Ilkogretim Online - Elementary Education Online, 2022; Vol 21 (Issue 3): pp. 195-214
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2022.03.13

Legalization Of Kuwaiti Version From Stanford–Binet Early Childhood Scale, Fifth Edition

Ahmed Nazzal Al-Rajhi , ra66di@yahoo.Com

Abstract

The current study aimed to standardize the Stanford Binet Scale, the fifth picture, SB5, for children in the Kuwaiti environment to reach building standards that are reliable in making appropriate decisions and to be an honest and codified tool in the Kuwaiti environment to determine the mental abilities of children in early childhood at the age of (4-7 years). The study population consisted of (92592), including (39,065) students in kindergarten and (53527) students in the first and second grades of primary school for the academic year 2019-2020, and the sample of the study amounted to (450) children in the age group (4-7 years). From all the governorates of Kuwait, the study sample was selected in a random cluster manner, and the Stanford-Binet Scale, Fifth Image (SB5), a standardized version, was used. The results indicated that both the difficulty coefficients and the discrimination coefficients for the items of the non-verbal side are statistically significant, and all of them indicate the effectiveness of the items in distinguishing between students of the higher and lower categories. during factor analysis; The highest percentage of variance explained was the liquid inference factor which explained 24% of the total variance, followed by the knowledge factor, then the quantitative inference factor, then the visuospatial processing factor, then the working memory factor. Finally, the results indicated that there were differences between the arithmetic averages according to the variable of classification of students (normal and people with disabilities), and all the differences were on the scale in favor of the ordinary, as evidenced by their higher arithmetic averages in all fields than the arithmetic averages for the disabled and this indicates the discriminatory ability of the test, and these results were discussed in The light of previous studies and theoretical literature.

Introduction

Measurement is one of the necessary things in human life, through which the phenomenon can be described and given a value that describes it as a quantitative description, which helps in determining this characteristic and its degree. The standards have been found according to the need for them, and special standards have been set for them that show the extent that they describe, and because humans differ in many characteristics and abilities, because of the importance of the capabilities that the

individual enjoys and their impact on many of his aspects, it was necessary to identify those abilities and find out about them. One of the most prominent abilities affecting the individual is his mental abilities, which are determined by the percentage of intelligence which prompted researchers and those interested in the field of education, specifically, special education, and psychology to search for special means and tools that measure mental abilities, called intelligence scales. One of the first of these scales is the Stanford-Binet scale, which first appeared in 1905 by the French scientist Binet and his assistant Simon. Then it was developed until it was used in the United States of America by Louis Terman and his students in 1916, where signs of sincerity and stability were found in the American environment, and then several revisions to the scale appeared that sought to address the gaps, criticisms, and shortcomings that were directed at the scale until it ended up in the fifth version, which is the subject of the current study which Gal Roid did with his team from 1996 to 2003, because of the importance of the IQ ratio and its impact on the different areas of the individual, and because the Binet Intelligence Scale enjoys high indications of sincerity and stability, the scale has been codified and applied to many different environments around the world.

Many studies and research conducted using the developed SB5 scale confirmed its importance and credibility in measuring intelligence, as it is a major criterion for various measures of intelligence. When the fifth version, improved from the fourth version of the Stanford Binet Scale was issued in 2003, which was built on the same theory, with the addition of some tests, the most important of which are the tests of measuring working memory, which came in line with recent developments in the field of measurement. Research is still being conducted on the Binet scale and its legalization in all environments, and its use to measure the mental abilities of different groups (Yassin and Shaheen, 2015, Al-Balaa, 2015, Muhammad 2013, Fancher & Rutherford, 2012, Faraj 2011, Taha and Farhan 2011, Faraj and Bakheet 2012, William, McIntosh, Dixon, Newton & Youman, 2010). In addition to comparing and evaluating copies, especially Arabic ones (Khattab, 2013).

Thus, the Stanford Binet Scale in its fifth edition was developed by Roid and released in 2003. In 2005, a modified version of Early Childhood (Early SB5) was released for Roid. The scale is a cognitive test of ability and intelligence used to diagnose developmental or intellectual deficiencies in young children 2-7 years old. The test measures five factors and consists of verbal and nonverbal tests. The five factors that were tested were: (the logical reasoning factor, the knowledge factor, the quantitative reasoning factor, the visual-spatial processing factor, and the working memory factor) (Roid 2005, Roid 2003).

accordingly; According to his knowledge, the researcher believes that there is no legalization of the Roid 2003 SB5 standard in the Kuwaiti environment. So, because of the great importance of this scale in the field of intelligence measurement, its credibility and

its global fame, and the importance of the age group covered by the thesis of (4-7) years, which is the early stage of childhood, the researcher decided to codify the Stanford Binet Scale, Fifth Edition (SB5), the original American version on the Kuwaiti environment. - which is the first scale of its kind codified in Kuwait-, after Arabizing it and taking the necessary legalization procedures to issue the Kuwaiti image of the Stanford Binh Fifth Scale.

Study Problem:

In light of the above, the problem of the study is determined in the following questions:

What is the effectiveness of the Kuwaiti version of the Binet Scale, the fifth version for the age group (4-7 years) in measuring IQ?

The following sub-questions emerge from the study problem:

1. What is the significance of the validity of the Kuwaiti version of the Stanford Binet Scale after applying it to the study sample in the Kuwaiti environment with the characteristics of a good and reliable test in measuring intelligence for the age group (4-7)?
2. What are the implications of the stability of the Kuwaiti version of the Stanford Scale between the fifth version of early childhood after applying it to the study sample in the Kuwaiti environment with the characteristics of a good and reliable test in measuring intelligence for the age group (4-7 years)?
3. What are the performance criteria extracted from SB5 after applying it to Kuwaiti environmental children aged (4-7 years)?

Study importance:

The importance of the study is reflected in the following two areas:

Theoretical importance:

- A standardization procedure for the Stanford Binet Intelligence Scale, fifth version, SB5, valid for the Kuwaiti environment; For the age group of 4-7 years.

Practical importance:

- Providing a Kuwaiti version of the Stanford Binet Scale that helps professionals, psychologists, parents, and decision-makers in diagnosing and classifying individuals and directing them towards the scientific path appropriate to their practical abilities.

Study limits and limitations

The limits of the study are as follows:

- Human limits: the study was limited to a sample of children in the age group (4-7 years) of normal and abnormal children with mental disabilities in kindergarten, first and second grades of primary school, and children with mental disabilities in the governorates of Kuwait.
- Time limits: The study was implemented in the academic year 2020/2021.
- The study is defined by the tool used and the response of the sample members.

Terminological and procedural definitions

The following are the terms used in the study:

1. Legalization: "It is to draw a clear and comprehensive picture of all the steps of the test, interpret its results, determine the appropriate place and time, and specify the appropriate terms. This is called objectivity and usability" (Al-Nimr 87:2018).

The researcher defined procedural legalization as "a process through which the measurement tool is modified to suit the local community, through appropriate modification and extraction of appropriate standards for the community."

2. The Stanford Binet Intelligence Scale, Fifth Edition SB5: It is a measure of the mental abilities of individuals, developed by Roid from the fourth version of the scale, for the age group from 2-85 years (Roid, 2003).
3. Early Childhood: The International Council defined early childhood children (2005) as "children from the moment of birth until the age of eight" (Karam El-Din 3: 2015).

   The researcher defines early childhood procedurally as "children whose ages range between (4-7) years, and they represent the study sample of normal children.
4. Intelligence: The researcher defines intelligence procedurally in this study as the degree obtained by the subject on the Stanford-Binet scale.

Theoretical framework and previous studies

First: theoretical literature

The process of measurement began in ancient times when humans needed to compare things or individuals, and the measurement continued in this way until the mid-nineteenth century in the United States of America when academic achievement was measured by oral exams, then it developed and varied ways of measuring it. In 1905, Alfred Binet and Simon published the first intelligence test, commissioned by the French Ministry of Public Education to find a measurement tool capable of separating the students who can learn from those who are unable to learn in schools to provide a quality education that suits their abilities. After that, psychometrics went hand in hand with psychological studies as an integral part of every human activity (Taha, 2006).

Concept of intelligence:

Intelligence is an intelligent source of language, clever and smart, and the stoked the fire, i.e inflamed and her flame intensified. Intelligence is the speed of acumen (Abadi, 2011). Terminology, although intelligence is one of the most studied terms and concepts in scientific research, and has received the attention of theorists and scientists, there is no specific agreement on the concept of intelligence. As its concept varies from one scientific to another, and even to the same scientific over time. As Spearman defined it (Sperman, 1945): it is the ability or factor and the general innate willingness that affects the mechanism of work of mental activity in all its forms and topics" (Al-Manqal 4-5 2013), and Wachsler (Wachsler 1983) defined it as "the general overall ability to perform an intentional action, think rationally, and interact with the environment efficiently." As for Binet (Binet, 1911): it consists of four capabilities, namely: understanding, innovation, criticism, the ability to direct thought processes in a particular aspect, and the continuity of engaging in it" (Al-Manqal, 2013).

Stanford Binet Scale, Fifth version SB5

In 2003, Roid developed the scale so that the Stanford Binet Fifth Edition (SB5) intelligence test became one of the most popular intelligence tests (Roid, 2003); The SB5 is a newly published multidimensional measure of intelligence based on the Cattell-Horn-Carroll (C-H-C) theory which is an integrative model in psychology. The Roid test presented results from the confirmation factor analysis in the technical evidence supporting the five factors of the tool (Becker 2003, William et al., 2010). With the progress of the development process on this scale. Roid developed the scale in 2005 to become the Stanford Binet Scale for Early Childhood (Early SB5), for the age group (2-7) years, and this designed test measures (the mental ability of children), and Roid (2005) derived this scale from the original image of the scale in its fifth form and made subtle adjustments to make it more suitable for early childhood.

Components of Beni Test Bag Fifth Version

The current version is based on the idea of the general factor, that is, the general mental ability, so he measures it in two main areas: the verbal domain and the non-verbal domain. In the fifth version, as mentioned (Faraj 2011), there are two entrance tests, in the verbal domain we measure it with the information test, which measures knowledge, and in the non-verbal domain, we measure it with chains of things that measure analytical inference. The first two tests have no levels because they were measured in the entrance, and the verbal test five levels, and the non-verbal test six levels, the sub-tests are called activities, these activities are called the test unit, and each unit consists of six tests, and each level measures four factors, and each factor on 6 questions for each level are fully applied, that

is, the entire page is applied regardless of the examinee's errors and inability to answer, and the first level contains only two factors and its score is only 4.

Second: Previous studies:

Previous Arab and foreign studies relevant to the present study will be presented in this chapter, namely:

Ruf (2003) conducted a study aimed at identifying the credibility of the SB5 scale, by evaluating it from the previous versions of the scale, and knowing the ability of the fifth version of the scale in detecting the gifted, the researcher used 60 studies published on the gifted. The sample for gifted students consisted of 5 cases of different degrees. The study used the fifth, fourth and third versions of the Stanford-Binet scale, in addition to different editions of the Wechsler and Woodcock-Johnson scale, and the cognitive abilities test. The study concluded that the fifth picture of the Stanford Binet Scale is an effective tool in detecting gifted people in different age groups and that the SB5 measures different mental abilities.

The Roid (2005) study aimed to review a specialized version of the Stanford Binet for the age group of (2-7) years in the United States of America to extract the Stanford Binet Scale, the fifth version of early childhood (Early SB5). There were indications about the validity of the scale, represented in the correlative validity between performance on the scale, and Wechsler's scale for preschoolers. Participation in the Gestalt Optical Kinetic Test Edition 2 (starting at age 4) and the Test Control Form (starting at age 2). In addition to the availability of indications about the stability of the scale using factor analysis, the study reached to extract criteria for the scale by calculating the arithmetic averages and standard deviations of all members of the study sample on the subscales and the total degree.

The study of Abu Abu El-Nil (2012) also aimed to examine the following hypotheses: There is a distinct knowledge page for the superior and the academically late (the study sample) on the Stanford-intermediate scale of the fifth version. There are statistically significant differences between the mean scores of (excellent, retarded) academically in the sub-tests, factors, and IQ ratios on the Stanford Binet Scale, fifth version. There are statistically significant differences between the average scores of superior students (males and females) in the sub-tests and the factors and IQ ratios on the Stanford Binet Scale - the fifth version. The sample consisted of 60 male and female students, and the sample was divided into (outstanding and late). The results showed that there is a distinct knowledge page for both the high achievers and the level of significance of the grades, and it was higher than the average and higher in several aspects that reach the levels of excellence and also the academic retarded (the study sample). The cognitive

page showed low scores in quantitative reasoning and visual processing, while average scores were achieved in several sub-tests on the Stanford-Binet Intelligence Scale, the fifth version. There are statistically significant differences between academically superior and backward students in the factors and sub-tests on the Stanford Binet Scale, fifth version. There are also statistically significant differences between superior males and females in the sub-tests on the Stanford-Binet scale in knowledge, quantitative non-verbal inference, and verbal fluid inference in favor of males, while the late females obtained higher percentages in the sub-tests of verbal fluid inference, non-verbal knowledge, and verbal working memory.

The Al-Balaa study (2015) aimed to work on conducting the statistical study on the Stanford Binet Scale - the fifth picture to ensure its validity and stability, as well as setting appropriate initial criteria for the Syrian society separated according to the age variable of the five measures of intelligence. As well as working on extracting a guide for the scale showing the method of applying and correcting the scale, and the scale tools have been modified to suit the community environment. The researcher followed the descriptive analytical approach with its theoretical and analytical sides. The sample consisted of 306 children aged 4-6 years. Some items of the scale were rearranged according to the results of the transactions. The Stanford Binet Scale - 5th version, the multi-level battery test, was used. The Raven test for colored progression matrices. After completing the study of the validity and stability of the scale and extracting the criteria, it was produced in its final form, accompanied by a guide to its use. It consisted of three booklets for items and an answer book that contains demographic variables. Differences were studied for each of the study variables, gender, age and residence, and significant differences were found in the age variable, which led to setting age standards, and differences were found in the residence variable, and there were no differences in the gender variable.

The Shabatat Study (2018), aimed to standardize a Jordanian version of the Stanford Binet Early Childhood Scale - the fifth version, and to extract performance criteria and verify the indications of validity and reliability, the sample consisted of 400 children. The study reached the significance of the content validity of the scale by presenting it to 10 arbitrators, and the percentage of agreement between them reached 90%, and calculating the correlative validity through its correlation with the Wechsler Intelligence Scale, the correlation coefficients of the paragraph with the total score of the branch were also calculated, and the correlations ranged between 0.38-0.65. Criteria for the scale were derived and the scores were converted into weighted scores with an average of 10 and a deviation of 3. and converting the weighted scores into degrees and IQ ratios with an average of 100 and a deviation of 15.

Commenting on previous studies:

By reviewing previous studies, it is noted that several studies have been conducted in several Arab countries to codify the Stanford-Binet Scale V, such as the Shabatat study (2018), which codified a Jordanian version of the Stanford-Binet scale for early childhood - the fifth version, and the studies of both Al-Bala'a (2015) and Al-Mutlaq (2015) which standardized the Stanford Binet Scale, the fifth version on the Syrian environment.

As for the current study, it aimed to standardize the scale on the Kuwaiti environment; Since no study has codified it on the Kuwaiti environment, it also compares the performance of students with mental disabilities with that of ordinary students. The current study is distinguished from previous studies in that it will codify the test on the Kuwaiti environment and on a certain age group of children, which is from 4-7 years, and by extracting criteria for this age group.

Method and procedure:

Study Methodology:

The study used the descriptive analytical method for its suitability to achieve the objectives of the study.

Study community:

The study population consisted of all age groups from 4-7 years, they are children in the first and second stages of kindergarten and the first and second grades in the primary stage, males and females (normal and people with mild mental disabilities) and their number is (39,065 students) for kindergarten and (53527 students)) for the primary stage with a total of (92,592 students) for the academic year 2019-2020.

Study sample:

The study sample included 450 children from all governorates of Kuwait, in the age group from 4-7 years, from all governorates of Kuwait. The study sample was selected in a cluster random manner, according to the governorate, consisting of (4) schools in each governorate, with (40) children from Males and females for almost every school.

Study tool:

Stanford Binet Scale Fifth version(SB5) Legalization version of the Kuwaiti environment

Scale description:

The Stanford-Binet Intelligence Scales is one of the first and most popular measures of mental abilities. The fifth version of the scale appeared in 2003 AD, which was prepared by Roid (2005), and it is an individually applied scale to measure intelligence and cognitive abilities. In this version, the total IQ is determined after measuring two domains,

verbal and non-verbal, with five factors in each domain (analytical reasoning, information, quantitative reasoning, visual-spatial processing, working memory), for each factor ten sub-tests.

Validation and reliability of tool:

1. Difficulty and Distinction:

The difficulty coefficients were calculated for each item of the scale for each sub-test (the non-verbal side and the verbal side), where the difficulty coefficients for the non-verbal side items ranged between (0.54 -0.92), and the difficulty coefficients for the verbal aspect paragraphs also ranged between (0.54 -0.92) and all of them are statistically significant. These coefficients indicate the effectiveness of the items in the gradation of their difficulty, and they are scientifically acceptable for verifying the validity of the scale items.

The discrimination coefficients were calculated for each item of the scale for each sub-test (the non-verbal aspect and the verbal aspect). It is noted that the discrimination coefficients for the non-verbal side items ranged between (0.24 -0.60) and the discrimination coefficients for the verbal side items also ranged between (0.24 -0.24), all of which are statistically significant.

2. Correlative validity

The correlative validity of the scale was verified in its initial form by calculating the Pearson correlation coefficient for performance on the Stanford-Binet Early Childhood Scale, Fifth Edition and the Kuwaiti version of the Wechsler Scale for Early Childhood, a research that was applied to a sample of (50) individuals from the age groups from 4-7, and the correlation coefficient was (0.877), which is a statistically significant value at the significance level (0.05).

Stability of tool:

To check the stability of the scale in the following ways:

A. Internal consistency method using Cronbach's alpha equation.
B. Half-segmentation method

3. Test and retest method.

4. Evaluators Agreement Method

Statistical processing:

Several statistical methods were used to answer the study questions, namely: Pearson correlation coefficient, arithmetic means, and standard deviations, one-way Anova analysis, test-retest, Cronbach's alpha equation, Spearman-Brown equation, etc...

Presentation of the results:

1. Difficulty and Distinction

The difficulty coefficients were calculated for each item of the scale for each sub-test, and the following table (1) shows those coefficients.

Table (1) Difficulty coefficients for each of the scale items in the two tests (the nonverbal side and the verbal side)

No. of item	Nonverbal side					Verbal side				
	Liquid inference	Knowledge	quantitative inference	visuospatial	working memory	Liquid inference	Knowledge	quantitative inference	visuospatial	working memory
1	0.62	0.62	0.74	0.70	0.54	0.72	0.74	0.74	0.72	0.74
2	0.74	0.72	0.70	0.72	0.82	0.66	0.92	0.76	0.62	0.74
3	0.92	0.92	0.72	0.60	0.74	0.62	0.56	0.74	0.74	0.76
4	0.56	0.60	0.60	0.80	0.74	0.74	0.72	0.70	0.92	0.74
5	0.72	0.72	0.80	0.68	0.76	0.92	0.66	0.72	0.56	0.70
6	0.66	0.62	0.68	0.68	0.74	0.56	0.54	0.62	0.72	0.72
7	0.54	0.74	0.62	0.68	0.70	0.72	0.82	0.74	0.66	0.60
8	0.82	0.62	0.74	0.76	0.72	0.66	0.74	0.92	0.54	0.80
9	0.74	0.74	0.92	0.62	0.60	0.54	0.62	0.56	0.82	0.68
10	0.74	0.92	0.56	0.76	0.80	0.82	0.74	0.72	0.74	
11	0.76	0.56	0.72	0.62	0.68	0.74	0.92	0.66	0.62	
12	0.74	0.72	0.66	0.74	0.68	0.62	0.62	0.54	0.74	
13	0.70	0.66	0.54	0.92	0.68		0.74	0.82	0.92	
14	0.72	0.54	0.82	0.56	0.76		0.92	0.74	0.56	
15	0.60	0.82	0.74	0.72	0.62		0.56	0.74	0.72	
16	0.80	0.74	0.74	0.66	0.76		0.72	0.76	0.66	
17	0.68	0.74	0.76		0.62		0.66	0.74	0.54	
18	0.68	0.76	0.74		0.74		0.54	0.70	0.82	
19	0.68				0.92		0.82			

20	0.76				0.56		0.74			
21	0.62									
22	0.76									
23	0.92									
24	0.56									

It is noted from the previous table that the coefficients of difficulty for the paragraphs of the non-verbal side ranged between (0.54 -0.92), and the difficulty coefficients for the verbal aspect paragraphs also ranged between (0.54 -0.92) and all of them are statistically significant. These coefficients indicate the effectiveness of the paragraphs in the gradation of their difficulty, and they are scientifically acceptable for verifying the validity of the paragraphs on the scale because the range of difficulty coefficients that are between (0.20) and (0.80) with an average of (0.50) can be acceptable (Musa Al-Nabhan, 2004).

Discrimination coefficients were calculated for each item of the scale for each sub-test

Table (2) Discrimination coefficients for each item of the scale in the two tests (the non-verbal side and the verbal side)

No. of item	**Nonverbal side**					**Verbal side**				
	Liquid inference	Knowledge	quantitative inference	visuospatial	working memory	Liquid inference	Knowledge	quantitative inference	visuospatial	working memory
1	0.60	0.44	0.28	0.44	0.28	0.56	0.36	0.44	0.24	0.44
2	0.36	0.48	0.44	0.24	0.36	0.28	0.46	0.24	0.60	0.44
3	0.56	0.46	0.24	0.40	0.44	0.60	0.32	0.28	0.36	0.24
4	0.32	0.36	0.40	0.32	0.44	0.36	0.56	0.44	0.36	0.28
5	0.56	0.40	0.32	0.32	0.24	0.36	0.28	0.24	0.32	0.44
6	0.28	0.60	0.32	0.40	0.28	0.32	0.28	0.60	0.56	0.24
7	0.28	0.36	0.60	0.48	0.44	0.56	0.36	0.36	0.28	0.40
8	0.36	0.60	0.36	0.24	0.24	0.28	0.44	0.56	0.28	0.32
9	0.44	0.36	0.26	0.36	0.40	0.28	0.60	0.32	0.36	0.32
10	0.44	0.36	0.32	0.46	0.32	0.36	0.36	0.56	0.44	
11	0.24	0.32	0.56	0.60	0.32	0.44	0.26	0.28	0.60	

12	0.28	0.56	0.28	0.36	0.40	0.60	0.60	0.28	0.36	
13	0.44	0.28	0.28	0.56	0.48		0.36	0.36	0.46	
14	0.24	0.28	0.36	0.32	0.24		0.56	0.44	0.32	
15	0.40	0.36	0.44	0.56	0.36		0.32	0.44	0.56	
16	0.32	0.44	0.44	0.28	0.56		0.56	0.24	0.28	
17	0.32	0.44	0.24		0.60		0.28	0.28	0.28	
18	0.40	0.24	0.28		0.36		0.28	0.44	0.36	
19	0.48				0.46		0.36			
20	0.24				0.32		0.44			
21	0.36									
22	0.36									
23	0.26									
24	0.56									

It is noted from the previous table that the discrimination coefficients for the non-verbal side items ranged between (0.24 -0.60) and the discrimination coefficients for the verbal side items also ranged between (0.24 -0.24), all of which are statistically significant. These coefficients indicate the effectiveness of the items in distinguishing between students from the higher and lower categories, and they are scientifically acceptable for verifying the validity of the scale items.

2. Correlative validity

The researcher calculated the correlation coefficients between the two scales. The results indicated that the correlation coefficient of the total score of the Stanford-Binet Early Childhood Scale, Fifth Edition, with the Kuwaiti version of the Wechsler Scale of Early Childhood reached (0.877), which is a statistically significant value at the significance level (0.05) which it indicates the existence of the correlative validity of the Stanford-Binet Early Childhood Scale, Fifth Edition, and the total degree correlation coefficient for the non-verbal side of the Stanford-Binet Early Childhood Scale, Fifth Edition, with the Kuwaiti version of the Wechsler Scale for Early Childhood, reached (0.868), which is a statistically significant value at the significance level (0.05). The correlation coefficient of the total degree of the verbal aspect of the Stanford-Binet Early Childhood Scale, fifth edition, with the Kuwaiti version of the Wechsler Scale for Early Childhood, reached (0.875), which is a statistically significant value at the significance level (0.05) which it indicates the existence of correlative validity of the verbal side of the Stanford-Binet Early Childhood Scale, Fifth Edition.

Constructivist Validity:

The constitutional validity of the Stanford-Binet Early Childhood Scale, Fifth Edition, was confirmed

a. Factor Analysis:

The results indicated that the values of the latent root of the five factors that make up the scale ranged between (41.626 - 12.444), and that the first factor (liquid inference) explained 24.061% of the total variance, which is the highest percentage of variance explained compared to other factors, and it is a high value if compared with the variance explained by the rest of the other factors while the second factor (knowledge) explained 12.355%, the third factor (quantitative inference) explained the rate of 11.068%, the fourth factor (visual-spatial) explained 8.247%, and the fifth factor (working memory) explained 7.193%, and the total percentage of what was explained by the five combined factors equals (62.924).

b. Efficacy of paragraphs

The correlation coefficients were calculated using the Pearson correlation coefficient between each paragraph and the dimension to which the paragraph belongs and between each paragraph and the total score of the non-verbal aspect to which the paragraph belongs, and between each paragraph and the total score of the scale. It was found that the correlation coefficients of the paragraphs related to the field of liquid inference with its scope ranged between (0.253 - 0.643), whereas, the correlation coefficients for items related to the domain of knowledge ranged between (0.307-0.642), and the correlation coefficients for items related to the domain of quantitative inference ranged between (0.386-0.613). The paragraphs' correlation coefficients for the field of visual-spatial inference ranged (0.308-0.655), and the correlation coefficients of the items in the field of working memory ranged between (0.287 - 0.615), and the correlation coefficients of the items with their non-verbal aspect ranged between (0.268 - 0.605). While the correlation coefficients of the paragraphs with the total score of the scale ranged between (0.249-0.599), and all of these coefficients are statistically significant at the significance level ($\alpha = 0.05$). The previous correlation coefficient indicates the effectiveness of the items and their belonging to the scale and the non-verbal aspect, and that they measure what they were designed to measure.

The correlation coefficients were also calculated using the Pearson correlation coefficient between each paragraph and the dimension to which the paragraph belongs, between each paragraph and the total score of the verbal side to which the paragraph belongs, and between each paragraph and the total score of the scale. 0.693), while the correlation coefficients of the items related to the field of knowledge ranged between (0.391-0.683), and the correlation coefficients of the items related to the field of

quantitative inference ranged between (0.339-0.628). The correlation coefficients of the items related to the field of visual-spatial inference ranged (0.451-0.625), and the correlation coefficients of the items related to the domain of working memory ranged between (0.412-0.645). The correlation coefficients of the paragraphs with their verbal aspect ranged between (0.344 - 0.631), while the correlation coefficients of the paragraphs with the total score of the scale ranged between (0.325-0.599), and all of these coefficients are statistically significant at the significance level (α = 0.05). The previous correlation coefficient indicates the effectiveness of the items and their belonging to the scale and the verbal aspect, and that they measure what they were designed to measure.

3. Discriminatory Validity

To verify the discriminatory validity and reach acceptable indications for the scale in its Kuwaiti form, this was done as follows:

a. Age group

It was found that there are apparent differences between the arithmetic averages and the standard deviations of the performance of the study sample members on the scale in its Kuwaiti form and for all its fields, with the non-verbal and verbal parts, and the total degree of the scale according to the age group variable. For example, the arithmetic mean of the age group (the oldest 7 years - 7 years and 11 months) was the highest, reaching (148.16), followed by the arithmetic average for the age group (6-7 years), which reached (132.32), then by the arithmetic average for the age group (5-6 years), which reached (110.09). Finally, the arithmetic mean for the age group (4 years-5 years) was (99.91). Also, apparent differences were found in all areas of the scale, both nonverbal and verbal, as shown in the previous table, to ensure that these differences are statistically significant at the significance level (α = 0.05) one-way analysis of variance (ANOVA) was applied, and the results of the analysis indicated that there were statistically significant differences at the significance level (α = 0.05).

b. Gender:

The arithmetic means and standard deviations of the performance of the study sample members on the scale were extracted, and the t-test was applied to the independent samples, and it was found that there were no statistically significant differences at the significance level (α = 0.05).

c. Student classification variable

To also calculate the discriminatory validity and to know the ability of the scale to discriminate between the categories of normal children and children with mild disabilities in the total degree of intelligence, the arithmetic averages and standard

deviations of the performance of the study sample members on the scale were extracted. The t-test was also applied to the independent samples, and the results indicated that there were statistically significant differences at the significance level (α = 0.05), between the arithmetic averages of the performance of the study sample members on the scale according to the student classification variable (normal, disabled).

2. What is the significance of the stability of the Stanford Scale, the fifth version of early childhood, after its application to the study sample in the Kuwaiti environment, with the characteristics of a good and reliable test for measuring intelligence for the age group (4-7 years)?

A. Internal consistency method using Cronbach's alpha equation

The results showed that the stability coefficient of Cronbach's alpha for the total score of the scale was (0.980), while the reliability coefficient of Cronbach's alpha for the total score for the non-verbal side was (0.963). The reliability coefficients for domains for the non-verbal aspect ranged between (0.847-0.774), while the reliability coefficient of Cronbach's alpha for the total score for the verbal aspect was (0.959). The reliability coefficients for domains of the verbal aspect ranged between (0.856-0.681), and all of these coefficients are high, and this indicates the existence of the scale's stability, which means that the scale has a statistically significant stability significance and can be relied upon in measuring intelligence for the age group.

B. Half-segmentation method

The reliability coefficient was calculated using the Spearman-Brown equation, and those coefficients were extracted for all its domains, with the non-verbal and verbal parts, and the total score of the scale. The results indicated that the Spearman-Brown reliability coefficient of the total score of the scale was (0.991). While the Spearman-Brown reliability coefficient for the total score for the non-verbal side was (0.996). The reliability coefficients for domains for the nonverbal side ranged between (0.877-0.774), while the Spearman-Brown reliability coefficient for the total score for the verbal side was (0.988), and the reliability coefficients for domains for the verbal side ranged between (0.982-0.721), and all of these coefficients are high, and this indicates the presence of Stability of the scale, which means that the scale has a statistically significant and reliable significance in measuring intelligence for the age group.

C. Test and retest method

It was found that the Pearson reliability coefficient for the total score of the scale was (0.982), while the Pearson reliability coefficient for the total score for the non-verbal aspect was (0.911), and the reliability coefficients for the domains for the non-verbal

aspect ranged between (0.968-0.887). While the Pearson reliability coefficient for the total score for the verbal aspect was (0.945). The reliability coefficients for the domains of the verbal side ranged between (0.982-0.918). All of these coefficients are high, and this indicates the stability of the scale, which means that the scale has a statistically significant significance and can be relied upon in measuring the intelligence of the age group.

D. Evaluators Agreement Method

The results showed that the Pearson reliability coefficient for the total score of the scale was (0.954), while the Pearson reliability coefficient for the total score for the non-verbal aspect was (0.925). The reliability coefficients for domains for the non-verbal side ranged between (0.901-0.831), while Pearson's reliability coefficient for the total score for the verbal aspect was (0.831). The reliability coefficients for the domains of the verbal side ranged between (0.895-985). All of these coefficients are high, and this indicates the stability of the scale, which means that the scale has a statistically significant significance and can be relied upon in measuring the intelligence of the age group.

3. What are the performance criteria extracted from SB5 after applying it to Kuwaiti environmental children aged (4-7 years)?

The results indicated that the percentile performance criteria that can be derived after the transformation of raw scores into percentile rank for the performance of the study sample members on the non-verbal aspect of the scale according to the age group variable, for example, in fluid inference, the IQ ratios based on the student's grades are low from the age group 4-5 years, reaching (7.00) and it is below average, as it ranged between (8.00 - 10.00), it is medium if it reached (12.00), and it is above average, as it ranged between (18.00-23.00), and it is high if it exceeded (23.00), and so on for the rest of the age groups.

T

able (3) Percentile standards of the raw scores for the performance of the study sample members on the total score for the non-verbal and verbal aspects and the total score for the scale according to the age group variable.

	Age	Percentile ranks						
		5	10	25	50	75	90	95
IQ		low	below average		mediu m	above average		High
Non-verbal sum	4 years-5 years	24.00	30.00	42.00	44.00	67.50	93.00	93.00
	From 5 years -6 years	33.25	36.00	50.00	63.00	78.00	87.00	87.00
	6-7 years old	52.75	56.00	57.75	69.00	87.00	96.00	96.00

	From 7 years - 7 years and 11 months	54.00	57.00	65.00	92.00	96.00	96.00	96.00
verbal sum	4 years-5 years	14.00	23.00	32.00	36.00	55.75	77.00	77.00
	From 5 years -6 years	24.80	27.00	40.00	48.00	61.00	65.00	65.00
	6-7 years old	41.00	43.00	50.00	58.00	67.00	77.00	77.00
	From 7 years - 7 years and 11 months	53.00	53.00	55.00	74.00	77.00	77.00	77.00
The total score of the scale	4 years-5 years	38.00	53.00	74.00	80.00	123.25	170.00	170.00
	From 5 years -6 years	58.05	63.00	87.00	113.50	139.00	152.00	152.00
	6-7 years old	95.40	98.00	112.00	123.00	154.00	173.00	173.00
	From 7 years - 7 years and 11 months	107.00	110.00	120.00	166.00	173.00	173.00	173.00

Student classification variable:

The results indicated that there were statistically significant differences at the significance level (α = 0.05), between the arithmetic averages of the performance of the study sample on the scale.

Discuss the results:

Discuss the results of the first question

The results indicated that the difficulty coefficients for the non-verbal side items ranged between (0.54-0.92) and the difficulty coefficients for the verbal side items ranged between (0.54-0.92). This result reflects good psychometric properties and is acceptable for the test; Which agrees with Ray (Al Nabhan, 2004). As for the results related to correlative validity, they indicate the existence of correlative honesty for the non-verbal as well as the verbal aspect. This result is similar to the results of the study by Shabat (2018) al-Mutlaq (2015) and Khrushcheh (2019).

Discuss the results of the second question

After the stability of the scale was calculated in four different ways, it was found that all the coefficients are high, and this indicates the existence of the stability of the scale, which means that the scale has a statistically significant significance and can be relied upon in measuring the intelligence of the age group. The high-reliability coefficient in this study may be because the test is long and has several items and components; As the number of

test items increases, the percentage of stability increases, and this result is similar to the results of Safwat Farag's study.

Discuss the results of the third question

The results of the question indicated that the percentile performance criteria can be derived after converting the raw scores to the percentile rank for the performance of the study sample members on the non-verbal as well as verbal aspects of the scale depending on the age group variable.

Recommendations:

Based on the findings of the study, the researcher recommends the following:

1. Adopting the fifth form of the Stanford Binh Scale in diagnosing the mental abilities of children aged 2-4 years, because it has high degrees of validity and stability.
2. Conducting a study to complete the legalization of the rest of the tests for the age groups that the study did not address.
3. Training psychologists in the Ministry of Education to apply the test.
4. Establishing a center for legalizing psychological measures and following up on global developments.

References

Abadi, Majduddin (2011). A Dictionary of the Ocean Dictionary arranged and documented: Khalil Shiha, 5th edition, Beirut: Dar Al Marfa.

Abu El-Nil, Mahmoud (2012) A comparative study of the knowledge page of the Stanford Intermediate Intelligence Scale, the fifth image, between a sample of autistic and non-psychic children. Childhood Studies Journal. Volume 15, Number 56, 67-77.

Al-Balaa, Abeer (2015). Stanford Binet Scale Fifth image - A field study to standardize the scale in Damascus governorate on a sample of (4-6) years old, unpublished doctoral thesis, Damascus University, Syria.

Khrushche, Ibaa (2019). Standardization of a Palestinian image of the Stanford Intermediate Intelligence Scale, fifth edition, for the age group from 6 to 12 years, in measuring and diagnosing mental ability (unpublished doctoral thesis). International Islamic Sciences University, Amman.

Khattab, Muhammad (2013). The Stanford-Binet Intelligence Scale The fifth picture: A comparative critical evaluation study of the two Arabic versions, Psychological Counseling Journal, Egypt, (35): 327-358.

Shabatat, Amani (2018). Standardizing a Jordanian image from the Stanford Binet Scale for Early Childhood - the fifth image -, unpublished doctoral thesis, International Islamic University of Sciences, Amman, Jordan.

Taha, Muhammad (2006). Human intelligence contemporary trends and critical issues. The world of knowledge series, 330.

Taha, Muhammad and Farhan, Abdel-Mawgod, Abu Al-Lail, Mahmoud (2011). Stanford-Binet Intelligence Scale, fifth picture, introduction to the Arabic version and the examiner's guide, Cairo: The Arab Foundation for Psychological Tests.

Faraj, Safwat (2011). Stanford Binet Scale Manual, Fifth Edition, Cairo: Anglo-Egyptian Library.

Farag, Kamal and Bakhit, Suhair (2011). The ability of Binet Intelligence Scale, Fifth Edition, to distinguish between hearing-impaired and normal children, Journal of Arab Studies in Psychology, Egypt, 10(4): 791-824.

Karam El-Din, Laila and Hadiya, Fouad and Muhanna, Ghada (2015). The ability of each of the cognitive assessment systems CAS and the cognitive page of the Stanford Binet Scale: The fifth picture to assess the cognitive abilities of a sample of people with special needs, Journal of Childhood Studies, 18 (68): 76-69.

Muhammad, Kamal (2013). The ability of the Stanford-Binet Intelligence Scale, Fifth Edition, to distinguish between normal people, people with learning disabilities, and people with mental disabilities, Journal of the College of Education, Ain Shams, Egypt, 3 (36): 239-204.

Al-Manqal, Muhammad (2013). Lectures in Psychology, 1st Edition, Damascus: Damascus University.

Nabhan, Musa (2004). Fundamentals of measurement in the behavioral sciences. Sunrise House.

Al-Namer, Essam (2018). Measurement and evaluation in special education, Amman: Dar Al-Yazuri for publication and distribution.

Noureddine, Amin Mohamed Sabry (2014). Psychometric characteristics of each of the two Arabic versions of the Intermediate Intelligence Scale: Fifth Edition on a sample of primary school students. Psychological Counseling Journal: Ain Shams University - Psychological Counseling Center, A. 37, 421-485.

Yassin, Hamdi and Shaheen, Hiam (2015). Linguistically retarded and normal children, a significant study in the light of the Stanford Binet Intelligence Scale, fifth picture, Journal of the College of Education, Ain Shams University, (16): 1-36.

Becker, K. A. (2003). History of the Stanford-Binet intelligence scales: Content and psychometrics. (Stanford-Binet Intelligence Scales, Fifth Edition Assessment Service Bulletin No. 1). Itasca, IL: Riverside Publishing.

Burtan, K. (2013). Standardization of an Australian Version of Stanford- Binet Intelligence Scales of Early Childhood Fifth Edition. Measurement and Evaluation in Counseling and Development, 73(21): 52-63.

Fancher, R. E., & Rutherford, A. (2012). Pioneers of psychology. New York, NY: W. W. Norton & Company, Inc.

Janzen, H., Obrzut, J., & Marusiak, C. (2003). Test review: Roid, G. H. (2003). Stanford–Binet intelligence scales, fifth edition (sb:v). Canadian Journal of School Psychology, 19, 235–244.

Roid, G. H. (2003). Stanford -Binet Intelligence Scales, Fifth Edition, Technical Manual. pro.ed, Austin, Texas.

Roid, G. H. (2005). Stanford Binet Intelligence Scales, Fifth Edition for Early Childhood, Interpretive Manual, Itasca, IL: Riverside Publishing.

Ruf, Deborah L (2003). Use of the SB5 in the Assessment of High Abilities. (Stanford-Binet intelligence scale, Fifth Edition Assessment Service Bulletin No 3). Itasca, IL: Riverside Publishing.

Williams, T. H., McIntosh, D. E., Dixon, F., Newton, J. H., & Youman, E. (2010). Confirmatory factor analysis of the Stanford-Binet Intelligence Scales, Fifth edition, with a high-achieving sample. Psychology in the Schools, 47(10), 1071-1083.