MANSOURA UNIVERSITY
FACULTY OF LELTRES
–

CONSTRUCTING A DIGITAL INTELLIGENCE SCALE (DIS) BASED ON ITEM RESPONSE THEORY (IRT) FOR ISRA UNIVERSITY STUDENTS

BY

Sami Salameh Almassarweh
Department of Psychology, Faculty of Arts, Isra University, Jordan

Hani S.A. Alkaldi
Department of Classroom Teacher, Faculty of Educational Sciences, Isra University, Jordan

Mohammad A. Abumaal
Department of Psychology, Faculty of Arts, Isra University, Jordan

Safia M. Jabali
Department of Child Education, Faculty of Educational Sciences, Isra University, Jordan

Huasm A. Alsarhan
Department of Psychology, Faculty of Arts, Isra University, Jordan

Journal of The Faculty of Arts- Mansoura University

77th ISSUE- OUG. 2025

CONSTRUCTING A DIGITAL INTELLIGENCE SCALE (DIS) BASED ON ITEM RESPONSE THEORY (IRT) FOR ISRA UNIVERSITY STUDENTS

Sami Salameh Almassarweh
Department of Psychology, Faculty of Arts, Isra University, Jordan

Hani S.A. Alkaldi
Department of Classroom Teacher, Faculty of Educational Sciences, Isra University, Jordan

Mohammad A. Abumaal
Department of Psychology, Faculty of Arts, Isra University, Jordan

Safia M. Jabali
Department of Child Education, Faculty of Educational Sciences, Isra University, Jordan

Huasm A. Alsarhan
Department of Psychology, Faculty of Arts, Isra University, Jordan

Abstract

This study develops a Digital Intelligence Scale (DIS) based on Item Response Theory (IRT), specifically employing Rasch's model, to assess the digital competencies of students at Isra University. Digital intelligence, which includes such skills as digital literacy, critical thinking, and responsible online behavior, is the core of the education system in this digital age. To meet the research goals, the study adopts a descriptive-analytical methodology. A simple sample of 347 students, both male and female, from various faculties at Isra University was randomly selected to ensure a representative cross-section of the institution. A 31-item scale was designed, evaluated, and validated to measure digital intelligence, with a focus on psychometric properties like unidimensionality. The findings indicate that the scale aligns with the assumptions of Rasch's model, confirming unidimensionality and a strong fit between the data and the model. The results show that participants' responses and scale items align closely with the model's expectations, with mean scores near zero and standard deviations approximating one. Additionally, the threshold values of the scale items demonstrate significant discriminatory power, effectively distinguishing different levels of digital intelligence among the students. This validated scale may also be applied beyond Isra University and provides a tool for the measurement of digital intelligence at higher education. Future research could further examine the scale's adaptability across diverse educational contexts and its role in supporting students' digital readiness and lifelong learning.

Keywords: Digital intelligence, item response theory, measurement, one-parameter model

ملخص البحث:

هدفت الدراسة لبناء مقياس في الذكاء الرقمي لدى عينة من طلبة جامعة الإسراء، وفق نظرية الاستجابة للفقرة (Item Response Theory). ولتحقيق هدف الدراسة تم صياغة (٣٩) فقرة بشكل أولي وبعد عرضها على مجموعة من المحكمين من ذوي الخبرة والاختصاص أصبح المقياس مكون من (٣١) فقرة. وقد تم تطبيق المقياس على عينة مؤلفة من (٣٤٧) طالباً وطالبة من طلبة جامعة الإسراء، وبعد إجراء التحليل- وفق نظرية استجابة الفقرة (IRT)- تكون المقياس من (٣١) فقرة بصورته النهائية. حيث بلغت قيمة معامل الثبات للمقياس (٠.٩١)، كما تمتع المقياس بدلالات متعددة للصدق. وتوصلت الدراسة لعدة نتائج منها: إن مقياس الذكاء الرقمي كان أحدي البعد، وكانت مطابقة أفراد عينة الدراسة وفقرات المقياس مطابقة للنموذج حيث كان متوسطها الحسابي قريب من الصفر وانحرافها المعياري قريب من الواحد الصحيح، وكان تقدير قيم العتبات المميزة لفقرات المقياس قدرة تمييزية واضحة.

الكلمات المفتاحية: الذكاء الرقمي، نموذج احادي البعد، النظرية الحديثة في القياس.

Introduction:

Digital intelligence (DI) is a novel term that refers to the ability to understand and process digital information efficiently and deals with and analyzes data and information using technology effectively and smoothly. Digital intelligence includes logistical thinking and the capability to solve problems facing students using digital tools and technologies that have become widespread in the age of technology. Therefore, digital intelligence has become a basic skill that all students must acquire to help them interact with the evolving digital world to achieve success and innovation in various fields of life (Altarawneh & Al-Ghammaz, 2023).

Developments in technology and digitization now require universities to use these modern technologies and programs and create educational systems compatible with the requirements of the digitization era. Penprase (2018) asserts the need to completely reconsider the vocabulary of academic curricula to enable students to understand and become aware of smart technologies and to diligently prepare students for this matter, which has become a requirement in various areas of life.

Therefore, if universities want their students to excel and develop their study skills to achieve high-quality outcomes, it is essential for students to master self-learning skills. Thus, the students can develop themselves, their personalities, skills, cultural and digital abilities, and self-reliance, along with the ability to make decisions and take responsibility (Al-Faleet, 2015; Yildizay & Leman, 2015). The modern era, described as the digital age and the era of the Fourth Industrial Revolution, have constituted a major and qualitative leap in the production and sharing of knowledge through digital tools, programs, and smart applications such as cloud computing (CC), the Internet of Things (IoT), and artificial intelligence (AI), which have allowed students to benefit from information and data. Therefore, universities must take advantage of these tools to achieve their goals and improve their educational outcomes (Sima et al., 2020, p.17).

Many studies, such as Solovieva et al. (2020), have shown that students have a low level of digital intelligence. Therefore, tools are needed to constantly measure students' digital intelligence to develop appropriate plans to address this weakness and modify study plans in the digital age. It is also necessary to build measures that have good statistical properties and can be relied upon when measuring this characteristic among students. In this regard, most studies have adopted the classical test theory or CTT (Spearman, 1904; Novick, 1966) in measurement, analysis, and verification of psychometric properties. This article, however, adopts the item response theory or IRT (Lord, 1952; Lord & Novick, 1968; Rasch, 1960), which aims to confront the weaknesses and shortcomings of the CTT and provides more reliable methods for estimating the degree of possession of the trait being measured and the features of the items such as their difficulty, distinctiveness, and the probability of the answer through guessing (Gruijter & Kamp, 2005).

Research Problem

Despite the abundant educational literature describing many different methods and procedures for constructing IRT-based scales, it has not received much attention at the level of Arab and Jordanian studies and research. The researchers indicate that the majority of scales used in Arab studies are developed based on the concepts of the CTT. Importantly, most of the education-based studies conducted on the concept of digital intelligence have not addressed the development of a scale that measures this trait among university students. This gap in the literature highlights a need for psychometrically sound instruments tailored to this population within the Arab and Jordanian academic context.

Research Significance

Researchers and scholars believe in the importance of digital intelligence because it is a basic and important requirement to help students engage in the digital world and use technology and digital techniques to help them possess learning skills and move largely towards self-learning. As a result, they are in urgent need to use these technologies to adapt to the digital age and process the huge amount of information available through digital platforms. Therefore, it is necessary to have tools and standards for digital intelligence based on theories that help universities constantly measure digital intelligence to develop appropriate plans to address this imbalance.

Therefore, given the scarcity of scales in this area and the urgent need for such scales, this study stands out as the first, within the limits of the researchers' knowledge, to develop a Digital Intelligence Scale (DIS) tailored for students at Isra University in Jordan, offering an objective tool for measurement on the subject of digital intelligence according to the Rasch model of the IRT. With this in mind, four research questions are formulated as follows:

RQ1: What are the unidimensional implications of the items of the digital intelligence scale?

RQ2: What are the implications of conformity of the empirical data of the digital intelligence scale to the one-dimensional partial estimate model?

RQ3: What are the implications of the validity and reliability of the digital intelligence scale based on item response theory?

RQ4: What are the estimates of the distinct threshold values for the item grading levels of the digital intelligence scale?

Research Limitations

The conclusions drawn from this study could be applied more widely, with consideration of the following limitations:

1. Human Limitations: This research paper is limited to a sample of students from Isra University.
2. Spatial Limitations: This research paper is conducted at Isra University in Amman, Jordan.
3. Temporal Limitations: This research paper is conducted in the second semester of the academic year 2023/2024.
4. Objective Limitations: This research paper is limited to one form of intelligence, which is digital intelligence, by applying the assumptions of the rating scale model based on item response theory. The research study is also limited to three available statistical programs specialized in developing scales based on the IRT, namely “SPSS & BIGSTEPS & Bilog-MG3”, respectively.

The structure of this paper comprises seven sections. The Introduction (this section) sets the stage by outlining the importance of digital intelligence (DI) and the need for a digital intelligence scale (DIS) based on Item Response Theory (IRT). The conceptual framework is presented in (2). The third section discusses existing studies on digital intelligence and how they have employed IRT models to construct a measurement scale. After that, the methodology of the current study is presented in (4). Following that, the results and discussion of study are discussed in (5). Finally, Sections 6 and 7 conclude with recommendations for future research investigations.

Conceptual Framework

The IRT theory, also known as Latent Trait Models (LTM), postulates that the performance of respondents can be predicted or explained in psychological tests and measures in light of a distinctive characteristic of this performance called “trait”. With the difficulty of observing this trait directly, it must be inferred or estimated from respondents’ observable performance on a set of scales or test items (Hambleton, Swaminthan, & Rogers, 1991). Various models of the theory have been developed, with the goal to determine the relationships between the respondent’s performance in the test and the trait that underlies and explains this performance.

The IRT

The IRT is based on a set of key assumptions, which are the assumption of unidimensionality, the assumption of local independence, the assumption of conformity to the item characteristic curve (ICC), and the assumption of freedom from speediness (Gruijter & Kamp, 2005). These four assumptions are explained as follows:

1. Unidimensionality

Dimension as a term refers to the number of latent traits responsible for respondents' performance on test items. The trait is a concept used to describe respondents and is a group of interconnected and overlapping behaviors that may occur together, as the trait is a theoretical, intangible concept. Therefore, defining the trait is one of the basic steps in behavioral measurement (Elliot, 1983). A measure or test is unidimensional when its items are homogeneous and measure essentially the same trait. This means that any of these items, which are graded in difficulty, require the same behavioral procedures and processes to be solved, but they differ from each other only in terms of the gradation of their difficulty (Kazim, 1996).

The assumption of unidimensionality is desirable for all test developers to improve and enhance the interpretation of test scores. However, this assumption is violated in education-based research, as many factors affect the respondent's performance, such as personality, and factors related to test application, such as the level of motivation, test anxiety, the ability to work quickly, and knowledge of the correct use of answer sheets (Hambleton & Swaminthan, 1985).

2. Local Independence

Local independence means that the respondent's response to the test items is statistically independent when the respondent's level is taken into account. That is, the respondent's response to one item should not affect another item, meaning that the estimate of the difficulty of any item does not depend on estimates of the difficulty of the rest of the items and does not depend on the ability of the respondents who answer it. Estimating the difficulty of any item also does not depend on estimating the ability of any other group of respondents to whom the test is administered, nor does it depend on the difficulty values of the items to which they answer. This shows that the ability of the respondent and the characteristics of the item are the only ones that affect the answer to the item. Therefore, obtaining any sequence of scores for a set of items is the result of multiplying the probabilities for all of these items (Allen & Yen, 1979).

3. Item Characteristic Curve (ICC)

All classical models and response models for test items agree on the assumption of the existence of a trait continuum, as it is possible to estimate the probability of a respondent answering a test item correctly if we know his or her position on this continuum. Psychometric models, whether classical or modern, rely on the assumption that the respondent's position on a latent trait continuum is an indicator of the probability that his or her answer will be a correct answer to any item in a test that measures this trait. However, they differ in how this position is determined and how it relates to the probability of the correct answer to the item. The item characteristic curve is one of the basic concepts in the IRT, as it represents the probability of the respondent answering a correct answer to an item as a function of ability (θ), and the probability of the respondent answering the item increases with the increase in the ability of the respondent (Croker & Algina, 1986; Fan, 1998).

4. Speediness

The implicit assumption of all users of item response models is that tests of model fit are not applied under the condition of response speed. The respondent who fails to answer the test items is due to his or her limited ability and not because of his or her failure in not reaching the test items. For example, when speed affects performance on the test, there are at least two traits measured in this test: speed of performance and the trait measured by the content of the test (Hambleton & Swaminthan, 1985).

The IRT is characterized by the fact that it enables us to (i) estimate the ability of the respondent so that it is statistically independent of the sample of specific vocabulary applied to the respondents as long as it is appropriate vocabulary and provided that all the vocabulary is included together on the same continuum so that it measures the same trait, (ii) obtain indicators of the item such as difficulty, discrimination, and guessing that are independent of the particular sample of respondents used to adjust the item as long as it is an appropriate sample, (iii) obtain a statistical coefficient that shows the adequacy of the estimate of each respondent's ability using the test items, as this coefficient can vary from one respondent to another, and (iv) replace the concept of reliability through equivalence coefficient with the concept of the statistical estimate and associated standard errors (Hambleton & Swaminthan, 1985).

A group of models, or LTMs, have emerged from this theory, which were used in developing scales and tests to obtain statistical indicators for the item that depends neither on the characteristics of the respondents and their estimates nor the difficulty of the items of the scale or test (Croker & Algina, 1986). Hence, the significance of invariance is given in estimating the features of items between different groups of respondents, as this significance is described as the most important characteristic in latent trait theory (Lord, 1980).

The Rasch Model

The One-Parameter Logistic Model (OPLM or 1PL), or what is called the Rasch Model, is considered one of the most widely applied models in IRT. A key advantage of this model is that when the data fits the model, the item parameters—such as their level of difficulty—can be estimated independently of the specific sample used. Also, respondents' abilities can be estimated independently of the items' level of difficulty (Hambleton, Swaminthan, & Rogers, 1991).

Several variations of the Rasch model have recently been developed, each designed to accommodate specific types of data, such as the Dichotomous Model, the Partial Credit Model, and the Rating Scale Model, which is applied to data derived from the rating scales developed by the scientist Andrich (Alastair & Hutchinson, 1987). The rating scale model is based on the concept that each item reflects a psychological or emotional response from the respondent, which is captured through their evaluation of that item. The model estimates this psychological or emotional response for each item based on the probability function it uses, as the only parameter that the rating scale model deals with is the difficulty parameter bi.

Over the past years, a group of LTMs have been developed, as these models differ in selecting the mathematical model for the item characteristic curve, or ICC. The most common of these models are logistic models and models that use the Normal Ogive Models. Each of these two types uses one parameter representing the degree of difficulty of the item, or two parameters, one representing the degree of difficulty and the other the degree of discrimination. Latent trait theory often uses the 1PL or Rasch model because it is distinguished by the fact that when the data aligns with this model, the item parameters, including their difficulty levels, can be estimated independently of the specific sample used. Also, the abilities of respondents can be estimated independently of the items' level of difficulty. In the Rasch model for dichotomous response items, the sum of the points obtained by the respondent on these items is statistically sufficient to estimate the ability of the respondent, and the sum of the respondents who answer the item correctly is statistically sufficient to estimate the degree of its difficulty (Masters, 1982).

The models developed from the Rasch model can be classified according to the scoring of the scale items into models for dichotomous scoring, which are used when the answer to the item is either zero or one. Another classification is Andrich's rating scale model, which is used when data are taken from a rating scale. Other classifications include the Partial Credit models, which are used when answering the item involves several steps, and Binomial Trial models, which require several independent attempts to answer the item (Wright & Masters, 1982). In the current research study, the dichotomous model is used in analyzing the data due to its suitability for answering the test items under research study.

Literature Review

In recent years, the development of reliable and valid assessment tools has become increasingly crucial in various fields of study. To provide context and support for the current study's approach, the following literature review highlights several key studies that have successfully utilized the Rasch model to develop psychometric scales. These examples illustrate the model's effectiveness and relevance, laying the groundwork for understanding its application in assessing digital intelligence.

A recent study by Saleh and AlAli (2024) explored 'psycho-computing', which looks at how humans interact with computers to understand and improve their use of computer applications. They developed a psycho-computing traits scale to measure the behaviors of computer users and computer science students. To ensure the validity and reliability of the scale, they first tested it on a group of 150 students. They then tested the final version on a larger group of 722 students to confirm its psychometric properties. The results showed that the scale is both valid and reliable, as supported by the Rasch model and confirmatory factor analysis. This means the scale is a trustworthy tool for assessing the psycho-computing traits of computer users and students.

In another study by Ayoub (2023), the Rasch model was applied to evaluate the psychometric properties of Hamdan Giftedness Test. This test consists of five cognitive ability assessments: (i) Verbal Ability I (Semantic Information Processing), Preknowledge in Science (Conceptual Information Processing), Verbal Ability II (Comprehension), These tests were designed for students in grades 4-12, divided into four age groups (grade 4-6, grade 7-8, grade 9-10, grade 11-12). The study involved 12,071 students (4,009 males and 8,062 females) from the seven governorates of the UAE. The psychometric analysis checked the test's validity and reliability, item calibration, and differential item functioning (DIF). It also used standardized residual variance and the largest standardized residual correlations to detect any dependent items. Additionally, the Rasch model was used to estimate difficulty coefficients using different methods to address missing values. The results included a percentiles/standards table that displayed raw scores and their corresponding estimated ability values according to the Rasch model. Finally, cut-off scores were established to identify gifted students.

Abo Al-ssel's (2022) study aimed to develop an achievement motivation scale for high school students in Damascus based on the IRT. The scale comprised 33 items. To ensure its content validity, the scale was presented to a panel of expert evaluators with relevant experience and specialization, and it was subsequently administered to a pilot sample of 400 high school students in Damascus. In addition, Rabi' Rashwan's scale was selected as the criterion measure, and the concurrent validity of the new scale was confirmed by calculating the correlation coefficient between the two measures, which was 0.85—indicating a strong association and supporting the criterion's validity.

Subsequently, the scale was administered to the main study sample of 1,200 high school students in Damascus. The results indicated that the achievement motivation scale developed by the researcher was unidimensional. The empirical data derived from the achievement motivation scale for the unidimensional partial credit model within IRT showed that the mean fit indices for both persons and for the *outfit* and *infit* statistics approached zero, while the standard deviation approximated one. Moreover, the estimated threshold parameters for the scale items demonstrated clear discriminative capacity and revealed distinct threshold levels. Finally, the findings related to the performance of Damascus high school students—as indicated by their transformed ability estimates in logits—revealed values on this scale that differed from those obtained with Rabi' Rashwan's achievement motivation measure.

The study by Al-Hawari and Al-Faqi (2021) aimed to determine the level of digital intelligence among a sample of faculty members and their assistants at Al-Azhar University, as well as to examine the relationships between digital intelligence, cognitive flexibility, and orientation toward a productive university among the same group. The researchers employed a descriptive-analytical approach, and the primary sample consisted of 267 faculty members and assistants during the second semester of the 2020/2021 academic year. They administered three self-developed scales: the Digital Intelligence Scale, the Cognitive Flexibility Scale, and the Productive University Orientation Scale.

The results indicated that the level of digital intelligence among these faculty members and assistants was low. Moreover, there was a statistically significant positive correlation between digital intelligence and cognitive flexibility, as well as between digital intelligence and orientation toward a productive university. Statistically significant differences were also observed in the mean scores on the Digital Intelligence Scale (both total score and dimensions) based on gender (male–female), with males scoring higher. However, no statistically significant differences were found based on college type—specifically, between humanities (theoretical) and scientific (practical) colleges. In addition, statistically significant differences emerged according to academic rank (teaching assistant, instructor, assistant professor, and professor). Finally, digital intelligence could be predicted from the scores on both the Cognitive Flexibility Scale and the Productive University Orientation Scale.

Using different study samples, Alkhaldi et al. (2021) developed a scale for self-confidence using the modern theory of measurement among secondary school students. The descriptive method is used to achieve the study objectives. The scale is applied to a random study sample of (1060) male and female students in the Mafraq Governorate in Jordan. The number of scale items in its final form is (39) after its validity is checked by the validators' validity, internal consistency, and the validity of the two different groups. The scale's items indicate that its logistical ability is between (-2.88) and (2.77), which indicates the possibility of applying it to students who have minimum and highest abilities.

Upon reviewing these studies, it is evident that while some have touched on parts of the current research, a comprehensive analysis of digital intelligence within the framework of modern measurement theory remains relatively scarce. Nevertheless, there are similarities with earlier research regarding some variables. Specifically, Alkhaldi et al. (2021), Abo Al-ssel (2022), Ayoub (2023), and Saleh and AlAli (2024) employed modern measurement theory, aligning with the methodological approach of the present study. Moreover, Al-Hawari and Al-Faqi (2021) investigated the variable of digital intelligence.

Regarding methodological design, this study employs a descriptive-analytical approach, which is consistent with the approaches adopted by Al-Hawari and Al-Faqi (2021) and Alkhaldi et al. (2021). However, the current study distinguishes itself in several key aspects. Firstly, it focuses specifically on digital intelligence among university students, a population that has not previously been targeted in this context. While Al-Hawari and Al-Faqi (2021) explored digital intelligence, they used faculty members and their assistants at Al-Azhar University as the target, who form a related but distinct population. Secondly, our study is distinct from Abo Al-ssel (2022), who employed an experimental design. Finally, the development of a researcher-designed scale distinguishes our approach from Ayoub (2023) that applied the Rasch model to analyze the psychometric properties of the Hamdan Giftedness Test.

Methodology

In this study, the descriptive-analytical research approach is adopted to achieve the research objectives. Additionally, the Rasch's model is used to construct a scale that measures the digital intelligence trait. This model is chosen because it is considered as a solid framework for selecting test items and standards, both now and in the future (Anastasi, 1982). The Rasch's model within the IRT provides reliable statistical indicators for the test and its items, which remain consistent regardless of the respondents sampled. This ensures a more dependable measurement with fewer errors (Nitko, 2001).

Research Population

The research population consists of (5324) students at Isra University, according to the statistics of the university's admission and registration department for the second semester of the academic year (2023/2024).

Research Sample

The research sample consists of a simple sample of (347) male and female students randomly selected from Isra University, with ages ranging from 18 to 24 years. Gender representation was nearly balanced, with 171 female students and 176 male students. Table 1 below provides a detailed breakdown of age and gender distribution within each faculty.

Table 1: Demographic Profile of University Student Participants

Faculty	Total Students	Age Range	Gender	
			Female	Male
Faculty of Arts	93	18-22	46	47
Faculty of Educational Sciences	87	18-22	42	45
Faculty of Engineering	84	18-24	43	41
Faculty of Pharmacy	83	18-23	40	43
Total	**347**		**171**	**176**

Research Procedures to Develop a Scale based on IRT

For the current research study, the digital intelligence scale is developed based on the IRT following the steps outlined by Hulin et al. (1983), as detailed below.

a. Review of previous education-based research related to digital intelligence

The process of measuring digital intelligence is difficult because it is inherently assuming characteristics. Therefore, there has been an urgent need to find scales that have good statistical properties for digital intelligence. In intelligence scales in general, the primary concern is reflected in collecting a comprehensive set of statements that capture all potential viewpoints related to the concept being studied. Each statement is designed to be clear, concise, and centered on assessing the student's skills rather than their knowledge of specific information (Thorndike & Hagin, 1989).

Researchers have drawn on existing literature to address this research problem, including the scale prepared by Sweilem, Ibrahim, & Al-Hadiri (2023), where they designed the scale using item response theory. The study by Al-Hawari and Al-Feki (2021) is also of high significance as it addressed the relationship of digital intelligence with cognitive flexibility and the trend toward a productive university. The current research study has also benefited from Ibrahim's study (2022), examining the analytical thinking and habits of mind as predictors of digital intelligence, and Al-Qahtani's study (2023), addressing the availability of digital competence for artificial intelligence among university students. In addition, some related books dealing with the concept of digital intelligence, such as Al-Daraissa's book (2021) are also reviewed.

b. Procedural formulation of the scale items

The suggestions from Al-Nabhan (2004) and Hendcson et al. (1987) for developing scales have been considered. These include writing scale items that clearly measure the intended trait, carefully reviewing keywords, especially nouns and adjectives, ensuring the scale is applied in the respondents' natural environment by the person responsible for the measurement, and conducting an exploratory study to test the scale items before final implementation. When drafting the statements, several key points need to be considered: avoiding phrases that refer to the past, steering clear of statements that could be interpreted as facts, ensuring clarity by avoiding ambiguous language, using simple and understandable words, focusing each statement on a single clear idea, and avoiding blanket words or generalizations such as always, never, each, and never, and being careful when using words such as only and merely.

Considering the steps taken earlier, a scale of (39) five-point Likert-type items is formulated, which is considered one of the most widely used scales due to its simplicity in design, application, and scoring. Using this scale, each response from the respondent is assigned a score ranging from 1 to 5. For the five response options, a score of 1 is given for a "very low degree", 2 for a "low degree", 3 for a "medium degree", 4 for a "high degree", and 5 for a "very large degree" for positive statements. For negative statements, the scoring is reversed, with the values adjusted accordingly to reflect the opposite direction.

Scale application

1. Testing the scale on an exploratory sample

A 39-item scale was reviewed by 12 experts with Ph.D. degrees in educational psychology, measurement and evaluation, guidance, and curricula. They assessed the clarity and wording of the items, evaluated how well each item fits within the scale, checked its alignment with the criteria for item formulation, and recommended any changes that could improve the coverage of the digital intelligence scale. The feedback, revisions, and suggestions provided by the validators are carefully considered, with items receiving an approval rating of 89% being selected. As a result, 6 items were either deleted, revised, or reworded. Consequently, the final version of the questionnaire consists of (31) instead of (39) items (see Table (1) in the Appendix).

After the scale was initially developed with 39 items, it was tested on an exploratory sample of 45 male and female students who were not part of the main research sample. Following the pilot study, 8 items were removed from the scale. The final version of the scale contains 31 items, as presented in Table (1) in the Appendix. The reliability coefficient is calculated using Cronbach's Alpha formula, and its value is (0.83), which indicates the reliability of the scale for its final application.

2. Applying the scale to the research sample

The scale was administered to 347 male and female bachelor's students at Isra University, randomly selected from the research population during the second semester of the 2023/2024 academic year. When using the 1PL (Rasch) model in IRT theory, at least 200 respondents are needed per group (Crocker & Algina, 1986).

Results and discussion

Results related to RQ1

'What are the unidimensional implications of the items of the digital intelligence scale?'

Testing the assumption of unidimensionality

The IRT models are based on the assumption that a single ability drives the respondent's performance on the scale, which is why they are referred to as unidimensional models. To verify this assumption, a factor analysis is performed on the scale data collected from 347 male and female students. The results of the factor analysis are attained using Principal Components Analysis (PCA) and rotation according to orthogonal rotation of varimax. The results of the analysis produced (7) factors, with the first factor explaining 15.817% of the variance, and the total of all factors accounting for 52.736% of the overall variance. Table (2) illustrates the latent root values "Eigenvalues", the percentage of explained variance for each factor, along with the cumulative percentage of explained variance.

Table 2: *Results of factor analysis of the response of male and female students to the digital intelligence scale.*

Factor No.	Eigenvalue	Percentage of Variance Explained%	Cumulative Explained Variance%
1	9.243	29.817	29.817
2	2.528	8.154	37.971
3	1.937	6.249	44.220
4	1.311	4.229	48.449
5	1.210	3.905	52.353
6	1.151	3.714	56.067
7	1.122	3.619	59.686

As illustrated in Table (2), the latent root value of the first factor is 9.243, and explains 29.817% of the total variance, which is a high value when compared with the latent root values of the rest of the factors. However, the latent root value of the second factor is 2.528 and explains 8.154% of the total variance, which is more than three times the amount explained by the second factor. Furthermore, it is observed that the percentage of variance explained by each of the other factors is quite similar. In other words, the proportions of variance explained by all factors, except the first, show a high degree of consistency. This indicates that the assumption of the scale's unidimensionality has been met, meaning that the scale effectively measures a single trait (Hulin et al., 1983; Hattie, 1985). The assumption of unidimensionality of the scale developed in this study is supported by visually representing the latent roots through the Scree plot method, as shown in Figure (1).

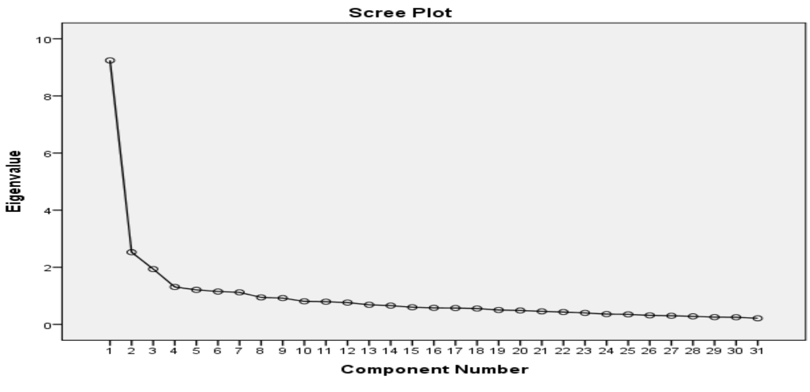

Figure 1: *The relationship between the factors attained from factor analysis and the amount of their latent roots taking the letter-l shape and indicating the unidimensionality of the digital intelligence scale*

Results related to RQ2

'What are the implications of conformity of the empirical data of the digital intelligence scale to the one-dimensional partial estimate model?'

To check the assumption of goodness-of-fit-test for the sample members' responses to the scale according to the rating scale model, the data are analyzed using the statistical software

BIGSTEPS.

On the subject of indicators for Persons-Fit, finding out the indicators for Persons-Fit for the research sample is conducted by estimating the ability of each student and the standard error in measuring this ability. An internal fit statistic, "INFIT", and an outfit statistic, "OUTFIT", are also calculated for each estimate. These statistics are more sensitive to unexpected behaviors from respondents than to items that deviate from their ability level. Table (3) displays the mean and standard deviation for each respondent's ability estimate, the measurement error (MODEL ERROR) associated with these abilities, the mean squares for infit and outfit statistics "Mean Square Infit & Outfit Statistic, MNSQ), and the values of the Standardized Information Weighted Fit Statistics for Persons (ZSTD).

Table 3: *Indicator of the Infit and Outfit statistics for the respondents and inter-respondent reliability of the digital intelligence scale*

Scale	**Number of Fit Respondents**	**(INFIT Statistic)**		**(OUTFIT Statistic)**		**Inter-respondent Reliability**
		(MNSQ)	(ZSTD)	(MNSQ)	(ZSTD)	
Digital Intelligence	347	1.01	-0.12	1.12	-0.25	0.87

As shown in Table (3), the mean values of the infit and outfit respondents' conformity index are close to zero, and their standard deviation is close to one, which means that the respondents in the research sample conform to the model.

Table 4: *Indicator of the Infit and Outfit statistics for the items and inter-item reliability of the digital intelligence scale*

Scale	**Number of Scale Items**	**(INFIT Statistic)**		**(OUTFIT Statistic)**		**Inter-item Reliability**
		(MNSQ)	**(ZSTD)**	**(MNSQ)**	**(ZSTD)**	
Digital Intelligence	347	1.01	-0.18	1.01	0.20	0.99

As shown in Table (4), the mean values of the infit and outfit items' conformity index are close to zero, and their standard deviation is close to one, which means that the digital intelligence scale items in the research sample conform to the model.

Results related to RQ3

'What are the implications of the validity and reliability of the digital intelligence scale based on item response theory?'

Face validity is used to check the digital intelligence scale validity by reviewing the 31-item scale in its final forms from, as some items are modified and some items are deleted. Validity is checked through internal consistency between the items by calculating the Pearson correlation coefficient between each item and the total number of items, where the values ranged between (0.37) and (0.81). Accordingly, all items are found statistically significant, indicating the validity of the items in the scale.

The validity of the scale is also ensured factorially by calculating the factor analysis of the scale scores, where seven factors are found, and the latent root value of the factors is (9.243, 2.528, 1.937, 1.311, 1.210, 1.151, 1.122), respectively which explained approximately (60%) of explained variance. The internal consistency reliability coefficient for the scale's items in its final form, consisting of (31) items, is also calculated using Cronbach's Alpha formula, which is (0.91), indicating that the scale has an excellent degree of reliability. Another way to assess the reliability coefficient of the scale, based on modern measurement theory, is through the information function. This function is derived by combining the individual item curves into a single curve, following the relationship outlined below:

$$I(\theta) = \sum_{g} Ig(\theta)$$

In this equation, I(θ) represents the amount of information provided by the scale at the ability level θ, and ∑ Ig(θ) is the sum of the information functions for the individual items at that ability level. As the number of items increases, the standard error (S.E.) of θ decreases, which in turn enhances the scale's ability to provide more precise information about θ, as shown in the following relationship:

$$I(\theta) = \frac{1}{\sqrt{S.E(\theta)}}$$

Importantly, what distinguishes the item response theory from the traditional theory of measurement is that the estimate of reliability in the traditional theory is linked to the sample. Item response theory, therefore, allows us to estimate the standard error of measurement at each ability level and shows how each item contributes to the overall effectiveness of the measurement. To carry out this analysis, statistical software (Bilog-MG3) is used to calculate the amount of information provided by the scale at each ability level and to plot the relationship between the values of the information function and the standard error of the estimate for the scale items.

On the other hand, the results show that the test provides the highest amount of information at a value of 8.523, while the maximum information value for each item is 0.141. Additionally, as the standard error of the estimate decreases, the amount of information provided by the scale increases, which aligns with the expectations set by the rating scale model in item response theory. The reliability of the scale is also checked using statistical software (BIGSTEPS), which provides two different estimates of the reliability coefficient: the reliability of the scale and the reliability of respondents, where the value of the Item Separation Index is (7.12). However, the value of the Person Separation Index is 3.11, and both values are more than 2, indicating that the reliability coefficient of the scale is (0.91), and the reliability coefficient of respondents is (0.81) (Wright & Masters, 1982). These high values clearly demonstrate that the scale items are effective in differentiating between respondents and their varying ability levels on the one hand, and the adequacy of the sample of respondents in separating between the scale items and identifying the attribute continuum that the items measure on the other.

Results related to RQ4

'What are the estimates of the distinct threshold values for the item grading levels of the digital intelligence scale?'

To address this question, the WINSTEPS statistical software is used to calculate the distinct threshold values for the grading levels of the Digital Intelligence Scale items. Table (٥) illustrates these results.

Table 5: *Distinctive threshold values for the levels of items in the digital intelligence scale*

Levels / Scale	1	2	3	4	5
Digital Intelligence Scale	-	-1.33	-0.39	0.37	1.38

As illustrated in Table (5), the distinct threshold values for the levels of the digital intelligence scale items reflect the distinctiveness of the research sample's performance on the scale.

Discussion

Regarding unidimensional implications and, the results indicate that the digital intelligence scale, in which saturation of its items was revealed on the factors calculated by factor analysis, measured one dimension. The value of the latent root of the first factor is equal to (9.243) and explained approximately (30%) of the explained variance, which is a high value compared to the rest of the latent roots, as the value of the latent root of the second factor is (2.528) and explained approximately (8%) of the total variance. In other words, the first factor explained approximately four times more than the second factor. Moreover, when the latent root of the first factor was divided by the second, its value was approximately (3.7), that is, greater than (2), which confirms that the digital intelligence scale measures only one dimension, and this is an indication that the assumption of unidimensionality has been achieved. This was reinforced through the Scree Plot by representing the latent roots graphically.

About the implications of matching the empirical data for the digital intelligence scale to the one-dimensional partial estimate model, a very large percentage of respondents matched the model used in estimating the features. All 31 items of the scale are also matched through infit and outfit indicators for the items and respondents in the research sample, and the inter-respondent reliability for respondents and items is high, indicating that the digital intelligence scale has excellent characteristics in terms of matching empirical data. Regarding the reliability of the scale, the inter-item reliability is almost one, and this indicates that the scale has good characteristics that enable it to be valid for field application in any future study related to digital intelligence. Therefore, this is evidence that the scale items match the research model used in this study.

Concerning validity and reliability, the reliability of the scale and the respondents are checked, as the value of the separation index between the items of the scale is 7.12, while the value of the separation index between the sample members is 3.11, and both values exceed (2). In other words, the reliability coefficient of the scale is (0.91) and the reliability coefficient of respondents is (0.81). These values indicate a high value of the reliability index, which indicates the adequacy of the items of the scale in separating between respondents, distinguishing between different levels of ability between respondents, and the adequacy of the sample of study respondents in separating between scale items. Face validity is assessed for the digital intelligence scale by having experienced and specialized faculty members from Jordanian universities review the initial 31-item scale to evaluate its relevance and appropriateness within the domain of digital intelligence, as some items are modified and some items are deleted. Validity is checked through internal consistency between the items by calculating the Pearson correlation coefficient between each item and the total number of items, and all items were found statistically significant, indicating the validity of the items in the scale. The validity of the scale is also ensured factorially by calculating the factor analysis of the scale scores, where seven factors are found. These seven factors explained approximately 60% of the explained variance, and thus the psychometric properties of the scale were verified.

On the topic of estimates of the values of the distinct thresholds for the grading levels of the items of the Digital Intelligence Scale, the results indicate that the reliability of the spacing of the items was close to one, and this is a result of the reliability of the spacing of the different thresholds for the items of the scale more than the reliability of the spacing of the items themselves. The items are similar, and the spacing is caused by the different thresholds and is achieved through the divergent values of the different thresholds. Therefore, it can be said that the differential threshold values were achieved for the grading levels of the items, which formed the items with a high discriminatory ability that can classify students with the presence of digital intelligence or the absence of digital intelligence among the students.

Conclusion

The current article constructed a digital intelligence scale (DIS) based on item response theory (IRT) for students at Isra University. The findings indicated that the digital intelligence scale is unidimensional model. The results also showed that research sample participants and the scale items were identical to the model, with their mean close to zero and their standard deviation close to one, as the distinct threshold values for the scale items have a clear discriminatory ability.

Recommendations

Based on the findings and discussion presented earlier, this study recommends applying the final form of the scale to identify the level of digital intelligence among university students because the scale has statistical properties such as indications of validity, reliability, and conformity of the data to the model used in this research study. Another key recommendation is to conduct a research study to predict digital intelligence using new concepts and variables such as academic achievement and learning motivation among Jordanian university students.

References

Abo Al-ssel, M. S. (2022). Binaa' Miqyās Dāfi' al-Injāz li-Ṭalabat al-Marhala al-Thānawiya fī Dimashq wa-fiqan li-Nazariyyat al-Istijāba lil-Faqra (IRT) [Constructing an Achievement Motivation Scale for Secondary Stage Students in Damascus City According to Item Response Theory (IRT)]. *Majallat Ittiḥād al-Jāmi'āt al-'Arabīyah wa-'Ilm al-Nafs* [Association of Arab Universities Journal for Education and Psychology], *14*(4), 194–228.

Alastair, P & Hutchinson, G. (1987). Calibrating graded assessment Rasch partial credit analysis of performance in writing. Language Testing, 4(1), 72-91.

Al-Daraissa, A. (2021). Digital education: Practical strategies and digital tools. Dar Al-Ayyam for Publishing and Distribution, Jordan.

Al-Faleet, J. (2015). Self-learning skills are necessary for postgraduate students in Palestinian universities in Gaza in light of the requirements of the knowledge society. Hebron University Research Journal. 10(2) p. 28-48.

Al-Hawari, J. F. Ismail, & Al-Faqi, M. M. Abdel-Razek. (2021). al-Dhakā' al-raqmi wa 'alāqatu-hū bi-l-murūnah al-ma'rifīyah wa-l-ittijāh naḥwa al-jāmi'ah al-muntağah lidā'at min a'ḍā' hay'at al-tadrīs wa-mu'āwinī-him bi-Jāmi'at al-Azhar [The digital intelligence and its relationship to cognitive flexibility and the attitude towards a productive university among a sample of faculty members and their assistants at Al-Azhar University (predictive difference study)]. Majallat al-Tarbiya, Jāmi'at al-Azhar [Al-Azhar Journal of Education], 40(192, Part 2), 1–64.

Alkhaldi, H., Alkhutaba, M., & Al-Dlalah, M. (2021). Building a self-confidence scale according to the item response theory for high school students in Jordan. Modern Applied Science, 15(3), 17-30.

Allen, M. & Yen, W. (1979). Introduction To measurement theory. California. Education, 20(4), 44-57. Brooks Cole Publishing Company Monterey.

Al-Nabhan, M. (2004). Fundamentals of measurement in behavioral sciences, first edition, Dar Al-Shorouk for Publishing and Distribution, Amman.

Al-Qahtani, K. (2023). The extent of the availability of digital competence for artificial intelligence among students of the Faculty of Education, University of Tabuk. Journal of the College of Education, 90(1), 483-539.

Altarawneh, M & Al-Ghammaz, S. (2023). The journey of e-learning technology from application to challenges: Evidence from Jordan. 11th International Conference on Information Technology: Cybersecurity Challenges for Sustainable Cities, ICIT-Proceeding, 531-536.

Anastasi, A. (1982). Psychological testing (6th ed.). New York: Macmillan.

Ayoub, A. E. A. (2023). Development of the Hamdan Giftedness Test: Psychometric Analysis Strategies According to Advanced Global Models (Rasch Model). Proceedings of the 2nd World Giftedness Center International Conference, Dubai, UAE, 16 - 19 Oct, 2023.

Cismaru, D., Gazzola, P., Ciochina, R., & Leovaridis, C. (2018). The rise of digital intelligence: Challenges for public relations education and practices. Kybernetes.

Croker, L. & Algina, J. (1986). Introduction to Classical and Modern Test Theory, New York G.B.S College publishing

Elliot, C. (1983). British ability scale manual, introductory handbook, Windsor. England: National Foundation for Education Research.

Fan, X. (1998). Item response theory and classical theory: an empirical comparison of their item/person statistics. Educational and Psychological Measurement, 58(3), 357 381

Gruijter, D & Kamp, L. (2005). Statistical Test Theory for Education and Psychology. www.leidenuniv.nl/-gruijterdnmde.

Hattie, J.(1985). Methodology review: assessing unidimensionality of tests and items. Applied Psychological Measurement, 9(2), 139-164

ambleton, R & Swaminthan, H. (1985). Item response theory: Principles & applications. Norwell, MA: Kluwer Nijhoff Publishing.

Hambleton, R., Swaminthan, H., & Rogers, J. (1991). Fundamental of Item Response Theory, Second Edition, Newbury Perk California.

Hendcson, M., Morris, I., & Fitz-Gibbon, C. (1987). How to Measure Attitudes, First Edition, Sage Publications, Newbury Park.

Hulin, C., Drasgow, F., & Parsons, C. (1983). Item Response Theory: Application to Psychological Measurement, First Edition, Dow Gone-Irwin.

Ibrahim, R. (2022). Analytical thinking and habits of mind as predictors of digital intelligence. Journal of the Faculty of Education, Assiut University, 38(12), 71-120.

Lord, F. M . (1980). Applications of item response theory to practical testing problems. Hillsdale, N.J: Lawrence Erlbaum.

Lord, F. M. (1952). A theory of test scores. Psychometrika Monograph No. 7.

Lord, F. M., & Novick, M. R. (1968). Statistical Theories of Mental Test Scores. Reading: Addison-Wesley.

Kazim, M. (1996). Using the Rasch model in constructing an achievement test in psychology and achieving objective interpretation of the results. Cairo: Anglo-Egyptian Library.

Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 2(1), 149-174.

Nitko, A. (2001). Educational assessment of student, (3rd ed). New Jersy: Merrill Prentice Hall.

Novick, M. R. (1966). The axioms and principal results of classical test theory. Journal of Mathematical Psychology, 3, 1-18

Penprase, B. (2018). The Fourth Industrial Revolution and Higher Education. In N. Gleason (Ed.), Higher Education in the Era of the Fourth Industrial Revolution, Palgrave Macmillan, 207-229.

Rasch, G. (1960). Probabilistic models for some intelligence and achievement tests. Copenhagen: Danish Institute for Educational Research; Chicago: MESA Press.

Saleh, S., & AlAli, R. (2024). Constructing and development of the psycho-computing traits scale: a psychometric analysis of computer users and students in the field of computer science. Cogent Social Sciences, 10(1).

Sima, V., Gheorghe, I. G., Subić, J., & Nancu, D. (2020). Influences of the Industry 4.0 revolution on the human capital development and consumer behavior: A systematic review. Sustainability, 12(10), 4035.

Solovieva, O., Palieva, N. A., Borozinets, N., Kozlovskaya, G., & Prilepko, J. (2020). Development of digital intelligence among participants of inclusive educational process. Journal of Educational Psychology-Propositos y Representaciones, 8(SPE2), 1-14.

Spearman, C. (1904). The proof and measurement of association between two things. American Journal of Psychology, 15, 72-101.

Thorndike, R & Hagin, E. (1989). Measurement and evaluation in psychology and education (translated by: Abdullah Zaid Al-Kilani and Abdul Rahman Adas) Amman, Jordanian Book Center.

Wright, B. & Master, G. (1982). Rating Scale Analysis. Rasch Measurement, First Edition, MESA Press, Chicago.

Yildizay, A. & Leman T. (2015). Development of the self-directed learning skills scale. International Journal of Lifelong Education, 34(6), 663- 679