگی نمرات شناختی پایین‌تری نسبت به نمونه آمریکایی است [۲۸] و نتیجه‌گیری کردند که تفاوت در میانگین نمرات، نیاز به هنجاریابی در هر جامعه را بیان می‌کند.

نتیجه‌گیری

آزمون بیلی ویرایش سوم به‌دلیل ویژگی‌های عالی روان‌سنجی و نظام امتیازدهی کمّی به‌عنوان ابزاری انتخابی برای مستندسازی دوره‌ای تکامل کودک شناخته می‌شود. برای افزایش کارایی، این آزمون داده‌های هنجارشده دقیقی را برای دوره نوزادی تا ۶ ماهگی فراهم ساخته است. در حال حاضر، نمرات ترازشده برای کودکان در این محدوده سنی به ازای هر ۱۰ روز افزایش سن ارائه شده است. با انجام این کار، میزان دقت بالایی را در زمینه ارزیابی کودکان طی این دوره رشد سریع ارائه می‌دهد.

نمرات رشدی این آزمون که جهت ترسیم روند پیشرفت کودک در ارزیابی‌های دوره‌ای استفاده می‌شوند، برای متخصصین بالینی بسیار مورد استفاده است. درنتیجه با توجه به مطالب فوق، هنجاریابی نسخه فارسی آزمون نه‌تنها تشخیص زودهنگام تأخیر تکامل را میسر می‌کند، بلکه خلأ وجود یک آزمون استاندارد طلایی به منظور مقایسه نتایج سایر آزمون‌های تکاملی با آن را برطرف می‌کند.

از نقاط قوت این مطالعه می‌توان به تعداد بالای نمونه (۱۷۴۴ نفری) و نمونه‌گیری در پهنه‌های جغرافیایی سراسر کشور و هنجاریابی در ۴۸ گروه سنی و در ۵ مقیاس شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت نام برد.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

پژوهش حاضر با مجوز کمیته اخلاق دانشگاه علوم توانبخشی و سلامت اجتماعی و کسب رضایت آگاهانه از والدین شرکت‌کننده انجام شد (۱۴۳۵۷/۱/ت/۸۰۱/۹۳).

حامی مالی

این مطالعه توسط دفتر سلامت جمعیت و خانواده وزارت بهداشت (اداره سلامت کودکان) حمایت مالی و پشتیبانی اجرایی شده است.

مشارکت نویسندگان

مفهوم‌سازی: فرین سلیمانی، نادیا آذری و روشنک وامقی؛ روش‌شناسی: فرین سلیمانی، نادیا آذری و آدیس کراسکیان؛ تحقیق: فرین سلیمانی، نادیا آذری، سید حامد برکتی و حمیدرضا لرنژاد؛ نگارش-اصل پیش‌نویس: فرین سلیمانی، نادیا آذری؛ نگارش و بررسی و ویرایش: همه نویسندگان.

تعارض منافع

بنابر اظهار نویسندگان این مقاله تعارض منافع ندارد.

تشکر و قدردانی

نویسندگان از همکاری دانشگاه علوم توانبخشی و سلامت اجتماعی و دانشگاه‌های علوم‌پزشکی شهید بهشتی، ایران، تهران، مشهد، رشت، تبریز، همدان، اهواز، شیراز، اصفهان، قم و زاهدان تشکر و قدردانی می‌کنند. همچنین از والدین کودکانی که در مطالعه شرکت کردند.

References

[1] National Research Council (US) and Institute of Medicine (US) Committee on Integrating the Science of Early Childhood Development. From Neurons to Neighborhoods: The science of early childhood development. Shonkoff JP, Phillips DA, editors. Washington (DC): National Academies Press (US); 2000. [PMID]

[2] Black J, Jones T, Nelson C, Greenough W. Neuronal plasticity and the developing brain. Handbook of child and adolescent psychiatry, vol 1. New York: Wiley, 1998. pp: 31-53. [Link]

[3] Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B, et al. Developmental potential in the first 5 years for children in developing countries. The Lancet. 2007; 369(9555):60-70. [DOI:10.1016/S0140-6736(07)60032-4]

[4] Williams J, Holmes CA. Improving the early detection of children with subtle developmental problems. Journal of Child Health Care. 2004; 8(1):34-46. [DOI:10.1177/1367493504041852] [PMID]

[5] Glascoe FP. Early detection of developmental and behavioral problems. Pediatrics in Review. 2000; 21(8):272-9; quiz 280. [DOI:10.1542/pir.21.8.272] [PMID]

[6] Blauw-Hospers CH, Hadders-Algra, M. A systematic review of the effects of early intervention on motor development. Developmental Medicine and Child Neurology. 2005; 47(6):421-32. [PMID]

[7] Cameron EC, Maehle V, Reid J. The effects of an early physical therapy intervention for very preterm, very low birth weight infants: A randomized controlled clinical trial. Pediatric Physical Therapy. 2005; 17(2):107-19. [DOI:10.1097/01.PEP.0000163073.50852.58] [PMID]

[8] Guralnick MJ. Early intervention for children with intellectual disabilities: Current knowledge and future prospects. Journal of Applied Research in Intellectual Disabilities. 2005; 18(4):313-24. [DOI:10.1111/j.1468-3148.2005.00270.x]

[9] Malekpour M. Low birth-weight infants and the importance of early intervention: enhancing mother-infant interactions a literature review. The British Journal of Development Disabilities. 2004; 50(99):78-88. [DOI:10.1179/096979504799103868.]

[10] Boyle CA, Decoufle P, Yeargin-Allsopp M. Prevalence and health impact of developmental disabilities in US children. Pediatrics. 1994; 93(3):399-403. [DOI:10.1542/peds.93.3.399] [PMID]

[11] Bloom B, Jones LI, Freeman G. Summary health statistics for U.S. children: National Health Interview Survey, 2002. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. 2002; 10(258):1-81. [Link]

[12] Wing L, Potter D. The epidemiology of autistic spectrum disorders: Is the prevalence rising? Mental Retardation and Developmental Disabilities Research Reviews. 2002; 8(3):151-61. [DOI:10.1002/mrdd.10029] [PMID]

[13] National Center for Education Statistics. The Condition of Education. United States, U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 2007. [Link]

[14] Council on Children With Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. Pediatrics 2006; 118(1):405-20. [DOI:10.1542/peds.2006-1231] [PMID]

[15] Palfrey JS, Singer JD, Walker DK, Butler JA. Early identification of children's special needs: A study in five metropolitan communities. The Journal of Pediatrics. 1987; 111(5):651-9. [DOI:10.1016/S0022-3476(87)80238-X]

[16] Barnett WS. Long-term effects of early childhood programs on cognitive and school outcomes. The Future of Children. 1995; 5(3):25-50. [DOI:10.2307/1602366]

[17] World Health Organization (WHO). Department of Maternal, Newborn, Child and Adolescent Health (MCA). [Link]

[18] No Authors. Developmental surveillance and screening of infants and young children. Pediatrics. 2001; 108(1):192-6. [DOI:10.1542/peds.108.1.192] [PMID]

[19] Shonkoff JP, Meisels SJ. Handbook of early childhood intervention. Cambridge: Cambridge University Press; 2000. [Link]

[20] Olade RA. Evaluation of the Denver developmental screening test as applied to African children. Nursing Research. 1984; 33(4):204-7. [DOI:10.1097/00006199-198407000-00005]

[21] Geisinger KF. Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment. 1994; 6(4):304-12. [DOI:10.1037/1040-3590.6.4.304]

[22] Lim HC, Chan T, Yoong T. Standardisation and adaptation of the Denver Developmental Screening Test (DDST) and Denver II for use in Singapore children. Singapore Medical Journal. 1994; 35(2):156-60. [PMID]

[23] Bayley N. Bayley scales of infant development and toddler development. London: Pearson; 2006. [DOI:10.1037/t14978-000]

[24] Vameghi R, Sajedi F, Kraskian Mojembari A, Habiollahi A, Lornezhad HR, Delavar B. Cross-Cultural Adaptation, Validation and Standardization of Ages and Stages Questionnaire (ASQ) in Iranian children. Iranian Journal of Public Health. 2013; 42(5):522-8. [PMID]

[25] Shahshahani S, Vameghi R, Azari N, Sajedi F, Kazemnejad A. Validity and Reliability Determination of Denver Developmental Screening Test-II in 0-6 Year-Olds in Tehran. Iranian Journal of Pediatrics. 2010; 20(3):313-22. [PMID]

[26] Cromwell EA, Dube Q, Cole SR, Chirambo C, Dow AE, Heyderman RS, et al. Validity of US norms for the Bayley Scales of Infant Development-III in Malawian children. European Journal of Paediatric Neurology. 2014; 18(2):223-30. [DOI:10.1016/j.ejpn.2013.11.011] [PMID] [PMCID]

[27] Yu YT, Hsieh WS, Hsu CH, Chen LC, Lee WT, Chiu NC, et al. A psychometric study of the Bayley Scales of Infant and Toddler Development - 3rd Edition for term and preterm Taiwanese infants.Research in Developmental Disabilities. 2013; 34(11):3875-83. [PMID]

[28] Godamunne P, Liyanage C, Wimaladharmasooriya N, Pathmeswaran A, Wickremasinghe AR, Patterson C, et al. Comparison of performance of Sri Lankan and US children on cogni-

tive and motor scales of the Bayley scales of infant development. BMC Research Notes. 2014; 7:300. [DOI:10.1186/1756-0500-7-300] [PMID] [PMCID]

[29] Moore T, Johnson S, Haider S, Hennessy E, Marlow N. Relationship between test scores using the second and third editions of the Bayley Scales in extremely preterm children. The Journal of Pediatrics. 2012; 160(4):553-8. [DOI:10.1016/j.jpeds.2011.09.047] [PMID]

[30] IRAN Statistical Center. [Iran-statistical-yearbook (Persian)]. Tehran: IRAN Statistical Center; 2020. [Link]

[31] Azari N, Soleimani F, Vameghi R, Sajedi F, Shahshahani S, Karimi H, et al. A psychometric study of the bayley scales of infant and toddler development in Persian Language Children. Iranian Journal of Child Neurology. 2017; 11(1):50-6. [DOI:10.22037/ijcn.v11i1.12056]

[32] Pena ED, Spaulding TJ, Plante E. The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. American Journal of Speech-Language Pathology. 2006; 15(3):247-54. [DOI:10.1044/1058-0360(2006/023)]

[33] Johnson S, Moore T, Marlow N. Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used? Pediatric Research. 2014; 75(5):670-4. [PMID]

[34] Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW; Victorian Infant Collaborative Group. Underestimation of developmental delay by the new Bayley-III Scale. Archives of Pediatrics & Adolescent Medicine. 2010; 164(4):352-6. [DOI:10.1001/archpediatrics.2010.20] [PMID]

[35] Steenis LJ, Verhoeven M, Hessen DJ, van Baar AL. Performance of Dutch Children on the Bayley III: A comparison study of US and Dutch Norms. PLoS One. 2015; 10(8):e0132871. [DOI:10.1371/journal.pone.0132871] [PMID] [PMCID]

This Page Intentionally Left Blank