Constructing an Achievement Test in Palestinian National Culture for Palestinian University Students Using the Rasch Unidimensional Model

[1] **Assoc. Prof. Dr. Mohammad Talab Dabous and** [2] **Assoc. Prof. Dr. Mohammad Muhiuddin Assaf**

[1] Al-Isteqlal University, Palestine.

Abstract. The study aimed to construct an achievement test in Palestinian national culture for university students using the Rasch unidimensional teacher model. To achieve the study's objective, an achievement test in Palestinian national culture was constructed consisting of (50) dichotomous items of multiple-choice type. The test was administered to a sample of (606) male and female students from Palestinian universities in Palestine during the second semester of the academic year 2022/2023. For data analysis, the statistical programs Winsteps and SPSS were used. The alignment of the items of the achievement test in Palestinian national culture with the assumptions of the Rasch unidimensional teacher model was verified, where all items were found to align with the model's assumptions. The test's validity and reliability were also confirmed, with the results indicating a person separation reliability of (0.83) and an item separation reliability of (0.99). Based on the results, the researchers recommended the use of the test in Palestinian universities to assess students' possession of Palestinian national culture.

Keywords: Achievement test, Item Response Theory, Rasch Unidimensional Model.

Background of the study and its importance:

Achievement tests are considered one of the most important tools for measurement and evaluation, and they are widely used. The word "test" is commonly used in the educational process. It is used in measurement and evaluation to determine an individual's level of possession of a specific characteristic through their responses to a sample of stimuli that represent the characteristic.

The process of constructing achievement tests has received significant attention from educators and decision-makers in the field of education. This is because they provide reliable data that can be relied upon in making sound decisions aimed at improving the processes of learning and teaching (Abu Jarad, 2011).

Murad and Suleiman (2002) have pointed out that achievement tests are a good means of learning and their results enhance behavior and raise the level of ambition among learners. They also improve the mastery of the learned material, which helps in transferring the impact of learning from the current situation to the future situation. Additionally, they improve the recall method due to the feedback they provide and direct students' attention towards achieving the desired teaching goals.

As for tests that are constructed based on the item response theory, they are characterized by a high degree of objectivity, and their results can be generalized. Furthermore, they provide the opportunity to compare diverse groups of individuals, allowing for accurate estimation and comparison of individuals who take the test with others (Mahmoud & Sabah, 2014).

Alimmat (2021) has indicated that tests in Arab countries are predominantly based on the traditional theory of measurement and evaluation, despite the numerous flaws associated with this theory. Despite the existence of the item response theory, which has gained the interest of researchers as it overcomes many problems in the traditional theory of measurement and evaluation.

The study of Palestinian history over the past decades reveals a wide interest in Palestinian national identity through school and university curricula, which reflect the reality and provide a clear image to students.

The item response theory has emerged to overcome many shortcomings in the traditional theory and has provided reliable methods for addressing fundamental issues in educational measurement, such as item calibration, building item banks, and constructing criterion-referenced tests (Hambelton & Swaminathan, 1985).

This theory assumes that individual performance or their interpretation in a specific psychological or educational test can be predicted in light of a characteristic or traits that determine the individual's response to the items. This characteristic is shared by all individuals (Molenaar & Hoijtink, 1996).

Based on this theory, a set of models known as latent trait models have been developed to determine the relationship between an individual's performance in the test, which can be directly observed, and the traits or abilities underlying and explaining this performance. Each model contains a mathematical equation that determines the relationship between an individual's performance on an item of the scale or test and their underlying ability that explains and interprets this performance.

One of the most famous and widely used models is the One-Parameter Logistic Model (Rasch Model). This model is associated with the Danish mathematician Georg Rasch. This model is considered one of the most important models in the item response theory for constructing and analyzing tests. It has been the subject of multiple research studies to verify its characteristics and its ability to construct tests and address the shortcomings resulting from the use of the classical approach in psychological and educational measurement (Zakaria, 2009).

This model is used to estimate the difficulty of items. It assumes that all items have the same discrimination value, and the estimation is at its minimum, meaning that the estimation is close to zero. It also assumes that the discrimination is constant for all test items, while the difficulty of the item takes variable values. The equation of this model is as follows:

$$Pi(\theta) = \frac{e^{D(\theta - bi)}}{1 + e^{D(\theta - bi)}} \qquad i = 1, 2, 3, \ldots n.$$

Where:

Pi(Θ): the probability that an individual with ability Θ will answer item i correctly

bi: the difficulty parameter of item i

Θ: individual ability

e: the natural logarithmic base, approximately equal to 2.718

D: the discrimination parameter, equal to 1.7

Any mathematical model relies on a number of assumptions regarding the data used by the model, which determine the mathematical relationships between the observed and unobserved constructs described by the model. Therefore, the item response theory is based on strong assumptions that need to be verified in the data in order to yield reliable results (Allam, 2005).

Hambleton and Swaminathan (1985) have clarified these assumptions, and the first assumption is the unidimensionality assumption. The item response theory assumes the existence of a single latent ability or trait that explains an individual's performance on the test, referred to as unidimensional models. To fulfill this assumption, there needs to be a dominant factor that influences performance on the test, which is referred to as the ability measured by the test. This assumption is verified through factor analysis of the items.

The second assumption is the local independence assumption, which means that the response to one item does not positively or negatively affect the response to any other item. The local independence assumption is fulfilled when the set of items is unidimensional (i.e., measuring a single ability or trait). If the unidimensionality assumption is not met, the item scores will be correlated at a certain ability level because the test items measure more than one ability.

The third assumption is the item characteristic curve (ICC) assumption, which refers to a mathematical function that relates the probability of answering an item correctly to the individual's measured ability.

The fourth assumption is the speediness assumption, where item response theory models assume that the speed factor does not play a role in answering test items. This means that examinees who fail to answer test items do so because of their limited ability, not because they lack sufficient time to access and answer the item.

Hambleton (1989) has pointed out that one of the key features of item response theory models is that their use makes item characteristics (difficulty, discrimination, guessing) independent of the sample of individuals used to estimate these characteristics. It also makes individual ability estimates independent of the test items used to obtain these estimates. Furthermore, it improves the accuracy and stability of results by identifying and removing individuals and items that do not conform to the model used.

Study Problem and Questions:

This study aims to construct a test for Palestinian national culture and employ the Rasch model in building this test. The study problem is defined in the following questions:

1. To what extent do the assumptions of the response theory paragraph in the Palestinian culture test achieve among

Palestinian university students according to the Rasch unidimensional logistic model?

2. What are the characteristics of the achievement test paragraphs in Palestinian national culture among Palestinian university students according to the unidimensional Rasch model?

3. What are the implications of the stability and validity of the achievement test in Palestinian national culture among Palestinian university students according to the unidimensional Rasch model?

4. What is the degree of alignment of the achievement test paragraphs in Palestinian national culture among Palestinian university students according to the unidimensional Rasch model?

Study Objectives:
The study aims to:

1. Construct an achievement test in Palestinian national culture according to one of the response theory paragraph models (unidimensional Rasch model) with the specifications of a good test in terms of psychometric properties.

2. Estimate the difficulty coefficient for the achievement test paragraphs in Palestinian national culture according to one of the response theory paragraph models (unidimensional Rasch model).

3. Calculate the reliability coefficients for the achievement test paragraphs according to the unidimensional Rasch model.

4. Match the achievement test paragraphs with the assumptions of the unidimensional Rasch model.

Study Significance
The importance of the study lies in the significance of the topic it addresses, which is the construction of the achievement test according to the unidimensional Rasch model in modern theory of

measurement and assessment (Item Response Theory) to overcome the difficulties faced by test developers according to the traditional theory of measurement and assessment. It also aims to provide achievement test items in Palestinian national culture that assist decision-makers in determining the level of proficiency attained by university students in their knowledge of Palestinian national culture. Building a test according to the Item Response Theory models makes the test more accurate, objective, and independent in measurement.

Furthermore, the importance of the study lies in determining the appropriate method and approach to increase the ability of statistics course instructors to identify suitable items in the statistics test in order to achieve accuracy and objectivity in the test.

Study Limitations:
The study is limited by the following boundaries:
1. Spatial limitation: This study is limited to Palestinian universities.
2. Temporal limitation: This study was conducted during the second semester of the academic year 2022/2023.
3. Human limitation: This study is limited to a sample of Palestinian university students.
4. Subject limitation: This study is limited to the construction of a test in Palestinian national culture according to one of the Item Response Theory models, specifically the unidimensional Rasch model.

Study Terminologies:
- **Achievement test:**
  A systematic method to determine students' level of achievement in knowledge and skills in a specific subject that has been previously learned, through their responses to a set of

test items that represent the content of the subject accurately (Abu Jodeh, 2018).

- **Palestinian national culture:**

  The researchers define it as the degree of Palestinian university students' knowledge about the history, geography, and general Palestinian culture from ancient times until now. It is measured through students' responses to a specific test that measures Palestinian national culture, which was developed by the researchers.

- **Item Response Theory:**

  A modern theory in psychological and educational measurement that determines the relationship between examinees' performance and the latent trait being measured, according to a specific mathematical function. This theory relies on several models called latent trait models, which establish the link between item performance and examinees' ability (Hambleton & Swaminathan, 1985).

- **Unidimensional Rasch model:**

  A logistic unidimensional logarithmic model proposed by George Rasch, which estimates the probability of an individual providing a correct response based on their ability and the difficulty parameter (Allam, 2005).

Previous Studies:

The aim of the study by Alimat (2022) was to construct a simulated reference test in mathematics using the three-parameter logistic model consisting of 35 multiple-choice items. The test was administered to a sample of 968 students, and the results showed a lack of fit for three items to the three-parameter logistic model. The person separation reliability coefficient was 0.83, and the test reliability coefficient was 0.833. The assumptions of the model

were met, and the study results indicated that the item parameter estimates (difficulty, discrimination, guessing) were acceptable.

Similarly, Al-Zubaidi (2018) aimed to construct a simulated reference achievement test in the Testing and Measurement course for students in the Diploma in Education program at Taif University using the Rasch model. The test consisted of 37 multiple-choice items with four alternatives. It was administered to a sample of 147 students. The results indicated that all items fit the assumptions of the Rasch model. The person reliability coefficient was 0.9, and the test reliability coefficient was 0.79.

In another study, Al-Mousawi (2017) aimed to develop a scale for assessing the quality of life of secondary school students using the Rasch model. The initial version of the scale consisted of 30 items with Likert scaling. It was administered to 442 secondary school students aged 15-17 years. The results indicated that 24 items matched the assumptions of the Rasch model with a single parameter, which is one of the suitable modern theories for the scale used. The final version of the scale consisted of 24 items with appropriate psychometric properties. The results showed that the developed scale provided the most information about the quality of life of adolescents with average ability, as the average ability value (0.50) was approximately equal to the average difficulty value of the scale items (0.57), which is consistent with the expected value in the measurement model.

Dabous (2016) aims to use the item response theory to construct a pool of items for a simulated reference test in mathematics using the two-parameter logistic model. The pool consisted of 50 dichotomous items and 10 polytomous items. The study sample included 502 students. The results confirmed the assumptions of the item response theory, and the analysis showed that all dichotomous items matched the two-parameter logistic model, while the polytomous items matched the graded response model.

Edwards and Alcock (2010) analyzed the results of a mathematics test in the UK using the Rasch model. They constructed a test consisting of 11 items and administered it to a sample of 164 first-year university students. The participants were asked to rank the items according to their difficulty level. The data obtained were analyzed using the Rasch model, and the results indicated a lack of fit for only six individuals after reordering the items according to the Rasch model.

Koh et al. (2006) aimed to verify the unidimensionality of the ARAT test items using the Rasch model. The test was administered to 351 patients in Taiwan. The results showed the effectiveness of 19 items in measuring all the test topics, as well as supporting the unidimensionality assumption of the previous items. The Mokken model was used for the analysis, which revealed a lack of fit for four items to the Rasch model.

Study Procedures

Study Methodology:

The researchers used the quantitative analytical method, which is expressed qualitatively or quantitatively in order to analyze test paragraphs according to the Rasch unidimensional logistic model using the Winsteps program.

Study Community:

The study population consists of all students of Palestinian universities located in the West Bank during the second semester of the academic year 2022/2023.

Study Sample:

The study sample consisted of (606) male and female students from various Palestinian universities located in the West Bank, and

Table (1) shows the distribution of the study sample by university and gender.

Table (1): Distribution of the Study Sample

Variable	Male	Percentage
National Success	135	22.3%
Palestine Technical "Khadouri"	126	20.8%
Bethlehem	58	9.6%
Jerusalem	97	16%
Independence	190	31.4%
Total	606	100%

Study Instrument:

The researchers followed the scientific steps in constructing an achievement test to measure the level of Palestinian national culture among university students, according to the following steps:

First: Determining the purpose of the test: The purpose of constructing the Palestinian national culture test is to measure the extent to which Palestinian university students possess Palestinian national culture in general, and their sufficient knowledge about the Palestinian issue in particular.

Second: Reviewing previous studies: The researchers reviewed a number of previous studies that focused on the construction and development of tests and scales, such as the study by Kassabah and Matarneh (2015), Alimat's study (2022), Alawneh's study (2018), and Al-Zoubi and Swalma's study (2017). They benefited from these studies in following the scientific steps discussed in the theoretical literature of these studies regarding the preparation of tests and scales based on the item response theory.

Third: Reviewing the courses related to the Palestinian issue: The researchers reviewed the courses related to the Palestinian issue

taught at various Palestinian universities. These courses have different names but share the content related to the Palestinian issue, which includes the study of Palestinian civilization and historical geography of Palestine, Palestine in the Arab Islamic era, the Palestinian Arab society, the Zionist program, Palestine and Arab relations, the Palestinian Liberation Organization, and the establishment of the Palestinian National Authority.

Fourth: Determining the test content: After reviewing the courses related to the Palestinian issue taught at Palestinian universities, the researchers determined the content of the test by formulating specific objectives that the test would measure. The test included three main topics: Palestinian history, Palestinian geography, and general Palestinian culture.

Fifth: Formulating the initial test items: After reviewing the literature related to the history of Palestine in general and the Palestinian issue in particular, as well as the courses related to the Palestinian issue taught at Palestinian universities, the researchers formulated 60 items related to Palestinian national culture. These items were of the multiple-choice type, with four alternatives for each item, one of which is correct. Each item was designed to measure a specific objective. The items were distributed across three dimensions: the first dimension was Palestinian history, the second dimension was Palestinian geography, and the third dimension was general Palestinian culture. Instructions for the test and the model answers were also written.

Sixth: Item validation: After formulating the initial items and providing model answers, the researchers presented them to a group of experts in Palestinian history, the Palestinian issue, and political science. The experts, who held doctoral degrees in their respective fields, were asked to express their opinions on the items, the extent to which each item was related to its corresponding dimension, the clarity of the items, the suitability of the alternatives for each item,

and the correctness of the answers. Based on the suggestions of the experts, the researchers made modifications and deletions to some items, resulting in a final test consisting of 54 items distributed across three dimensions: Palestinian history, Palestinian geography, and Palestinian culture.

Seventh: Pilot testing of the items: The researchers applied the initial version of the test to a sample of 40 students selected from Palestinian universities. The students were asked to carefully read the test instructions and answer all the items. The purpose of the pilot testing was to identify any ambiguous items, assess the difficulty and discrimination of each item, and determine the time needed to complete the test. The time taken by each student to answer the items ranged from 70 to 130 minutes, with an average of 100 minutes, or two minutes per item.

Eighth: Finalizing the test: The researchers corrected the answers of the pilot sample and studied the difficulty and discrimination indices of each item according to the traditional measurement and evaluation theory. They found that there were items with discrimination indices below 0.25 and items with difficulty indices of 0.22. These items were deleted, resulting in a final test consisting of 50 items distributed across three dimensions: Palestinian history (16 items), Palestinian geography (17 items), and Palestinian culture (17 items). The researchers also ensured the reliability of the test and found that it had a reliability coefficient of 0.853, which is high and suitable.

Ninth: Applying the test to the study sample: The researchers administered the Palestinian national culture test to a sample of 606 students from various Palestinian universities in the West Bank. During the administration, it was ensured that each student answered all the test items.

Tenth: Scoring the test: The researchers scored the test papers according to the key of the model answers. The students' answers

were then entered into the computer using the Statistical Package for the Social Sciences (SPSS) software. Each correct answer was given a score of 1, while incorrect answers were given a score of 0. The total score for each examinee on the test was the sum of their correct answers, with a maximum score of 50.

Eleventh: Data analysis: The researchers analyzed the data using the Rasch model with the Winsteps software.

Study Results and Discussions

Results Related to the First Research Question:

To what extent do the assumptions of the Response Theory regarding the item in the Palestinian culture test among Palestinian university students conform to the unidimensional logistic monomodel (Rasch model)?

To answer this question, the assumptions of the Response Theory for the item were verified, which are:

1. Unidimensionality Assumption:

To verify the unidimensionality assumption for the binary graded test data, a first-order factor analysis was conducted for the items with Promax oblique rotation to extract the factors responsible for performance in the test models. The Eigen Value and the explained variance ratio for each factor extracted and the cumulative percentage of explained variance were calculated. Table (2) shows the results of the first-order factor analysis.

Table (2): Results of the first-order factor analysis for the sample responses on the binary graded Palestinian culture test items.

Accumulated Percentage of	The percentage of explained variance	The Latent	The Factor
19.520	19.520	6.832	1
24.703	5.183	1.814	2
29.095	4.392	1.537	3
33.020	3.925	1.374	4

36.540	3.520	1.232	5
40.003	3.464	1.212	6
43.362	3.359	1.176	7
46.508	3.146	1.101	8
49.505	2.997	1.049	9
52.397	2.893	1.012	10
55.276	2.879	1.008	11

The table (2) shows that the results of the first-order factor analysis for the two-factor model revealed the presence of (11) factors explaining a total of (55.276%) of the total variance. The analysis results in table (2) also indicate that the eigenvalue of the first factor is (6.832), which is high compared to the other factors, and that the remaining eigenvalues of the other factors are low and close together, indicating the presence of a dominant factor that favors unidimensionality for the purposes of analysis using item response models. As for the variance explained by each factor, we find that the percentage of variance explained by the first factor is (19.520%) of the total variance, while the percentage of variance explained by the second factor is (5.183%) of the total variance. It is noted here that the percentage of variance explained by the first factor is high compared to the percentage of variance explained by the second factor. It is also observed that the differences in the percentages of variance explained by each of the remaining factors are small and very close together, indicating relative symmetry and quasi-stability in the percentages of variance explained by all factors except the first factor, which suggests the achievement of unidimensionality in the test data.

By observing the graphical representation of the Scree plot for the factors with their eigenvalues in figure (1), it is evident that there is a shift in the curve's slope at the second factor, while the slope remains close for the rest of the factors, also suggesting the presence

of a dominant factor that can be used to infer unidimensionality for the purpose of estimating item and person parameters.
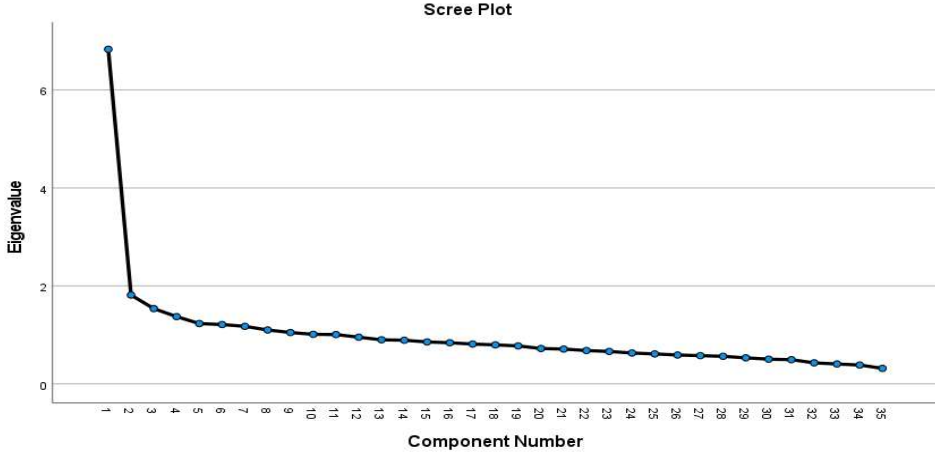

Figure 1. The graphical representation of the latent root values of the first-degree factors for the dichotomous National Culture Test items.

2. Local Independence Assumption:

Since the unidimensionality assumption is equivalent to the local independence assumption, verifying the unidimensionality assumption is sufficient to infer the fulfillment of the local independence assumption. It is worth mentioning that Hambelton and Swaminathan (1985) have stated that verifying the unidimensionality assumption is equivalent to the second assumption, which is the local independence assumption. Similarly, Warm (1978) believes that the unidimensionality assumption includes the local independence assumption and not vice versa.

3. Speededness Assumption:

The researchers ensured that examinees who fail to answer the test items do so due to limitations in their ability and not because the time is insufficient to reach and answer the item.

Results Related to the Second Research Question:

What are the characteristics of the achievement test items in Palestinian national culture among Palestinian university students according to the Item Response Theory (Rasch Model)?

To determine the item characteristics for each item of the dichotomous Palestinian National Culture Test items according to the Item Response Theory, the single-parameter model (Rasch model) was adopted. The Winsteps program was used to calculate the difficulty and standard error parameters in measuring these values. Additionally, the maximum value of the information function $I(\theta)max$ and the corresponding ability $(\theta)max$ were calculated for each item. Table 3 shows the item analysis results.

Table (3): Difficulty and Standard Error Parameters, Maximum Value of Information Function $I(\theta)max$, and Corresponding Ability for the Maximum Value of Information Function $(\theta)max$.

Corresponding Ability for the Maximum Value of Information Function $(\theta)max$	Maximum Value of Information Function $I(\theta)max$	Standard Error	Difficulty Coefficient	The Paragraph Title	The Paragraph
-0.7	0.199	0.17	-0.74	A1	1
-1.1	0.09	0.33	-1.2	A2	2
-0.2	0.106	0.2	-0.33	A3	3
-0.4	0.149	0.17	-0.5	A4	4
-0.9	0.17	0.22	-0.97	A5	5
-0.5	0.187	0.17	-0.58	A6	6
0	0.053	0.26	-0.13	A7	7
-0.3	0.332	0.12	-0.47	A8	8
-1	0.339	0.17	-1.1	A9	9

العدد التاسع – حزيران – June / 2024

0.1	0.214	0.14	0.05	A10	10
-0.5	0.322	0.12	-0.56	A11	11
-0.1	0.267	0.12	-0.19	A12	12
-0.5	0.416	0.11	-0.56	A13	13
-0.3	0.238	0.13	-0.34	A14	14
-0.1	0.32	0.11	-0.11	A15	15
0.1	0.254	0.12	-0.03	A16	16
0.9	0.513	0.11	0.75	B1	17
-0.1	0.364	0.11	-0.11	B2	18
0	0.071	0.23	-0.05	B3	19
1.7	0.045	0.52	1.55	B4	20
-0.2	0.185	0.15	-0.32	B5	21
-0.3	0.11	0.2	-0.43	B6	22
-0.3	0.336	0.12	-0.42	B7	23
0.1	0.108	0.19	0.05	B8	24
-0.1	0.386	0.1	-0.24	B9	25
1.1	0.114	0.26	1.03	B10	26
0.1	0.614	0.08	-0.06	B11	27
-0.3	0.231	0.13	-0.31	B12	28
0.3	0.161	0.16	0.21	B13	29
0.6	0.262	0.14	0.51	B14	30
-0.3	0.223	0.14	-0.36	B15	31
-0.5	0.206	0.15	-0.56	B16	32
0.5	0.275	0.13	0.46	B17	33
-0.7	0.174	0.19	-0.78	C1	34
-0.1	0.295	0.12	-0.26	C2	35
0.5	0.202	0.16	0.43	C3	36
-0.4	0.384	0.11	-0.49	C4	37

العدد التاسع – حزيران – June / 2024

-0.5	0.401	0.11	-0.65	C5	38
0.3	0.136	0.17	0.22	C6	39
-1	0.073	0.33	-1.07	C7	40
-0.7	0.123	0.22	0.8	C8	41
0	0.144	0.16	-0.02	C9	42
0.3	0.102	0.2	0.19	C10	43
0.3	0.102	0.2	0.23	C11	44
-1.1	0.315	0.19	-1.17	C12	45
-0.1	0.264	0.12	-0.19	C13	46
-0.5	0.319	0.12	-0.61	C14	47
-0.4	0.187	0.16	-0.53	C15	48
-0.5	0.384	0.11	-0.59	C16	49
-0.2	0.121	0.18	-0.29	C17	50
		0.167	-0.217	Arithmetic Mean	

The table (3) shows that the range of item difficulty coefficients varies between (-1.2) and (+1.55), with an average value of (-0.217), and a standard error ranging from (0.08) to (0.52) with an average value of (0.167). Paragraph (2) obtained the lowest difficulty coefficient value according to the Rasch model, which was (-1.2), indicating that it is the easiest item. On the other hand, paragraph number (20) obtained the highest difficulty coefficient value of (+1.55), indicating that it is the most difficult item. The remaining items were of moderate difficulty.

A study by Jamhawi (2000) indicated that items with difficulty coefficients ranging from (-1.5) logits to (+1.5) logits are considered within an acceptable range of difficulty coefficients. Therefore, the remaining items are of moderate difficulty. As shown in table (3), the remaining test items consider individual differences in sample levels and cover a wide range.

Results Related to the Third Research Question:

What are the implications of validity and reliability of the achievement test in Palestinian national culture among Palestinian university students according to the Rasch model?

Test Validity: The validity of the test was verified through two types of validity evidence. The first type relates to the validity of the judges, where the final version of the test consisting of (50) items was presented to a group of expert judges who indicated that the test measures what it is intended to measure. The second type of validity evidence is construct validity, where the researchers used factor analysis to confirm the validity of the test, as shown in the previous table (2). Factor analysis was used to analyze the students' responses to identify the factors measured by the test. The first-order factor analysis was conducted for the items using Promax oblique rotation to extract the factors responsible for performance in the test. The eigenvalue and the explained variance ratio for each extracted factor and the cumulative variance ratio were calculated. The results indicated the presence of a dominant factor that favors unidimensionality for the purposes of analysis using item response models.

Test Reliability Estimation: The reliability coefficient in item response theory refers to the accuracy of individual and item locations on the latent trait being measured. The degree of this accuracy can be determined by calculating the separation index between items. The WinSteps software was used to extract the person separation reliability coefficient, item separation reliability coefficient, item separation index, and person separation index. The results are shown in table (4).

Table (4): Grading reliability for ability estimates (individuals) and grading reliability for the item parameter

separation index	reliability	

2.21	0.83	individuals
10.43	0.99	items

It is evident from Table (4) that the reliability coefficient for individuals reached (0.83) and the separation index was (2.21), indicating that the reliability coefficient for individuals was (0.99) and the separation index was (10.43). These values related to the separation index between individuals, known as the Person Separation Index, reflect the ratio between the standard deviation of the estimated values of individuals' abilities and the difficulty of the items, as well as the average standard error. An increase in these values above (2) is considered evidence that the sample of individuals is sufficient for distinguishing between individuals' abilities and items (Masters and Wright, 1982). Consequently, the reliability is high and the separation index is greater than 2, confirming the test's ability to highlight individual differences among individuals and its precision in measurement.

Results Related to the Fourth Research Question:

What is the degree of alignment of the items of a cultural achievement test in Palestinian national culture among Palestinian university students according to the Rasch unidimensional model?

The WinSteps program was used to extract the internal and external alignment values of the items. Table (5) shows the results of the alignment of the items of the cultural achievement test in Palestinian culture for the Rasch bidimensional model.

Table (5): Testing the alignment of the Palestinian Culture Achievement Test items for the single-teacher model.

External correspond OUT.MSQ	internal correspond IN.MSQ	The Paragraph Title	The Paragraph
1.299	1.0892	A1	1
0.6585	0.5175	A2	2

0.8993	0.8555	A3	3
1.6054	1.1805	A4	4
0.9977	1.0069	A5	5
0.7754	0.887	A6	6
1.04	1.0215	A7	7
0.9006	0.9408	A8	8
0.934	0.9714	A9	9
1.038	1.042	A10	10
0.8563	0.9003	A11	11
4.1448	1.6858	A12	12
0.9824	0.9605	A13	13
1.1828	1.0829	A14	14
1.4011	1.1249	A15	15
0.9205	0.9514	A16	16
0.6964	0.628	B1	17
0.9297	0.9095	B2	18
0.9767	0.9817	B3	19
1.3157	1.1101	B4	20
1.0032	1.0177	B5	21
1.0118	1.0107	B6	22
0.6376	0.7917	B7	23
0.9093	0.9111	B8	24
0.8638	0.8965	B9	25
0.7015	0.8419	B10	26
0.7862	0.877	B11	27
0.9284	0.954	B12	28

العدد التاسع – حزيران – June / 2024

1.0502	1.0342	B13	29
1.3092	1.131	B14	30
0.8232	0.8698	B15	31
0.7803	0.8548	B16	32
1.0491	1.0415	B17	33
1.3095	1.2009	C1	34
0.8695	0.8923	C2	35
1.1006	1.0526	C3	36
0.893	0.9244	C4	37
0.9639	0.9557	C5	38
1.6007	1.2641	C6	39
1.4062	1.078	C7	40
1.1462	1.103	C8	41
1.5023	1.1419	C9	42
1.049	1.0147	C10	43
1.254	1.0799	C11	44
0.8678	0.887	C12	45
1.3065	1.1681	C13	46
0.9089	0.9471	C14	47
1.3206	1.1182	C15	48
1.065	1.0133	C16	49
0.8514	0.9075	C17	50

It is evident from Table (5) that all items showed acceptable internal and external consistency values within the range of -2 to +2. Based on this, all items are considered to be consistent with the unidimensional Rasch model.

Discussion of Study Results:

This study aimed to develop a national cultural achievement test in Palestinian culture in light of Palestinian national knowledge, which Palestinian university students are required to possess. The researchers referred to the issues of the Palestinian cause and the Palestinian history studied in Palestinian universities in order to construct test items. The researchers used modern measurement theory to analyze test items according to the assumptions of the unidimensional Rasch model.

Recommendations:

Based on the study results, the researchers recommend the following:

1. Adopting the final version of the Palestinian national culture test to measure the extent of Palestinian university students' knowledge of the Palestinian cause, given the test's validity and reliability indicators according to the item response theory.
2. Utilizing the methodology used in test construction to develop tests for general courses in Palestinian universities.
3. Verifying the characteristics of the Palestinian national culture test using a two-parameter model and a three-parameter model.

References

[1]   Abu Jarad, Hamdi. (2011). The extent of teachers' commitment to analyzing the results of achievement tests and their relationship with their attitudes towards them. Journal of Al-Azhar University - Gaza, Series of Human Sciences, 13(2), 89-106.

[2]   Abu Jodeh, Shaza Ibrahim. (2018). The effectiveness of using the Rasch model in constructing a criterion-referenced achievement test in mathematics for ninth grade students in Jordan. (Published Master's thesis). Amman Arab University, Jordan.

[3]   Alimat, Mahdi Mohammad Fahad (2022). Constructing a criterion-

referenced test in mathematics based on the three-parameter logistic model. Journal of the Open University of Jerusalem for Educational and Psychological Research and Studies. Volume 13, Number 38. 46-62.

[4] Alimat, Mahdi. (2021). Constructing a mathematics achievement test based on the item response theory (the two-parameter model). Journal of An-Najah University for Research (Humanities), Volume 35, Number 7.

[5] Allam, Salah El-Din. (2005). Test item response models: unidimensional and multidimensional and their applications in psychological and educational measurement. 1st ed. Cairo: Dar Al Fikr Al Arabi.

[6] Al-Musawi, Nu'man Muhammad Saleh (2017). Using the Rasch Model to Develop a Tool for Measuring Quality of Life among High School Students in the Kingdom of Bahrain. Arab Childhood Journal, Issue 70, Volume 17.

[7] Al-Zubaidi, Mohammed bin Hussein Yahya (2018). Constructing a Narrative Achievement Test in the Testing and Measurement Course Using the Rasch Unidimensional Model for Educational Diploma Students at the College of Education, Taif University. International Educational Journal. Volume 7, Issue 2. pp. 40-54.

[8] Dabous, Mohammad Taleb (2016). Using item response theory to build a pool of items for a criterion-referenced test in mathematics with binary and polytomous items based on the two-parameter logistic model. Journal of An-Najah University for Research (Humanities), Volume 1, Number 7.

[9] Edwards, H. and Aslcock, I. (2010). Using Rasch analysis to identify uncharacteristic responses to undergraduate assessment. Teacher Mathematics and its Application, (29), pp:165-175.

[10] Hambleton, R. K. (1989). Principles and selected applications of item response theory. In R. L. Linn (Ed.), Educational measurement (pp. 147–200). Macmillan Publishing Co, Inc; American Council on Education

[11] Hambleton, R.K. Swaminthan, H. (1985). Item Response Theory: Principles and Application, Kluwer. Nijhoff Publishing, Boston.

[12] Koh, Chia-Lin,et,al. (2006). Validation of the Action Research Arm. Test Using Item Response Theory in Patients after Stroke, Retrieved 22/5/2013 from: journal of Rehabilitation Medicine.

[13] Mahmoud, Lamia Rafik Ismail. (2014). Constructing an achievement test

العدد التاسع – حزيران / June / 2024

in geometry for fifth-grade elementary school students using the Rasch model. (Published Master's thesis). Al-Hashemite University, Jordan.

[14] Masters. G. and wright B.(1982). Rating scale analysis. Ahand book for research measurement Chicago: MESA press.

[15] Zakaria, Ali. (2009). Psychometric properties of the Otis-Lennon Mental Ability Test estimated according to classical measurement and Rasch model. (Unpublished Master's thesis). Umm Al-Qura University, Mecca, Saudi Arabia