Psychometric Properties OF THE Wechsler Intelligence Scale-Iv Arithmetic Test FOR Elementary School Students IN THE Kingdom OF Saudi Arabia

Dr. Hamad Abdullatif Al-Hulaibi

Department of Curriculum and Instruction, College of Education, Imam Abdulrahman Bin Faisal University, Dammam 31451, Saudi Arabia
haalhalaibi@iau.edu.sa

Abstract:

This study aimed to identify the psychometric properties of the arithmetic test of the Wechsler Intelligence Scale-4 among primary school students in the Kingdom of Saudi Arabia. The study sample included (50) primary school students, aged (6-12) years, from schools in the Eastern Province of the Kingdom of Saudi Arabia. The sample was selected intentionally, and the (arithmetic) subtest of the Wechsler Intelligence Scale-4 was applied.

The results showed that the validity indications represented by Content Validity were (80%), construct validity ($\alpha = 0.01$, and $\alpha = 0.05$), and the factorial validity indications among students were between (0.410 - 0.918), and the factorial analysis validity indications were (100%) of the variance in the total score of the test, and the discriminant validity indications according to gender reached the value of the statistic (T) (2.610) in favor of males, and the discriminant validity indications according to age reached the value of the statistic (T) (3.580) in favor of the student group aged (6-8 years). The results showed that the Reliability implications represented by the internal scale Reliability implications of the Cronbach Alpha test = (0.791), and the Reliability implications of the split-half Cronbach Alpha reliability coefficient for the first part were (0.776) and the second part were (0.852). The value of the correlation coefficient between the first and second parts was (0.692) and the value of the Reliability coefficient split-half using the Guttman Split-Half Coefficient test was (0.815). The study recommends conducting more studies and research related to conducting studies related to the impact of both verbal intelligence and mathematical problem-solving skills and working on developing the solution steps for students.

Keywords: Wechsler scale, Arithmetic, Elementary Stage.

Introduction

The Wechsler Intelligence Scale-4, which focuses on academic accomplishment and the concept of intelligence, is regarded as one of the key measures for evaluating the capacity to predict mental talents. One of the most crucial subtests on the Wechsler Intelligence Scale-4 for determining mental capacity and pinpointing pupils' academic difficulties is the arithmetic subtest. As a result, the current study developed a mental ability assessment tool.

Research Problem

Knowing the psychometric properties of the scale is important, especially with regard to the concept of arithmetic among students. Accordingly, the study problem was formulated in the following main question:

- What are the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test for primary school students in the Kingdom of Saudi Arabia?

The following sub-questions arise from the research problem, which deals with the psychometric properties:

1- What are the implications of the Validity of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

2- What are the implications of the Reliability of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

Research Importance

First: The theoretical importance of this research lies in:

- Creating curiosity about pupils' cognitive capacities.
- Making clear how achievement and prediction ability are related.
- To the best of the researcher's knowledge, this is the first study in the Arab world to examine predictive validity using Wechsler scale math skills.

Second: The practical importance:

- Providing a tool specific to the Saudi environment to measure the ability to predict mathematical achievement.
- Enriching the field of education to provide a concept of the ability to predict using intelligence scales.
- Helping researchers benefit from the results of this study in dealing and communicating with students.
- Benefiting from the results of this study in making appropriate decisions in predicting student achievement.

Research Goals

To create diagnostic criteria for the Wechsler-4 Scale's arithmetic subtest to confirm the degree to which it shares the same arithmetic characteristics.

Research Justification

- The scarcity of Arab studies that dealt with research on achievement and its relationship to mental ability.
- The scarcity of Arab studies that focused on the Wechsler-4 scale through the subtest (Arithmetic).

Research Limitations:

Research limits include the following:

- **Age limits:** The application of this study was limited to primary school students of the age group (6-12) years in the school stage.
- **Spatial limits:** The application of this study was limited to the Kingdom of Saudi Arabia (Central, Eastern Region) through primary schools.
- **Temporal limits:** Data for this study were collected during the academic year 2023.

Research Delimitations

- The results of this study are determined by the limited sample to be studied.

- The researcher was not able, through their research, to reach and identify other cases in schools in the Kingdom of Saudi Arabia.

Definitions of Terms

1- Primary stage: It is one of the most important stages of education in the child's academic life, as it constitutes the first stage in the formal education process. This stage involves teaching the child the basics of reading, writing, and arithmetic, in addition to enhancing his basic knowledge.

2- Arithmetic: A set of abstract knowledge resulting from logical deductions applied to various mathematical objects such as sets, numbers, shapes, structures, and transformations. Mathematics also deals with the study of topics such as quantity, structure, space, and change. There is no generally agreed-upon definition of the term.

3- Wechsler scale-4: the subject's score on the Wechsler intelligence scale for the sub-test (Arithmetic).

Theoretical framework and previous studies

According to (Prinitera, Weiss, Rolfhus, Saklofske, 2005), the Wechsler Intelligence Scale, or WISC-IV, is utilized for several objectives and in a variety of ways in clinical evaluations, research, and other kinds of

assessments. its power and capacity to yield useful data in a variety of evaluations, including neuropsychological evaluations, a field that was just being started when Wechsler started creating his scales. The assessment method created by Dr. Edith Kaplan and her associates served as the foundation for the introduction of the WISC-III system as a process tool. The Wechsler Scale-IV's early adoption of a process approach is one of its main advances. This invention gives the scale clinical and diagnostic sensitivity, develops intervention strategies, and broadens the scope of psychological processes evaluated by Wechsler. It also offers a practical clinical instrument. Numerous studies have been carried out using clinical groups to help the clinician evaluate the findings. There is more balance because the WISC-IV has been successful in including tasks that demonstrate fluid intelligence. This application offers the chance to track the variations in scores among intelligence subtests as well as the crystallization of intellect.

The Wechsler Intelligence Scale (WISC-IV, 2003) is a commonly used test with good psychometric qualities, according to (Bodin, Pardini, Burns, Stevens, 2009). In various clinical samples, the four-factor factor structure—verbal comprehension, perceptual, working memory, and speed of execution—has been confirmed; however, it has also been observed that the general ability level's impact should not be undervalued when analyzing factor index results.

First: Literature Review

First : Prediction of the Wechsler-4 scale

According to (Bodin, Pardini, Burns, Stevens, 2009), the Wechsler Intelligence Scale (WISC-IV, 2003) is a widely used test with high psychometric properties. The four-factor factor structure—verbal comprehension, perceptual, working memory, and speed of execution—has been validated in a number of clinical samples; however, it has also been noted that the influence of general ability level should not be underestimated when examining factor index data.

Using a commonality analysis, (Schneider, Flanagan, colleagues, 2015) discovered a relationship between mathematics achievement and the Wechsler subtests that measure comprehension, knowledge, short-term memory, and, to a lesser extent, visual processing variables.

According to (Wechsler, 2015), the following factors should be considered while forecasting the idea of cognition and knowledge in mathematics:

-To a lesser degree, processing speed and short-term memory should also positively predict quantitative knowledge, but fluid thinking and absorptive knowledge should strongly and favorably predict quantitative knowledge.

-Even after knowledge has been added, fluid thinking and absorptive knowledge should still be able to predict quantitative knowledge, and their contribution should outweigh that of short-term memory and processing speed.

- The majority of quantitative knowledge's predictive effects are mediated through fluid reasoning and absorptive knowledge because of its primarily indirect or only indirect impact.

Drawing on the aforementioned, (Flanagan, Kaufman, 2009) suggest that arithmetic is one of the WISC-IV's subtests. The subtest was included in the Wechsler Interpretation Guidelines as a measure of working memory (a specific skill associated with the CHC's short-term memory capacity), while Keith et al. (2006) classed it as a measure of fluid reasoning. The arithmetic subtest, however, had higher factor loadings on the quantitative knowledge factor than on the fluid reasoning component when factor-analyzed with measures of quantitative knowledge (Phelps, McGrew, Knopik, Ford, 2005; Woodcock, 1990). According to Wechsler (2002), there are comparable issues with the mathematical reasoning subtest. According to certain cognitive assessment experts (Flanagan, Alfonso, Ortiz, Dynda, 2013), it was created to gauge mathematical aptitude; however, it measures fluid reasoning rather than numeric knowledge. To comprehend how cognitive capacities could explain mathematics achievement, it is crucial to look into where these two activities sit in the CHC taxonomy given the intricacy of these subtests.

These subtests of the Wechsler 4 (WISC-IV) scale can be used to accurately assess an individual's strengths and shortcomings, according to (Gregory, 2004).

Second: Previous studies

The study "Structural and incremental validity of the Wechsler Adult Intelligence Scale—Fourth Edition with a clinical sample" by (Nelson, Canivez, Watkins, 2013) sought to investigate the Wechsler Adult Intelligence Scale—Fourth Edition (WAIS-IV; Wechsler, 2008a) structural and incremental validity. There were (300) people in the sample. The findings demonstrated that:

- a direct hierarchical model of four first-order factors and a general intelligence factor best reflected the structure of the WAIS-IV, according to confirmatory factor analysis. -The most prevalent and overall variance across the subtests was explained by the general intelligence factor.
- Analyses of incremental validity showed that the full-scale intelligence quotient (FSIQ) typically accounted for moderate to significant amounts of the variation in academic performance.
- For all measures of academic achievement, the first-order factors combined accounted for significant achievement variance beyond that explained by the (FSIQ).

- Individual factor index scores contributed small amounts of achievement variance.

The study "Measurement of Working Memory on the Wechsler Adult Intelligence Scale: Should We Subtract Arithmetic?" (Harrison, Beal, Armstrong, 2024) sought to determine whether working memory issues could be used as evidence to support the diagnosis of a neuropsychological disorder or acquired brain injury, given the significance of memory and its relationship to mathematics among students. Though it is considered that all of the subtests that make up the Working Memory Index (WMI) of the WAIS-IV measure the same construct, performance on the arithmetic subtest may be excessively impacted by issues with fundamental math skills, particularly in individuals with neurodevelopmental disorders. The current study looked at the correlation between tested mathematical aptitude and working memory scores for a sample of (605) students who were seeking psychoeducational evaluations to confirm that they needed academic adjustment. The results showed:

- The different combinations of the scores of the three WAIS-IV working memory subtests to the composite WMI scores were significantly different.

- Diagnostic groups had the highest subtest scores on the letter-digit sequence and digit span subtests and the lowest on arithmetic. Furthermore, mathematics achievement scores were more strongly correlated with arithmetic scores than with the digit span or letter-digit sequence subtest scores. - Students with learning disabilities performed significantly worse on arithmetic than students with ADHD.

And (Chen, Hua, 2020) conducted a study titled "Selecting tetradic short forms of the Taiwan Wechsler Adult Intelligence Scale IV" aimed to investigate the tetradic short forms of the Wechsler Factor-Based Adult Intelligence Scale - Fourth Edition (WAIS-IV) using a normative sample from Taiwan Wechsler IV. The sample included (1105) normal adults aged (16-90) years. Various psychological characteristics, time constraints, and estimation quality were compared among (90) tetradics using linear equation procedures. Among the tetradics. The results showed:

- The performance of the information-visual puzzle-digit span-digit symbol group was higher than the other groups in terms of overall estimation quality and time saved.

- The similarity-visual puzzle-digit span-digit symbol, information-matrix reasoning-digit span-digit symbol, and information-visual puzzle-digit sequence-digit symbol groups obtained the most efficient estimations.
- For clinicians who value the utility of block design, the combination of information-block design-digit range-digit symbols has been found to provide high-quality estimates.

- The selected quartiles of the IQ scale for high-intelligence adults had large misclassification rates; the short forms of four factors tend to underestimate the full-scale IQ of high-

intelligence adults. Therefore, these short forms should be used with caution and for screening purposes only.

And the purpose of the study "The effect of cognitive, personality, and background factors on the WAIS-III arithmetic subtest" (Karzmark, 2009) was to examine the arithmetic subtest as a gauge of working memory, concentration, or distraction avoidance. Numerous additional factors have been proposed to affect math achievement, though. To further define what the arithmetic subtest measures, the current study set out to investigate these variables. (118) people who were referred for neuropsychological testing were included in the sample. The findings revealed:

- A strong association between arithmetic and other subtests of the Working Memory Index (WMI).
- Arithmetic also showed a significant association with arithmetic skills and verbal memory.
- Moderate contributions to arithmetic performance were found for most other cognitive measures. Measures of anxiety and background factors, such as difficulty learning arithmetic, were weakly related to arithmetic scores.

- Although arithmetic may be considered a measure of concentration or working memory, many other factors influence it, and its specificity as a measure of concentration is limited.

And the study "The effects of Wechsler Intelligence Scale for Children—Fourth Edition cognitive abilities on math achievement" was carried out by (Parkin, Beaujean, Wechsler, 2012). Using the subtests of the Wechsler Individual Achievement Test, Second Edition (WIAT-II; Wechsler, 2002) and the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 2003a), the study sought to model structural equations to investigate the effects of the third-layer (general intelligence) and second-layer (understanding and knowledge, fluid reasoning, short-term memory, processing speed, and visual processing) factors of the Cattell–Horn–Carroll (CHC) cognitive abilities. There were (550) in the sample. We contrasted

models that used only third-layer variables, solely second-layer factors, and third-layer plus second-layer components to predict quantitative knowledge. The results showed:
- The model that relies only on the third-class factor to predict quantitative knowledge is the model that best fits the data.

And (Stroud, Blommers, Lauber, 1957) carried out a study titled "Correlation analysis of WISC and achievement tests" to confirm that all or various groups of WISC subtests were useful in predicting performance on the Iowa Basic Skills Test Set's spelling, math, and reading comprehension tests. (775) pupils in grades (3–6) were included in the sample. Selected from a region of (20) counties in the state of Iowa, (621) of these students had access to the Stanford Binet Intelligence Tests. Every student scored an average of (89.0) on the Wechsler IQ scale. The findings indicated:

- The Stanford Binet IQ and the WISC full-scale IQ had a (0.94) correlation. A correlation was found between each accomplishment test result and the WISC full-scale IQ.
- The subtests most useful in predicting school achievement were arithmetic, vocabulary, block design, and object assembly.

- The overall score was the most revealing, however, for predicting academic achievement.

And to ascertain the quantity and kinds of factors that the (Peabody Individual Achievement Test, PIAT) measures, a study titled "Correlational and factor analysis of the Peabody Individual Achievement Test and the WISC-R" (Wikoff, 1978) examined the subtest results of the (PIAT) and the WISC-R test in a sample of (180) children between the ages of (6 and 17). When the PIAT test was examined separately, the findings revealed:

- There were two factors when the PIAT test was analyzed alone. Reading recognition, reading comprehension, and spelling loaded heavily on the word recognition factor, while mathematics and general information had medium to high

loadings on the school-related knowledge factor.
- When the PIAT test was analyzed with the WISC-R subtests, a word recognition factor was found in addition to the three factors usually mentioned for the WISC-R test. The school-related knowledge factor in the PIAT test was absorbed by the other factors.
- General information loaded heavily on the verbal comprehension factor, and mathematics loaded heavily on the freedom from distraction factor. Implications for interpreting the PIAT are discussed.

And (Abu Drei, Al-Rausan, 2021) conducted a study entitled "Psychometric Properties of the Jordanian Version of the Wechsler-4 Intelligence Scale for Children for the School Stage" which aimed to identify the indications of the validity and reliability of the Wechsler-4 Intelligence Scale for Children for the Primary School Stage. The study sample included primary school students, numbering (418) male and female students, who were selected randomly. The Jordanian version of the scale was applied, and then the data were processed statistically. The results indicated the availability of indications of content validity, represented by an agreement rate of (80%), and construct validity by calculating the correlation coefficients between the paragraph and the dimension, ranging between (0.414 - 0.824), and concurrent validity with the Goodenough-Harris scale for drawing a man (0.946) and achievement (0.887), and Pearson correlation coefficients for verbal comprehension (0.414-0.824) and sensory perception (0.734-0.922) and working memory (0.743-0.930) and speed of execution (0.643-0.954), and the correlation coefficients for all sub-paragraphs ranged between (0.602-0.823). The stability of the scale was also indicated by using the Cronbach's alpha method (0.928), the re-test method, the stability coefficient was (0.888), and the raters' agreement method, where the correlation coefficient was (0.890), and the raters' stability according to the Holsti method were (86.5%).

As for the Wechsler-4 scale and its ability to identify mental performance by measuring some aspects of the brain using mathematics, (Levine al et al, 2000) conducted a study aimed at

"Assessment of strategic self-regulation in traumatic brain injury: its relationship to injury severity and psychosocial outcome" and the sample included (n = 9) and (the modified Wechsler scale for assessing cognitive performance, the cube drawing test to measure visual perception, a test of awareness and a test of spontaneous memory, the Boston test for naming objects, the word association test to measure speech-language skills, a reading test, a color naming test, and a test for arithmetic) were used, and the results of the study showed:

- The presence of difficulties in learning mathematics despite the normal cognitive performance of these children.
- Visual-spatial skills range from low to high levels.

According to the aforementioned, the Wechsler Intelligence Scale, fourth edition, has been used by numerous researchers and evaluators in numerous countries and has been translated into numerous languages. This is because of its significance in identifying and diagnosing mental ability through indications of validity and reliability that support its use. A study titled "Long-Term Stability of the Wechsler Intelligence Scale for Children—Fourth Edition" (Watkins, Smith, 2013) sought to compare the content validity of the Wechsler Intelligence Scale-4 when it was administered to children in an Irish population. There were (794) Irish youngsters in the study group, of which (494) were boys and (300) were girls. Using confirmatory factor analysis, the findings revealed:
- The Wechsler Test-4 is applicable.
- There are no individual differences between males and females.
The reliability coefficient values are (0.63) and the validity coefficient is (0.86), which indicates the validity of applying the scale to Irish children.

Methods and Procedures
This section provides a detailed description of the study subjects, the method of selection, and a description of the Wechsler Intelligence Scale as a study tool. It also describes methods for verifying the validity and reliability of these tests,

as well as the study method and data statistical analysis.

Research approach:
This study is based on the use of the descriptive analytical approach because it is suitable for the study related to the psychometric properties of the arithmetic test from the Wechsler Intelligence Scale (4) for primary school students - a field study on school students in the Kingdom of Saudi Arabia.

Population and Sampling
The Wechsler Intelligence Scale (4) was administered to 50 male and female primary school pupils, ages ranging from 6 to 12, who were chosen by the researcher from among the restricted number of students in Saudi Arabian schools. To administer the arithmetic test as part of the Wechsler Scale, which consists of fifteen subtests, these pupils were specifically selected from among the students who demonstrated the highest level of mathematical achievement "on the recommendation of teachers."

Table 1. Distribution of subjects (N=50)

variable	Categories	Frequency	Ratio
Student gender	Male	25	50.0
	Female	25	50.0
	Total	50	100.0
Age	years 8-6	20	40.0
	years 12-9	30	60.0
	Total	**50**	**100.0**

Table (1) shows that (50%) of the study sample members are males, (50%) of the sample members are females, (40%) of the sample members are (6-8) years old, and (60%) of the sample members are (9-12) years old.

The study tools
The first tool: The Wechsler Children's Intelligence Scale (WISC-4), which (Abu Drei, 2017) performed Standardization a Jordanian Version for Deaf via Sign Language.
Standard Description: The fourth version of the Wechsler Children's Intelligence Scale appeared in the year (2003) by (Williams, Weiss and Rothes, 2003). Evidence of validity, reliability, and standards was available in its primary form. The scale consists of 15 sub-tests:
- Block Design, Similarities, Digit Span, Picture Concepts, Coding, Vocabulary, Letter-Number Sequencing, Matrix Reasoning, Comprehension, Symbol Search, Picture Completion, Cancellation, Information, Arithmetic, Word Reasoning. (Wechsler, 2003).

Divided into four sub-measures:
1. (Verbal Comprehension Index, VCI) scale includes the following sub-tests: (Similarities, Vocabulary, Comprehension, Information, and Word Reasoning).
2. (Perceptual Reasoning Index, PRI) scale includes the following sub-tests: (Block Design, Picture Completion, Matrix Reasoning, and Picture Concepts).
3. (Working Memory Index, WMI) scale includes the following sub-tests: (Digit Span, Letter-Number Sequencing, and Arithmetic).
4. (Processing Speed Index, PSI) scale includes the following sub-tests: (Coding, Symbol Search, and Cancellation).
5. Full-Scale IQ (FSIQ) standard and includes tests: (VCI, PRI, WMI, PSI).

The Wechsler (4) arithmetic test contains (34) questions and each question is answered by solving it in a closed answer within a multiple-choice or an essay answer. The student gets a

mark (0) if his answer is wrong and a mark (1 if the student's answer is correct. The final score ranges between (0 – 15) marks. Intelligence levels are classified on the Wechsler (4) intelligence scale test by applying the following equation (highest value minus lowest value / divided by the number of levels, "which are three levels": low, medium, and high). Thus, the low level is from (0 – less than 5.00), the medium level is from (5 – less than 10.00), and the high level is from (10.00 – 15.00) marks.

Study Procedures

To achieve the objectives of the study, the following measures were taken:

- The pilot study was applied to the experimental sample (n = 5) to verify the extent of linguistic formulation and application, and correction procedures.

- The data were processed statistically according to the appropriate methods mentioned in the study.
- The test was applied to categories according to the gender variable (males, females).
- Indications of the validity and reliability of the scales were reached.

Statistical treatment

Following data entry, the Statistical Package for the Social Sciences (SPSS) program was used to process the data to analyze it and provide answers to the research questions. The frequencies and percentages were then extracted to determine the distribution of the study sample individuals based on demographic characteristics. The following methods were employed:

- Pearson Correlation Test.

- Using Factor Analysis and extracting the values of the data suitability for the sample using the KMO   test.
- Using the Varimax rotation test.

- Using the Cronbach Alpha test.

- Using the split-half test to verify the validity of the construct and the stability of the study tools.
- Using the Independent Samples T-test to verify the discriminant validity of the (arithmetic) test from the Wechsler Intelligence Scale (4) for primary school students in schools in the Kingdom of Saudi Arabia.

Results and Discussion

The study's findings were sorted into the following categories:

The First question: What are the implications of the Validity of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

The validity of the Wechsler-4 scale (arithmetic) test was verified. The following was used:

First: The validity of the Wechsler (4) arithmetic test among primary school students in the Kingdom of Saudi Arabia.

1- Content Validity

The validity of the content of the arithmetic test was verified from the original version of the Wechsler Scale (4), by presenting the Arabic version to some specialists related to special education, educational sciences, measurement and evaluation, and mathematics teachers in schools in the Kingdom of Saudi Arabia to review it and suggest any amendments so that it is compatible with the target group of students within the age group (6-12 years). And to verify the reliability of applying the arithmetic test to students within the primary stage and its suitability for the study individuals by a percentage of (80%).

2- Second: Construct validity

After administering the Wechsler Intelligence Scale (4) arithmetic exam to a study sample of fifty male and female primary school kids in the Kingdom of Saudi Arabia, its internal consistency was confirmed. The next table No. (2) Displays the Pearson correlation coefficient between the question score and the overall score of the

arithmetic test since it was computed between the question score and the test's total score.

Table (2) Pearson's correlation coefficient between the question score and the total score of the arithmetic test from the Wechsler Intelligence Scale (4), where (N= 50)

Arithmetic Test							
Paragraph number	(R)	Paragraph number	(R)	Paragraph number	(R)	Paragraph number	(R)
1	**.572	11	**.576	21	**.397	31	.462*
2	**.554	12	**.493	22	.365*	32	.631**
3	**.582	13	**.525	23	.489**	33	.637**
4	**.606	14	**.583	24	.548**	34	.656**
5	**.564	15	**.590	25	.559**		
6	**.512	16	**.648	26	.600**		
7	**.556	17	**.474	27	.372*		
8	**.472	18	**.559	28	.598**		
9	**.471	19	**.588	29	.640**		
10	**.607	20	**.644	30	.399*		

**There is statistical significance at the significance level ($\alpha = 0.01$)
*There is statistical significance at the significance level ($\alpha = 0.05$)

Table (2) shows that values of the correlation coefficients between the question score and the total score of the arithmetic test from the Wechsler Intelligence Scale (4) are statistically significant at the level ($\alpha = 0.01$, and $\alpha = 0.05$), which means that there is an acceptable degree of internal consistency for the arithmetic test from the Wechsler Intelligence Scale (4) among primary school students in the Kingdom of Saudi Arabia.

3- The factorial validity of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students (n=50):

The factorial validity of the arithmetic test was verified from the Wechsler Intelligence Scale (4), where the values of the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) were extracted. The values of the Chi-square coefficient (Chi2) were extracted, as well as the saturation coefficients and the statistical significance for each question of the test for primary school students, which were as follows:

Table (3) Factor Analysis of the Arithmetic Test from the Wechsler Intelligence Scale (4) for Elementary School Students (n = 50)

Question Number	Saturation coefficient	Chi square ($\chi 2$)	Kaiser-Meyer-Olkin (KMO)	Statistical significance
1	.770	828.424	0.831	*0.00
2	.769			

3	.753			
4	.723			
5	.721			
6	.666			
7	.739			
8	.681			
9	.554			
10	.758			
11	.765			
12	.616			
13	.616			
14	.717			
15	.696			
16	.741			
17	.602			
18	.803			
19	.792			
20	.819			
21	.682			
22	.918			
23	.847			
24	.740			
25	.710			
26	.650			
27	.410			
28	.413			
29	.564			
30	.551			
31	.531			
32	.603			

| 33 | .610 | | | |
| 34 | .578 | | | |

*: Significant at significance level (0.05)

Table (3) shows that arithmetic test items from the Wechsler Intelligence Scale (4) for primary school students in the Kingdom of Saudi Arabia had saturation coefficients ranging from (0.410 to 0.918), which is higher than the standard value of 0.40. The KMO value reached (0.831), which is higher than the standard value of (0.50), and the Chi-square value reached (828.424), with a statistical significance of less than (0.05). These findings show that all test questions were accepted and that they are significant and key components of the Wechsler Intelligence Scale

(4) for primary school students in government schools. The test has 34 acceptable questions, as shown in Figure (1).

The researcher used the factor analysis test using the Varimax rotation method to determine the importance of each question from the arithmetic test questions from the Wechsler Intelligence Scale (4) for primary school students in schools in the Kingdom of Saudi Arabia.

Table (4) Varimax factor analysis of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students in the Kingdom of Saudi Arabia

| | Iinitial Eigenvalues | | |
Question	Total	Variance (%)	Total Variance (%)
1	8.137	23.932	23.932
2	5.511	16.210	40.143
3	3.317	9.756	49.898
4	2.557	7.521	57.419
5	2.262	6.653	64.072
6	1.697	4.990	69.062
7	1.476	4.342	73.405
8	1.369	4.026	77.430
9	1.259	3.702	81.133
10	1.150	3.382	84.515
11	1.023	3.010	87.524
12	.863	2.537	90.061
13	.618	1.817	91.879
14	.461	1.355	93.234
15	.435	1.279	94.513
16	.364	1.072	95.585
17	.338	.994	96.580
18	.264	.777	97.357
19	.198	.582	97.939
20	.188	.553	98.492
21	.135	.396	98.888
22	.107	.315	99.203
23	.098	.288	99.491
24	.066	.195	99.686

25	.045	.132	99.818
26	.029	.084	99.902
27	.018	.054	99.956
28	.010	.030	99.986
29	.005	.014	100.000
30	3.958E-17	1.164E-16	100.000
31	-2.656E-16	-7.811E-16	100.000
32	-4.054E-16	-1.192E-15	100.000
33	-9.103E-16	-2.677E-15	100.000
34	-1.245E-15	-3.663E-15	100.000

Table (4) showed that all the questions of the arithmetic test from the Wechsler Intelligence Scale (4) for primary school students in schools in the Kingdom of Saudi Arabia explained (100%) of the variance in the total score of the test, and Figure No. (1) Illustrates this.
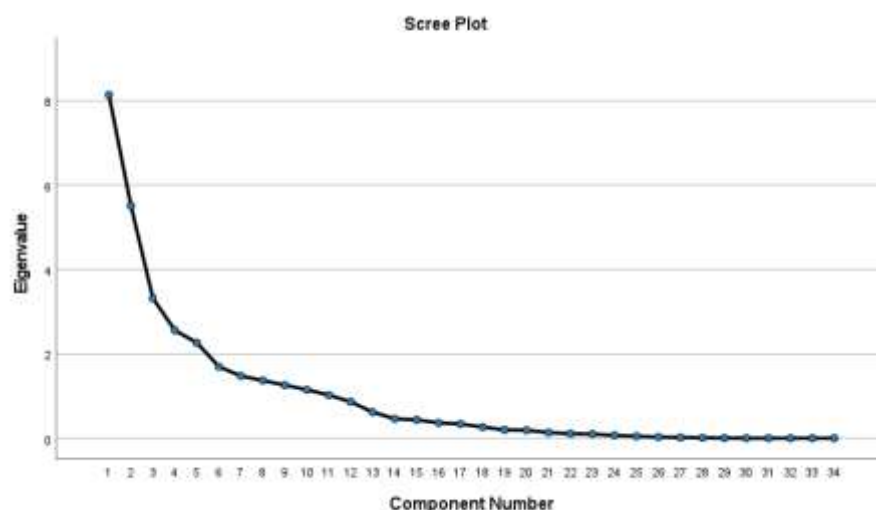

Figure No. (1) Components of the arithmetic test from the Wechsler Intelligence Scale (4) for primary school students in the Kingdom of Saudi Arabia

4- Discriminant validity

The researcher verified the discriminant validity of the arithmetic test of the Wechsler Intelligence Scale (4) among primary school students in schools in the Kingdom of Saudi Arabia by using the Independent Samples T-test. The following are the results:

1- Discriminant validity of the arithmetic test of the Wechsler Intelligence Scale (4) among primary school students in schools in the Kingdom of Saudi Arabia, regardless of gender.

Table (5) Independent Samples T-test to verify the discriminant validity of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students, depending on gender.

	Gender	Number	Arithmetic Mean	Standard deviation	Degrees of Freedom	Value (t)	Statistical significance

Arithmetic Test	Male	25	12.75	0.89			
	Female	25	9.74	0.88	48	2.610	*0.009

*: Statistically significant at the significance level (0.05).

Table (5) showed that the value of the statistic (T) reached (2.610) on the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students, depending on gender, which is a statistically significant value at the significance level (α = 0.05). The differences in the level of achievement on the arithmetic test were in favor of the male category. From the previous results, it is clear that the arithmetic test of the Wechsler

Intelligence Scale (4) for primary school students has acceptable discriminant validity indicators for the current study.

2- Discriminant validity of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students in schools in the Kingdom of Saudi Arabia, depending on age.

Table (6) Independent Samples T-test to verify the discriminant validity of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students in the Kingdom of Saudi Arabia according to age

	Age	Number	Arithmetic Mean	Standard deviation	Degrees of Freedom	Value (t)	Statistical significance
Arithmetic Test	6-8 years	20	12.341	0.94			
	9-12 years	30	8.457	0.64	48	3.580	*0.000

*: Statistically significant at the significance level (0.05).

Table (6) showed that the value of the statistic (T) reached (3.580) on the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students according to age, which is a statistically significant value at the significance level (0.05 = α), and the differences were in favor of the student's age group (6-8 years). From the previous results, it is clear that the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students has acceptable discriminant validity indications for the current study and that the older the students are, the more difficult the test becomes.

The Second question: What are the implications of the Reliability of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

The reliability of the Wechsler Intelligence Scale (4) arithmetic test among primary school students in schools in the Kingdom of Saudi Arabia for the purposes of the current study:

1- Cronbachs Alpha

The reliability of the Wechsler Intelligence Scale (4) arithmetic test for primary school students was verified by estimating the internal reliability coefficient of the scale by applying the Cronbach's alpha equation, where the scale was applied to the study individuals (n = 50), and the Cronbach's value for the total score on the arithmetic test was = (0.791), which indicates that the Wechsler Intelligence Scale (4) arithmetic test for primary school students has acceptable reliability coefficient values for the current study. (Hair et al., 2010).

2- Split-Half

The Wechsler Intelligence Scale (4) arithmetic test's reliability was confirmed among elementary school pupils using the split-half test, which was split into two sections. The reliability coefficient of each section was determined independently, as was the correlation coefficient between the two sections of the questions, and the reliability coefficient of the entire test was determined. The outcomes are as follows:

Table (7) Reliability coefficients of the arithmetic test of the Wechsler Intelligence Scale (4) among primary school students in the Kingdom of Saudi Arabia using the split-half method

Cronbach's alpha	Part One	Value	0.776
		Paragraphs	17
	Part Two	Value	0.852
		Paragraphs	17
	Total number of paragraphs		34
Correlation coefficient between the two parts			0.692
Spearman-Brown Coefficient		Equal Length	0.818
		Unequal Length	0.818
Guttman Split-Half Coefficient			0.815

Table (7) shows that the value of Cronbach's alpha coefficient for the first part of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students was (0.776), the value of the Cronbach's alpha coefficient for the second part of the arithmetic test was (0.852), the value of the correlation coefficient between the first and second parts was (0.692), and the value of the split-half coefficient using the Guttman Split-Half Coefficient test was (0.815), which indicates that the reliability coefficients are acceptable for the current study and that the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students has good and acceptable reliability indicators.

3- Difficulty and discrimination coefficients

The researcher determined the difficulty and discrimination coefficients for each test question for the study participants (n = 50) primary school students in order to determine the test's reliability. The correct response for a paragraph with a discrimination coefficient greater than (0.30) was deemed highly appropriate, while a paragraph with a discrimination coefficient between 0.20 and 0.30 was deemed acceptable. A paragraph with a discrimination coefficient lower than 0.20 or with a negative degree was deemed unacceptable, inappropriate, and incomprehensible.

If the difficulty coefficients for the paragraph that obtains coefficients range between (0.30-0.70), they are appropriate and good paragraphs, and if the difficulty coefficients are less than the value (0.30), they are considered difficult paragraphs, and if they are negative, they are rejected paragraphs. The following table (8) shows the results of the difficulty and discrimination coefficients for each paragraph of the test:

Table (8) Difficulty and discrimination coefficients on the arithmetic test items of the Wechsler Intelligence Scale (4) for primary school students in the Kingdom of Saudi Arabia

Paragraph	Difficulty factor	discrimination coefficient	Paragraph	Difficulty factor	discrimination coefficient
1	0.733	.341	18	0.400	.638

2	0.467	.232	19	0.400	.805
3	0.500	.298	20	0.467	.494
4	0.467	.430	21	0.467	.634
5	0.333	.496	22	0.467	.700
6	0.533	.505	23	0.500	.569
7	0.600	.551	24	0.500	.562
8	0.555	.320	25	0.533	.509
9	0.567	.548	26	0.433	.516
10	0.600	.595	27	0.777	.350
11	0.633	.403	28	0.466	.222
12	0.333	.587	29	0.551	.354
13	0.433	.590	30	0.556	.340
14	0.433	.612	31	0.350	.401
15	0.400	.000	32	0.556	.501
16	0.500	.681	33	0.605	.414
17	0.433	.590	34	0.530	.380

Table (8) shows that the exam contains both challenging and clear and unique questions and that the values of the discrimination and difficulty coefficients are within the acceptable range for the study participants. As a result, the test's (34) paragraphs are appropriate for measuring the study participants' age range.

Discussion
1- What are the implications of the Validity of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

The results of the study showed the availability of indications of validity in the Wechsler-4 scale for the subtest (arithmetic), through Content Validity at a rate of (80%), and indications of construct validity/internal consistency, as the values of the correlation coefficients between the question score and the total score of the (arithmetic) test from the Wechsler-4 scale reached statistical significance at the level of ($\alpha = 0.01$, and $0.05 = \alpha$), and indications of factorial validity, the saturation coefficients on the (arithmetic) test items among students ranged between (0.410 - 0.918), which is higher than the standard value (0.40), and the KMO value reached (0.831), which is higher than the standard value (0.50), and the Chi-square value reached (828.424), with a statistical significance of less than (0.05). Also, the indications of the validity of the factor analysis using the Varimax orthogonal rotation method to determine the importance of each question of the (arithmetic) test explained (100%) of the variance in the total score of the test, and the indications of discriminant validity according to gender through (Independent Samples T-test) where the value of the statistic (T) reached (2.610) in favor of males, and also the indications of discriminant validity according to age where the value of the statistic

(T) reached (3.580) in favor of the category of students aged (6-8 years).

- Researcher's interpretation: Any of the aforementioned techniques are regarded as having good and acceptable validity coefficients. This suggests that the scale's high validity indicators support its application in the Saudi context. The test gets harder as pupils get older since the scale likewise gradually increases the difficulty level from easy to challenging.

2- What are the implications of the Reliability of the psychometric properties of the Wechsler Intelligence Scale-4 arithmetic test among primary school students in the Kingdom of Saudi Arabia?

The results of the study showed the availability of the internal scale reliability indications through the use of the Cronbach's Alpha test, where the value of the Cronbach's Alpha coefficient for the total score on the arithmetic test = (0.791). Also through the split-half reliability indications, where the value of the Cronbach's Alpha reliability coefficient for the first part of the arithmetic test of the Wechsler Intelligence Scale (4) for primary school students reached (0.776), and the value of the Cronbach's Alpha reliability coefficient for the second part of the arithmetic test reached (0.852), while the value of the correlation coefficient between the first and second parts reached (0.692), and the value of the split-half reliability coefficient using the Guttman Split-Half Coefficient test reached (0.815).

- Researcher's interpretation: Any of the aforementioned techniques' reliability coefficients are regarded as good and acceptable. This suggests that the scale has consequences for reliability that support its application in the Saudi context.

Recommendations:

- Educational recommendations:
  - Educating educators in all fields regarding pupils' mathematical proficiency.

- Using mathematics problems in both written and oral ways.
- Developing the steps for the students' solutions.
- Assessing pupils by determining their strong and weak points in the greatest number of mathematical problems and solution steps.
- Suggestions for research purposes.
- Performing additional study and studies on pupils and how they affect mathematics.
- Research the effects of both mathematical problem-solving abilities and verbal intelligence.

References

1. Abu Drei, S. (2017). Standardization A Jordanian Version of The Wechsler Intelligence Scale for Children-Adapted for Deaf Via Sign Language From (6-16, 11) years (Doctoral dissertation, Phd dissertation. Amman, Jordan: The Word Islamic Science and Education University).
2. Abu Drei, Samer, Al-Rusan, Farouk. (2021). Psychometric properties of the Jordanian version of the Wechsler Intelligence Scale-4 for children in the school stage. Jordanian Journal of Applied Sciences, Humanities Series, Vol. () 27, No. (2). Jordan.
3. Bodin, D., Pardini, D. A., Burns, T. G., & Stevens, A. B. (2009). Higher order factor structure of the WISC-IV in a clinical neuropsychological sample. Child Neuropsychology, 15(5), 417-424.
4. Chen, H., & Hua, M. S. (2020). Selecting tetradic short forms of the taiwan wechsler adult intelligence scale IV. Assessment, 27(7), 1633-1644.
5. Flanagan, D. P., & Kaufman, A. S. (2009). Claves para la evaluación con WISC-IV. Editorial El Manual Moderno.
6. Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., & Dynda, A. M. (2013). 12 Cognitive Assessment: Progress in Psychometric Theories of Intelligence, the Structure of Cognitive Ability Tests, and Interpretive Approaches to Cognitive Test Performance. The Oxford handbook of child psychological assessment, 239.
7. Gilbert, K., Benson, N. F., & Kranzler, J. H. (2023). What Does the Digital Administration Format of the Wechsler Intelligence Scale for Children-(WISC-V) Measure?. Contemporary School Psychology, 27(4), 623-633.

8. Gregory, R. J. (2004). Psychological testing: History, principles, and applications (4th ed.). Boston: Allyn & Bacon.

9. Hair, J. F; Black, W. C; Babin, B. J; Anderson, R. E & Tatham, R. L, (2010). "Multivariate Data Analysis ", 7th edition., New York:

10. Harrison, A.G., Beal, A.L., Armstrong, I.T. et al. Measurement of Working Memory on the Wechsler Adult Intelligence Scale: Should We Subtract Arithmetic?. Psychol. Inj. and Law 17, 55–65 (2024). https://doi.org/10.1007/s12207-024-09499-3

11. Karzmark, P. (2009). The effect of cognitive, personality, and background factors on the WAIS-III arithmetic subtest. Applied Neuropsychology, 16(1), 49-53.

12. Levine, B., Dawson, D., Boutet, I., Schwartz, M. L., & Stuss, D. T. (2000). Assessment of strategic self-regulation in traumatic brain injury: its relationship to injury severity and psychosocial outcome. Neuropsychology, 14(4), 491.

13. Nelson, J. M., Canivez, G. L., & Watkins, M. W. (2013). Structural and incremental validity of the Wechsler Adult Intelligence Scale–Fourth Edition with a clinical sample. Psychological Assessment, 25(2), 618–630. https://doi.org/10.1037/a0032086

14. Parkin, J. R., & Beaujean, A. A. (2012). The effects of Wechsler Intelligence Scale for Children—Fourth Edition cognitive abilities on math achievement. Journal of School Psychology, 50(1), 113-128.

15. Phelps, L., McGrew, K. S., Knopik, S. N., & Ford, L. (2005). The General (g), Broad, and Narrow CHC Stratum Characteristics of the WJ III and WISC-III Tests: A Confirmatory Cross-Battery Investigation. School Psychology Quarterly, 20(1), 66.

16. Prifitera, A., Weiss, L. G., Rolfhus, E., & Saklofske, D. H. (2005). The WISC-IV in the clinical assessment context. In WISC-IV clinical use and interpretation (pp. 3-32). Academic Press.

17. Schneider, W. J., & Flanagan, D. P. (2015). The relationship between theories of intelligence and intelligence tests. Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts, 317-340.

18. Stroud, J. B., Blommers, P., & Lauber, M. (1957). Correlation analysis of WISC and achievement tests. Journal of Educational Psychology, 48(1), 18–26. https://doi.org/10.1037/h0047378

19. Watkins M ; Smith L(2013). Long-Term Stability of the Wechsler Intelligence Scale for Children—Fourth Edition. American Psychological Association. Vol. 25, No. 2, 477– 483.

20. Wechsler, D. (2005). Escala Wechsler de inteligencia para niños-IV. Manual moderno.

21. Wikoff, R. L. (1978). Correlational and factor analysis of the Peabody Individual Achievement Test and the WISC-R. Journal of Consulting and Clinical Psychology, 46(2), 322–325. https://doi.org/10.1037/0022-006X.46.2.322.

22. Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. Journal of Psychoeducational Assessment, 8(3), 231-258.