Iran J Psychiatry 2026; 21: 1: 43-54

Original Article

Standardization of Developmental Indicators for the Assessment of Learning for Persian Children

Meysam Sadeghi[1], Malahat Amani[2*], Shadi Zolfaghari[3]

Abstract

Objective: The Developmental Indicators for Assessment of Learning–Fourth Edition (DIAL-4) is a widely used screening tool grounded in developmental theory, designed to identify potential delays across motor, language, conceptual, self-help, and socio-emotional domains prior to school entry. While the DIAL-4 has been standardized in several countries, its norms may not fully capture the cultural, linguistic, and environmental characteristics that shape developmental trajectories of Iranian children. To address this gap, the present study aimed to establish culturally appropriate norms for the DIAL-4 in Iran, providing a valid and reliable framework for early identification and intervention.
Method: This cross-sectional study standardized the Developmental Indicators for Assessment of Learning–Fourth Edition (DIAL-4) for Iranian children aged 2 years 6 months to 5 years 11 months. Districts in both cities were first selected using cluster sampling, considering socioeconomic and cultural diversity. Within these districts, kindergartens were randomly chosen. The final sample comprised 678 children. Trained examiners administered the Persian version of DIAL-4 individually. Raw scores on motor, concepts, and language subscales were converted into standardized developmental scores (M = 100, SD = 15), following established psychometric guidelines to generate culturally appropriate norms.
Results: The findings indicated that the motor, concepts, and language subscales differed significantly across age groups ($p < 0.05$). Based on standardized scoring criteria, values below 70 were classified as indicative of severe deficits or developmental delays, scores between 70 and 84 reflected mild delays, scores from 85 to 115 were considered within the normative range, scores between 115 and 129 indicated above-average performance, and scores of 130 or higher represented an advanced level of functioning. When comparing the standardized developmental scores of Iranian children with those of American children, the results demonstrated that Iranian children consistently obtained lower scores across all three subscales.
Conclusion: This study confirmed the suitability of the DIAL-4 for Iranian children and showed significant differences in motor, concepts, and language skills across age groups. Iranian children scored lower than American children on all subscales, highlighting the need for culturally specific norms to ensure accurate developmental assessment.

Key words: *Child Development; Developmental Disabilities; Diagnosis; Standardization*

1. Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran.
2. Department of Psychology, University of Mohaghegh Ardabili, Ardabil, Iran.
3. Allameh Tabatabai University, Tehran, Iran.

***Corresponding Author:**
Address: Department of Psychology, University of Mohaghegh Ardabili, Ardabil, Iran, Postal Code: 5619635875.
Tel: 98-9143580277, Email: m.amani@uma.ac.ir

Article Information:
Received Date: 2025/07/30, Revised Date: 2025/09/30, Accepted Date: 2025/11/01

Developmental delay refers to a significant lag in achieving age-appropriate milestones in one or more domains, including motor, cognitive, language, social–emotional, or adaptive skills, often defined as performance two or more standard deviations below the mean (1, 2). A systematic review reported an overall prevalence of 14.6%, with variability due to differences in assessment methods (3). Importantly, if professionals use only their clinical judgment to diagnose these cases, only 20–30% of such children would be diagnosed before school age (4).

Although several tools have been designed and applied for developmental screening of children, the psychometric properties of only a limited number of them have been examined and reported as satisfactory in Iran. Among these are: the ASQ-3 (5), the Vineland Social Maturity Scale (VSMS) (6), and the Early Learning Observation and Rating Scale (7), all of which are exclusively questionnaire-based and completed by parents. The Peabody Picture Vocabulary Test (PPVT) (8) measures only verbal learning, while the NEPSY test (9) provides a comprehensive assessment of children's neuropsychological functioning. The Bayley Scales of Infant and Toddler Development (BSID-III) (10) are mainly used to evaluate the development of cognitive, language, motor, visual, adaptive, and socio-emotional skills in infants and children aged 1 to 42 months. The Denver Developmental Screening Test (DDST) (11) measures personal-social, fine and gross motor, and language skills in preschool-aged children.

A study investigating developmental delay in Iranian children aged 1 to 42 months using the BSID-III and Ages and Stages Questionnaire, Third Edition (ASQ-3), highlighted culturally sensitive adaptations for assessment tools (12). In Iran, limited data exist on the prevalence and patterns of developmental delays in preschool-aged children, largely due to the absence of validated and standardized screening instruments. This gap poses a significant public health concern, as early identification of developmental delays is critical for implementing timely interventions that can reduce long-term educational and social difficulties. Barriers such as limited access to early childhood education, disparities in healthcare between urban and rural areas, and insufficiently trained professionals further complicate early detection in Iran, underscoring the urgent need for standardized developmental screening tools adapted to this context.

The Developmental Indicators for Assessment of Learning–Fourth Edition (DIAL-4) is among the most widely used screening tools internationally, designed for children aged 2:6 to 5:11 years. DIAL is based on the information processing model, which assumes that the child's performance depends on auditory or visual inputs as well as motor or linguistic outputs. To achieve the appropriate output, the questions give the child the opportunity to demonstrate the performance that is expected of them (13). Also, for assessment, the theoretical and conceptual model of DIAL relies on an eclectic approach and focuses on combining all aspects of learning to obtain a more comprehensive assessment of the child's functional level (14). It screens motor, concepts, and language skills through direct assessment and supplements these with parent/teacher questionnaires on self-help and social-emotional development. Compared with other tools, the DIAL-4 offers several advantages particularly relevant to Iran: its relatively short administration time (~30 minutes), which is feasible in resource-constrained settings with limited trained professionals; its integration of both examiner observations and parent reports, which increases reliability in diverse cultural contexts; and its performance-based assessments that simultaneously evaluates children's self-help and socio-emotional skills through parent- and teacher-report questionnaires. In addition to as a screening tool, DIAL not only identifies children at risk of developmental delays but can also predict potential future academic difficulties (15).

DIAL-4 was normed on a nationally representative sample and provides standardized scores and percentile ranks by age (16). The motor area assesses the skills of throwing, jumping, hopping, building blocks, thumb coordination, cutting, and writing. The concept area assesses the child's ability to point to body parts, color knowledge, counting, concepts, sorting, and shape recognition.

The language area assesses knowledge of personal information, articulation, objects and actions, letters, and sounds, rhyming and I spy, and problem solving. The self-help development area assesses everyday skills such as dressing and feeding through a parent/teacher questionnaire, and the social-emotional development area assesses the skills needed to build relationships and how the child feels about themselves through a parent/teacher questionnaire (15, 17).

Mardel-Czudnowski and Goldenberg (16) examined the psychometric properties of the DIAL-4 in a sample of 1400 normal children (70% English-speaking and 30% Spanish-speaking). Internal consistency reliability was calculated for each area (separated by age at 6-month intervals for the English version and 1-year intervals for the Spanish version). Split-half reliability coefficients for the total score were 0.95 for the English version and 0.96 for the Spanish version. Test-retest coefficients ranged from 0.64 (motor area for the English version) to 0.95 (language area for the Spanish version). Inter-rater agreement ranged from 0.89 to 0.98. Correlations between the DIAL-4 total score and areas were moderate to high, ranging from 0.74 for the motor area in the English and Spanish versions to 0.90 for the concepts and language area in the Spanish version. The correlations between the third and fourth editions of the DIAL were 0.85 for the total score, 0.84 for concepts, 0.75 for language, 0.58 for motor, 0.66 for self-help, and 0.66 for social-emotional development. The correlations

between the cognitive-linguistic domain of the Early Screening Profiles (ESP) and the concepts and language areas of the DIAL-4 were 0.59 and 0.61, respectively, and the correlations were reported to be weak for the remaining domains. The correlation between the total score of the DIAL-4 and the general conceptual ability domain of the Differential Ability Scales was 0.73, indicating that the DIAL-4 is an appropriate screening tool for detecting possible developmental delays. Both the English and Spanish versions of the DIAL-4 have moderate to strong internal consistency of 0.80 across all areas (16). Studies by Pearson Education, Incorporated studies have provided strong evidence for the content and construct validity of the DIAL-4 (16, 18). The DIAL-4 was compared with the DIAL-3 and other screening instruments, providing moderate to strong correlations, and suggesting convergent validity (16, 18). Currently, the DIAL-4 is widely used compared to its previous editions, and few tools are as easy, inexpensive, and effective as this test while providing an engaging experience for children. Overall, the DIAL-4 provides valid and reliable measures for screening for developmental delays in children's developmental skills in preschool and kindergarten. No major shortcomings have been identified regarding the screening efficacy, theoretical and empirical support for the scales, the quality of the norming sample, and the norming process of the DIAL-4. However, these findings may not generalize directly to Iran due to cultural and linguistic differences that can influence test performance—for example, language-specific items, culturally embedded tasks, or parental perceptions reflected in questionnaire responses. These factors highlight the importance of translation, cultural adaptation, and pilot testing to ensure the tool’s validity and reliability in Iranian children. The present study aimed to standardize the DIAL-4 for Iranian children by developing age-specific norms. Establishing these norms provides clinicians, educators, and policymakers with a valid and reliable tool for early identification of developmental delays in Iran.

Materials and Methods

The present study employed a descriptive and psychometric design aiming at standardizing the DIAL-4 for Iranian preschool children. The statistical population included children aged 2 years and 6 months to 5 years and 11 months attending kindergartens in Tehran and Bojnourd during the study period. The selection of Tehran and Bojnourd was intentional. Tehran, as the capital, provides access to children from diverse ethnic, cultural, and socioeconomic groups, while Bojnourd represents smaller urban and semi-rural contexts, ensuring regional diversity and enhancing the applicability of the norms across different Iranian settings.

A multistage cluster sampling method was applied. In the first stage, districts in both cities were stratified to capture geographical diversity representation. In the second stage, kindergartens were randomly selected from each stratum to ensure inclusion of children from diverse socioeconomic and demographic backgrounds. In the third stage, children whose parents completed the consent form and whose children were willing and prepared to participate in the testing were selected as the sample. The final sample comprised 678 children. The required sample size was estimated based on the precision method for normative studies. Considering an expected standard deviation of 15 points for the test scores, a maximum allowable error of ± 3 points, and a 95% confidence level ($Z = 1.96$), the minimum required sample size was calculated as $n = (Z \times \sigma / d) 2 = (1.96 \times 15 / 3) 2 \approx 97$ participants per group. Given that the study involved seven age groups, the total sample size was determined to be approximately 678 participants. This estimation was cross-checked using G*Power 3.1. G*Power was applied assuming a small effect size ($\eta^2 = 0.10$) to ensure sufficient sensitivity to detect even subtle group differences. This estimation confirmed that the achieved sample provided adequate statistical power ($1-\beta = 0.80$) for detecting group differences while ensuring the desired level of precision for normative data. In practice, the distribution of participants across age groups was not perfectly equal due to recruitment constraints. The younger age groups (2:6–3:5 years) were slightly under-represented, whereas the older groups (5:0–5:11 years) were slightly over-represented

Inclusion criteria included the age range of 2 years and 6 months to 5 years 11 months, parental consent to participate in the research and testing of their child, and proficiency in Persian, the language of test administration. Exclusion criteria in the study included the child's lack of cooperation in testing, history of diagnosed developmental disorders, and presence of significant medical or neurological conditions that could affect performance. These criteria ensured the sample represented a normative population, suitable for standardization and valid norm development.

Developmental Indicators for Assessment of Learning—Fourth Edition (DIAL-4): The DIAL-4 is a screening test comprising skills and rating forms to identify children aged 2 years and 6 months to 5 years 11 months who need rehabilitation. The DIAL-4 is designed to screen a large number of children efficiently. Children are placed in individual performance assessment centers in the areas of motor, language, and concepts, and questionnaires from teachers and parents are used to measure children's self-help skills and social-emotional development. In Iran, the DIAL-4 has been previously validated through concurrent validity with the Denver Developmental Screening Test, and high internal consistency was reported (Cronbach’s alpha of 0.94 for motor, 0.95 for concepts, and 0.94 for language subtests) (19). In this study, the Persian version of the DIAL-4 was used.

Procedure

After obtaining institutional approvals, trained examiners administered the DIAL-4. Examiners underwent a structured training program, including practice sessions under supervision, to ensure standardized administration. Testing was conducted in quiet rooms within the kindergartens to minimize distractions and provide a child-friendly environment. Each child was individually assessed in motor, language, and concepts subtests.

In the norming process, raw scores for each subscale (motor, concepts, and language) were first summarized within each of the seven age groups by calculating the mean and standard deviation. Raw scores were then directly transformed into standardized developmental scores with a mean of 100 and a standard deviation of 15, consistent with established psychometric practices in instruments such as the Wechsler and Stanford–Binet scales. This standardization enables the use of widely accepted cut-off points: scores below 70 indicate severe developmental delays, 70–84 represents mild delays, 85–115 denotes the normal range, 116–129 reflect above average, and ≥ 130 indicates precocious development. The norms were developed using a representative sample ($n = 678$) selected through multistage cluster sampling across Tehran and Bojnourd, capturing children from diverse socio-economic backgrounds.

Ethical Considerations

In accordance with the ethical guidelines of the 1964 Helsinki Declaration for research on human participants, the participants in this study were children. Parental consent was obtained for the test and they were assured that the test results would not be used in the decision-making process for screening, classifying, or labeling children and that the test results would only be used for research purposes. No harm or psychological pressure was inflicted on children during the testing process. Assessments were conducted in supportive, child-friendly environments, breaks were provided if needed, and participation was voluntary. Data confidentiality and anonymity were strictly maintained.

Results

Demographic findings showed that 312 (46%) were girls and 366 (54%) were boys. The children ranged in age from 2 years and six months to 5 years and 11 months. In the 2:6–2:11 age group, there were 34 girls and 31 boys. In the 3:0–3:5 group, the numbers increased to 43 girls and 32 boys, and in the 3:6–3:11 group to 47 girls and 35 boys. In the 4:0–4:5 range, the number of boys rose noticeably to 55 compared to 35 girls, while in the 4:6–4:11 group, there were 34 girls and 48 boys. The highest counts were observed in the 5:0–5:5 group with 62 girls and 82 boys, and similarly in the 5:6–5:11 group with 57 girls and 83 boys. Overall, although the distribution is relatively balanced in the younger ages, the number of boys becomes noticeably higher than girls in the older age groups.

Among the participants, 35 (5.2%) had mothers with a lower diploma, 188 (27.7%) had mothers with a diploma, 331 (48.8%) had mothers with associate and bachelor's degrees, and 124 (18.3%) had mothers with master's degrees and higher. Among the participants, 43 (6.3%) had fathers with a lower diploma, 232 (34.2%) had fathers with a diploma, 277 (40.9%) had fathers with associate and bachelor's degrees, and 126 (18.6%) had fathers with master's degrees and higher.

Table 1 presents the mean scores and standard deviations of the DIAL-4 subscales across the seven age groups. In the motor domain, the items include both gross motor skills (e.g., catching, jumping, hopping, and jumping rope) and fine motor skills (e.g., block building, cutting, and copying shapes), which together provide a comprehensive assessment of children's motor development. The concepts domain consists of tasks such as identifying body parts, recognizing colors, counting, understanding spatial positions (e.g., up, down, left, right), vocabulary, and sequencing shapes, all of which reflect cognitive and perceptual growth. The language domain includes speech production, letter/sound recognition, rhyming, expressive language, and intelligibility, thereby capturing key aspects of early language development. As shown in Table 1, the scores on each of the subscale items were separated into seven age groups. Differences between age groups on the subscales were examined using one-way analyses of variance (ANOVAs). In the motor subscale, significant differences were found across age groups for all items and the total score (e.g., Item 1: $F(6, 671) = 13.21$, $P < 0.001$, $\eta^2 = 0.10$; Item 2: $F(6, 671) = 86.11$, $P < 0.001$, $\eta^2 = 0.18$; Item 7: $F(6, 671) = 16.77$, $P < 0.001$, $\eta^2 = 0.12$; Total score: $F(6, 671) = 126.14$, $P < 0.001$, $\eta^2 = 0.22$). Post-hoc analyses using Tukey's HSD revealed that younger children (ages 2:6–2:11 and 3–3:5) scored significantly lower than older groups (ages 4:6–4:11 through 5:6–5:11), while adjacent age groups often did not differ significantly, reflecting gradual motor development. In the concepts subscale, all items and the total score showed significant differences between age groups (e.g., Item 2: $F(6, 671) = 40.19$, $P < 0.001$, $\eta^2 = 0.15$; Item 4: $F(6, 671) = 54.41$, $P < 0.001$, $\eta^2 = 0.16$; Total score: $F(6, 671) = 70.83$, $P < 0.001$, $\eta^2 = 0.19$). Tukey's HSD indicated that performance improved significantly with age, especially between the youngest children (2:6–2:11) and those older than 4 years, consistent with expected perceptual growth. In the language subscale, significant age-group differences were also observed across all items and the total score (e.g., Item 1: $F(6, 671) = 40.80$, $P < 0.001$, $\eta^2 = 0.14$; Item 3–actions: $F(6, 671) = 31.75$, $P < 0.001$, $\eta^2 = 0.12$; Total score: $F(6, 671) = 66.53$, $P < 0.001$, $\eta^2 = 0.17$). Post-hoc comparisons demonstrated that younger groups performed significantly lower than children aged 4:6 years and above, particularly on expressive language and rhyming tasks, reflecting rapid gains in early language acquisition. A statistical significance threshold of $P <$

0.05 was adopted for all analyses. Overall, the results confirmed that performance on motor, concepts, and language tasks increases steadily with age, aligning with expected developmental trajectories.

Table 1. Mean and Standard Deviation of Subscale Scores Across Age Groups

	Subscale	2:6 to 2:11		3 to 3:5		3:6 to 3:11		4 to 4:5		4:6 to 4:11		5 to 5:5		5:6 to 5:11	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Motor	Item1	2.38	1.11	2.72	1.29	2.75	1.21	3.38	0.96	3.71	1.02	3.80	0.94	3.86	1.10
	Item2	0.76	0.83	2.34	1.36	3.02	1.57	3.87	1.54	4.66	0.73	4.84	0.44	4.93	0.38
	Item3	0.90	0.62	1.79	1.05	2.25	1.27	3.04	1.23	3.66	1.34	4.02	1.28	4.65	0.75
	Item4	0.95	1.16	2.27	1.48	2.70	1.28	3.53	1.41	4.28	1.39	4.53	0.87	4.63	0.87
	Item5	0.09	0.30	0.90	0.86	1.40	1.10	2.23	1.32	2.94	1.55	3.58	1.37	4.21	1.05
	Item6	0.71	0.95	1.60	1.09	2.05	1.28	2.57	1.29	3.74	1.16	4.28	1.04	4.69	0.70
	Item7	0	0	0	0	0	0	0	0	0.38	1.34	1.02	1.75	1.89	2.17
	total	5.79	2.82	11.62	4.64	14.20	5.26	18.72	5.09	23.38	5.22	26.09	4.90	28.86	4.50
Concept	Item1	1	0.44	1.41	0.87	1.62	0.77	1.76	1.10	2.12	1.17	2.55	1.24	2.92	1.29
	Item2	0.66	1.01	1.79	1.35	2.65	1.64	3.70	1.33	3.79	1.32	4.06	1.06	4.10	1.09
	Item3	0.92	0.49	1.16	0.65	1.80	0.99	2.23	0.81	2.89	1.18	3.51	1.01	3.83	1.06
	Item4	0.23	0.43	0.55	0.76	1.02	0.86	1.31	1.25	1.97	1.13	2.55	1.42	3.69	1.16
	Item5	0.57	0.59	1.20	0.55	1.32	0.81	1.51	0.89	1.80	0.66	2	1.15	2.37	1.44
	Item6	0.80	0.92	1.76	1.37	2.25	1.44	3.31	1.58	3.89	1.42	4.57	0.90	4.66	0.82
	Item7	1	0.83	1.34	1.11	1.85	1.25	1.82	1.72	2.69	1.97	3.02	1.94	3.93	1.77
	total	5.19	3.10	9.04	4.39	12.77	5.16	15.06	5.79	18.25	6.11	21.53	6.08	25.56	5.90
Language	Item1	0.47	0.67	1.18	0.95	1.62	0.74	2.19	1.15	2.23	0.66	2.61	0.77	3.12	1.01
	Item2	3.95	1.35	4.37	1.41	4.72	0.59	4.74	0.64	4.79	0.87	4.86	0.49	5	0
	Item3ob	0.71	0.72	0.93	0.59	1.70	1.11	1.87	0.92	2.28	1.16	2.73	1.23	3.15	1.25
	Item3 ac	0.61	0.80	1.18	1.15	2.25	1.53	2.75	1.70	3	1.62	3.64	1.32	3.92	1.39
	Item4	0	0	0	0	0	0	0	0	0.23	0.98	0.72	0.95	0.76	0.94
	Item5	0	0	0	0	0.12	0.46	0.27	0.68	0.33	0.73	1.29	1.30	1.77	1.46
	Item6	0.33	0.48	0.97	1.12	1.45	1.21	2.59	1.58	2.78	1.39	3.31	1.45	3.72	1.36
	total	6.09	2.98	8.65	3.91	11.95	4.13	14.25	4.84	15.25	4.43	19.21	4.82	21.50	4.82

Table 2 presents the comparison of DIAL-4 subscale scores between girls (n = 312) and boys (n = 366). Independent samples t-tests indicated no statistically significant gender differences in the motor, concepts, language, or total scores (all $P > 0.05$). Effect sizes were negligible across domains, suggesting that developmental performance was comparable between boys and girls. These results support the use of unified age-based norms for both genders, consistent with the study's objective of establishing broadly applicable developmental benchmarks for Iranian preschool children.

Table 2. Comparison of Girls and Boys on DIAL Subscales Using Independent t-Test

	Girls		Boys			
Subscales	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**t**	**P**
motor	21.51	8.77	20.57	8.32	1.02	0.30
concepts	17.74	8.84	17.41	8.13	0.37	0.71
language	15.72	6.90	15.57	6.13	0.21	0.83
total	54.98	23	53.57	21.46	0.59	0.55

Tables 3-5 present the normative conversion of raw scores into developmental scores with a mean of 100 and a standard deviation of 15 for the motor, concepts, and language domains. Developmental scores were calculated directly from raw scores using a linear transformation, following established psychometric methods applied in instruments such as the Wechsler and Stanford–Binet scales:

$$DS = 100 + 15 \times \frac{(X - M_{raw})}{SD_{raw}}$$

where X is the child's raw score, Mraw is the mean raw score for the age group, and SDraw is the standard deviation of raw scores. This transformation places children's performance on a standardized scale that is directly comparable across domains and age groups. Cut-off points were defined according to psychometric standards and the DIAL-4 guidelines: scores below 70 indicate severe developmental delay, scores between 70–84 indicate mild delays, scores between 85–115 fall within the average range, scores between 116–129 represent above-average performance, and scores ≥ 130 indicate precocious development. These thresholds provide clinically meaningful benchmarks for identifying children at risk of developmental delays and for distinguishing between typical and atypical performance.

Table 3. Normalization of Motor Area Scores Across Age Groups

Raw Scores	Standardized Developmental Score							Raw Scores
	2:6 to 2:11	3 to 3:5	3:6 to 3:11	4 to 4:5	4:6 to 4:11	5 to 5:5	5:6 to 5:11	
1	75	63	64	45				1
2	81	66	67	48	40			2
3	86	70	70	51	43			3
4	92	74	73	55	46			4
5	98	77	76	58	49	40		5
6	104	81	79	61	52	43		6
7	109	85	81	64	54	46		7
8	115	88	84	67	57	49		8
9	121	92	87	71	60	52		9
10	126	96	90	74	63	55		10
11	132	99	93	77	66	57		11
12	138	103	95	80	69	60		12
13	144	107	98	83	72	63	42	13
14	149	110	101	87	75	66	45	14
15	155	114	104	90	78	69	49	15
16	161	118	107	93	80	72	53	16
17	166	121	109	96	83	75	56	17
18	172	125	112	100	86	78	60	18
19	178	129	115	103	89	81	64	19
20	183	132	118	106	92	89	68	20
21		136	121	109	95	84	71	21
22		139	123	112	98	87	75	22
23		143	126	116	101	92	79	23
24		147	129	119	104	95	83	24
25		150	132	122	106	98	86	25
26		154	135	125	109	101	90	26
27		158	138	128	112	104	94	27
28		161	140	132	115	107	98	28
29		165	143	135	118	110	101	29
30		169	146	138	121	113	105	30
31		172	149	141	124	116	109	31
32		176	152	145	127	119	112	32
33		180	154	148	130	121	116	33
34		183	157	151	132	124	120	34
35		187	160	154	135	127	124	35

Table 4. Normalization of Concepts Area Scores Across Age Groups

Raw scores	Standardized Developmental score							Raw scores
	2:6 to 2:11	3 to 3:5	3:6 to 3:11	4 to 4:5	4:6 to 4:11	5 to 5:5	5:6 to 5:11	
1	80	73	66	63	59	53		1
2	84	77	69	66	62	56		2
3	89	81	72	69	64	58		3
4	93	85	75	71	67	61		4
5	98	89	78	74	69	63	43	5
6	103	93	81	76	72	65	46	6
7	107	97	84	79	74	68	49	7
8	112	100	87	82	76	70	51	8
9	117	104	90	84	79	73	54	9
10	121	108	92	87	81	75	57	10
11	126	112	95	89	84	77	60	11
12	131	116	98	92	86	80	63	12
13	135	120	101	95	89	82	66	13
14	140	124	104	97	91	85	68	14
15	145	128	107	100	94	87	71	15
16	149	132	110	102	96	90	74	16
17	154	136	113	105	99	92	77	17
18	159	139	116	108	101	94	80	18
19	163	143	119	110	104	97	82	19
20	168	147	122	113	106	99	85	20
21	172	151	124	115	108	102	88	21
22	177	155	127	118	111	104	91	22
23	182	159	130	121	113	106	94	23
24		163	133	123	116	19	96	24
25		167	136	126	118	111	99	25
26		171	139	128	121	114	102	26
27		175	142	131	123	116	105	27
28		179	145	133	126	119	108	28
29		182	148	136	128	121	110	29
30			151	139	131	123	113	30
31			154	141	133	126	116	31
32			156	144	135	128	119	32
33			159	146	138	131	122	33
34			162	149	140	133	125	34
35			165	152	143	135	127	35

Table 5. Normalization of Language Area Scores Across Age Groups

Raw Scores	Standardized Developmental Score							Raw Scores
	2:6 to 2:11	3 to 3:5	3:6 to 3:11	4 to 4:5	4:6 to 4:11	5 to 5:5	5:6 to 5:11	
1	73	70	61	61	52	46		1
2	79	74	65	64	55	50		2
3	84	79	69	67	59	53		3
4	89	83	72	70	62	56	44	4
5	94	87	76	73	66	59	47	5
6	100	92	79	76	69	62	51	6
7	105	96	83	79	72	65	54	7
8	110	100	86	82	76	68	57	8
9	115	105	90	85	79	71	60	9
10	120	109	93	88	83	74	63	10
11	126	113	97	91	86	78	67	11
12	131	118	100	94	90	81	70	12
13	136	122	104	97	93	84	73	13
14	141	126	107	100	97	87	76	14
15	146	130	111	104	100	90	80	15
16	152	135	114	107	104	93	83	16
17	157	139	118	110	107	96	86	17
18	162	143	121	113	110	99	89	18
19	167	148	125	116	114	102	92	19
20	172	152	128	119	117	152	96	20
21		156	132	122	121	109	99	21
22		161	135	125	124	112	102	22
23		165	139	128	128	115	105	23
24		169	142	131	131	118	108	24
25		174	146	135	134	121	112	25
26		178	149	138	137	124	115	26
27			153	142	140	127	118	27
28			156	145	143	130	121	28
29			160	148	147	134	124	29
30			163	152	150	137	128	30
31			167	155	153	140	131	31
32			170	159	156	143	134	32
33			174	162	159	146	137	33
34			177	166	162	149	141	34
35			181	169	165	152	144	35

The motor area consists of 23 practical items, and the minimum raw scores in this subtest is 1 and the maximum is 35. Table 3 shows the normative scores for the motor area on the developmental score for different age groups from 2 years and 6 months to 5 years and 11 months and shows the relationship between the raw scores and the standardized developmental score (with a mean of 100 and a standard deviation of 15). With increasing age, a higher raw score is required to achieve a given standardized developmental score. In this table, the cut-off points for interpreting the scores are usually considered psychometric standards: a standardized developmental score below 70 indicates a severe motor delay, between 70 and 84 mild delays, between 85 and 115 in the normal range, between 115 and 129 above average, and above 130 in the precocious range. These cut-off points help to accurately assess the motor development of children based on their age and performance. In Table 3, for each age group, a raw score that is equivalent to a standardized developmental score

of approximately 100 is considered the norm score. This score indicates "average" motor performance and is appropriate for the child's age. The approximate raw scores corresponding to a standardized developmental score of 100 are given below for each age group:

- 2 years and 6 months to 2 years and 11 months: A raw score of approximately 5 is equivalent to a standardized developmental score of approximately 100; therefore, the norm score in this group is 5.
- 3 years to 3 years and 5 months: A raw score of approximately 11 is approximately equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 11.
- 3 years and 6 months to 3 years and 11 months: A raw score of approximately 14 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 14.
- 4 years to 4 years and 5 months: A raw score of approximately 18 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 18.
- 4 years and 6 months to 4 years and 11 months: A t score of 23 is equivalent to a standardized developmental score of 100; the norm score for this group is 23.
- 5 years to 5 years and 5 months: A t score of 26 is equivalent to a standardized developmental score of 100; so, the norm score for this group is 26.
- 5 years and 6 months to 5 years and 11 months: A t score of 29 is equivalent to a standardized developmental score of 100; thus, the norm score for this group is about 29.

The concepts area consists of 58 practical items, and the minimum raw score in this subtest is 0 and the maximum is 35. In Table 4, for each age group, a raw score that is equivalent to a standardized developmental score of approximately 100 is considered the norm score. This score indicates "average" conceptual development and is appropriate for the child's age. The approximate raw scores corresponding to a standardized developmental score of 100 are given below for each age group:

- 2 years and 6 months to 2 years and 11 months: A raw score of approximately 5 is equivalent to a standardized developmental score of approximately 100; therefore, the norm score in this group is 5.
- 3 years to 3 years and 5 months: A raw score of 8 is approximately equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 8.
- 3 years and 6 months to 3 years and 11 months: A raw score of 13 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 13.
- 4 years to 4 years and 5 months: A raw score of 15 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 15.
- 4 years and 6 months to 4 years and 11 months: A raw score of 18 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 18.
- 5 years to 5 years and 5 months: A raw score of 20 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 20.
- 5 years and 6 months to 5 years and 11 months: A raw score of 25 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is about 25.

The language area consists of 52 practical items, and the minimum raw score in this subtest is 0 and the maximum is 35. In Table 5, for each age group, a raw score that is equivalent to a standardized developmental score of approximately 100 is considered the norm score. This score indicates “average” language development and is appropriate for the child’s age. The approximate raw scores corresponding to a standardized developmental score of 100 are given below for each age group:

- 2 years and 6 months to 2 years and 11 months: A raw score of approximately 6 is equivalent to a standardized developmental score of approximately 100; therefore, the norm score in this group is 6.
- 3 years to 3 years and 5 months: A raw score of 8 is approximately equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 8.
- 3 years and 6 months to 3 years and 11 months: A raw score of 12 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 12.
- 4 years to 4 years and 5 months: A raw score of 14 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 14.
- 4 years and 6 months to 4 years and 11 months: A raw score of 15 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 15.
- 5 years to 5 years and 5 months: A raw score of 18 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is 18.
- 5 years and 6 months to 5 years and 11 months: A raw score of 21 is equivalent to a standardized developmental score of 100; therefore, the norm score for this group is about 21.

Discussion

This study standardized the DIAL-4 for Iranian children aged 2 years and 6 months to 5 years 11 months, establishing age-specific developmental norms for the motor, concepts, and language subscales. The findings demonstrated systematic differences across age groups, confirming the developmental sensitivity of the DIAL-4

and supporting its use as a culturally adapted screening tool in Iran.

In this study, "standardized developmental scores" (mean = 100, SD = 15) were used to categorize developmental status. Scores below 70 indicate severe developmental delay, scores between 70 and 84 reflect mild delay, and scores between 85 and 115 represent the normative range. Scores between 116 and 129 are considered above average, and those above 130 indicate advanced development.

The performance of Iranian children in this study was compared with that of American children reported in the Mardel-Czudnowski and Goldenberg (15). The results indicated notable differences in standardized developmental scores between Iranian and American children across all age groups. In the youngest group (2:6–2:11 years), Iranian children scored lower than American peers in motor, concepts, and language areas, with the greatest deficit observed in concepts. For the 3–3:5-year group, motor development was comparable to American children, while deficits remained in language and concepts, particularly in concepts. Children aged 3:6–3:11 years showed lower scores than American peers in all three domains, with the largest gap again in concepts.

In the 4–4:5 and 4:6–4:11-year groups, Iranian children continued to score lower than American children across all areas, with the greatest weakness shifting to language development by 4:6–4:11 years. Similarly, in the 5–5:5 and 5:6–5:11-year groups, scores in motor, concepts, and language were below those of American children, with the most pronounced deficit in language. Overall, while motor development approached parity with American norms in some age groups, Iranian children consistently exhibited lower performance in concepts and language, highlighting areas for targeted developmental support.

These disparities may reflect cultural and contextual factors. In Iran, preschool programs often emphasize socialization and play, with limited structured instruction in early literacy or phonological awareness. The absence of systematic training in phonological awareness, letters, and sound recognition may explain the marked differences in language scores, especially after age 4 years and 6 months, when these skills become more prominent in the DIAL-4. Similarly, the weaker performance in concepts reflects limited exposure to structured early numeracy and problem-solving activities in preschool curricula. These findings align with cross-cultural research indicating that early educational content and parental practices significantly influence developmental outcomes (20). In Iran, kindergartens typically do not provide formal training in letter recognition or phonological skills due to curricular guidelines discouraging pre-literacy instruction. This contrasts with U.S. preschool curricula, which frequently incorporate structured literacy and numeracy activities, contributing to stronger performance in concepts and language domains (20). These findings highlight the importance of culturally specific norms for developmental screening tools.

Limitation

This study had several limitations. First, the sample was drawn from Tehran and Bojnourd, which limits generalizability to other regions and cultural groups in Iran. Future research should replicate the study across diverse provinces, including rural and minority populations. Second, the sample included only typically developing children from mainstream kindergartens, limiting applicability to children with special needs. Future work should examine the psychometric properties and norms of the DIAL-4 in children with developmental disorders to expand its utility. Third, one limitation of this study was the uneven distribution of participants across age groups. Although the total sample (N = 678) met the target for adequate precision, younger groups were slightly under-represented and older groups slightly over-represented, which may slightly affect the stability of normative estimates. Fourth, teacher and parent report questionnaires of DIAL were not utilized in the present study for the standardization process. Future studies are recommended to include these complementary reports to obtain a more comprehensive assessment of children's developmental functioning across different settings and informants. Future studies should also evaluate the predictive validity of the DIAL-4 norms by following children longitudinally to assess how early developmental scores relate to later academic outcomes. Investigating the impact of interventions informed by the DIAL-4 on developmental trajectories would strengthen its practical value. Developing training modules for DIAL-4 administrators in Iran could enhance reliability and ensure effective nationwide implementation. The standardized DIAL-4 norms developed in this study have important applications for educators, clinicians, and policymakers. They provide a reliable framework for screening developmental delays in preschool settings, enabling timely referral and intervention.

Conclusion

This study standardized the DIAL-4 for Iranian children aged 2 years and 6 months to 5 years 11 months, providing age-specific standardized scores for motor, concepts, and language subscales. These norms fill a critical gap in developmental assessment in Iran. The findings revealed that Iranian children scored lower than American children, particularly in concepts and language, highlighting cultural and educational influences on developmental outcomes. By establishing localized norms, this study enables accurate identification of developmental delays and supports timely interventions, contributing to improved developmental outcomes and public health in Iran.

Acknowledgment
We would like to express our sincere gratitude to all the teachers, administrators, parents, and their children who participated in this research.

Funding
This study was conducted with financial support from the Cognitive Science and Technologies Council under grant number CT1402102260.

Conflict of Interest
None.

Author's Contributions
Malahat Amani and Meysam Sadeghi were the executors of the research project and were responsible for the design and implementation of the research. Shadi Zolfaghari, played a role in the process of data collection and training of examiners. Malahat Amani was also responsible for writing the article.

References
1. Harm M, Hope M, Household A. American psychiatric association, 2013, diagnostic and statistical manual of mental disorders, 5th edn, washington, dc: American psychiatric association anderson, j, sapey, b, spandler, h (eds.), 2012, distress or disability?, lancaster: Centre for disability research. Arya. 2013;347:64.
2. Organization WH, UNICEF. Global report on children with developmental disabilities: from the margins to the mainstream. World Health Organization: Geneva, Switzerland. 2023.
3. Sajedi F, Doulabi MA, Vameghi R, Baghban AA, Mazaheri MA, Mahmodi Z, et al. Development of Children in Iran: A Systematic Review and Meta-Analysis. Glob J Health Sci. 2016;8(8):51251.
4. Rydz D, Srour M, Oskoui M, Marget N, Shiller M, Birnbaum R, et al. Screening for developmental delay in the setting of a community pediatric clinic: a prospective assessment of parent-report questionnaires. Pediatrics. 2006;118(4):e1178-86.
5. Shamsaei F, Shokohi Z, Farhadian M, Sadeghian E. Comparison of Social Skills of Primary School Children With and Without Parents With Mental Disorders. Avicenna Journal of Nursing and Midwifery Care. 2019;27(2):97-106.
6. Zamyad A, Yasemi M, Vaezi SA. Preliminary standardization of vineland adaptive behavior scale in urban and rural population of Kerman. Iranian Journal of Psychiatry and Clinical Psychology. 1996;2(4):44-55.
7. Amani M, Reihani L, Rad ASI. Determinig Psychometric Properties of Persian Version of the Early Learning Observation and Rating Scale.Iraninan Journal of Pediatric nursing. 2024;10(2):11-24.
8. Salehi Zahabi S, Ghorbani A, Jalilehvand N, Kamali M. Development and determine the Psychometric properties of Picture perceptive Objective Vocabulary Test for normal Persian-speaking 6-13 years-old children. Journal of Modern Rehabilitation. 2016;9(6):159-67.
9. Abedi A, Malekpour M, Oraizi H, Faramarzi S, Paghale SJ. Standardization of the neuropsychological test of NEPSY on 3-4 years old children. Iranian Journal of Psychiatry and Clinical Psychology. 2012;18(1), 52-60.
10. Soleimani F, Azari N, Vameghi R, Sajedi F, Shahshahani S, Karimi H, et al. Are the Norms of Bayley Screening Test Appropriate for Persian Language Children? Iran J Child Neurol. 2018;12(2):91.
11. Shahshahani S, Vameghi R, Azari N, Sajedi F, Kazemnejad A. Has Denver developmental screening test-II appropriate validity and reliability for screening of developmental disorders in 0-6 years old children in Tehran city? Payesh. 2011; 10 (4) :469-475.
12. Karimi M, Emarati A, Nafei Z, Shamsi F, Gashty Mazar N, Akbarian E. Unveiling developmental delays in early childhood: insights from a comparative study of the Bayley Scales (BSID-III) and the ASQ-3 in Iran. BMC pediatrics. 2025;25(1):158.
13. Mardell-Czudnowski C, Goldenberg D. DIAL-R: Developmental Indicators for the Assessment of Learning--Revised: American Guidance Service; 1990.
14. Mardell-Czudnowski C, Goldenberg D. Developmental indicators for the assessment of learning—revised 1983 Edison. NJ Childcraft Education [manual].
15. Mardell C, Goldenberg D. Developmental Indicators for the Assessment of Learning. (No Title). 1998.
16. Coughlan KA. Test Review: Developmental Indicators for the Assessment of Learning, Fourth Edition (DIAL-4), by C. Mardell-Czudnowski & D. S. Goldenberg. J Psychoeduc Assess. 2015;33(3):291-5.
17. Allin MJ. The effect of the Missouri preschool program on school readiness: Lindenwood University; 2020.
18. Moodie S, Daneri P, Goldhagen S, Halle T, Green K, LaMonte L. Early Childhood Developmental Screening: A Compendium of Measures for Children Ages Birth to Five. OPRE Report 2014-11. US Department of Health and Human Services. 2014.
19. Sadeghi M, Amani M, Abbassi H, Zolfaghari S. Psychometric characteristics of Developmental Indicators for the Assessment of Learning (DIAL). Educational Measurement, 2025; 62(4), in press.
20. Melhuish E, Ereky-Stevens K, Petrogiannis K, Ariescu A, Penderi E, Rentzou K, et al. A review of research on the effects of Early Childhood Education and Care (ECEC) upon child

development. Curriculum Quality Analysis and Impact Review of European Early Childhood Education and Care. European Early Childhood Education and Care Publications. 2015.