http:// ijp.mums.ac.ir
Original Article (Pages: 15426-15438)

A Comparison Study of the norms of Tehran and Iran in the Performance of 1-42-month-old children on the Bayley Scale

Farin Soleimani [1], * Adis Kraskian [2], Mehdi Noroozi [3], Ali Khakshour [4], Nadia Azari [5]

[1] MD, Pediatrician, Research Professor, Pediatric Neurorehabilitation Research Center, Speech & Language Pathology Department, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

[2] PhD, Assistant Professor, Department of Psychology, Karaj Branch, Islamic Azad University, Karaj, Iran.

[3] PhD, Assistant Professor of Epidemiology, Substance Abuse and Dependence Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

[4] Associate Professor of Pediatrics, Department of Pediatrics, School of Medicine, Akbar Hospital, Mashhad University of Medical Sciences, Mashhad, Iran.

[5] MD, Pediatrician, Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

Abstract

Background: There is a wide variety of sociocultural and environmental background characteristics in different geographical places of Iran. The aim of this study was to investigate the rates of developmental delay in Tehran in comparison to the norm of Iran, using the Bayley III measure.

Methods: This cross sectional study was a part of a national study conducted in Iran, between 2014 and 2016. During the study period, five hundred fifty Persian speaking children in Tehran were included. The sampling was in proportion to the population of children covered in each region. First, the differences between the scaled scores, based on the norms of Tehran and Iran were calculated and then, a one sample Multivariate Analysis of Variance (MANOVA) was used, which showed a significant difference between the scaled scores. Next, we used a univariate analysis to find which scales these significant differences were generated from. Finally, the rate of children with low scores (<-1SD, and <-2 SD) were compared by means of the McNemar analyses.

Results: The numbers of male participants were 310 (54.5%). The mothers in the sample of Tehran had higher educational levels in comparison to those in the sample of Iran (P= 0.001). Considering Iran's norm compared to Tehran's norm leads to significantly fewer rates of delay, on cognition scale (<-1SD; 11.6% to 19.8%) and fine motor scale (<-1SD; 15.1% to 21.1%)(<-2SD; 1.6% to 3.3%) respectively (p <0.01). The differences in estimation rate were somewhat age-dependent. The greatest difference between the norm of Tehran and Iran was in the age group of "25 months 16 days to 42 months 15 days".

Conclusion: In some developmental domains, the norm of Iran in comparison to that of Tehran indicates a lower rate of children with developmental delay.

Key Words: Bayley Scales of Infant and Toddler Development, Child development, Infants, Toddlers.

* Please cite this article as: Soleimani F, Kraskian A, Noroozi M, Khakshour A, Azari N. A Comparison Study of the norms of Tehran and Iran in the Performance of 1-42-month-old children on the Bayley Scale. Int J Pediatr 2022; 10 (2):15426-15438. DOI: **10.22038/IJP.2021.57436.4509**

**** Corresponding Author:***

Adis Kraskian, PhD, Assistant Professor, Department of Psychology, Islamic Azad University, Branch of Karaj, Karaj, Iran. Email: adiskraskian@yahoo.com

Received date: Apr.30,2021; Accepted date:Dec.11,2021

1- INTRODUCTION

In the first three years of life, the child has a rapid development, and the factors affecting this development have a lasting effect on his/her health throughout life (1). Investing in a child's development in the first 3 years of life is very important for the prevention of developmental consequences, because of the rapid growth of the brain and the flexibility of the brain during this period (2). It has been proven that children with healthy development have a high level of education, employment and subsequent income later in life (2-7), while poorly developed children later have higher unemployment and crime rates and teen pregnancies (4, 8, 9).

An estimated 200 million children living in low- and middle-income countries (LMICs) do not fulfil their developmental potentials. It has negative effects on academic achievement, adult income capacity and parents' ability for their future children (10). Black et al. showed that forty-three per cent (43 %) of children under the age of 5 do not reach their developmental potential due to environmental factors around the world (11).

Measuring the appropriateness of the child's developmental context in LMICs is vital for identifying the particularly susceptible people. It helps to measure the dimensions of the issue, to differentiate across people and sets, and to board and estimate interventions, and then measure their progress (or worsening) during the intervention. Accordingly, it is important to choose a valid and practical method for measuring children's development. Since interventions are more effective in the first years of a child's life (12), it is essential for children younger than 3 years of age to achieve developmental measures.

The structure of child development is multifaceted and includes cognitive, motor, and social-emotional domains (13). Standard tools provide a way to measure the development of a child and compare it to a standard norm. Certainly in this context, for clinical and research purposes, the Bayley Scales of Infant and Toddler Development, third edition (Bayley III) is one of the most widely used standard measures for assessing child development (14) . The Bayley III is a test that measures development from neonatal period to 42 months. The validity of child development tools created in high-income countries may not apply to LMICs, as they are different in cultural backgrounds and the official education provided for the people (15). There is an adapted, valid and reliable form of the Bayley III for Persian speaking children (16); and norms of the Bayley III in Iran, though only in Tehran, the capital of Iran, have been determined in the previous studies. (17, 18).

Due to the wide variety of background characteristics in different geographical places of Iran, there is a concern that the regional differences may lead to under or over estimation of developmental delay by Iranian norms. Because we only had the norms of Tehran, we decided to compare them with the norm previously obtained from Iran in general, to find out whether the capital-based Iranian norm is appropriate for the detection of developmental delay in other cities of Iran.

The aim of this study was to investigate the norm of developmental delay in 1-42-month-old children in Tehran in comparison to the norm of Iran.

This cross sectional study was a part of a national study conducted in Iran, 2014-2016, entitled "the standardisation and cut off point determination of the Persian version of Bayley Scales for assessing the development of 1-42-month-old infants and toddlers in Iran" (18).

2- MATERIALS AND METHODS

2-1. Study design and population

This cross sectional study was a part of a national study conducted in Iran, from 2014 to 2016. During the study period, 550 Persian speaking children (Tehran sample) were investigated. Tehran was divided into three geographical regions according to the health-care services including; north and east, central and south, and west, which collectively cover 60% of the health-care visits and 98% of vaccinations from birth until the 5 years of age. The samples were selected in proportion to children covered in each region and were representative of the 1-42-month-old infants.

2-2. Methods

Trained examiners identified the children who met the specified inclusion criteria for the sample. The parents were asked to provide information pertaining to the length of gestation and any difficulties at birth to ensure that only typically developing children born at 36 to 42 gestational weeks would be included in the sample. The families, who agreed to participate, signed informed consent forms. The questionnaire also requested information about the child's age, gender, main speaking language, and the educational level of the parents. Then, the Bayley III was evaluated by the examiners. Examiners were selected based on the following criteria: qualification in professional therapy or psychology, and at least two years of experience in working with children. The Bayley Administration Manual was used for teaching, which included two sections: theoretical and practical training. During the official performance of the Bayley III, three days of practical training were also administered for the examiners in a course of one-week. After completing the training course, a practical training exam was administered, which was monitored and approved by some experts. The Bayley III takes into account the age of each child (month-day) as well as the preterm births. These two factors, along with the child's ability in a series of tasks, using special toy kits, play a role in producing independent and standardised scores. We use all scales of Bayley III including; cognitive, receptive and expressive communication, as well as fine and gross motor in this study.

The data were collected by the researchers and raw and scaled scores were calculated; if they were equal to or less than -2 SD, the developmental delay was hypothesised and the child was referred for the follow-up and more accurate assessment.

2-3. Measuring tools: validity and reliability

The Bayley III, published in 2006, is a worldwide measure for developmental assessment of one to 42 months of age. Bayley III is the modified version of the Bayley II and assesses the child's three developmental skills: cognitive, language, and motor (19, 20). The cognitive scale evaluates information processing, play skills, calculation and number skills. The language scale consists of receptive and expressive scales that are used to assess communication skills. The motor scale consists of fine and gross motor subscales.

Age range in the Bayley III is defined in 17 age groups, and scaled scores are derived from the raw scores. The range of the scaled scores is from 1 to 19, with a standard deviation (SD) of 3 and a mean value of 10. Therefore, a scaled score of 10 in each scale indicates mean functioning in that age group, and scaled scores 7, and 4 show 1 and 2 standard deviations below the mean, respectively (14). Bayley III is standardised in the US population, whose native language is English.

The adapted version of the Bayley III, used in the present study, had given accurate consideration to the cultural background of the Persian speaking children (16). The

forward-backward translation method was used to prepare the Persian version, and to increase its application in Iranian culture, modifications were made to the items, especially in the scales of receptive and expressive communication for compatibility with grammar and language development in Persian-speaking children, and also in the illustrations of the stimulus book, which are explained as follows. Changes in the receptive communication scale: the games were replaced with the usual games played in the Iranian culture. Given the words' frequency in the similar period of language development in Persian-speaking children, the word "candy" was replaced with "cake", and "bird" with "fish", and the illustrations were appropriately changed, and the tool "cup" was replaced with "handled glass". Given that there is a vowel point to indicate possession in Persian language, this form of pronoun was also added to the instructions, and the simpler and more popular form of the continuous tense, namely 'to have + present tense' was used. Changes were also made to the pronouns. In the expressive communication scale, changes were made with respect to expressing continuous verb tenses, use of plural words, signs of possession in Persian language, present tense verbs and signs of future verbs.

In this study the validated version of the Bayley III in Persian speaking children was used (16). Also we used the Iranian norm obtained in a study on 1,744 healthy children, from health-care centres in 10 provinces in Iran (18). Iran was divided into ten geographical regions including Tehran, Isfahan, Qom, Shiraz, Hamadan, Tabriz, Mashhad, Ahvaz, Zahedan, and Rasht. Sampling was in proportion to the population of children covered in each region, according to the yearbook of The Statistical Center of Iran (21) and representative regarding to gender, and age groups. Also, we used norms of Tehran for cognitive, language, and motor scales of the Bayley III, obtained through an investigation of 1, 674 healthy children from health-care centers and representatives regarding gender and age groups (17).

2-4. Inclusion and exclusion criteria

A typically developing child was defined as any child who was born without any significant medical complication, and was not currently diagnosed with or received any treatment (including medication) for mental, physical, or behavioural difficulties (8). The exclusion criteria included receiving rehabilitation services, diagnosed with one or more of the following: attention deficit and hyperactivity disorder, autistic spectrum disorders, chromosomal abnormality, congenital infections, mental deficiency, severe sensory impairments, respiratory disorders, inborn errors of metabolism, low birth weights, prematurity, and/or intraventricular haemorrhage, currently admitted to a hospital.

2-5. Ethical consideration

All parents were informed about the objectives of the research. This study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation Sciences with the decision number of USER.116, dated 12/10/2014. The study conformed to the Helsinki Declaration and good clinical practice guidelines. All parents provided written informed consent.

2-6. Data Analyses

Data was recorded and analysed with the SPSS software version 20.0. Categorical data was represented by frequency and percentage, and continuous data by mean ±standard deviation. Compatibility of continuous data to normal distribution was tested with the Shapiro-Wilk Test. The T-test or the Mann-Whitney U test was used in the comparison of the means according

to their suitability for normal distribution. A Chi-square test was used in comparing the categorical data. The $p<0.05$ value was accepted as the level of significance.

First, differences between the scaled scores of the sample from Tehran (550 participants) on all scales, based on the norms of Tehran and Iran were calculated. Second, a one sample Multivariate analysis of variance (MANOVA) was used to test for inflation of type 1 error, and to control the mean differences over all scales. When this MANOVA showed a significant difference between the scaled scores based on the norms of Tehran and Iran, we referred to a univariate analysis to see in which scales these significant differences were created. Because the average difference may depend on age, in the next step the MANOVA for all scales was performed separately for each age group. The results were evaluated and interpreted according to Cohen (24) indicating that 0.06 or less represents a small effect, 0.07-13 medium effect and 0.14 or higher large effect (η p2). Finally, the rate of children with low scores; scaled scores <7 (< -1SD), and <4 (<-2SD) based on the norms of Tehran and Iran, were compared by means of the McNemar analyses.

3- RESULTS

A slightly over half (n=310, 54.5 %) of the participants were male. Mothers from the Tehran sample had higher levels of education than those from other regions of Iran (**Table.1**).

Table-1: The demographic characteristics of the children and the mothers' educational level in the samples of Iran (N=1744) and Tehran (N=550)

Characteristics		Iran sample	Tehran sample	P value
Gender Frequency (%)	Boys	903(52)	310(54.5)	P= 0.06 Chi-square test
	Girls	836(48)	240(45.5)	
Mother Educational level * Frequency (%)	Low	1014(58)	247(45)	P= 0.001 Chi-square test
	Moderate	712(41)	248(45)	
	High	18(1)	55(10)	

*'Low educational level' refers to special education, primary school, or pre-vocational secondary education (< 12 years); 'Moderate educational level' refers to senior general secondary education, pre-university education, or secondary vocational education (13–16 years); 'high educational level refers to higher professional education or university (17+ years).

Table 2 shows the number of participants in the 17 age groups in Tehran (N=550) and Iran (N=1744).

Table 3 shows the univariate results of MANOVA and the mean differences that make information on the standard deviations according to the effect sizes between the scaled scores of Tehran's sample, based on the norms of Tehran and Iran. The scales of cognition, receptive and expressive communication, and gross motor have significant differences with effect size (ηp2) of .07, .31, .43 and .33 respectively (**Table 3**).

The mean difference and Partial eta squared values between scaled scores based on the norms of Tehran and Iran in different age groups are presented in **Table 4**. The smallest mean difference (equal to .00) was found for cognition for age group; 5 months 16 days-6 months 15 days. The largest mean difference of 3.3 was found for fine motor for age group; 39 months 30 days-42 months 15 days, which was more

than 1 SD based on scaled score. The effect sizes regarding the multivariate analyses included displays in the third column in **Table 4**.

Large effect sizes were found for the differences between the scaled scores based on the norms of Tehran and Iran, but not consistently for particular subscales or for definite age groups (**Table 4**).

Table-2: Number of participants in the 17 age groups in the samples of Tehran (N=550) and Iran (N=1744)

Age groups*	Tehran sample	Iran sample
	N (%)	N (%)
A: 16 days-1 month 15 days	32(5.8)	104(6)
B: 1 months 16 days-2 months 15 days	27(4.9)	100(5.7)
C: 2 months 16 days-3 months 15 days	36(6.5)	103(5.9)
D: 3 months 16 days-4 months 15 days	27(4.9)	99(5.7)
E: 4 months 16 days-5 months 15 days	28(5.1)	100(5.7)
F: 5 months 16 days-6 months 15 days	5(0.9)	61(3.5)
G: 6 months 16 days-8 months 30 days	19(3.4)	89(5.1)
H: 9 months 0 days-10 months 30 days	23(4.2)	84(4.8)
I: 11 months 0 days-13 months 15 days	45(8.2)	97(5.6)
J: 13 months 16 days-16 months 15 days	46(8.3)	117(6.7)
K: 16 months 16 days-19 months 15 days	41(7.4)	111(6.4)
L: 19 months 16 days-22 months 15 days	34(6.2)	103(5.9)
M: 22 months 16 days-25 months 15 days	28(5.1)	104(6)
N: 25 months 16 days-28 months 15 days	40(7.3)	108(6.2)
O: 28 months 16 days-32 months 30 days	52(9.4)	151(8.7)
P: 33 months 30 days-38 months 30 days	57(10.3)	146(8.4)
Q: 39 months 30 days-42 months 15 days	10(1.8)	67(3.8)

* The age range of 1 to 42 months presumed in Bayley III is distributed into 17 age groups (A to Q)

Table-3: MANOVA results comparing the norms of Iran and Tehran in different scales (N=550)

Scales	Mean difference	P value	95% CI	df	ηp2*
Cognition	-0.75	0.001	[-0.85,0.66]	1	0.07
Receptive Communication	-0.87	0.001	[-0.96,0.77]	1	0.31
Expressive Communication	-0.88	0.001	[-0.97,0.79]	1	0.43
Fine Motor	-0.27	0.76	[-0.42,-0.12]	1	0.04
Gross Motor	-0.40	0.001	[-0.49,-0.31]	1	0.33

Note. The Mean difference is calculated by the scaled score based on Tehran's norms minus the scaled score based on Iran's norms. Mean differences < 0 indicate that the score based on Iran's norm was higher than the scaled scores based on Tehran's norm. Mean differences > 0 indicate that the scaled score based on the sample from Iran is lower than the scaled score based on the sample from Tehran.

* ηp2 (effect size) of 0.06 or less represents a small effect, 0.07-13 medium effect and 0.14 or higher large effect.

Table-4: MANOVA results comparing the norms of Tehran and Iran based on age groups (N=550).

Age Groups *	Cognition		Receptive Communication		Expressive Communication		Fine Motor		Gross Motor	
	Mean difference (SD)	ηp2**	Mean difference (SD)	ηp2**	Mean difference (SD)	ηp2**	Mean difference (SD)	ηp2**	Mean difference (SD)	ηp2**
A	1.18(0.64)	0.11	-0.43 (0.3)	0.33	0.5(0.21)	0.11	1.03(0.44)	0.22	0.09(0.03)	0.05
B	0.7(0.54)	0.43	-1.74 (0.7)	0.33	0.07(0.55)	0.21	0.22(0.3)	0.32	0.92(0.15)	0.11
C	-0.69(0.98)	0.32	-0.72(0.23)	0.43	1.22(0.34)	0.02	0.69(0.15)	0.22	0.52(0.11)	0.32
D	-0.4(0.74)	0.14	-0.33(0.48)	0.16	0.48(0.11)	0.32	-1.45(0.89)	0.22	-0.14(0.10)	0.16
E	-0.21(0.87)	0.30	0.14(0.10)	0.04	1.2(0.88)	0.22	1.43(0.76)	0.34	-0.14(0.001)	0.33
F	0.0(0.8)	0.05	0.8(0.12)	0.30	-0.73(0.09)	0.24	-0.4(0.15)	0.08	-0.2(0.02)	0.21
G	0.73(0.08)	0.21	-0.47(0.16)	0.09	1.27(0.04)	0.04	1.054(0.94)	0.33	-0.81(0.22)	0.44
H	-0.17(0.57)	0.53	-0.608(0.18)	0.43	1.043(0.97)	0.05	-0.43(0.17)	0.11	-0.13(0.14)	0.22
I	-0.77(0.56)	0.02	-1.57(0.19)	0.02	1.55.(0.94)	0.12	0.44(0.21)	0.12	-0.52(0.11)	0.10
J	-1.56(0.06)	0.34	-1.82(.66)	0.04	-1.54(0.77)	0.08	0.73(.`9)	0.24	0.82 (0.17)	0.04
K	-0.68(0.77)	0.19	-0.75(0.04)	0.12	-0.34(0.98)	0.45	-0.90(0.33)	0.43	1.12(0.12)	0. 22
L	-1.55(2.11)	0.02	1.29(0.77)	0.20	-0.88(0.22)	0.04	-0.75(0.55)	0.10	1.05(0.19)	0.42
M	-1.107(.31)	0.43	-1.46(0.86)	0.13	-0.75(0.44)	0.03	-0.35(0.27)	0.21	0.96(0.17)	0.22
N	-0.82(0.78)	0.01	-1.55(0.77)	0.07	-0.66(0.65)	0.04	-0.47(0.16)	0.03	-0.87(0.17)	0.27
O	-1.73(0.87)	0.04	-1.173(0.89)	0,04	-1.00(0.19)	0.14	-0.59(0.22)	0.11	0.67(0.44)	0.16
P	-0.86(0.54)	0.05	0.66(.48)	0.05	-0.84(0.55)	0.27	1.87(0.33)	0.20	0.32(0.19)	0.22
Q	-0.03(0.94)	0.01	-1.3(.77)	0.17	-0.80(0.14)	0.32	3.3(0.32)	0.16	0.3 (0.11)	0.32

Note. The Mean difference is calculated by the scaled score based on Tehran's norms minus the scaled score based on Iran's norms. Bold effect sizes are statistically significant (p <0 .01).

* The age range of 1 to 42 months presumed in Bayley III is distributed into 17 age groups (A to Q)

** ηp2 (effect size) of 0.06 or less represents a small effect, 0.07-13 a medium effect, and 0.14 or higher a large effect.

Effect sizes for fine and gross motor subscales were generally large for all age groups. For most variables, the differences go up and down among age groups; this usually means that it is not a "real difference" but a sample /normal spread induced difference. For most of these small effect sizes, the 0 falls within the confidence intervals indicating that no significant difference exists between the scaled scores based on norms of Tehran and Iran.

McNemar analysis detected that the rates of children with low scores considering the norms of Iran were different from those based on the norm of Tehran (**Table 5**). It means that in cognition (<-1SD; 11.6% (Iran); 19.8% (Tehran)) and fine motor (<-1 SD; 15.1% (Iran); 21.1% (Tehran)) (<-2 SD; 1.6% (Iran); 3.3% (Tehran)) scales, fewer children were significantly determined as developmentally delayed, when using Iran's norm in comparison to Tehran's norm, respectively ($p < 0.01$). In addition we performed McNemar analysis for the four age groups; (0–6 months 15 days, 6 months 16 day -13 months 15 days, 13 months 16 day -25 months 15 days, 25 months 16 day -42 months 15 days) (**Table 5**). The results of this classification showed that the most significant difference is between the sub-groups of 25 months 16 day and 42 months 15 days.

4- DISCUSSION

The present study aimed at investigating the norm of 1-42-month-old children's developmental delay in Tehran in comparison to that of Iran. According to Tehran's norm, the rate of children with mild cognition and fine motor scales, and with moderate to severe on fine motor scales are significantly fewer in comparison to the rate of children in Iran's norm.

Furthermore, the results showed that the difference in estimation rates based on these two norms is somewhat age-dependent. The greatest differences between Tehran's and Iran's norms were in the age groups of "25 months 16 day - 42 months 15 days".

The findings show differences in the rate of performance and growth of children not only in different populations (25-29), but also in different geographical regions of a country. On the other hand we must consider that norms take the spread into account and only children <-2SD are statistically "abnormal". If a child lives in a less enabling environment, the child should not be seen as abnormal (1). Children with normal development in less enabling environments do not presumably reach their optimal capacity due to the contextual circumstances (7).

An important explanation for the difference between the norms is related to the background characteristics of the population in different regions of Iran. In Iran, there are large differences between the population of children, including differences in socio-cultural and environmental contexts. The same Persian version of the Bayley III had been used for assessing the norms of Tehran and Iran; therefore, the difference cannot be attributed to the test. To develop Bayley III norms in Iran, the standardisation methods had been performed in both Tehran's and Iran's investigations. Based on the report of the yearbook of the Statistics Center of Iran (21), both samplings had the same population based on gender, and age groups. To the real and experimental issues that researchers typically address, early developmental assessment requires a number of conceptual notifications. Development happens in the circumstance of social and biological progression that affects the ability of a toddler. Development can be viewed as cascades, in which growth of one area affects the growth of other areas (30).

Table-5: Rate of children with low scores based on Tehran to Iran norms.

Age groups	Iran norm <-1SD %	Tehran norm < -1SD %	Iran norm < -2SD %	Tehran norm < -2SD %
1) All age groups				
Cognition	11.6	19.8*	2,1	3.1
Receptive Communication	19.8	19.8	3.4	3.4
Expressive communication	20.9	20.9	3.6	3.4
Fine Motor	15.1	21.1*	1.6	3.3*
Gross Motor	19.8	18.5	2.9	2.2
2) 4 Age groups				
a. 0–6 months 15 days				
Cognition	11.9	11.9	0	0
Receptive communication	1.7	6.9*	1.7	1.7
Expressive communication	3.4	3.4	0	0
Fine Motor	8.5	13.6*	0	1.7
Gross Motor	16.9	11.9*	0	1.7
b. 6 months 16 day -13 months 15 days				
Cognition	15.9	26.1*	1.4	2.9
Receptive Communication	20.4	20.4	1.4	1.4
Expressive communication	21.7	21.7	1.4	1.4
Fine Motor	15.9	24.3*	1.4	5.8*
Gross Motor	24.6	24.6	1.5	1.5
c. 13 months 16 day -25 months 15 days				
Cognition	12	19.5*	2.8	3.6
Receptive Communication	23.2	23.2	5.6	5.6
Expressive communication	26.8	26.8	4.9	4.9
Fine Motor	17.6	22.5*	2.1	4.2*
Gross Motor	22.6	21.7	4.2	3.5
d. 25 months 16 day -42 months 15 days				
Cognition	10.6	19.2*	2.9	4.8*
Receptive Communication	20.2	20.2	3.8	3.8
Expressive communication	24	24	4.8	4.8
Fine Motor	15.9	23.1*	1.9	2.4
Gross Motor	16.3	23.6*	1.9	3.4*

Note. Scaled scores of <-1SD reflect "mild developmental delay" and scaled scores of <-2SD indicate "moderate to severe developmental delay".

* $p < 0.01$

A second important point to evaluate development is that growth does not just happen linearly within a specific domain. For example, it is probable that growth in the emotions lead to growth in the cognition, and vice versa (31). Children from low socio-cultural circumstances may perform poorly on the Bayley III for several reasons. It may be due to the lack of an enabling home environment (such as books and toys suitable for the child), which are typically a major part of Bayley III tools.

Previous studies have shown that children from different ethnic backgrounds significantly differ in their developments in terms of motor (32) and language skills (33), even in the same country. Thus, the

differences in the normative patterns of Tehran and Iran in regard to such factors may play a role in the difference between the norms.

In this study, the mothers in the sample of Tehran had significantly higher educational levels than those in Iran's sample. The level of the mothers' education is another important factor related to the developmental consequence. The correlation between mother's education and child's development is well established theoretically and experimentally (34-35). In a study conducted by Steenis et al., in the Netherlands, an analysis of the relationship between the level of maternal education and scale scores showed that there was a significant difference between the scores of children with mothers having low, medium and higher educational levels. Moreover, as they increased in their age, the children of mothers with higher education had higher grades in their cognitive and communication skills than children of low-educated mothers (36). This is consistent with previous studies, which have shown that children of parents with lower levels of education have poorer language skills than children of parents with higher levels of education (37). Thus, the difference in the educational levels of the samples of Tehran and Iran might also have contributed to the difference between the norms.

In a study conducted by Chineta et al., Australian children were compared to the US norms in five developmental domains, using Bayley III measures. The mean scores were higher in the cognitive, fine motor, receptive and expressive communication scales, and there were no significant differences in the gross motor scale. The researchers concluded that children with mild developmental delay may not be diagnosed when using US norms (26). A number of other studies, which assessed the developments in relation to the reference norms, found average scores higher than the norm (27, 38). Likewise, in our study, the rate of "mild developmental delay" was significantly different based on the norms of Tehran and Iran. This suggests that the results of these tests may be influenced by socio-cultural factors along with the levels of maturity and performance of the individuals.

The strength of this study was that; the comparison between the norms of Tehran and Iran has been done in different age groups and five scales with adequate numbers in each group by a representative sample.

4-1. Study Limitations

The limitation of the study was that, we didn't examine the developmentally delayed children by a clinical evaluation. On the other hand, the sample from Tehran was found to be more educated than that from Iran.

5- CONCLUSION

Our findings revealed that, in some developmental domains, the norm of Iran in comparison to that of Tehran indicates a lower rate of children with developmental delay. The research highlights the importance of cross-cultural validation, which is fundamental for the different screening instruments.

6- ACKNOWLEDGMENTS

We would like to thank the University of Social Welfare and Rehabilitation Sciences and Ministry of Health and Medical Education for their support in implementing of study. The authors are grateful for all children and parents who participated in this study.

7- CONFLICT OF INTERESTS

We would like to thank the Ministry of Health and Medical Education for financial support. The authors are grateful

for all children and parents who participated in this study.

8- REFERENCES

1. Almond D, Currie J. Human Capital Development before Age Five. In: Ashenfelter O, Card D. Handbook of Labour Economics. vol 4B: Elsevier. 2011, 1315-486.

2. Knudsen EI, Heckman JJ, Cameron JL, Shonkoff JP. Economic, neurobiological, and behavioural perspectives on building America's future workforce. Proc Natl Acad Sci USA. 2006 Jul 5; 103(27):10155-10162.

3. Cunha F, Heckman JJ, Schennach SM. Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica. 2010; 78(3):883-931.

4. Heckman, J., Moon, S.H., Pinto, R., Savelyev, P. and Yavitz, A., Analysing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. Quantitative Economics. 2010; 1:1-46.

5. Attanasio O, Cattan S, Fitzsimons E, et al. Estimating the Production Function for Human Capital: Results from a Randomised Controlled Trial in Colombia. American Economic Review. 2020; 110(1):48-85.

6. Manning M, Patterson J. LIFETIME EFFECTS: The High/Scope Perry preschool study through age 40. Childhood Education. 2006; 83(2):121.

7. Alloway TP, Alloway RG. Investigating the predictive roles of working memory and IQ in academic attainment. J Exp Child Psychol. 2010; 106(1):20-9.

8. Walker SP, Wachs TD, Grantham-McGregor S, Black MM, Nelson CA, Huffman SL, et al. Inequality in early childhood: risk and protective factors for early child development. Lancet. 2011; 378(9799):1325-38.

9. Gertler P, Heckman J, Pinto R, Zanolini A, Vermeersch C, Walker S, et al. Labour market returns to an early childhood stimulation intervention in Jamaica. Science. 2014; 344(6187):998-1001.

10. Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B; International Child Development Steering Group. Developmental potential in the first 5 years for children in developing countries. Lancet. 2007; 369(9555):60-70.

11. Black MM, Walker SP, Fernald LCH, Andersen CT, DiGirolamo AM, Lu C, et al. Early childhood development coming of age: science through the life course. Lancet 2017; 389(10064):77-90.

12. Heckman JJ. Policies to foster human capital. Research in Economics. 2000; 54(1):3-56.

13. Walker SP, Wachs TD, Gardner JM, Lozoff B, Wasserman GA, Pollitt E, et al. Child development: risk factors for adverse outcomes in developing countries. Lancet. 2007; 369(9556):145-57.

14. Bayley N. Bayley scales of infant and toddler development: 2006; PsychCorp, Pearson.

15. Sternberg RJ, Nokes C, Geissler PW, Prince R, Okatcha F, Bundy DA, et al. The relationship between academic and practical intelligence: a case study in Kenya. Intelligence. 2001; 29(5):401-18.

16. 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. Iran J Child Neurol. 2017; 11(1):50-56.

17. Soleimani F, Azari N, Vameghi R, Sajedi F, Shahshahani S, Karimi H et al. [Cut off point determination of the Persian version of Bayley Scales of Infant and toddlers Development III test in 1-42 months old children in Tehran city]. Research Report.Tehran, Iran; University

of Social Welfare & Rehabilitation Sciences, Pediatric Neurorehabilitation Research Center, 2015 (Persian).

18. Soleimani F, Azari N, Kraskian A, Vameghi R, Barakati SH, Lornegad H, et al. [The standardisation and cut off point determination of the Persian version of Bayley Scales for assessing the development III of 1-42-month-old infants and toddlers in Iran]. Research Report.Tehran;Iran:University of Social Welfare & Rehabilitation Sciences, Pediatric Neurorehabilitation Research Center, 2016 (Persian).

19. Bos AF. Bayley-II or Bayley-III: what do the scores tell us? Developmental medicine and child neurology. 2013; 55(11):978-9.

20. Lowe JR, Erickson SJ, Schrader R, Duncan AF. Comparison of the Bayley II Mental Developmental Index and the Bayley III Cognitive Scale: are we measuring the same thing? Acta Paediatr. 2012; 101(2):e55-8.

21. Iran-Statistical-Yearbook. 2020. Available from: URL; http:// www.amar.org.ir/english/Iran-Statistical-Yearbook.

22. Peña ED, Spaulding TJ, Plante E. The composition of normative groups and diagnostic decision making: shooting ourselves in the foot. Am J Speech Lang Pathol. 2006; 15(3):247-54. doi: 10.1044/1058-0360(2006/023). PMID: 16896174.

23. Anderson PJ, Burnett A. Assessing developmental delay in early childhood - concerns with the Bayley-III scales. Clin Neuropsychol. 2017; 31(2):371-381.

24. Cohen J. Statistical power analysis. Current directions in psychological science. 1992 Jun; 1(3):98-101.

25. Westera JJ, Houtzager BA, Overdiek B, van Wassenaer AG. Applying Dutch and US versions of the BSID-II in Dutch children born preterm leads to different outcomes. Dev Med Child Neurol. 2008; 50(6):445-9.

26. Chinta S, Walker K, Halliday R, Loughran-Fowlds A, Badawi N. A comparison of the performance of healthy Australian 3-year-olds with the standardised norms of the Bayley Scales of Infant and Toddler Development (version-III). Arch Dis Child. 2014; 99(7):621-4.

27. Walker K, Badawi N, Halliday R, Laing S. Brief Report: Performance of Australian Children at One Year of Age on the Bayley Scales of Infant and Toddler Development (Version III), the Educational and Developmental Psychologist. 2010; 27:1, 54-58.

28. Yu YT, Hsieh WS, Hsu CH, Chen LC, Lee WT, Chiu NC, Wu YC, Jeng SF. A psychometric study of the Bayley Scales of Infant and Toddler Development - 3rd Edition for term and preterm Taiwanese infants. Res Dev Disabil. 2013 Nov; 34(11):3875-83.

29. Krogh MT, Væver MS, Harder S, Køppe S. Cultural differences in infant development during the first year: A study of Danish infants assessed by the Bayley-III and compared to the American norms, European Journal of Developmental Psychology. 2012; 9:6, 730-736, DOI: 10.1080/17405629.2012.688101.

30. Blair BL, Perry NB, O'Brien M, Calkins SD, Keane SP, Shanahan L. Identifying developmental cascades among differentiated dimensions of social competence and emotion regulation. Dev Psychol. 2015; 51(8):1062-73.

31. Calkins SD, Bell MA. Child development at the intersection of emotion and cognition. American Psychological Association; 2010.

32. Kelly Y, Sacker A, Schoon I, Nazroo J. Ethnic differences in achievement of developmental milestones by 9 months of

age: The Millennium Cohort Study. Dev Med Child Neurol. 2006; 48(10):825-30.

33. Duncan AF, Watterberg KL, Nolen TL, Vohr BR, Adams-Chapman I, Das A, et al. Effect of ethnicity and race on cognitive and language testing at age 18-22 months in extremely preterm infants. J Pediatr. 2012; 160(6):966-71.e2.

34. Carneiro P, Meghir C, Parey M. Maternal Education, Home Environments, and the Development of Children and Adolescents, Journal of the European Economic Association. 2013; 11:123–160.

35. Attig M, Weinert S. What Impacts Early Language Skills? Effects of Social Disparities and Different Process Characteristics of the Home Learning Environment in the First 2 Years. Front Psychol. 2020; 8:11:557751.

36. 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; 12:10(8):e0132871.

37. Marton K, Gehebe T, Pazuelo L. Cognitive Control along the Language Spectrum: From the Typical Bilingual Child to Language Impairment. Semin Speech Lang. 2019; 40(4):256-271.

38. Hickey A, Froude EH, Williams A, Hart T, Summers J. Performance of Australian children on the Miller assessment for preschoolers compared with USA norms. Australian Occupational Therapy Journal. 2000; 47: 86-94.