Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 8, S. 430-438, 2, 2024

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 8, pp. 430-438, 2, 2024

Research Article

Comparison of Artificial Intelligence-Assisted Adaptive Testing Methods in Terms of Advantages and Disadvantages

Ahmet Hakan İNCE[*1], Serkan ÖZBAY [2]
[1] *Gaziantep University, Engineering Faculty, Electrical & Electronics Engineering, Gaziantep, TÜRKİYE*
[2] *Gaziantep University, Engineering Faculty, Electrical & Electronics Engineering, Gaziantep, TÜRKİYE*
Email of corresponding author: a.hakanince@gmail.com, sozbay@gantep.edu.tr

(Received: 11 March 2024, Accepted: 12 March 2024)

(4th International Conference on Innovative Academic Studies ICIAS 2024, March 12-13, 2024)

ATIF/REFERENCE: İnce, A. H. & Özbay, S. (2024). Comparison of Artificial Intelligence-Assisted Adaptive Testing Methods in Terms of Advantages and Disadvantages. *International Journal of Advanced Natural Sciences and Engineering Researches, 8(2), 430-438.*

Abstract – Adaptive tests aim to measure individuals' ability levels and knowledge levels on a subject in the shortest and most accurate way. Unlike classical tests, adaptive tests measure the participant's knowledge level in a more accurate and short way by creating the test according to the participant's ability level, instead of asking all the questions to each participant. In this text, adaptive test creation with Item Response Theory (IRT) and Knowledge Space Theory (KST)), which are important in the field of measurement, evaluation and teaching are discussed. The model structures of these theorems, their roles in measurement, evaluation and teaching processes, their purposes of use, and their differences with each other have been evaluated. In addition, one parameter logistic model (Rasch Model), two parameter logistic Model (2PL), 3 parameter logistic Model (3PL), which are the most well-known models of Item Response Theory, are discussed and the advantages and disadvantages of each model compared to the other model are explained. As a result of relevant research, it is shown that all models offer different approaches and make different contributions to the fields of measurement, evaluation and teaching.

Keywords – Item Response Theory, Adaptive Test, Traditional Test, Knowledge Space Theory, Rasch Model, 2PL Model, 3PL Model.

I. INTRODUCTION

Classical tests, also known as traditional tests, are standard tests in which the same questions are asked to each participant, regardless of their skill level and knowledge level. In this testing model, asking the same questions to each participant is not sufficient to measure the actual (real) performance of the participants. With the classical test model, it is not possible to find people with different knowledge levels and abilities in the same test. At this point, adaptive tests have been developed today to analyze the actual performance of the participants and accurately determine their skill levels. Unlike the classical test model, adaptive tests are prepared according to the knowledge and ability levels of the participants. Thus, each participant's actual skill levels and knowledge levels are measured more accurately by asking different questions. Additionally, adaptive testing allows participants to see their strengths and weaknesses at each stage of the test [1-3]. Adaptive tests are used not only to measure the knowledge level and ability level of the participants but also to prepare a learning process on a topic according to the knowledge level of each participant. In this text, the basic principles of adaptive tests created with Item Response Theory (IRT) and

Knowledge Space Theory (KST), general areas of their usage, learning processes and the effects of these methods on test efficiency are examined. In addition, the advantages and disadvantages of the one parameter logistic model (Rasch Model), two parameter logistic model and three parameter logistic model, which are the most frequently used methods of Item Response Theory in the literature, are mentioned. Basic algorithms of Item Response Theory and Knowledge Space Theory, their working principles and auxiliary algorithms that are generally used together in the literature are also discussed. By explaining the purposes of use, working principles and differences of the most common algorithms used in adaptive test preparation, it helps in choosing the algorithm in the adaptive test preparation process.

II. THE ADAPTIVE TEST METHODS

1. ITEM RESPONSE THEORY

Item Response Theory is a method that examines the statistical properties of test items, such as difficulty level and discrimination coefficient, and relates these properties to the ability levels of the participants. This theorem is one of the most widely used theorems in the process of determining participants' multiple choice test questions and their ability levels. IRT is based on the assumption that the probability of a test question being answered correctly by participants with a certain ability level can be related to the difficulty level of the question and the ability level of the participant. For this reason, IRT models use mathematical equations to define the effects of the questions on the test and to ensure a more accurate and reliable evaluation of the test. The most frequently used Item Response theories in the adaptive test preparation process include the One Parameter Logistic Model (Rasch model), Two Parameter Logistic Model (2PL) and Three Parameter Logistic Model (3PL). Each model creates certain advantages and disadvantages in the adaptive test preparation process, depending on the available data and the purpose of the test preparation [4-6]. IRT enables tests to measure participants' ability levels more accurately, reliably and sensitively. In addition, IRT is one of the best methods used to select test items, adjust the difficulty levels of the tests, and measure the abilities of the participants more precisely.

1.1 Item Response Theory Rasch Model (1PL)

One of the most commonly used methods in adaptive testing processes prepared using Item Response Theory is the "one parameter model (rasch model)". This model is based on calculating the probability of how a question will be answered by the participants using equations 1 and 2, with the difficulty level of the question in the test and the student's ability level of the participant as the only parameters. In this model, each question has a difficulty level and each participant has a certain skill level. By combining the skill level and the difficulty level of the question in the same equation, it predicts how you will answer the question and decides which question will be the next question. The difference between the participant's skill level and the difficulty parameter of the question is expressed by the logit function. Another advantage of the Rasch model is that test items and participants' ability levels can be estimated independently. In addition, since the Rasch model includes only the difficulty parameter of the question as a single parameter, it offers simpler, faster and more stable calculations in applications. Another advantage of containing a single parameter is that comparison between tests can be made more easily. As a result, the Rasch model's inclusion of a single parameter and consisting of simple equations has made it one of the most widely used IRT methods in multiple-choice adaptive test models [7]. Correctly determining the parameters that affect test performance is important in calculating the performance of participants taking the test. Equation (1) and Equation (2) represent the same equations and both are used in the literature [8].

The equation for the Rasch(1PL) model is as follows:

$$prob\ (x = 1 \mid \theta, b) = \frac{e^{(\theta - b)}}{1 + e^{(\theta - b)}} \quad (1)$$

$$prob\ (x = 1 \mid \theta, b) = \frac{1}{1 + e^{(b-\theta)}} \quad (2)$$

where θ is current ability of test taker, b is difficulty of the question. $prob\ (x = 1 \mid \theta, b)$ means that probability of correct answer that the student has theta (θ) ability and question has b difficulty level.

1.2 Item Response Theory Two Parameter Logistic Model (2PL)

In the two-parameter logistic model of Item Response Theory, the discrimination coefficient parameter is added in addition to the difficulty level of the questions and the ability levels of the participants. Adaptive testing is designed so that all parameters work in harmony and related to each other.

Question Difficulty Parameter (b): This parameter is generally indicated with the letter "b" in the literature and indicates the difficulty level of the test question. The difficulty level of the question represents the skill level of the participant taking the test at the point where the probability of answering the question correctly is 50% on the graph of the probability of answering the question correctly (3).

Question Discrimination Parameter (a): The question discrimination parameter expresses how effectively a question can discriminate between participants of different ability levels. This parameter is generally represented by the letter "a" in the literature. A higher discrimination parameter of a question indicates that question is more effective in distinguishing participants with different ability levels. If the question discrimination parameter is low, that question is less effective in distinguishing participants at different ability levels. In two parameter logistic models, the probability of participants answering a question correctly is modeled using the logistic function [9].
The equation for the 2PL model is as follows:

$$prob\ (x = 1 \mid \theta, b, a) = \frac{1}{1 + e^{a*(b-\theta)}} \quad (3)$$

where θ is current ability of test taker and "a" represents the discrimination parameter of the question, "b" represents the difficulty parameter of the question, $prob\ (x = 1 \mid \theta, b, a)$ means probability of test taker correct answer at 'a','b', θ.

Advantages and Disadvantages of Two Parameter Model (2PL) Compared to One Parameter Model (1PL) (Rasch Model)

Advantages of 2PL Model:

More Flexible Model: While the one-parameter model considers only the difficulty parameter of the questions, the two-parameter model also considers the difficulty level of the questions and the discrimination coefficient of the questions. In this sense, the evaluation of two parameters allows the test to be analyzed more comprehensively and thus increases the sensitivity and quality of the test.

More Accurate Predictions: While the one-parameter model tries to estimate the participant's ability level by using a single parameter, the two-parameter logit model allows the two-parameter participants to choose more appropriate questions according to their ability level by combining the ability level of the participants with the difficulty level of the question and the discrimination coefficient of the question [10].

Disadvantages of 2PL Model:

More Complex Model: While the one-parameter model tries to predict how the participant answers a question with the logistic function consisting of the single parameter, the difficulty level of the question and the skill level of the participant, the two-parameter model uses the difficulty level of the questions, the discrimination coefficient of the questions and the ability levels of the participant taking the test to answer the question in the test. Tries to guess how to answer. This makes the two-parameter logic model more complex and makes analysis difficult.

Higher Data Requirements: Since the one-parameter model includes the difficulty level parameter of the problem, it requires less data than the initial values. The difficulty level of the two-parameter model problem requires more data to accurately calculate the discrimination coefficients of the problem. In addition, the suitability of the data set used is important in calculating these parameters. This makes it significantly difficult to calculate the parameters accurately and precisely in the two-parameter model.

1.3 Item Response Theory Three Parameter Logistic Model (3PL)

Item Response Theory is based on three parameters in preparing the test by adding a three-parameter logit model, the difficulty level of the question, and the estimation parameter as well as the discrimination coefficient of the question. The prediction parameter represents the probability that the test taker answers a question correctly at random. The estimation parameter is generally denoted by the letter "c" in the literature. A high prediction parameter of a question means that the participant has a high probability of guessing the answer to the question, even if the participant does not know the answer to the question. If the estimation parameter is low, it means that the problem is difficult to predict and solve. While this parameter is assumed to be zero when creating an equation in a two-parameter logit model, it is added to the equation with a coefficient "c" in a three-parameter logit model equation (4). By adding the prediction parameter to the equation, it enables the performance of the test participants to be calculated in more detail [10]. Considering that the predicticition parameters of the questions in the test are high in the adaptive test preparation process, it is preferred to use the 3-parameter model.

$$prob\ (x = 1 \mid \theta, b, a, c) = \mathrm{c} + (1 - \mathrm{c}) * \frac{1}{1 + e^{a*(b-\theta)}} \quad (4)$$

where θ represents current ability of test taker, "b" is difficulty of current question, "a" is discrimination parameter of the question, "c" is the guessing parameter of the question. Guessing parameter means that the probability that participants will answer the item randomly. $prob\ (x = 1 \mid \theta, b, a)$ means probability of test taker correct answer at 'a','b','c', θ.

Advantages and Disadvantages of Three Parameter Model (3PL) Compared to Two Parameter Model (2PL)

Advantages:

More Comprehensive Model: While the 2-parameter logit model uses the difficulty level discrimination coefficient of the question in the test, the 3-parameter model performs a more comprehensive analysis by adding the estimation coefficient of the question in addition to these parameters. This allows better analysis of the participants' performance and increases the reliability of the test.

More Precise Analysis: The 3-parameter logit model creates an prediction parameter for each test question. Thus, it enables a more precise analysis of the accuracy of the answer given by the participant taking the

test according to the ability level. It helps in more precise evaluation by distinguishing whether the answers of the people taking the test are due to real ability or random answers. It significantly increases the accuracy and reliability of the test and more precisely distinguishes the ability levels of the participants taking the test. In addition, the difficulty level of the question and the discrimination coefficient are evaluated more precisely, allowing the test to be better calibrated.

Disadvantages:

More Complex Model: Adaptive tests created with three parameters have a more complex equation system than the two-parameter model. Additionally, these parameters make it difficult to predict as a whole. For this reason, it makes it difficult to use and apply this model in the adaptive test preparation process. Analyzing the relationships between model parameters and their connections with ability level is more difficult than the two-parameter model.

Higher Data Requirement: Since the 3-parameter logit model contains 3 parameters, it needs much more data to estimate these parameters than the 2-parameter model. Analyzing the data used in the preparation of the adaptive test well and estimating the parameters correctly in the 3-parameter model significantly affects the efficiency of the test. It may require extra data collection, especially in determining the estimation parameters "c" of the questions in the test. Compatibility of the data set used with the test parameters increases the accuracy of parameter estimation. In this sense, the size and quality of the data set play an important role in determining the parameters in the 3-parameter model. In cases where less data is available or data collection is difficult, a 2-parameter logit model or a one-parameter logit model is generally preferred instead of a 3-parameter logit model.

Each model has its own advantages and disadvantages in certain situations, so model selection depends on the goals and requirements of the test [10, 11].

Methods Used to Calculate The Parameters in Item Response Theory

We discussed Item Response Theory, models and parameter of the models in the previous section. The parameters included in these models, what these parameters mean, the effects of the parameters on the test and the use of the parameters in the equation are explained in the previous section. In this section, the basic working principles of Maximum Likelihood Theory and Bayesian estimation methods, which are one of the most used methods in the literature to calculate these parameters, are briefly explained. Thus, information is given about the basic working principles of the two most common methods used in calculating parameters in the adaptive test preparation process with Item Response Theory.

Maximum Likelihood Estimation (MLE): MLE is a statistical estimation method used to estimate the parameters of a specified model [12]. First of all, the probability distribution of the determined model is expressed as a function. Two equations are used in the literature for Maximum Likelihood Estimation (5) and (6). Since the solution of these equations in product form is more difficult and complex, the use of equations in logarithmic and sum form makes it more advantageous. The estimation of the parameters is done sequentially in an iterative manner, with one parameter being fixed and the other parameter being variable. For example, the initial ability level is determined and after the difficulty level of the question is estimated with the MLE method, the difficulty level of the problem is kept constant and the ability level is tried to be estimated by substituting it in the equation. In the 2-parameter model, while all parameters are calculated respectively, one parameter is expressed in the equation as a variable, while the other parameters are assigned a constant value and values that maximize equation (5) or (6) are tried to be found. One of the methods used to find the values that maximize these equations is the Gradient Descent method. As the number of parameters increases, it makes the solution of the resulting equation more difficult, and when the available data are not compatible with the equation, test parameters may be calculated incorrectly. At

this point, the size and accuracy of the data are of great importance in calculating the test parameter values correctly.

$$L(\theta) = \prod_{i=1}^{N}(P_i(\theta)^{X_i} + (1\text{-}P_i(\theta)^{1-X_i})) \quad (5)$$

$$L(\theta) = \sum_{i=1}^{N}[\, X_i \,.\log\,(P_i(\theta)) + (1 - X_i).\log\,(1 - P_i(\theta))] \quad (6)$$

Bayesian Estimation: One of the methods used in parameter calculations in adaptive testing processes prepared with Item Response Theory is the Bayesian estimation method based on Bayes theorem. First of all, preliminary distributions that provide information about the possible values of the parameters are determined. Posterior distributions are then calculated based on prior distributions and observed data using Bayes' theorem. These posterior distributions represent model parameters calculated according to available data. Finally, the estimated values of the parameters are obtained using summary statistics of the expected value posterior distributions [12].

2. KNOWLEDGE SPACE THEORY (KST)

Knowledge Space Theory (KST) is one of the most effective methods used in adaptive learning processes. Based on Knowledge Space Theory, asking questions is one of the important theories used to determine not only the student's ability level but also how much knowledge the student has about a subject, where the test taker has deficiencies in the subject, and which subjects he should study. Here is an explanation of the use of this theory in adaptive testing:

Defining the Knowledge Field: firstly a particular subject is created all subheadings and concepts. All concepts related to the subjects that make up this test and the relationship between these concepts are determined. For example, the topic of solving mathematical equations includes basic mathematical concepts and the connections between these concepts. The Knowledge Space Theory is a model that structures all information elements regarding a particular subject and the relationships between these moments. All knowledge elements and concepts are represented in the form of nodes, and the relationships between them are indicated by arrows.

Determination of Student Knowledge Status: Knowledge field refers to a cluster that represents the knowledge status of students. A student's knowledge status represents which knowledge elements and concepts of a subject student knows or does not know. Before the student starts the test, the initial knowledge level is determined based on previous data or by looking at the class situation according to expert opinion. According to the student's correct or incorrect answers to the test, student’s knowledge status is updated and questions about other knowledge items are presented to the test taker. Thus, an adaptive test is prepared by determining which concepts will be tested according to the student's knowledge and ability level.

Adaptive Test Design: All knowledge elements and concepts about a topic are extracted. The connections between these concepts and elements are determined. Expert opinions or some algorithms are used to determine these links. Once these connections are determined, tree-shaped structures are created and evaluated. The structures created represent the student's learning process.

Application of the Test: An adaptive test that includes all knowledge items and concepts related to a subject is applied to students. Depending on the right or wrong answers given by the student taking the test, the student moves to questions on the next more difficult or simpler topic. After each question, the student proceeds on the tree model appropriate to test taker. Thus, an adaptive test specific to each student is performed.

Analysis of Results and Update of the Test: After the adaptive test created on a subject with knowledge space theory is presented to a certain number of students, the test knowledge items and concepts and the relationships between them are updated. The relationships between these concepts are carried out by the expert or by the algorithms used in this field. Thus, the relationship between the subject knowledge items and the learning tree are updated. Thus, the reliability and sensitive evaluation criteria of the test are increased.

As example of knowledge space theory, a knowledge structure is denoted as a pair (Q, K), where Q represents a set of items, and K is a collection of subsets of Q. As evident, the elements of K signify potential knowledge states, indicating the combinations of items to which an individual may belong. Within a knowledge structure, the collection K must encompass the empty set and Q.
Q = {a,b,c,d} (Each item represented by one letter)
K = {{}, {a}, {d}, {a,b}, {a,d}, {a,b,c}, {a,b,d}, Q} (Each subset of Q is the knowledge states, for example the state {a,b,d} means the mastery of the items a, b and d [13].

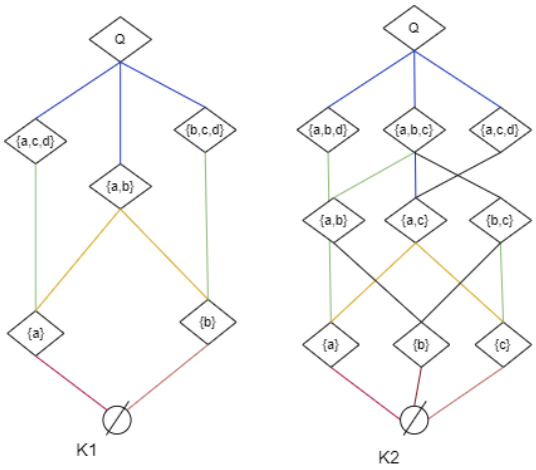

Figure 1. Example flowchart of knowledge structures

Adaptive tests designed using Knowledge Space Theory more precisely assess a student's knowledge status and provide an appropriate test based on the student's ability level. Instead of just determining the student's ability level, it also shows which of the knowledge elements that make up a subject as a whole are missing or insufficient. This provides more effective measurement and evaluation of the student. In this way, the student not only learns the ability level in a subject, but also learns the deficiencies regarding the subject in detail.

In the adaptive test preparation process created using Knowledge Space Theory, many algorithms are used to determine the relationship between knowledge items and concepts of the relevant subject. One of the most popular algorithms is the K2 algorithm. K2 algorithm is one of the probabilistic graphical modeling algorithms used in the field of artificial intelligence. K2 algorithm is an algorithm that formally models the dependencies of the subheadings of the topic and which topic should be learned first or which question should be answered first. With this algorithm, the connections between the concepts and information items in the training data set to be used in the test preparation process can be graphically displayed. The K2 algorithm creates connections between concepts in an iterative process. In the K2 algorithm, information items or concepts are expressed by nodes. It is assumed that initially the nodes are sorted [14]. Other algorithms such as the Genetic algorithm or expert opinion can be used to rank the nodes. Then, the maximum number of node parents that each node can take is determined. The first node is considered to have no parent. The parent list is empty when nodes are added to the first created network. After the node is added, it accepts the nodes that came before it as parent and its K2 algorithm score is checked. It is decided whether a parental relationship will be created or not by looking at the situation in the score. Adding parents for each node continues until the score increases and the maximum number of parents is reached.

K2 algorithm is one of the most suitable algorithms for determining parent-child relationships between concepts and knowledge items for adaptive testing created with Knowledge Space Theory. However, it must be carefully assessed whether the properties of the data set to be used correspond to the requirements of the algorithm. To evaluate the accuracy, effectiveness and usefulness of the algorithm, appropriate tests must be performed and analyzed in detail.

Comparing Knowledge Space Theory (KST) and Item Response Theory (IRT) in Adaptive Test Preparation

Item Response Theory (IRT) and Knowledge Space Theory (KST) offer different approaches to the preparation of adaptive tests. KST is used to model the relationships between sub-knowledge elements and concepts required for learning a subject, and adaptive tests are created to analyze the student's knowledge of the subject in more detail based on the relationships between these concepts [15]. Adaptive tests created with IRT are used to numerically estimate the ability levels of participants by examining the statistical properties of the questions in the test. While KST is generally used to evaluate students' knowledge about a subject and to monitor the student's learning process on that subject, IRT generally aims to determine the effect of the parameters affecting the test on the test and the ability level of the participants according to their answers to the questions. In terms of test design, while KST uses graphical modeling by creating sub-information items and concepts about a subject, IRT uses a method based on the selection of the appropriate question and numerical calculation of the ability level based on statistical features such as the difficulty level of the questions, discrimination parameter and prediction parameter. IRT graphical modeling does not consider parameters such as the subheading of a topic and the learning process. IRT focuses on calculating the ability level of the participants in the most precise way by selecting the most appropriate question according to the participant's ability level. While KST focuses on the learning process, IRT focuses on the selection of questions based on the most appropriate test and difficulty level [16]. While KST is used to establish learning and measurement processes in fields such as education, psychology and other social sciences, IRT is widely used in graded testing, psychometrics and measurement evaluation.

III. DISCUSSION AND CONCLUSION

In this text, Item Response Theory and Knowledge Space Theory, which are the most common theories used in the preparation of adaptive tests, are discussed. Adaptive tests are one of the effective measurement methods used to analyze students' learning processes and what they have learned. Unlike classical tests, not all students are asked the same questions. An adaptive test is prepared according to the student's ability level by asking questions according to each student's ability level. Item Response Theory one parameter logit model (Rasch model) creates a test for the student by connecting the difficulty levels of the questions in the test to the student's skill level, whereas the two parameter logit model creates an adaptive test for a student by connecting the level of difficulty and discrimination of questions to the students' skill level. The three-parameter logit model aims to produce a deeper and more precise adaptive test by adding the difficulty level of the questions, the differential ratio of questions, and the probability ratio that the questions can be predicted to be solved. Since the Rasch model is simple and easy to analyze, it is one of the most preferred methods in preparing adaptive tests in the literature. While it provides more precise measurement as the number of parameters increases, it becomes more difficult to analyze and calculate the parameter values as a whole. The basic principles of Maximum Likelihood Estimation and Bayesian Estimation methods, which are among the most commonly used methods in the literature for Item Response Theory parameter calculations, are explained. Maximum Likelihood Estimation is used in two forms in the literature. Its first usage is in multiplication form. Since the usage of this form makes calculation difficult, the use of the second form" the logaritmic summation form" is preferred. In calculating parameter values with Maximum Likelihood Estimation, the parameter to be calculated is calculated through an iterative solution process, keeping the other parameters constant. For this iterative calculation process, the Gradient Descent method, which is one of the most widely used methods in the literature, can be preferred. The data set to be used in Maximum Likelihood calculation must be compatible. Otherwise, errors or parameter values cannot be

found in parameter calculations. Another most common method used in the Adaptive Test Preparation Process is Knowledge space theory. Knowledge Space theory (KST) is a method that fully deals with all sub-topics (subheadings) of subject which is necessary to learn the subject. Before preparing the adaptive test, all subheadings of a topic are created as a cluster. There are many methods used in the literature to determine the relationships between these subheadings. One of the most widely used methods is K2 algorithm. K2 algorithm takes all the subheadings and concepts of a subject within the parent-child relationship and makes graphical modeling as the form of Family Tree. Before making graphical modeling with the subheadings of the subject using the K2 algorithm, the subheadings and concepts must be listed. This sorting can be done by an expert or by using algorithms such as the genetic algorithm in the literature. The beginning starts with a subheading and this subheading is considered not to be a parent. Then, as each subheading is added to the graphical model, the K2 score is checked to determine whether a parental relationship can be established. By adding/removing all subheadings to the graphical modeling, the K2 score is checked and the best graphical modeling is obtained. While Knowledge Space Theory uses graphical modeling, Item Response Theory generally proceeds with numerical calculations. While item response theory aims to evaluate students statistically, Knowledge Space Theory shows students' level of knowledge about the subjects as well as statistical evaluation. While Knowledge Space Theory aims to measure the level of knowledge on a subject, item response theory aims to score students by taking an exam according to their ability levels. While Item Response Theory questions can include more than one topic depending on their difficulty level, Knowledge Space Theory divides a topic into all subheadings and prepares questions to show which subject the student is deficient in by addressing the concepts one by one. In this text, Item response theory and its models, which are among the most commonly used methods in the adaptive test preparation process in the literature, and Knowledge Space Theory are discussed together. The usage areas, purposes of use, and basic working principles of these theories are explained and compared with each other In addition, the most commonly used auxiliary algorithms of these theories in the literature are mentioned The advantages and limitations of each model have a significant impact on the effectiveness and accuracy of adaptive testing. Item Response Theory and Knowledge Space Theory usage purposes, basic working principles, advantages and disadvantages of the models are discussed together, with the aim of providing the reader with which theorem to use according to their needs and creating an auxiliary resource.

REFERENCES

1. Gershon, R. C. (2005). Computer adaptive testing. Journal of applied measurement, 6(1), 109-127.
2. Weiss, D. J. (1985). Adaptive testing by computer. Journal of consulting and clinical psychology, 53(6), 774.
3. Van der Linden, W. J., & Glas, C. A. (2000). Computerized adaptive testing: Theory and practice: Springer.
4. Segall, D. O. (2005). Computerized adaptive testing. Encyclopedia of social measurement, 1, 429-438.
5. Baker, F. B. (2001). The basics of item response theory: ERIC.
6. DeMars, C. (2010). Item response theory: Oxford University Press.
7. Reeve, B. B.. & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire item and scale properties. Assessing quality of life in clinical trials: Methods of practice, 2, 55-73.
8. Stemler, S. E., & Naples, A. (2021). Rasch measurement v. item response theory: Knowing when to cross the line. Practical Assessment, Research & Evaluation, 26, 11.
9. van der Linden, W. J., & Hambleton, R. K. (2013). Handbook of modern item response theory: Springer Science & Business Media.
10. Embretson, S. E., & Reise, S. P. (2013). Item response theory: Psychology Press.
11. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory (Vol. 2): Sage.
12. Baker, F. B., & Kim, S.-H. (2004). Item response theory: Parameter estimation techniques: CRC press.
13. Doignon, J.-P., & Falmagne, J.-C. (2016). Knowledge spaces and learning spaces.
14. Behjati, S., & Beigy, H. (2020). Improved K2 algorithm for Bayesian network structure learning. Engineering Applications of Artificial Intelligence, 91, 103617.
15. Doble, C., Matayoshi, J., Cosyn, E., Uzun, H., & Karami, A. (2019). A data-based simulation study of reliability for an adaptive assessment based on knowledge space theory. International Journal of Artificial Intelligence in Education, 29(2), 258-282.
16. Muñoz-Merino, P. J., Novillo, R. G., & Kloos, C. D. (2018). Assessment of skills and adaptive learning for parametric exercises combining knowledge spaces and item response theory. Applied Soft Computing, 68, 110-124.