O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA INNOVATSIYALAR VAZIRLIGI HUZURIDAGI BILIM VA MALAKALARNI BAHOLASH AGENTLIGI

# AXBOROTNOMA

# BULLETIN

3

2025

Toshkent

**“AXBOROTNOMA”**
O‘zbekiston Respublikasi Oliy ta’lim, fan va innovatsiyalar vazirligi huzuridagi Bilim va malakalarni baholash agentligi ilmiy-uslubiy jurnali
Bir yilda 4 marta chiqadi.
O‘zbekiston Matbuot va axborot agentligida 2007-yil 19-aprelda qayta ro‘yxatdan o‘tkazilgan.
Guvohnoma № 0247

**Ta’sischi:**
Bilim va malakalarni baholash agentligi



Bosishga ruxsat etildi: 09.12.2025
Shartli bosma tabog‘i: 5,4
Nashriyot hisob tabog‘i: 5,3
Adadi 100. Buyurtma №2-A - 25

Bilim va malakalarni baholash agentligi bosmaxonasida chop etildi.

**Nashrga tayyorlovchilar:**
A. Normurodov
**Muharrir:**
N.Abduraxmanova
**Kompyuterda sahifalovchi:**
Y. To‘rayev
**Tahririyat manzili:**
100084, Toshkent shahri,
Bog‘ishamol ko‘chasi, 12-uy.



## MUNDARIJA



## CONTENTS

## KIRISH

“Axborotnoma” ilmiy-uslubiy jurnalining ushbu sonida pedagogik o‘lchovlar bo‘yicha ilmiy-uslubiy izlanishlar olib boruvchi mutaxassislar tomonidan olib borilgan tadqiqotlar natijalari haqida to‘rtta maqola berilgan.

Birinchi maqolada test ballarining ichki muvofiqligi va ishonchliligini ta’minlashda, baholashda Cronbach alfa ko‘rsatkichining ahamiyati o'rganilgan. Keng qamrovli tahlil orqali test natijalarining javoblar matritsasi Guttman namunasiga qanchalik mos kelishi ko‘rsatilib, Kronbax alfa yordamida xulosa chiqarishda topshiriqlar sonining muhimligi ta’kidlangan. Ilmiy tadqiqot uchun biologiya fanidan 423 nafar o‘quvchi ishtirokida test sinovi o‘tkazilgan va testning validligi hamda ishonchliligi statistik usullar bilan tahlil qilingan. Tadqiqotda Guttman namunasining batafsil tahlili, uning test tahlilidagi nimani anglatishi hamda Kronbax alfani turli testlar uchun qanday talqin qilinishi keng yoritilgan.

Ikkinchi maqolada huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 – o‘quv yilida bakalavriatga kirish test sinovlari natijalarining klassik test nazariyasi asosida olingan statistik ko‘rsatkichlari - o‘rtacha qiymati, mediana, moda, o‘lchashning standart xatoligi va distraktorlari o‘rganilgan. Tahlil orqali test sinovlarida foydalanilgan test topshiriqlarining distraktorlari samarali ishlayotganligi aniqlangan.

Uchinchi maqolada ilmiy tadqiqot uchun biologiya fanidan o‘tkazilgan test sinovi natijalari asosida test ma’lumoti va xarateristikasi chiziqlari Rash modeli va ikki parametrli modellar bilan tahlil qilingan. Tadqiqotda har ikki model bo‘yicha testning qobiliyat darajasiga ko‘ra ma’lumot funksiyalari taqqoslangan. Natijalar orqali testning validligi va ishonchliligini oshirishda model tanlashning ahamiyati ko‘rsatilgan. Ikki parametrli model bo‘yicha hisoblangan test ma’lumot chiziqlari yuqori qobiliyatlar uchun ko‘proq ma’lumot berishi ko‘rsatilgan.

To‘rtinchi maqolada fizika fanining optika bo‘limini o‘qitish jarayonida talabalarning bilim va ko‘nikmalarini baholashning zamonaviy usullari va mezonlari tahlil qilingan. Test, laboratoriya ishi va portfolio asosida baholash usullari yoritilgan. Shuningdek, ta’lim sifatini oshirishda diagnostika va refleksiyaning o‘rni ham ko‘rsatilgan. Talabalarda optika bo‘limini innovatsion yondashuv asosida o‘zlashtirish darajasini baholash mezonlari va shakllanganlik darajalari ko‘rib chiqilgan.

# INTERNAL CONSISTENCY IN ACHIEVEMENT TESTS: ROLE OF CRONBACH'S ALPHA AND GUTTMAN PATTERN

**M.J. Ermamatov, A.B. Normurodov, A.R. Sattiev**

*Scientific and Educational Practical Center Under the Agency for Assessment of Knowledge and Competences, a.normurodov@uzbmb.uz*

**Abstract.** This paper explores the significance of Cronbach's alpha in evaluating the internal consistency and reliability of test scores, particularly in educational assessments. Through a comprehensive analysis, we highlight how closely the response matrix of test results aligns with the Guttman pattern, emphasizing the importance of the number of items when making inferences using Cronbach's alpha. We conducted a field test with 423 students to evaluate a newly developed biology assessment tool, ensuring the reliability and validity of the test through robust statistical methods. The study presents a detailed examination of the Guttman pattern, its implications for test analysis, and the interpretation of Cronbach's alpha in different test scenarios. We also compare the descriptive statistics and reliability measures of simulated data sets with normal distribution and actual test data, providing insights into the practical applications and limitations of Cronbach's alpha in real-world settings. The findings underscore the importance of careful item selection and test design to enhance the reliability and interpretability of educational assessments

**Keywords:** assessment, classical test theory, achievement tests.

## 1. Introduction

Analyzing test results is essential for comprehending information, making informed decisions, and deriving meaningful conclusions. This analysis can help teachers assess the reliability of their tests, evaluate student achievement, and identify areas for improvement in future exams. By accurately interpreting test scores, teachers can draw valid conclusions from the data and better prepare for post-exam discussions with students. Furthermore, this process helps evaluate the quality of the tests, assess their readiness for release, and track their developmental progress.

It is imperative to develop instruments that measure accurately. To advance this goal, numerous investigations have been conducted [1-4], and various statistical formulas have been introduced to quantify measurement properties. Among these, Cronbach's alpha [5] stands out as a metric that specifically assesses

the reliability of tests, particularly their internal consistency. For example, Beckmann and Jastrowski Mano [6] demonstrated excellent internal consistency for their measure with Cronbach's alpha ranging from 0.94 to 0.97, while Bharara and Duncan [7] reported a total scale alpha of 0.943, indicating strong interrelatedness of items. Chen et al. [8] similarly documented good internal consistency across interventions with alpha values between 0.67 and 0.91.

Reliability refers to the extent to which a test consistently measures the same concept across different instances or sets of items. One key aspect of reliability is internal consistency, which measures the degree to which the individual items of a test are consistent with one another in assessing the same construct. High internal consistency indicates that the test items are all measuring the same underlying trait or ability, contributing coherently to the overall assessment. For example, in a math test designed to assess algebraic skills, high internal consistency would suggest that the test items focus on algebra rather than unrelated topics like geometry or calculus.

Internal consistency is usually evaluated by examining the correlations among test items. Strong correlations between items suggest that they all measure the same construct. In contrast, low correlations could signal that the items assess different constructs or that the test lacks focus.

Reliability, and particularly internal consistency, is crucial because it indicates the consistency of the test results. A test with high reliability consistently produces the same outcomes when administered under similar conditions, reflecting a stable measurement of the intended construct. If a test lacks internal consistency, its results may be unreliable, meaning the test could produce inconsistent outcomes, and the scores may not accurately reflect the test-taker's ability in the targeted area. In contrast, validity concerns whether the test measures what it is intended to measure. It refers to the accuracy of the test in assessing the specific concept, trait, or ability it is designed to evaluate. Therefore, while reliability (including internal consistency) ensures that the test consistently measures a construct, validity confirms that the test actually measures that construct accurately. During the test development process, ensuring reliability involves selecting and crafting items that consistently measure the same construct. To ensure validity, test developers must also confirm that the items effectively assess the intended concept without measuring irrelevant or unrelated constructs.

Internal consistency is essential in educational testing because it ensures that the test items collectively measure the same underlying construct,

contributing to the reliability of the assessment. Reliability refers to the degree to which test results are free from errors – unwanted variations that can influence the outcome of a test. Errors can arise from various sources, such as differences in test-takers' performance due to external factors (e.g., anxiety, fatigue) or inconsistencies in how test items are interpreted. These errors can undermine the accuracy and consistency of test scores.

By ensuring internal consistency, test developers can reduce the likelihood of such errors and enhance the reliability of the test. When test items are internally consistent, they are less likely to introduce random errors that could distort the test results. This allows educators and test developers to confidently interpret test scores, knowing that the results reflect the specific ability or knowledge area the test is designed to measure, with minimal error.

One of the most common measures of internal consistency is Cronbach's alpha [5]. This statistic provides a numerical estimate of how well the test items measure the same construct. A high Cronbach's alpha (typically above 0.7) indicates good internal consistency, meaning the items reliably assess the same concept. For example, a reading comprehension test with a Cronbach's alpha of 0.85 suggests that the items consistently measure the construct of reading comprehension. The correct interpretation of Cronbach's alpha is crucial for accurately assessing the reliability of scales, a point thoroughly discussed by Cortina [4]. Cronbach's alpha is often misunderstood as a simple measure of internal consistency; however, Cortina emphasizes that it is influenced by multiple factors, including both the number of items on a scale and the inter-item correlations. A higher number of items can inflate alpha even if the inter-item correlations are relatively low, which may give a false sense of reliability.

Moreover, the shape of the score distribution and the pattern of inter-item correlations significantly influence the interpretation of Cronbach's alpha [1]. Alpha assumes tau-equivalence, meaning all items have equal true-score variances and measure the same underlying construct. When this assumption is violated, alpha can still appear high, which may be misleading. As Schmitt [3] cautions, a high alpha does not guarantee unidimensionality. Multi-dimensional tests – those measuring multiple constructs – can also produce inflated alpha values due to the number of items or item redundancy. While our analysis focuses on alpha, we recognize that additional evidence (e.g., dimensionality analysis) would be necessary to confirm the internal structure of the test. We note this as a limitation and encourage future work to incorporate such methods for a more comprehensive evaluation.

Given these limitations, researchers have explored alternative reliability coefficients that address the shortcomings of alpha. One such measure is McDonald's Omega, which incorporates factor analysis into the estimation of reliability [9]. Omega is particularly advantageous when the assumption of tau-equivalence is violated, as it provides a more accurate representation of the reliability in scales where items have differing factor loadings. Omega distinguishes between general and specific factors, making it more appropriate for scales that may be multidimensional in nature.

Recent research, such as the work of Orcan [10], has examined and compared the estimates provided by Cronbach's alpha and McDonald's Omega. These studies suggest that while alpha remains a popular metric due to its simplicity, omega offers a more nuanced and potentially more reliable estimate, particularly when factor structure is a concern. As a result, many researchers are now advocating for the use of McDonald's Omega as a complement or even an alternative to Cronbach's alpha, especially in scales where the unidimensionality assumption may not hold.

In psychometric literature, reliability refers to the consistency or stability of measurement. It encompasses several forms, including test–retest reliability (the stability of scores over time), inter-rater reliability (the consistency of scores across different raters), and internal consistency reliability, which is the focus of this paper. Internal consistency refers to the degree to which items within a test measure the same underlying construct. It is typically assessed using statistical indices such as Cronbach's alpha or McDonald's omega. While all types of reliability relate to measurement consistency, internal consistency specifically evaluates the coherence among items within a single test administration.

In this paper, we aim to demonstrate that Cronbach's alpha indeed reflects how closely the matrix of responses from test results aligns with the Guttman pattern [11]. Guttman's contribution to the classical test theory is thoroughly examined and discussed in reference [12]. Additionally, we will explore why the number of items is crucial when making inferences using Cronbach's alpha. To this end, we will analyze the results from a field test in biology to assess the quality of the test before its formal application.

## 2. Method

### 2.1. Study Design

The study was designed to assess the internal consistency and reliability of a newly developed biology assessment tool using Cronbach's alpha and the Guttman pattern. The focus was on evaluating how well the test results aligned with the Guttman pattern, particularly in educational assessments. In addition to reliability, the study also aimed to assess the validity of the tool. To determine validity, the test was evaluated using several established methods. First, content validity was assessed by reviewing the test items for relevance and coverage of the intended biological concepts, with input from subject matter experts. Next, construct validity was examined by analyzing how well the test scores correlated with other measures of similar constructs, such as performance on related biology assessments. Finally, the test's criterion-related validity was evaluated by comparing its results with real-world outcomes, such as student grades or performance in biology-related tasks.

### 2.2. Sample or Study Group

The study was conducted with a total of 423 students from various educational backgrounds. The participants included 179 students from academic lyceums, 185 from secondary schools, and 59 from training centers. This diverse sample was chosen to ensure that the test results would be representative of different student populations.

### 2.3. Data Collection Tools

The primary data collection tool was a 48-item biology assessment test specifically designed for this study. The items in the test were selected to represent varying levels of difficulty, aligning with the theoretical Guttman pattern. While the Guttman pattern offers an idealized model for how test items should correlate based on their difficulty levels, it is important to note that achieving perfect alignment with this pattern is rare in practice. The literature suggests that, in reality, tests that attempt to follow a Guttman pattern typically exhibit only moderate levels of alignment due to various factors, such as individual differences in test-taker performance and the complexity of real-world assessments. Therefore, while the theoretical Guttman pattern guided the design of this test, the expectation was that the data would show some deviation from this idealized model, reflecting the inherent complexities of educational assessments.

### 2.4. Data Collection Process

The test was administered across the three different educational settings – academic lyceums, secondary schools, and training centers. The process involved distributing the test to the students and collecting their responses, which were then used for statistical analysis. The students' responses were scored, with correct answers marked as '1' and incorrect answers as '0'.

### 2.5. Data Analysis Procedures

The data collected from the test was analyzed using Cronbach's alpha to determine the internal consistency of the test. The study also examined the alignment of the response matrix with the Guttman pattern. Descriptive statistics such as mean, median, mode, variance, skewness, and kurtosis were calculated to understand the distribution of the test scores. The analysis also involved comparing the real test data with simulated datasets to evaluate the reliability and interpretability of the results across different groups. This phase was crucial for evaluating the test's reliability, as indicated by measures such as Cronbach's alpha, and its validity across a representative student population. The scale of the test allowed for robust statistical insights, ensuring that the findings were well-grounded and reflective of the test's performance in realistic settings.

## 3. Results

### 3.1. Guttman pattern as a foundation

Louis Guttman [11], a prominent psychometrician, conceptualized the ideal test as one in which a person succeeds on all items up to a certain level of difficulty, and then fails on all items that are more difficult. When individuals and items are ordered by their raw scores, this produces a dataset characterized by a "Guttman pattern" [11]. Such data sets are not typically analyzable using Rasch analysis because each person or item in turn becomes an extreme score. In simpler terms, items in a Guttman pattern are arranged such that a person who agrees with or correctly answers a particular item will also agree with or answer all less difficult items correctly. This pattern suggests a unidimensional underlying trait.

The use of the Guttman response pattern in this study serves a primarily illustrative and pedagogical purpose. As a theoretical ideal, the Guttman pattern represents a perfectly ordered item-response matrix, where each item

strictly increases in difficulty and each respondent answers all easier items correctly before missing harder ones. While such a structure can demonstrate the upper boundary of internal consistency (e.g., Cronbach's alpha approaching 1), it is not a realistic or desirable goal in practical test design. Real-world assessments rarely exhibit such deterministic patterns due to variability in human performance, guessing behavior, and multi-dimensional traits. Thus, the Guttman matrix in this study is used not as a test design model but as a benchmark to highlight how item ordering and inter-item correlations can influence reliability coefficients.

To illustrate, consider a hypothetical scenario where 20 students each answer 20 questions. Correct answers are scored as '1' and incorrect ones as '0'. The resulting data, presented in Table 1, exemplifies a perfect Guttman pattern. Although such patterns are not analyzable using the Rasch model, they can be assessed through various other methods. It is crucial to recognize that interpreting extreme scores within classical test theory is challenging. Extreme scores suggest either an infinite ability level or extreme difficulty. For instance, a student who fails to answer any of the items on the test might succeed on a hypothetically easier 21st question, not included in the original test. Conversely, a student who answers all the provided questions correctly might fail an additional, more challenging question. Moreover, if all students answer a particular question correctly or if no students answer it, does this make the question infinitely easy or difficult, or are there potential respondents not yet tested who could affect this outcome? Can a scenario in which each successive student answers one more question than the last accurately represent a real class, or is it possible to structure the questions such that the first is answered by one student, the second by two students, and so on? Such considerations underscore the importance of the number and distribution of both questions and test-takers.

*Table 1*

**Hypothetical response pattern of 20 students (Guttman pattern)**

| Item / person | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 17 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 18 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 20 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 21 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Let's examine the components of the Cronbach's alpha formula, detailed at [13]:

$$\boldsymbol{\alpha = \frac{k}{k-1}\left(1 - \frac{S^2}{\sum s^2}\right)}$$

where k is the number of items, $S^2$ is variance of total score, $\boldsymbol{\sum s^2}$ is the sum of variances of the students' responses to each item.

In the matrix of responses, units appear in the upper triangle, and zeros appear in the lower triangle when the abilities of the respondents and the difficulties of the questions are arranged in increasing order. This structured arrangement highlights the predictable response patterns aligned with the Guttman scale. Notably, the mean and median of the total scores across this dataset are both 10, although the mode of the total score remains undefined due to the diversity of response patterns.

As we analyze the variances in the responses of respondents to each item (denoted as $s^2$), we observe that these variances increase towards items of average difficulty and decrease

symmetrically thereafter, as illustrated in Figure 1. This variance pattern reflects the concentration of errors around the median difficulty level, where respondent performance begins to diverge.

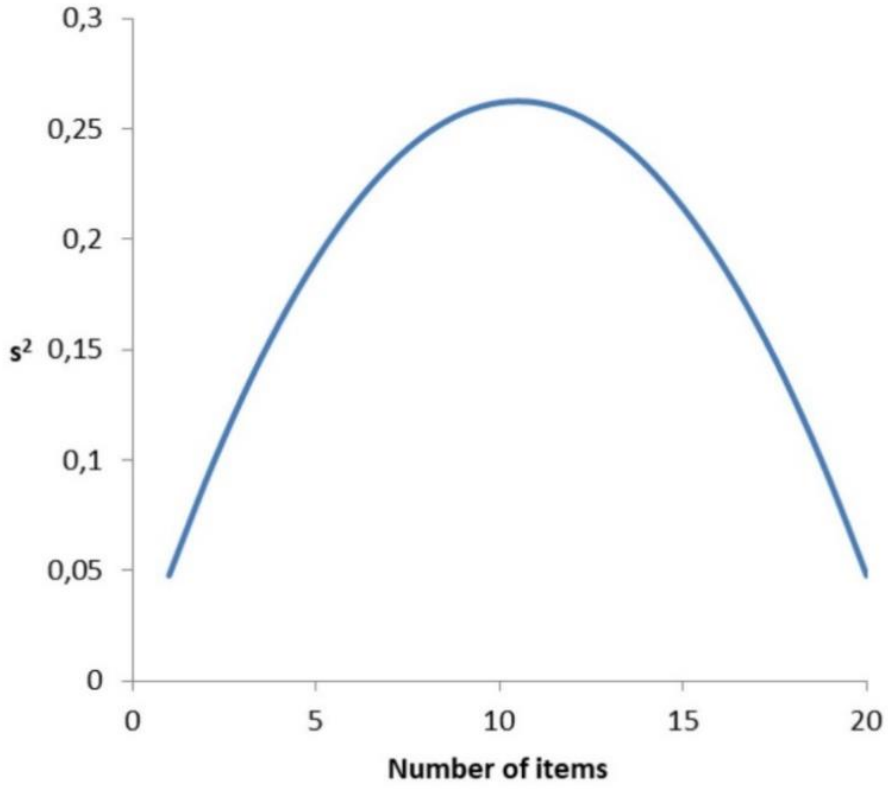



Figure 1. Variation of $s^2$ (item variance) with items arranged in increasing order of difficulty

The variance of the total scores ($S^2$) for such a pattern is calculated to be 38.5. Given these statistical properties, the calculated Cronbach's alpha for this pattern is exceptionally high at 0.95, indicating a high level of internal consistency within the test. This high reliability suggests that the test consistently measures some underlying construct. However, it is important to note that reliability does not necessarily guarantee validity. While the test is reliable in terms of internal consistency, further steps were taken to evaluate its validity. Validity concerns whether the test measures what it is intended to measure - i.e., the specific biological concepts it is designed to assess. The validity of the test was assessed through expert reviews of the content (content validity), as well as by examining the correlation between the test scores and other established measures of biological knowledge (construct validity).

These assessments, alongside the high reliability, support the validity of the test in measuring the intended biological constructs. This high alpha

value underscores the reliability of the test in assessing the trait it is intended to measure, albeit within the confines of a structured and idealized response pattern as depicted. Let's see the level interpretation usually people use for the Cronbach's alpha in Table 2.

Upon examining formula, it becomes apparent that the number of questions, k, plays a critical role in determining the value of Cronbach's alpha.

*Table 2*

**Level interpretation for Cronbach's alpha**

| Alpha | Level interpretation |
|---|---|
| < 0.5 | Unacceptable |
| > 0.5 | Poor |
| > 0.6 | Questionable |
| > 0.7 | Acceptable |
| > 0.8 | Good |
| > 0.9 | Excellent |

This component influences all other variables in the formula, affecting the overall calculation of reliability. Specifically, a smaller variance in responses to each individual question ($s^2$) and a larger variance of the total scores ($S^2$) contribute to an increase in Cronbach's alpha. This dynamic indicates that a diverse range of total scores across the test, combined with consistent individual responses, enhances the reliability measure.

Moreover, it is essential to consider the implications of an exceptionally high Cronbach's alpha, approaching 1. Such a high value typically suggests an ideal response pattern, as depicted in Table 1. While this might indicate excellent internal consistency, it can also hint at redundancy among items or a lack of sufficient challenge across the test's scope. In practical settings, an alpha value very close to 1 might not always signify an optimal test but could point to potential issues in test design that need further investigation.

As mentioned previously, achieving response patterns like those depicted in Table 1 is generally unrealistic in practical settings. Even with carefully selected groups and test items, it is highly improbable to achieve such a perfect pattern. A Cronbach's alpha value of 0.95, as calculated for the scenario involving 20 questions and 21 respondents, might itself be regarded as suspicious in real-life situations. Such a high value typically suggests an

unusually consistent pattern of responses that may not reflect the diversity of an actual test-taking population.

In more realistic scenarios, one might encounter patterns that approximate but do not exactly match the Guttman pattern. For instance, the upper triangle of the response matrix may contain a few 'ones', while the lower triangle predominantly contains 'zeros'. Consider a more idealized scenario where, instead of having one student scoring '0', another '1', and so on, we have 100 students scoring '0', another 100 scoring '1', etc. In this situation, the variance of the total score would be significantly higher, potentially leading to an even larger Cronbach's alpha. This occurs because the presence of larger groups of students scoring at the same level results in a pattern where 'zeros' and 'ones' occupy larger blocks within the matrix. Mathematically, this configuration suggests a higher probability of achieving a smaller sum of item variances $\sum s^2$ and a larger total score variance $S^2$, thereby inflating the Cronbach's alpha value. Similar reasoning can apply to the selection and number of items included in the test.

Another aspect of the ideal patterns discussed so far is that the total scores from such patterns are uniformly distributed. However, in real-world situations, the distribution of total scores often approximates a normal distribution, especially when the sample size is large and the test measures a trait or ability that naturally varies within the population, as described by the Central Limit Theorem. This phenomenon is common with standardized tests and various types of assessments, where individual variations and a large number of questions can average out to produce a bell-shaped curve. Nevertheless, it's important to recognize that not all real-world situations will perfectly follow a normal distribution. The actual distribution can be influenced by the nature of the group being tested, the type of test, and its alignment with the abilities being assessed. For instance, if a test is too easy or too hard for the group being tested, a skewed distribution might be observed instead.

To understand these nuances in distribution, the concepts of skewness and kurtosis are employed. Skewness measures the degree of asymmetry in a distribution; a distribution with most data concentrated on the left and a longer right tail is said to be right-skewed or positively skewed, while one with the peak toward the right and a longer left tail is left-skewed or negatively skewed. Kurtosis, on the other hand, describes the 'tailedness' of the distribution. For kurtosis, we refer to the definition provided by Kevin P. Balanda and H.L. MacGillivray in "Kurtosis: A Critical Review," published in The American Statistician in May 1988 [14], which can also be found at [15]. In the context of this paper, we consider excess kurtosis as defined in

the aforementioned reference. For the Guttman pattern discussed, the skewness is 0, indicating no asymmetry, and the kurtosis is 1.2, suggesting a relatively low level of peakedness.

Let's now examine a simulated pattern for 500 individuals and 20 items, where the total scores are normally distributed. The frequency distribution of the total scores is presented in Figure 2, with bins chosen at a step size of 1. The solid line represents the theoretical normal distribution, which has the same mean, median, and mode as the total scores from the simulated data.

Descriptive statistics for this dataset are presented in Table 3. In a perfect normal distribution, the mean, median, and mode are identical, with both skewness and kurtosis equaling zero. The total score from the simulated data closely approximates this normal distribution profile. In contrast, analysis of the Guttman pattern data reveals a variance of 38.5, with both mean and median at 10, skewness at 0, and kurtosis at -1.2; the mode was undetermined. Additionally, the Cronbach's alpha for the Guttman pattern stands at 0.95, signifying excellent reliability.

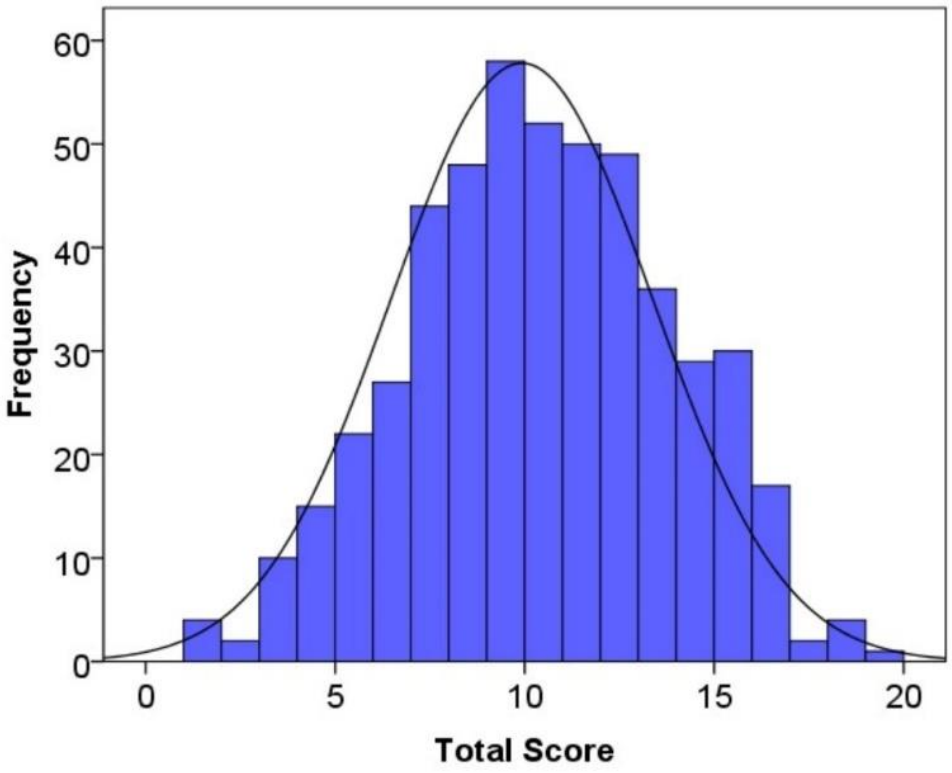



Figure 2. Frequency Distribution of Total Scores from the Simulated Data for 500 Individuals and 20 Items

The alpha for the simulated dataset is 0.75, indicating that the reliability of the test results from this 20-item test form is 'acceptable'. With

careful selection of items and targeted learning strategies, there is potential to significantly improve this reliability score. Typically, however, a Cronbach's alpha between 0.75 and 0.8 is expected for a test form consisting of 20 items. Descriptive statistics for this dataset are presented in Table 3.

While a higher Cronbach's alpha is often interpreted as an indicator of greater internal consistency, it is important to recognize that alpha can also increase due to factors that may not reflect improved test quality. As noted by Cortina [4], alpha is sensitive not only to inter-item correlations but also to the number of items, meaning that adding similar or overlapping items can artificially inflate alpha. In such cases, a high alpha may suggest item redundancy rather than meaningful coherence.

*Table 3*

**Descriptive statistics for the simulated dataset**

| | |
|---|---|
| Number of students | 500 |
| Number of Items | 20 |
| Mean | 9.93 |
| Median | 10 |
| Mode | 9 |
| Variance | 11.90 |
| Standard deviation | 3.45 |
| Skewness | -0.061 |
| Kurtosis | -0.41 |
| Range | 18 |
| Minimum | 1 |
| Maximum | 19 |

Furthermore, if added items are not conceptually relevant to the construct, this can lead to construct irrelevance, where the scale appears statistically reliable but lacks substantive validity. Therefore, alpha should be interpreted with caution, alongside considerations of content representativeness and unidimensionality. For comparison with Guttman pattern, we present the variation of $s^2$ (item variance) with items arranged in increasing order of difficulty in Figure 3. This visualization illustrates how the dependency

changes when transitioning from a Guttman pattern to a normal one.

As shown in Figure 3, the $s^2$ for the simulated normal distribution, and consequently the sum of $s^2$ (3.38), has slightly decreased compared to that of the uniform distribution (3.67), a change attributable to the increase in the number of students from 21 to 500. However, the decrease in Cronbach's alpha for the simulated normal distribution, which is 11.90, as opposed to 38.5 for the uniform distribution, is due to the smaller $S^2$.

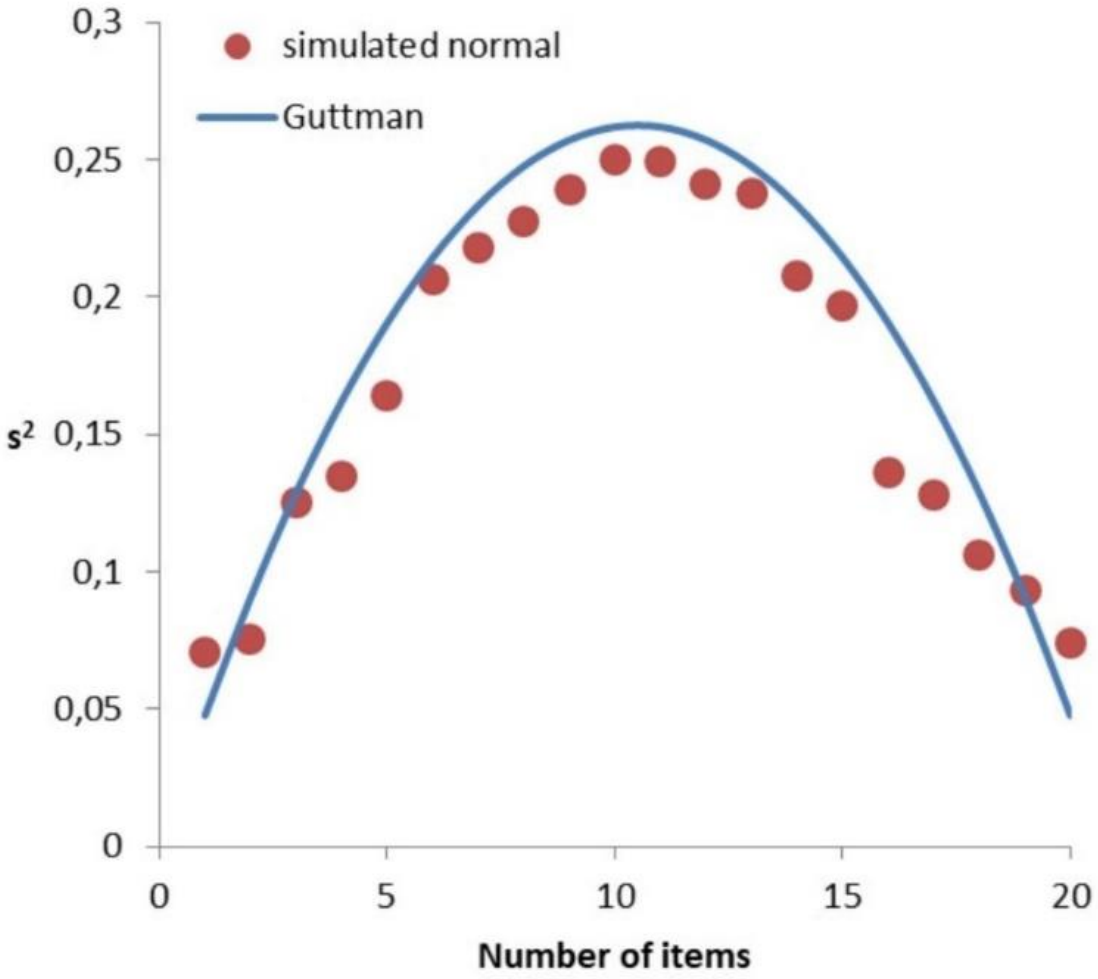



Figure 3. The variation of $s^2$ (item variance) with items arranged in increasing order of difficulty for simulated normal distribution

### 3.2. Normal distribution with 48 items

Before delving into the analysis of the test results, we think it is necessary to clarify what constitutes a typical Cronbach's alpha for scores from a test form with 48 items that are normally distributed. Let's now examine a simulated pattern for 500 individuals using 48 items, where the total scores are normally distributed. The frequency distribution of the total scores is presented in Figure 4, with bins set at a step size of 1. The solid line

represents the theoretical normal distribution curve, which has the same mean, median, and mode as the total scores from the simulated data.

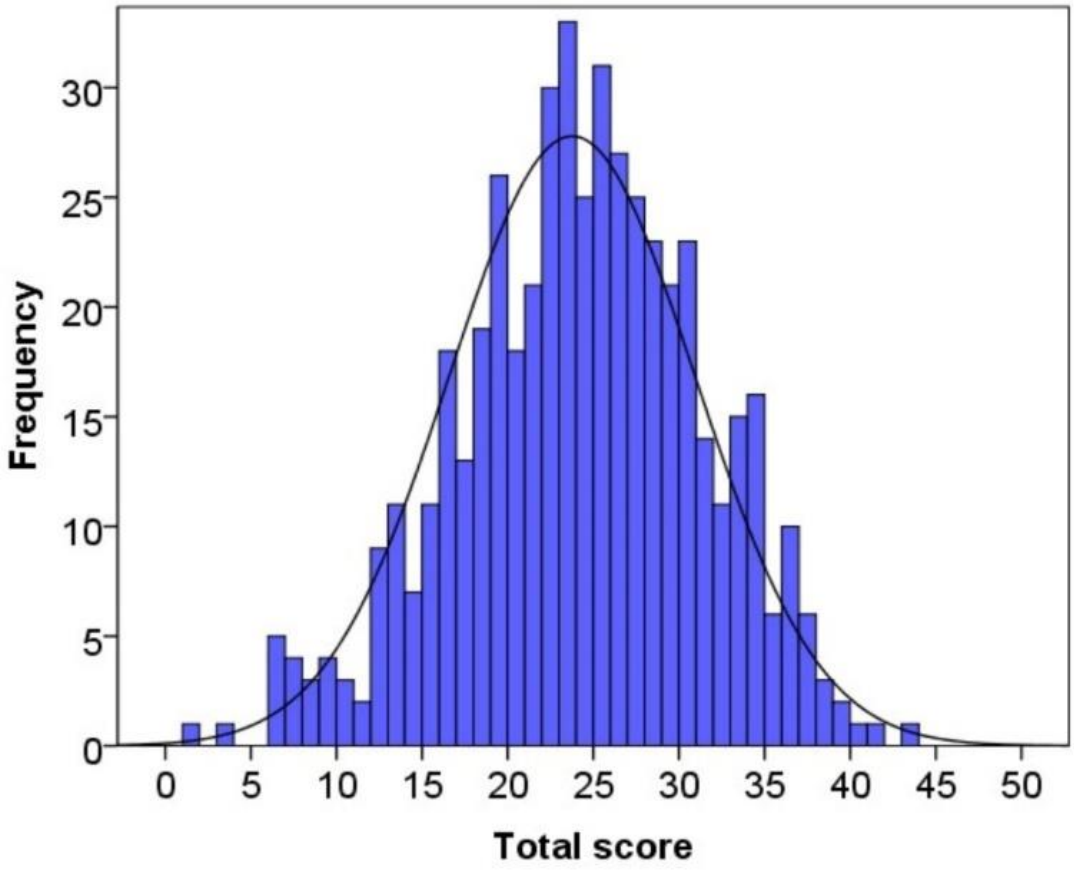



Figure 4. Frequency Distribution of Total Scores from the Simulated Data for 500 Individuals and 48 Items

Descriptive statistics for this dataset are presented in Table 4. As with the earlier 20-item dataset, the total scores from this simulated data closely approximate a normal distribution profile. The alpha for this simulated dataset is 0.86, indicating that the reliability of the test results from this 48-item test form is very close to 'excellent'. However, with careful selection of items and targeted learning strategies, there is potential to significantly improve this reliability score. Typically, a Cronbach's alpha around 0.9 is expected for a test form consisting of 48 items. Notably, the $s^2$ for the simulated normal distribution, and consequently the sum of $s^2$ (8.02), has increased compared to the 20-item dataset due to the larger number of items. However, the increase in Cronbach's alpha is attributed not to this individual item variance but rather to the enhanced effect of the total variance ($S^2$) which is a function of the larger number of items (48).

*Table 4*

**Descriptive Statistics for the Simulated Dataset for 500 Individuals and 48 Items**

| | |
|---|---|
| Number of students | 500 |
| Number of Items | 48 |
| Mean | 23.76 |
| Median | 24 |
| Mode | 24 |
| Variance | 51.50 |
| Standard deviation | 7.18 |
| Skewness | -0.23 |
| Kurtosis | -0.07 |
| Range | 42 |
| Minimum | 1 |
| Maximum | 43 |

### 3.3. Comparative analysis of descriptive statistics for real test results

Having established an understanding of the descriptive statistics of the Guttman pattern and the datasets simulated for 20 and 48 items, we now turn our attention to the analysis of a biology field test designed for 9$^{th}$ – grade students. We prepared a test form consisting of 48 items, anticipating a typical Cronbach's alpha for this number of items. The test was administered to 179 students from academic lyceums, 185 students from secondary schools, and 59 students from training centers.

The frequency distribution of the total scores is presented in Figure 5, with bins set at a step size of 1. The solid line represents the theoretical normal distribution curve, which has the same mean, median, and mode as the total scores from the biology test data.

Descriptive statistics for this dataset are presented in Table 5. Although the mean (27.32) and median (29) of the distribution are relatively close, the mode of this distribution (12) is significantly lower. Visually, as observed in Figure 4. This multifaceted pattern is quantitatively supported by Table 5, where the skewness value indicates a symmetrical distribution, while the kurtosis value resembles that of the uniform distribution discussed earlier in relation to the Guttman pattern.

This distribution pattern may be linked to the three distinct groups (academic lyceums, secondary schools, and training centers) to which the test was administered. Comparing this with simulated data, the variances $s^2$ (item score variance) and $S^2$ (total score variance) show notable differences: $s^2$ increased to 10.45 and $S^2$ to 132.39, compared to 7.18 and 51.50, respectively, in the simulated data for the same number of items. Comparing this with simulated data, the variances $s^2$ (item score variance) and $S^2$ (total score variance) show notable differences: $s^2$ increased to 10.45 and $S^2$ to 132.39, compared to 7.18 and 51.50, respectively, in the simulated data for the same number of items.

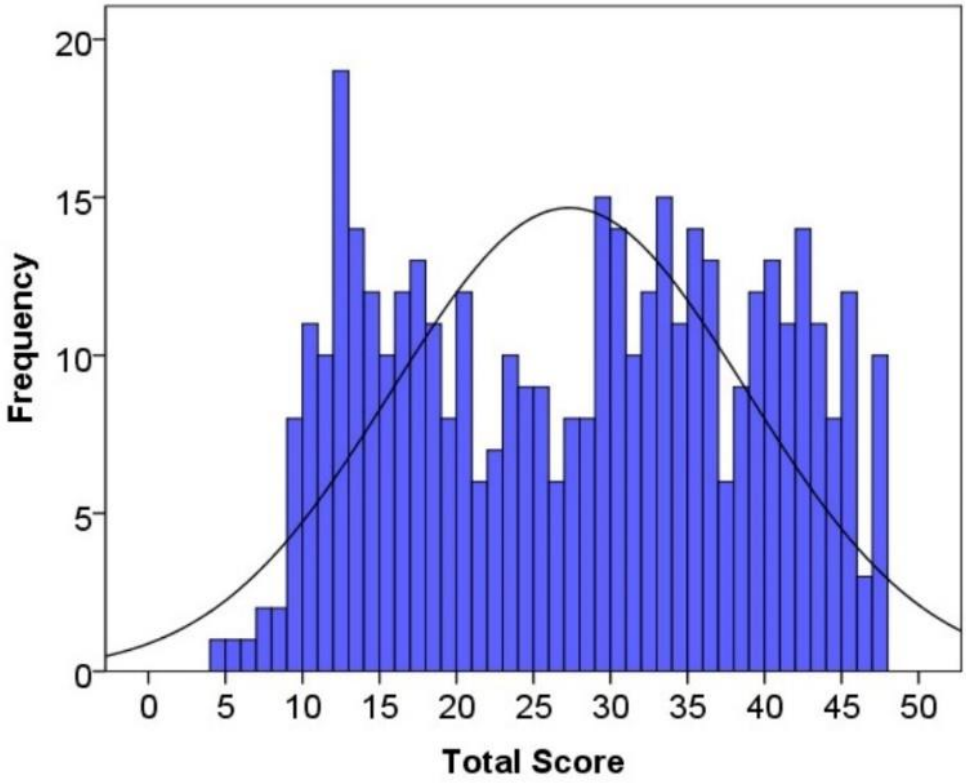



Figure 5. Frequency Distribution of Total Scores for the Dataset from the biology test for 423 Individuals and 48 Items

*Table 5*

**Descriptive Statistics for the Dataset from the biology test**

| | |
|---|---|
| Number of students | 423 |
| Number of Items | 48 |
| Mean | 27.32 |
| Median | 29 |
| Mode | 12 |

| | |
|---|---|
| Variance | 132.39 |
| Standard deviation | 10.45 |
| Skewness | -0.04 |
| Kurtosis | -1.24 |
| Range | 43 |
| Minimum | 4 |
| Maximum | 47 |

These changes suggest that $s^2$ is moving in a direction that could slightly decrease Cronbach's alpha, which is not desirable. Ideally, we would prefer a smaller variance among items, given the sufficient number of items available, as this typically stabilizes the alpha estimate. A higher item variance indicates inconsistency in how individual items measure the underlying ability, which could destabilize the reliability measure provided by Cronbach's alpha. As observed in Fig. 1 and 3, item variance values are larger in the middle part of the graph where the items with medium difficulty are located. This implies that the test form used in the real test contains fewer items with "stronger and weaker" difficulty and more items with medium difficulty compared to both the simulated normal and Guttman distribution. Conversely, excessively small item variances may indicate that the test form contains an imbalance of "easy", "hard" or a majority of "easy" and "hard" items with few "medium" difficulty items. Table 6 presents the percentage of correct answers for each item in the simulated normal distribution from 500 individuals and 48 items, the Guttman pattern similar to that in Table 1 for 48 items, and the real test administered to 423 individuals with 48 items. Analysis of this table reveals distinct patterns in item difficulty across the different testing scenarios. Notably, the real test dataset shows variations in item difficulty compared to the simulated distributions, with a greater proportion of items falling within the medium difficulty range.

*Table 6*

**Percent of correct answer to each item for simulated normal distribution from 500 persons and 48 items, Guttman pattern like in Table 1 for 48 items and real test for 423 persons and 48 items**

| No | Guttman | Real | Normal | No | Guttman | Real | Normal |
|---|---|---|---|---|---|---|---|
| 1 | 97,96 | 88,89 | 93,6 | 25 | 48,98 | 60,76 | 45 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 95,92 | 82,03 | 92,2 | 26 | 46,94 | 59,57 | 44,6 |
| 3 | 93,88 | 81,09 | 89,6 | 27 | 44,90 | 59,57 | 44,2 |
| 4 | 91,84 | 79,91 | 89,4 | 28 | 42,86 | 59,34 | 43,4 |
| 5 | 89,80 | 76,60 | 88,2 | 29 | 40,82 | 58,63 | 37,2 |
| 6 | 87,76 | 74,23 | 87,2 | 30 | 38,78 | 58,16 | 33 |
| 7 | 85,71 | 72,10 | 87 | 31 | 36,73 | 57,92 | 32,2 |
| 8 | 83,67 | 70,45 | 84,6 | 32 | 34,69 | 52,96 | 30,8 |
| 9 | 81,63 | 69,98 | 83,4 | 33 | 32,65 | 51,77 | 26 |
| 10 | 79,59 | 69,74 | 80,4 | 34 | 30,61 | 49,65 | 26 |
| 11 | 77,55 | 68,09 | 79 | 35 | 28,57 | 49,17 | 25,8 |
| 12 | 75,51 | 66,67 | 78,2 | 36 | 26,53 | 47,99 | 24,2 |
| 13 | 73,47 | 66,67 | 76,8 | 37 | 24,49 | 47,04 | 22,8 |
| 14 | 71,43 | 66,19 | 76,6 | 38 | 22,45 | 45,86 | 18,6 |
| 15 | 69,39 | 66,19 | 72,6 | 39 | 20,41 | 43,97 | 18 |
| 16 | 67,35 | 65,25 | 71 | 40 | 18,37 | 39,95 | 15,2 |
| 17 | 65,31 | 65,01 | 70,4 | 41 | 16,33 | 36,64 | 14,4 |
| 18 | 63,27 | 65,01 | 64,8 | 42 | 14,29 | 34,99 | 12,6 |
| 19 | 61,22 | 64,78 | 64,2 | 43 | 12,24 | 34,52 | 10,6 |
| 20 | 59,18 | 63,83 | 63,2 | 44 | 10,20 | 29,08 | 10,2 |
| 21 | 57,14 | 62,88 | 58,6 | 45 | 8,16 | 28,13 | 9,6 |
| 22 | 55,10 | 62,17 | 55,6 | 46 | 6,12 | 21,75 | 9,4 |
| 23 | 53,06 | 61,23 | 54,8 | 47 | 4,08 | 20,80 | 6,2 |
| 24 | 51,02 | 60,99 | 48,8 | 48 | 2,04 | 13,48 | 6 |

Though the significant increase in $S^2$ favourably impacts the alpha value, consequently yielding a Cronbach's alpha of 0.94, slightly higher than that of the simulated normal distribution, this discrepancy may be attributable to the three distinct groups, potentially leading to a more dispersed total score distribution. To explore this assumption further, we can examine the total score distributions of these distinct groups separately. Figures 6-8 display the total score frequencies, with bins set at a step size of 1 for these

groups. The solid lines represent the theoretical normal distribution curve, which have the same means, medians, and modes as the total scores from the biology test data obtained from lyceums, schools, and training centers, respectively. To address potential heterogeneity among test-takers from different educational backgrounds, we conducted a stratified analysis by institution type – academic lyceums, secondary schools, and training centers – as shown in Figures 6–8 and Tables 7–9.

Descriptive statistics for the datasets from lyceums, schools and training centers separately are presented in Table 7-9.

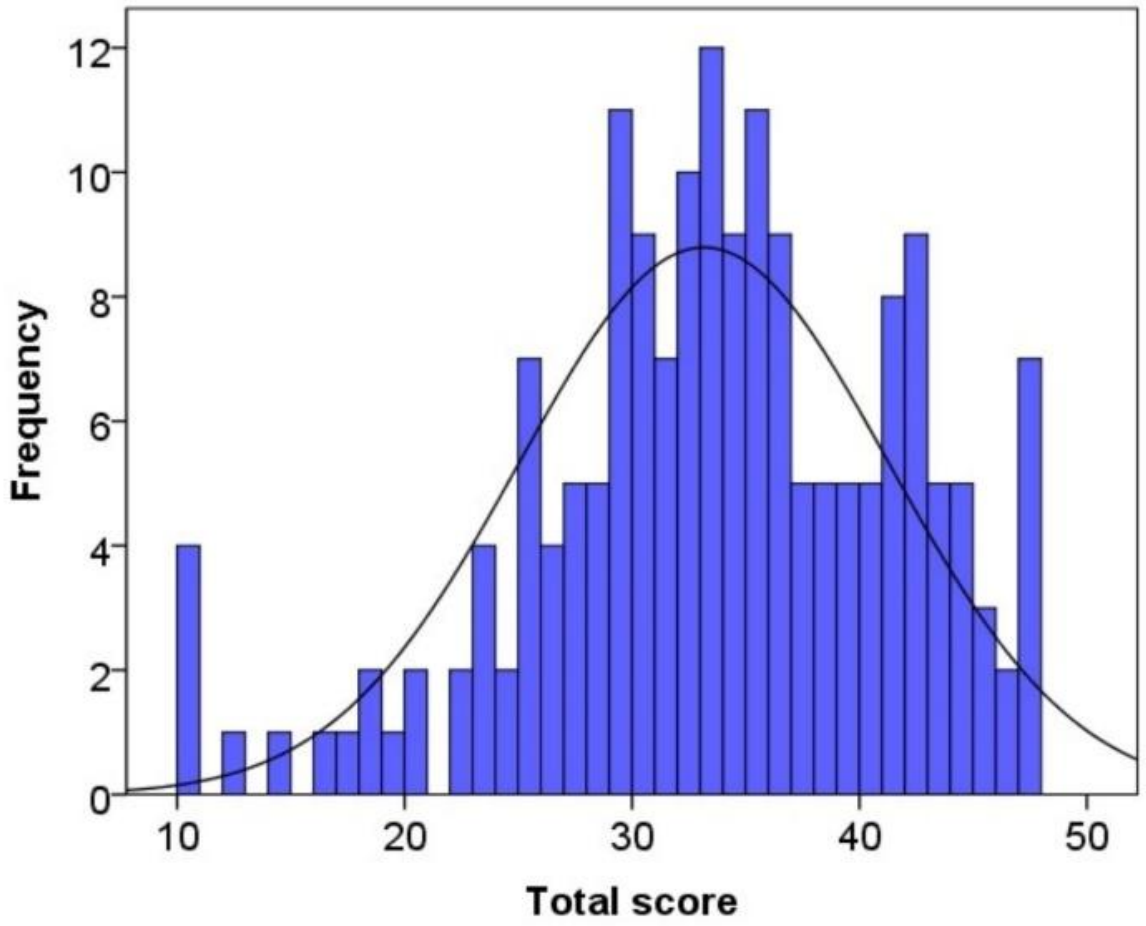



Figure 6. Frequency Distribution of Total Scores for the lyceum for the Dataset from the biology test for 179 Individuals and 48 Items

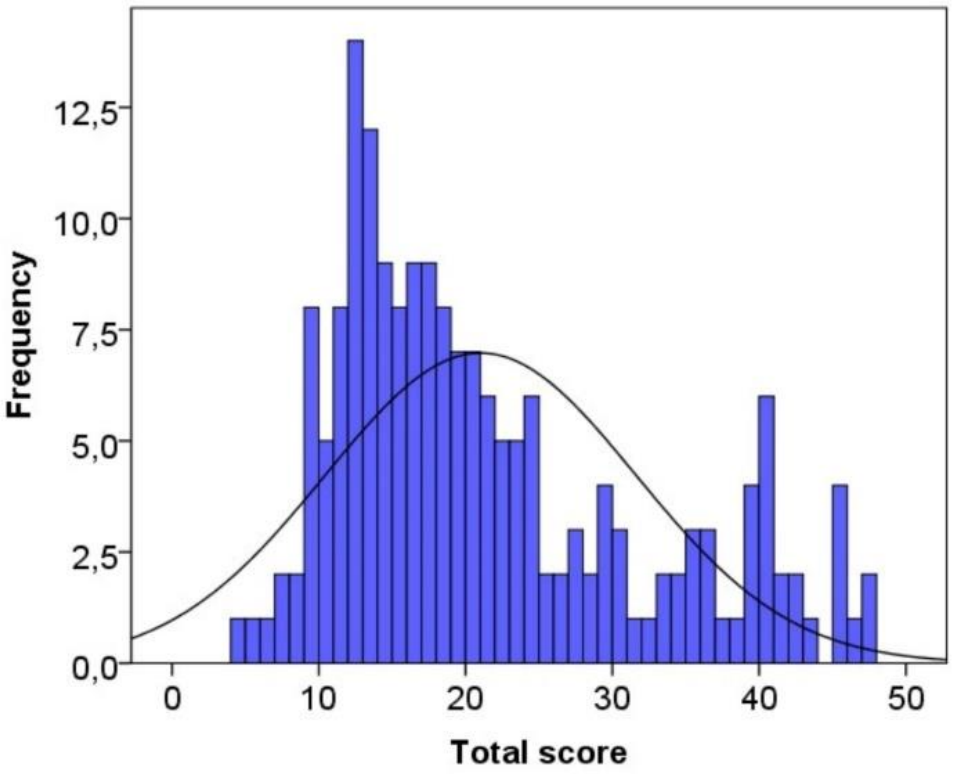



Figure 7. Frequency Distribution of Total Scores for the schools for the Dataset from the biology test for 185 Individuals and 48 Items

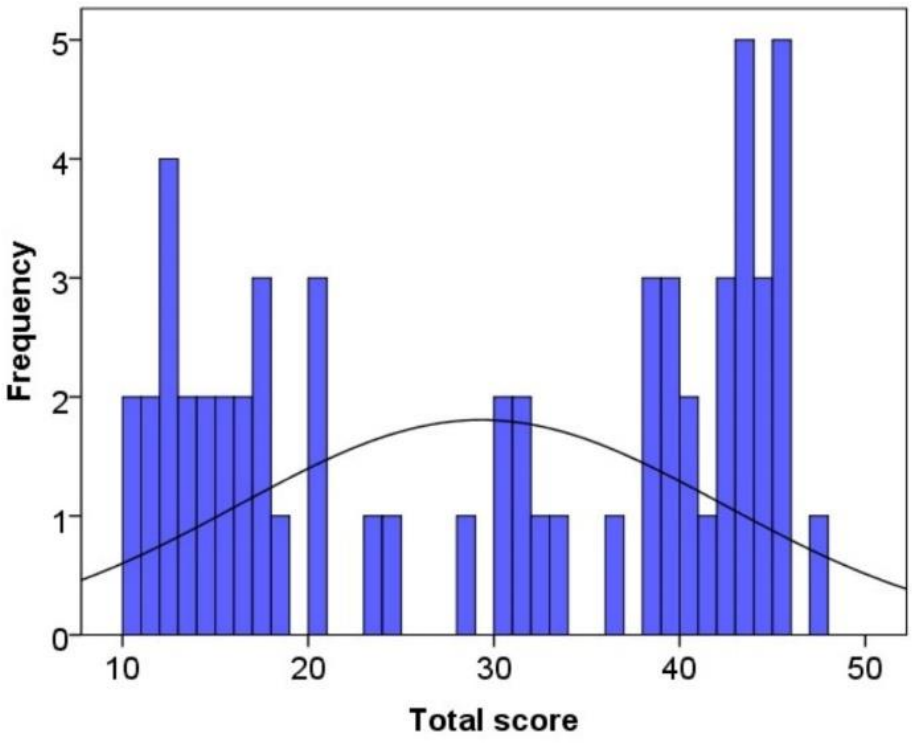



Figure 8. Frequency Distribution of Total Scores for the training centers for the Dataset from the biology test for 59 Individuals and 48 Items

*Table 7*

**Descriptive Statistics for lyceums for the Dataset from the biology test**

| | |
|---|---|
| Number of students | 179 |
| Number of Items | 48 |
| Mean | 33.16 |
| Median | 33 |
| Mode | 33 |
| Variance | 65.97 |
| Standard deviation | 8.12 |
| Skewness | -0.57 |
| Kurtosis | -0.40 |
| Range | 37 |
| Minimum | 10 |
| Maximum | 47 |

*Table 8*

**Descriptive Statistics for schools for the Dataset from the biology test**

| | |
|---|---|
| Number of students | 185 |
| Number of Items | 48 |
| Mean | 21.02 |
| Median | 18 |
| Mode | 12 |
| Variance | 111.92 |
| Standard deviation | 10.58 |
| Skewness | 0.86 |
| Kurtosis | -0.28 |
| Range | 43 |
| Minimum | 4 |
| Maximum | 47 |

*Table 9*

**Descriptive Statistics for training centers for the Dataset from the biology test**

| | |
|---|---|
| Number of students | 59 |
| Number of Items | 48 |
| Mean | 29.32 |
| Median | 31 |
| Mode | 43 |
| Variance | 169.81 |
| Standard deviation | 13.03 |
| Skewness | -0.175 |
| Kurtosis | -1.67 |
| Range | 37 |
| Minimum | 10 |
| Maximum | 47 |

As is seen from Fig. 6 and Table 7 for the lyceums the total score distribution is more likely to the normal distribution, since the mean, median and mode for this distribution are almost the same. Variance (65.97) and standard deviation (8.12) more looks like to those of (51.50, 7.18, respectively) simulated data for 48 items and 500 persons. As a result, Cronbach's alpha 0.89 is also close to that of the simulated normal distribution for 48 items and 500 persons. From these one can imagine "zeros" and "ones" at lower and upper triangles as compared to the Guttman pattern for 48 items – number of "ones" in the lower triangle are more than "zeros" at the upper triangle. Minimum of the total score for the lyceums is 10, maximum is 47 and the range is 37. The students, who would answer less than 10 items correctly, do not exist in this group. This is understandable since the lyceums are specialized in biology.

Slightly higher absolute value of -0.57 of the skewness as compared to the simulated normal distribution (-0.24) shows slightly left skewed distribution - more data concentrated on the right and a longer tail on the left. This also mean for this group the difficulty of test form is a little bit "easier" than those of the "medium" difficulty of the test form for the whole groups. Absolute value of kurtosis -0.40 is much larger than that of the simulated normal distribution (-0.07) and much less than that of the Guttman pattern (-1.2). The frequency

distributions at the total score 10 and 47 affect to the value of kurtosis. These are seen visually in Fig. 6, too.

As seen from Figure 7 and Table 6, the total score distribution for the schools is more likely to be right-skewed. The mean, median, and mode for this distribution are not the same. The variance (169.18) and standard deviation (13.03) are significantly larger than those of the simulated data for 48 items and 500 persons (51.50 and 7.18, respectively). Consequently, Cronbach's alpha is 0.96, higher than that of the simulated normal distribution for 48 items and 500 persons. In this case, the increased kurtosis leads to greater dispersion in the total scores, resulting in a higher Cronbach's alpha value. The minimum total score for the schools is 4, the maximum is 47, and the range is 43, which is larger than in the case of the lyceums. There are no students who scored less than 4, indicating that the schools are not specialized in biology. However, some students answered up to 47 questions correctly, similar to the lyceums.

A higher positive skewness value (0.86) compared to the simulated normal distribution (-0.24) indicates a right-skewed distribution, with more data concentrated on the left and a longer tail on the right. This suggests that the test form is slightly more challenging for this group compared to the overall medium difficulty. The absolute value of kurtosis (-0.28) is larger than that of the simulated normal distribution (-0.07) but less than that of the Guttman pattern (-1.2). This is visually evident in Figure 7.

As seen from Figure 6 and Table 7, the total score distribution for the lyceums is closer to a normal distribution, with the mean, median, and mode being nearly the same. The variance (65.97) and standard deviation (8.12) are more similar to those of the simulated data for 48 items and 500 persons (51.50 and 7.18, respectively).

Consequently, Cronbach's alpha is 0.89, close to that of the simulated normal distribution for 48 items and 500 persons. This suggests that there are more "ones" in the lower triangle compared to the Guttman pattern for 48 items. The minimum total score for the lyceums is 10, the maximum is 47, and the range is 37. No students scored less than 10, which is understandable given that the lyceums specialize in biology.

A slightly higher absolute value of skewness (-0.57) compared to the simulated normal distribution (-0.24) indicates a slightly left-skewed distribution, with more data concentrated on the right and a longer tail on the left. This suggests that the test form is slightly easier for this group compared to the overall medium difficulty. The absolute value of kurtosis (-0.40) is much larger than that of the simulated normal distribution (-0.07) and much less than that of the Guttman pattern (-1.2). This is visually evident in Figure 6.

As seen from Figure 8 and Table 7, the total score distribution for the training centers is closer to a uniform distribution than a normal distribution. The distribution contains three sub-distributions that also resemble uniform distributions, likely due to the presence of beginners, intermediate, and advanced groups, as training centers accept students throughout the year. The mean, median, and mode are not the same. The variance (111.92) and standard deviation (10.58) are larger than those of the simulated data for 48 items and 500 persons (51.50 and 7.18, respectively). Consequently, Cronbach's alpha is 0.92, higher than that of the simulated normal distribution for 48 items and 500 persons. The increased skewness contributes to a higher dispersion in the total scores, resulting in a higher Cronbach's alpha value. The minimum total score for the training centers is 10, the maximum is 47, and the range is 37. No students scored less than 10, which is understandable given that the training centers specialize in biology.

A smaller absolute value of skewness (-0.176) compared to the simulated normal distribution (-0.24) indicates a sufficiently symmetric distribution, with slightly more data concentrated on the left. The absolute value of kurtosis (-1.67) is larger than that of the simulated normal distribution (-0.07) and even larger than that of the Guttman pattern (-1.2), indicating a much flatter peak and thinner tails compared to a normal distribution. This is visually evident in Figure 8.

## Discussion and conclusion

The comparative analysis of the different test patterns reveals several critical insights regarding the reliability and validity of assessment tools. Firstly, the Guttman pattern exemplifies an ideal, yet unrealistic, response distribution that results in a high Cronbach's alpha. This ideal pattern serves as a useful benchmark for understanding the upper limits of internal consistency within a test. However, real-world tests rarely achieve such perfection, reflecting more complex and varied patterns of student responses. The real test data, derived from a biology field test, shows a distribution that deviates from the idealized Guttman pattern but still maintains a significant degree of internal consistency as indicated by its Cronbach's alpha. This suggests that while perfect patterns are unattainable, a well-designed test can still exhibit strong reliability. The mean, median, and mode discrepancies, along with the multimodal distribution observed, hint at underlying differences in the test-taker population, likely attributable to the distinct groups (academic lyceums, secondary schools, and training centers) involved in the study.

Furthermore, the analysis underscores the importance of item difficulty distribution in shaping the

reliability of a test. The real test's item difficulties, predominantly falling in the medium range, suggest that balanced item selection is crucial for maintaining internal consistency without inflating Cronbach's alpha artificially. This balance ensures that the test measures a broad spectrum of abilities while avoiding redundancy or overly homogeneous item difficulties.

The observation that medium-difficulty items exhibit higher variance suggests that such items contribute most to test discriminability, particularly among test-takers of average ability. This pattern aligns with psychometric principles, where items located near the population mean offer the greatest information. These findings can inform item selection strategies by highlighting the value of including a balanced distribution of item difficulties, with emphasis on moderately challenging items to enhance measurement precision.

Comparative statistics from simulated normal distributions with 20 and 48 items reinforce the understanding that increasing the number of items generally enhances reliability, provided the items are well-calibrated. The significant increase in Cronbach's alpha observed in the 48-item simulated dataset confirms this, aligning with the expectations for tests of this length.

The study highlights the importance of analyzing test patterns and understanding the components influencing reliability metrics like Cronbach's alpha. While ideal patterns like the Guttman scale offer theoretical benchmarks, practical test design must account for the inherent variability in real-world data. By carefully selecting items and considering their difficulty distribution, educators and test designers can develop reliable assessment tools that accurately measure student abilities and provide meaningful insights into their performance.

The findings from the biology field test, supported by simulated data, highlight the importance of test structure in achieving high reliability. A well-designed test can produce strong reliability metrics, but excessively high Cronbach's alpha values can raise concerns beyond those discussed by Cortina [4] and Nunnally & Bernstein [1]. While a high alpha is often seen as a positive indicator of internal consistency, values that are too high may also signal potential issues such as item redundancy or imbalanced item difficulty.

Redundancy occurs when multiple items are essentially measuring the same construct in highly similar ways, leading to inflated alpha scores without necessarily improving the overall quality of the test. This lack of diversity in item content can diminish the test's ability to capture the full breadth of the underlying construct. Similarly, an imbalance in item difficulty – where too many items are either too easy or too difficult for the target population – can lead to an artificially high alpha,

masking deeper issues with the test's discriminatory power across different ability levels.

Therefore, excessively high alpha values should prompt a careful review and adjustment of the test. It is essential to ensure that the items are varied enough to cover the complexity of the domain being assessed while avoiding unnecessary overlap that may inflate reliability metrics without adding value to the measurement process.

Ultimately, the goal is to design balanced and reliable assessments that not only achieve strong psychometric properties but also reflect the diversity and complexity of the population being tested. This requires a thoughtful approach to test development, ensuring that the test captures a wide range of abilities and nuances while avoiding overemphasis on specific subdomains or redundant item content. Achieving this balance enhances the test's ability to provide meaningful and accurate information about the constructs of interest, leading to more valid inferences about the tested population.

Future research could further explore the impact of different test-taker populations on reliability metrics and investigate methods for optimizing item selection to enhance both reliability and validity. Additionally, applying these findings across various subjects and educational levels would help generalize the conclusions and support the development of robust assessment tools in diverse educational contexts.

## References


1. J. Nunnally & L. Bernstein. Psychometric theory (3rd ed.). McGraw-Hill. 1994.
2. R. Cohen & M. Swerdlik. Psychological testing and assessment. McGraw-Hill Higher Education. 2010.
3. N. Schmitt. Uses and abuses of coefficient alpha. Psychological Assessment, 8(4), 350-353, 1996.
4. J. M. Cortina. What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78(1), 98-104, 1993.
5. L. J. Cronbach. Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334, 1951.
6. E. A. Beckmann & K. E. Jastrowski Mano. Initial development and validation of the School Anxiety Inventory–College Version (SAI-CV). Psychology in the Schools, 60, 2540–2563, 2023. https://doi.org/10.1002/pits.22879.
7. G. Bharara & S. Duncan. Preliminary development and validation of the positive school transition readiness survey (PSTRS). Psychology in the Schools, 61, 1217–1237, 2023. https://doi.org/10.1002/pits.23108.
8. F. Chen et al. Enhancing children's well-being using a Malaysian-adapted version of Super Skills for Life. Psychology in the Schools, 61, 3894–3907, 2024. https://doi.org/10.1002/pits.23258.
9. R. P. McDonald. Test theory: A unified treatment. Lawrence Erlbaum Associates, 1999.
10. F. Orcan. Comparison of Cronbach's alpha and McDonald's omega for ordinal data: Are they different? International Journal of Assessment Tools in Education, 10(3), 709–722, 2023. https://doi.org/10.21449/ijate.1271693.
11. L. Guttman. The basis for scalogram analysis. In S. A. Stouffer et al. (Eds.), Measurement and prediction, Princeton University Press, pp. 60-90, 1950.

12. D. W Zimmerman, R. H. Williams, B. D. Zumbo & D. Ross. Louis Guttman's Contributions to Classical Test Theory. International Journal of Testing, 5(1), 81–95, 2009. https://doi.org/10.1207/s15327574ijt0501_7.

13. B. Griffin. Calculating Cronbach's alpha in Excel. Psychology of Business. 2023. https://psychologyofbusiness.beehiiv.com/p/calculating-cronbachs-alpha-excel.

14. K. P. Balanda & H. L. MacGillivray. Kurtosis: A critical review. The American Statistician, 42(2), 111–119, 1998.

15. S. Brown. Measures of shape: Skewness and kurtosis. 2011. https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/Simon?action=AttachFile&do=get&target=kurtosis.pdf. 2011.

# O‘ZLASHTIRISH TESTLARIDA ICHKI MUVOFIQLIK: KRONBAX ALFA VA GUTTMAN NAMUNASINING ROLI

**M.DJ. Ermamatov, A.B. Normurodov, A.R. Sattiyev**

*Bilim va malakalarni baholash agentligi huzuridagi Ilmiy-o‘quv amaliy markazi, 100084, Toshkent sh., Bog‘ishamol k., 12, a.normurodov@uzbmb.uz*

**Qisqacha mazmuni.** Ushbu maqolada test ballarining ichki muvofiqligi va ishonchliligini ta’minlashda, baholashda Cronbach alfa ko‘rsatkichining ahamiyati o'rganildi. Keng qamrovli tahlil orqali test natijalarining javoblar matritsasi Guttman namunasiga qanchalik mos kelishi ko‘rsatilib, Kronbax alfa yordamida xulosa chiqarishda topshiriqlar sonining muhimligi ta’kidlandi. Ilmiy tadqiqot uchun biologiya fanidan 423 nafar o‘quvchi ishtirokida test sinovi o‘tkazildi va testning validligi hamda ishonchliligi statistik usullar bilan tahlil qilindi. Tadqiqotda Guttman namunasining batafsil tahlili, uning test tahlilidagi nimani anglatishi hamda Kronbax alfani turli testlar uchun qanday talqin qilinishi keng yoritilgan. Shuningdek, normal taqsimotga ega simulyatsiya qilingan va haqiqiy test ma’lumotlarining tavsif statistikasi hamda ishonchlilik qiymatlari taqqoslanib, Kronbax alfaning amaliy qo‘llanilishi va cheklovlari muhokama qilingan. Natijalar topshiriqlarni ehtiyotkorlik bilan tanlash va testni puxta loyihalash baholashning ishonchliligini hamda talqin qilinish imkoniyatlarini oshirishda naqadar muhim ekanini ko‘rsatdi.

**Kalit so‘zlar:** baholash, klassik test nazariyasi, o’zlashtirish testlari.

# HUQUQSHUNOSLIK FANIDAN O‘TKAZILGAN TEST SINOVI NATIJALARI TAHLILI

**A.R. Sattiyev, M.Dj. Ermamatov**

*Bilim va malakalarni baholash agentligi huzuridagi Ilmiy-o‘quv amaliy markazi, 100084, Toshkent sh., Bog‘ishamol k., 12.*

**Qisqacha mazmuni.** Ushbu maqolada huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 – o‘quv yilida bakalavriatga kirish test sinovlari natijalarining klassik test nazariyasi asosida olingan statistik ko‘rsatkichlari - o‘rtacha qiymati, mediana, moda, o‘lchashning standart xatoligi va distraktorlari o‘rganilgan. Tahlil orqali test sinovlarida foydalanilgan test topshiriqlarining distraktorlari samarali ishlayotganligi aniqlandi.

**Kalit so‘zlar:** test topshiriqlari, klassik test nazariyasi, distraktorlar.

## 1. Kirish

Bilimlarni baholashda eng samarali va keng qo‘llaniladigan pedagogik o‘lchov vositalaridan biri – bu test hisoblanadi. Standart test natijalari yordamida turli guruhlardagi har xil qobiliyat darajasidagi talabgorlarning bilim darajasini solishtirish, o‘quv dasturini o‘zlashtirish ko‘rsatkichlarini aniqlash, pedagoglar hamda ta’lim muassasalari faoliyati samaradorligini baholash, ularning yutuq va kamchiliklarini tahlil qilish imkoniyati yaratiladi. Shu sababli standart test variantlarini ishlab chiqish jarayonida aprobatsiya test sinovlari bilan test natijalarini sifatini ilmiy asoslash muhim ahamiyat kasb etadi. Buning natijasida test topshiriqlari sifatini takomillashtirish va baholashning ishonchliligi va validligini ta’minlash mumkin bo‘ladi [1]. Turli xil guruhlardagi talabgorlarning bilimlarini baholashda standart test variantidan [2] yoki bir-biriga parallel bo‘lgan test variantlaridan foydalaniladi. Bitta test variantidan foydalanilganda baholashdagi standartlikni ta’minlash nisbatan oson, lekin bir nechta variantlardan foydalanilganda esa turli xil guruhlarga turli xil qiyinlikdagi variantlar to‘g‘ri kelib qolishini hisobga olmaslik adolatlilik tamoyilining buzilish muammosiga olib kelishi mumkin. Shu sababli parallel test variantlaridan foydalanish bu muammolarni qisman bartaraf etadi. Testlarning imkon qadar parallelligini ta’minlash uchun test spetsifikatsiyasiga qat’iy rioya qilish talab etiladi [3]. Pedagogik o‘lchovlarning nazariy asoslariga ko‘ra, har bir test

topshirig‘ini dastlab, baholashdagi sifat mezonlariga to‘liq amal qilgan holda tuzilganligini aniqlash uchun eks-pertizadan o‘tkaziladi. Shundan so‘ng aprobatsiya test sinovlari o‘tkazilib, olingan natijalar asosida test topshiriqlarining statistik xarakte-ristikalari aniqlanadi. Ushbu jarayon testlarning sifat ko‘rsatkichlarini baholash, ularning ishonchliligi va aniqligini tashxislash imkonini beradi. Aprobatsiya sinovlari real testlar to‘g‘risida ilmiy asoslangan xulosalar chiqarish uchun muhim manba bo‘lib xizmat qiladi. Statistik tahlil natijalari esa test topshiriqlariga qo‘yilgan asosiy talablarning bajarilish darajasini aniqlash imkonini beradi [4-5].

Test topshiriqlarining dis-traktorlari tahlili ham ularning baholash maqsadlariga muvofiqligi haqida muhim ma’lumot beradi. Test topshirig‘i savoldan, to‘g‘ri javobni va distraktorlarni o‘z ichiga olgan muqobil javoblardan iborat bo‘lib, bunda samarali distraktorlarni tanlash muhim o‘rin tutadi [6]. Odatda mutaxassislar muqobil javobli test topshiriqlarining distraktorlarini tuzish uchun savol tuzishga nisbatan ozroq vaqt sarflaydilar, lekin muqobil javoblarni tanlash ham malaka va tajriba talab qiladi.

Distraktorlarning to‘g‘ri javobdan farqlanishi bilan birga to‘g‘ri javobga o‘xshagan bo‘lishi maqsadga muvofiqdir. Tadqiqotlarda dis-traktorlik vazifasini bajarmaydigan javoblar ham uchrab turadi va bu odatda juda kam tanlangan distraktorlardir. Bu borada mezon sifatida 5 foizdan kam va 95 foizdan ko‘p tanlangan distraktorlarni ishlamaydigan distraktorlar deb olinadi [7-8], shuning uchun bunday distraktorlarni o‘zgartirish tavsiya qilinadi, lekin test topshiriqlarining umumiy ball bilan korrelyatsiyasi yaxshi bo‘lganda ular juda yuqori o‘zlashtish darajasiga va juda past o‘zlashtirish darajasiga ega bo‘lgan talabgorlarni aniqlab berishi mumkin, odatda bunday talabgorlar juda kam bo‘ladi.

Ushbu maqolada huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlari natijalarining klassik test nazariyasi asosida qilingan tahlil natijasida olingan ma’lumotlar solishtirilgan va o‘rganilgan. O‘tkazilgan test sinovlarida foydalanilgan test varianlarining statistik xarakte-ristikalari - o‘rtacha qiymati, mediana, moda, o‘lchashning standart xatoligi haqida xulosalar keltirilgan. Bundan tashqari mazkur test sinovlarida foydalanilgan test topshiriqlarining distraktorlari ham o‘rganib chiqilgan.

## 2. Natijalar

Huquqshunoslik fanidan 2025-2026 - o‘quv yilida o‘zbek va rus tili guruhlari uchun bakalavriatga kirish test sinovi natijalari klassik test nazariyasi bilan tahlil qilindi. Olingan natijalar orqali joriy o‘quv yilida mazkur fandan o‘zbek va rus tili guruhlarida test sinovida foydalanilgan test variantlaridagi test top-shiriqlarining statistik xarakte-ristikalari - o‘rtacha qiymati, mediana, moda, o‘lchashning standart xatoligi aniqlandi.

1-jadvalda 2025-2026 – o‘quv yilida huquqshunoslik fanidan (o‘zbek va rus guruhlari uchun) bakalavriatga kirish test sinovlari natijalarining yuqorida aytib o‘tilgan qiymatlari solishtirilgan.

Test topshiriqlari o‘zlashtirilishi kerak bo‘lgan o‘quv dasturlari bo‘yicha tuzilganligini inobatga olganda, o‘rtacha qiymat taxminan 15 ball ekanligi yaxshi ko‘rsatkich hisoblanadi, chunki odatda o‘rtacha qiymat test variantidagi test topshiriqlari 30 ta bo‘lganda 15 ball atrofida bo‘lishi maqsadga muvofiqdir. Bunday holda test variantining barcha qobiliyat darajalari haqida beradigan ma’lumoti ko‘p bo‘ladi va ballar taqsimoti qanchalik normalligi haqida xulosalarning ishonchliligi yuqori bo‘ladi.

Jadvaldan o‘zbek va rus guruhlari uchun foydalanilgan test variantlari asosida olingan natijalarining qiymatlari (yuqorida aytib o‘tilgan kattaliklari) o‘zbek guruhlariga nisbatan rus guruhlarida biroz farqli ekanligini ko‘rish mumkin. Rus tili guruhlari natijalarining o‘rtacha qiymatlari va medianasi o‘zbek tili guruhlariga nisbatan biroz past ekanligini ko‘rish mumkin. Lekin xatolik doirasida ular bir-biriga juda yaqin.

Test sinovlarida foydalanilgan variantlar bo‘yicha tahlil natijalari bir-biriga yaqinligi ularning parallelligi imkon qadar ta’minlangan deyish mumkin. Biroq shuni ta’kidlab o‘tish lozimki, test variantlarning pa-rallelligini to‘liq ta’minlash imkoni yo‘q bo‘lib, bunday muammolarni hal etish uchun bir necha usullar ishlab chiqilgan [9].

*1-jadval*

**2025-2026- o‘quv yilida huquqshunoslik fanidan (o‘zbek va rus guruhlari uchun) bakalavriatga kirish test sinovlari natijalari haqida statistik ma’lumotlar**

| Kun | Smena | Variant | O‘rta qiymat | | Mediana | | Moda | | O‘lchashning standart xatoligi | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | *o‘zbek guruhi* | *rus guruhi* | *o‘zbek guruhi* | *rus guruhi* | *o‘zbek guruhi* | *rus guruhi* | *o‘zbek guruhi* | *rus guruhi* |
| 1 | 1 | 1 | 15,42 | 14,19 | 15,0 | 14,0 | 11 | 14 | 2,44 | 2,48 |
| | | 2 | 15,70 | 15,32 | 15,0 | 15,0 | 10 | 16 | 2,42 | 2,47 |
| | | 3 | 15,49 | 14,72 | 15,0 | 14,0 | 12 | 11 | 2,41 | 2,43 |
| 1 | 2 | 1 | 16,12 | 14,01 | 15,0 | 13,0 | 11 | 9 | 2,40 | 2,43 |
| | | 2 | 15,80 | 14,25 | 15,0 | 13,0 | 11 | 8 | 2,39 | 2,42 |
| | | 3 | 15,74 | 14,47 | 15,0 | 14,0 | 10 | 11 | 2,45 | 2,45 |

Statistik ko‘rsatkichlari yaxshi bo‘lishi uchun test topshiriqlari qobiliyat darajalarini ajrata olishi, test topshiriqlari sonidan tashqari o‘zlashtirish darajasi yuqori bo‘lgan abituriyentlar soni ham yetarli darajada ko‘p bo‘lishi kerakligini ta’kidlab o‘tish lozim. Tayyorgarlik darajasi past bo‘lgan abituriyentlarning ko‘pligi ularning qobiliyat bo‘yicha yaxshi ajralmasligiga va buning natijasida statistik ko‘rsatkichlarning pasayib ketishiga olib keladi. Tayyorgarlik darajasi yuqori bo‘lgan abituriyentlar ko‘p bo‘lishini va aksincha, o‘zlashtirish past bo‘lgan abituriyentlar juda kam miqdorda bo‘lishi hisobiga statistik ko‘rsatkichlar yuqori bo‘lishini milliy sertifikat bo‘yicha test sinovlari natijalarining tahliliga doir maqolalarda ko‘rish mumkin [4,5].

Huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 – o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining distraktorlari o‘rganib chiqildi. 2-3-jadvallarda o‘zbek tili va 4-5-jadvallarda esa rus tili guruhlari uchun (kunlar va variantlar kesimida) test sinovida foydalanilgan test topshiriqlarining distraktorlari haqida ma’lumotlar keltirilgan bo‘lib, ushbu jadvallarda test topshiriqlariga abituriyentlarning tanlagan muqobil javoblari foizi keltirilgan. Bu yerda: A*-

kalit, B, C va D distraktorlar, NA-belgilanmagan yoki ikki va undan ortiq belgilangan javoblardir. Endi mazkur jadvallar uchun alohida to‘xtalib o‘tamiz. 2-jadvalda o‘zbek guruhlari uchun (1-kun 1-smena) 1-, 2- va 3-variantlarda foydalanilgan test topshiriqlarining barchasining dis-traktorlari yaxshi ishlayotganligi ko‘rinadi. 3-jadvaldan esa o‘zbek guruhlari uchun (1-kun 2-smena) 1-variantda foydalanilgan 1 ta test topshirig‘ining (22-test topshirig‘i) C distraktorini abituriyentlarning 3,86 foizi tanlaganlar. 2-variantda foy-dalanilgan 3 ta test topshiriqlari (4, 22-va 28-test topshiriqlari) ning mos ravishda C, B va D distraktorlarini abituriyentlarning mos ravishda 4,36, 3,41 va 4,43 foizi tanlaganlar. 3-variantda foydalanilgan test top-shiriqlaridan esa 1 ta test topshirig‘i (22-test topshirig‘i)ning B distraktorini abituriyentlarning 3,35 foizi tan-laganlar. O‘zbek guruhlari uchun 1- va 2-smenalarda foy-dalanilgan test topshiriqlarining distraktorlarini 95 foiz va undan yuqori tanlagan abituriyentlar kuzatilmadi.

*2-jadval*

**2025-2026- o‘quv yilida huquqshunoslik fanidan bakalavriatga kirish test sinovlari natijalari haqida statistik ma’lumotlar (*o‘zbek guruhi, 1-kun, 1-smena*)**

| Test topshiriqlari raqami | Kalit | Distraktorlar | | | NA, % |
|---|---|---|---|---|---|
| | A, % | B, % | C, % | D, % | |
| **1-variant** | | | | | |
| 1 | 52,41 | 14,04 | 18,28 | 14,97 | 0,30 |
| 2 | 48,29 | 21,65 | 12,99 | 16,90 | 0,16 |
| 3 | 64,70 | 19,65 | 8,27 | 7,05 | 0,33 |
| 4 | 47,22 | 18,79 | 18,46 | 15,20 | 0,33 |
| 5 | 53,15 | 12,43 | 16,27 | 17,74 | 0,40 |
| 6 | 46,10 | 15,97 | 22,54 | 15,09 | 0,30 |
| 7 | 42,21 | 23,07 | 19,00 | 15,41 | 0,30 |
| 8 | 68,82 | 10,50 | 12,15 | 8,24 | 0,28 |

| | | | | | |
|---|---|---|---|---|---|
| 9 | 53,53 | 12,69 | 16,23 | 17,18 | 0,37 |
| 10 | 29,45 | 12,69 | 22,14 | 35,44 | 0,28 |
| 11 | 26,87 | 45,19 | 13,29 | 14,39 | 0,26 |
| 12 | 57,37 | 14,37 | 17,72 | 10,31 | 0,23 |
| 13 | 67,05 | 8,54 | 15,04 | 9,03 | 0,33 |
| 14 | 42,44 | 18,88 | 10,36 | 27,94 | 0,37 |
| 15 | 51,64 | 26,59 | 13,60 | 7,94 | 0,23 |
| 16 | 69,17 | 11,62 | 11,04 | 7,94 | 0,23 |
| 17 | 22,70 | 34,44 | 20,88 | 21,54 | 0,44 |
| 18 | 62,93 | 12,25 | 11,29 | 13,27 | 0,26 |
| 19 | 53,08 | 16,14 | 11,34 | 19,14 | 0,30 |
| 20 | 37,46 | 12,46 | 12,53 | 37,18 | 0,37 |
| 21 | 66,05 | 17,23 | 8,96 | 7,45 | 0,30 |
| 22 | 56,27 | 15,95 | 16,58 | 10,83 | 0,37 |
| 23 | 38,23 | 14,46 | 10,85 | 36,11 | 0,35 |
| 24 | 62,68 | 9,08 | 18,42 | 9,52 | 0,30 |
| 25 | 42,33 | 37,53 | 9,22 | 10,66 | 0,26 |
| 26 | 40,70 | 17,44 | 23,54 | 18,04 | 0,28 |
| 27 | 75,27 | 9,69 | 7,64 | 7,01 | 0,40 |
| 28 | 68,68 | 8,31 | 5,45 | 17,30 | 0,26 |
| 29 | 46,73 | 16,62 | 20,77 | 15,62 | 0,26 |
| 30 | 51,99 | 16,60 | 17,69 | 13,46 | 0,26 |
| **2-variant** | | | | | |
| 1 | 51,11 | 25,05 | 7,09 | 16,50 | 0,25 |
| 2 | 52,16 | 14,02 | 10,78 | 22,76 | 0,27 |
| 3 | 52,30 | 9,33 | 24,04 | 13,96 | 0,37 |
| 4 | 39,44 | 25,65 | 18,33 | 16,24 | 0,34 |

| | | | | | |
|---|---|---|---|---|---|
| 5 | 53,74 | 22,65 | 12,61 | 10,84 | 0,16 |
| 6 | 46,76 | 16,40 | 18,23 | 18,33 | 0,27 |
| 7 | 67,88 | 9,72 | 10,09 | 12,13 | 0,18 |
| 8 | 27,96 | 17,41 | 43,15 | 11,23 | 0,25 |
| 9 | 53,69 | 17,82 | 13,75 | 14,34 | 0,39 |
| 10 | 35,39 | 19,61 | 29,22 | 15,60 | 0,18 |
| 11 | 26,31 | 19,79 | 12,10 | 41,41 | 0,39 |
| 12 | 61,70 | 13,13 | 11,37 | 13,57 | 0,23 |
| 13 | 69,82 | 9,11 | 7,14 | 13,75 | 0,18 |
| 14 | 29,90 | 50,65 | 8,99 | 10,18 | 0,27 |
| 15 | 42,94 | 12,10 | 28,96 | 15,67 | 0,32 |
| 16 | 58,38 | 9,31 | 6,31 | 25,74 | 0,25 |
| 17 | 67,99 | 6,54 | 12,26 | 13,02 | 0,18 |
| 18 | 70,19 | 7,69 | 13,32 | 8,53 | 0,27 |
| 19 | 58,06 | 12,31 | 15,19 | 14,12 | 0,32 |
| 20 | 62,73 | 13,54 | 11,16 | 12,38 | 0,18 |
| 21 | 68,47 | 13,11 | 6,77 | 11,51 | 0,14 |
| 22 | 52,18 | 13,13 | 17,34 | 16,98 | 0,37 |
| 23 | 45,09 | 23,95 | 17,43 | 13,13 | 0,39 |
| 24 | 60,22 | 15,69 | 15,35 | 8,56 | 0,18 |
| 25 | 53,56 | 21,62 | 8,46 | 16,20 | 0,16 |
| 26 | 52,44 | 17,30 | 18,58 | 11,44 | 0,25 |
| 27 | 61,73 | 11,76 | 10,71 | 15,51 | 0,30 |
| 28 | 48,64 | 27,84 | 13,73 | 9,54 | 0,25 |
| 29 | 49,94 | 12,83 | 12,13 | 24,64 | 0,46 |
| 30 | 54,98 | 15,08 | 14,87 | 14,66 | 0,41 |
| **3-variant** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 1 | 54,68 | 21,30 | 6,93 | 16,74 | 0,35 |
| 2 | 54,01 | 10,07 | 23,39 | 12,32 | 0,21 |
| 3 | 67,66 | 8,90 | 14,67 | 8,67 | 0,09 |
| 4 | 29,44 | 18,81 | 35,90 | 15,46 | 0,40 |
| 5 | 51,34 | 20,51 | 12,44 | 15,44 | 0,28 |
| 6 | 48,97 | 14,69 | 17,04 | 19,11 | 0,19 |
| 7 | 24,37 | 29,57 | 25,88 | 19,90 | 0,28 |
| 8 | 67,64 | 13,95 | 11,58 | 6,58 | 0,26 |
| 9 | 47,11 | 9,83 | 27,53 | 15,37 | 0,16 |
| 10 | 23,23 | 18,00 | 44,66 | 13,90 | 0,21 |
| 11 | 40,13 | 21,00 | 20,37 | 18,16 | 0,35 |
| 12 | 66,47 | 9,88 | 11,21 | 12,25 | 0,19 |
| 13 | 50,57 | 8,86 | 10,14 | 30,27 | 0,16 |
| 14 | 42,34 | 27,57 | 15,69 | 14,09 | 0,30 |
| 15 | 52,31 | 6,49 | 6,72 | 34,34 | 0,14 |
| 16 | 70,10 | 9,83 | 10,07 | 9,81 | 0,19 |
| 17 | 60,01 | 7,25 | 15,62 | 16,86 | 0,26 |
| 18 | 67,66 | 7,77 | 16,81 | 7,49 | 0,28 |
| 19 | 58,71 | 15,67 | 10,72 | 14,74 | 0,16 |
| 20 | 63,92 | 10,21 | 12,02 | 13,76 | 0,09 |
| 21 | 60,66 | 15,58 | 11,88 | 11,74 | 0,14 |
| 22 | 73,36 | 10,74 | 9,42 | 6,35 | 0,14 |
| 23 | 45,92 | 14,32 | 17,67 | 21,86 | 0,23 |
| 24 | 54,41 | 9,86 | 18,18 | 17,39 | 0,16 |
| 25 | 38,39 | 17,76 | 31,69 | 12,02 | 0,14 |
| 26 | 43,73 | 31,69 | 10,65 | 13,72 | 0,21 |
| 27 | 63,43 | 15,67 | 9,04 | 11,58 | 0,28 |

| | | | | | |
|---|---|---|---|---|---|
| 28 | 44,92 | 8,37 | 24,39 | 21,97 | 0,35 |
| 29 | 35,76 | 19,74 | 26,69 | 17,58 | 0,23 |
| 30 | 52,13 | 16,76 | 17,11 | 13,76 | 0,23 |

*3-jadval*

**2025-2026- o‘quv yilida huquqshunoslik fanidan bakalavriatga kirish test sinovlari natijalari haqida statistik ma’lumotlar (*o‘zbek guruhi, 1-kun, 2-smena*)**

| Test topshiriqlari raqami | Kalit | Distraktorlar | | | NA, % |
|---|---|---|---|---|---|
| | A, % | B, % | C, % | D, % | |
| **1-variant** | | | | | |
| 1 | 62,86 | 13,33 | 10,40 | 13,18 | 0,23 |
| 2 | 55,35 | 14,04 | 7,36 | 22,87 | 0,38 |
| 3 | 46,75 | 12,88 | 18,40 | 21,67 | 0,30 |
| 4 | 45,29 | 14,46 | 21,33 | 18,63 | 0,30 |
| 5 | 64,29 | 13,03 | 14,27 | 8,22 | 0,19 |
| 6 | 40,97 | 20,77 | 22,31 | 15,70 | 0,26 |
| 7 | 52,57 | 9,61 | 21,70 | 15,88 | 0,23 |
| 8 | 36,73 | 13,41 | 16,33 | 33,16 | 0,38 |
| 9 | 26,77 | 34,21 | 16,11 | 22,53 | 0,38 |
| 10 | 51,67 | 14,16 | 20,32 | 13,82 | 0,04 |
| 11 | 38,08 | 11,87 | 9,95 | 39,84 | 0,26 |
| 12 | 57,90 | 26,77 | 7,32 | 7,70 | 0,30 |
| 13 | 62,56 | 8,00 | 17,69 | 11,49 | 0,26 |
| 14 | 55,91 | 9,50 | 9,61 | 24,86 | 0,11 |
| 15 | 66,13 | 14,68 | 9,16 | 9,80 | 0,23 |
| 16 | 54,30 | 14,34 | 6,20 | 25,01 | 0,15 |

| | | | | | |
|---|---|---|---|---|---|
| 17 | 52,08 | 16,30 | 12,65 | 18,74 | 0,23 |
| 18 | 69,92 | 10,78 | 10,29 | 8,75 | 0,26 |
| 19 | 42,96 | 14,31 | 16,75 | 25,54 | 0,45 |
| 20 | 55,88 | 11,75 | 23,24 | 8,82 | 0,30 |
| 21 | 59,86 | 10,25 | 9,39 | 20,32 | 0,19 |
| 22 | 74,46 | 15,06 | **3,83** | 6,42 | 0,23 |
| 23 | 32,37 | 20,43 | 11,00 | 36,05 | 0,15 |
| 24 | 61,96 | 16,67 | 10,55 | 10,40 | 0,41 |
| 25 | 57,08 | 18,51 | 11,08 | 13,14 | 0,19 |
| 26 | 56,82 | 13,18 | 17,16 | 12,73 | 0,11 |
| 27 | 78,71 | 8,64 | 6,83 | 5,67 | 0,15 |
| 28 | 69,28 | 11,57 | 12,77 | 6,31 | 0,08 |
| 29 | 52,50 | 28,01 | 9,09 | 10,21 | 0,19 |
| 30 | 47,88 | 19,04 | 14,31 | 18,48 | 0,30 |
| **2-variant** | | | | | |
| 1 | 49,91 | 14,01 | 16,01 | 19,78 | 0,29 |
| 2 | 44,28 | 20,40 | 9,22 | 25,88 | 0,22 |
| 3 | 28,17 | 14,70 | 38,44 | 18,40 | 0,29 |
| 4 | 60,87 | 8,28 | **4,36** | 26,03 | 0,47 |
| 5 | 56,66 | 18,51 | 11,91 | 12,56 | 0,36 |
| 6 | 56,33 | 21,85 | 11,43 | 10,13 | 0,25 |
| 7 | 31,14 | 20,00 | 24,72 | 23,74 | 0,40 |
| 8 | 47,59 | 13,25 | 6,61 | 32,16 | 0,40 |
| 9 | 34,05 | 21,05 | 29,44 | 15,03 | 0,44 |
| 10 | 55,50 | 12,99 | 17,24 | 13,83 | 0,44 |
| 11 | 32,30 | 18,73 | 16,44 | 32,34 | 0,18 |
| 12 | 51,40 | 11,87 | 22,94 | 13,54 | 0,25 |

| | | | | | |
|---|---|---|---|---|---|
| 13 | 55,21 | 5,48 | 33,36 | 5,63 | 0,33 |
| 14 | 50,13 | 18,37 | 17,10 | 14,16 | 0,25 |
| 15 | 60,29 | 11,91 | 17,60 | 9,95 | 0,25 |
| 16 | 76,59 | 8,60 | 5,44 | 9,15 | 0,22 |
| 17 | 32,67 | 41,16 | 17,57 | 8,38 | 0,22 |
| 18 | 73,94 | 7,22 | 9,98 | 8,49 | 0,36 |
| 19 | 38,51 | 17,50 | 25,88 | 17,79 | 0,33 |
| 20 | 59,42 | 16,33 | 10,85 | 13,18 | 0,22 |
| 21 | 51,40 | 16,77 | 12,20 | 19,42 | 0,22 |
| 22 | 71,80 | **3,41** | 16,81 | 7,73 | 0,25 |
| 23 | 53,25 | 15,64 | 12,12 | 18,77 | 0,22 |
| 24 | 65,08 | 16,12 | 9,62 | 8,93 | 0,25 |
| 25 | 62,07 | 15,03 | 8,64 | 13,97 | 0,29 |
| 26 | 48,89 | 10,38 | 10,85 | 29,40 | 0,47 |
| 27 | 64,79 | 9,95 | 14,48 | 10,38 | 0,40 |
| 28 | 71,94 | 11,07 | 12,30 | **4,43** | 0,25 |
| 29 | 55,03 | 14,77 | 16,84 | 12,74 | 0,62 |
| 30 | 45,52 | 22,00 | 16,52 | 15,72 | 0,25 |
| **3-variant** | | | | | |
| 1 | 46,92 | 20,59 | 15,50 | 16,82 | 0,18 |
| 2 | 50,12 | 16,14 | 8,05 | 25,37 | 0,32 |
| 3 | 42,18 | 17,99 | 19,42 | 20,16 | 0,25 |
| 4 | 60,74 | 23,30 | 10,47 | 5,31 | 0,18 |
| 5 | 56,82 | 14,75 | 13,50 | 14,61 | 0,32 |
| 6 | 39,08 | 20,48 | 21,38 | 18,88 | 0,18 |
| 7 | 47,56 | 14,82 | 17,71 | 19,56 | 0,36 |
| 8 | 62,42 | 8,69 | 9,55 | 19,13 | 0,21 |

| | | | | | |
|---|---|---|---|---|---|
| 9 | 33,02 | 25,19 | 9,69 | 31,92 | 0,18 |
| 10 | 50,20 | 14,78 | 22,12 | 12,75 | 0,14 |
| 11 | 28,29 | 14,18 | 38,51 | 18,92 | 0,11 |
| 12 | 59,35 | 23,26 | 9,37 | 7,66 | 0,36 |
| 13 | 66,01 | 12,43 | 11,90 | 9,48 | 0,18 |
| 14 | 37,48 | 43,96 | 10,33 | 7,87 | 0,36 |
| 15 | 65,91 | 7,02 | 11,19 | 15,64 | 0,25 |
| 16 | 61,38 | 16,92 | 11,69 | 9,98 | 0,04 |
| 17 | 51,76 | 12,97 | 14,07 | 20,66 | 0,53 |
| 18 | 62,84 | 17,06 | 8,73 | 11,15 | 0,21 |
| 19 | 52,90 | 14,46 | 14,61 | 17,60 | 0,43 |
| 20 | 54,97 | 13,22 | 16,85 | 14,82 | 0,14 |
| 21 | 64,80 | 13,15 | 11,51 | 10,30 | 0,25 |
| 22 | 56,22 | **3,35** | 25,58 | 14,75 | 0,11 |
| 23 | 42,39 | 10,72 | 35,59 | 11,08 | 0,21 |
| 24 | 71,18 | 7,77 | 14,11 | 6,77 | 0,18 |
| 25 | 56,75 | 23,98 | 9,33 | 9,80 | 0,14 |
| 26 | 47,81 | 11,19 | 23,23 | 17,53 | 0,25 |
| 27 | 65,94 | 8,37 | 16,28 | 9,19 | 0,21 |
| 28 | 54,19 | 10,40 | 7,37 | 27,89 | 0,14 |
| 29 | 46,06 | 8,94 | 36,16 | 8,66 | 0,18 |
| 30 | 52,94 | 17,14 | 14,89 | 14,86 | 0,18 |

4-jadvaldan rus guruhlari uchun (1-kun 1-smena) 1- va 3-variantlarda foydalanilgan test topshiriqlarining barchasining distraktorlari yaxshi ishlayotganligi, 2-variantda foydalanilgan 1 ta test topshirig‘ining (14-test topshirig‘i) C distraktorini abituriyentlarning 4,33 foizi tanlaganlar. 5-jadvaldan esa rus guruhlari uchun (1-kun 2-smena) 1-variantda foydalanilgan 1 ta test topshirig‘ining (18-test topshirig‘i) B

distraktorini abituriyentlarning 4,58 foizi tanlaganlar. 2-variantda ham foydalanilgan 1 ta test topshirig‘ining (24-test topshirig‘i) D distraktorini abituriyentlarning 3,92 foizi tanlaganlar. 3-variantda foydalanilgan test topshiriqlaridan esa 1 ta test topshirig‘i (8-test topshirig‘i) ning D distraktorini abituriyentlarning 4,98 foizi tanlaganlar. Rus guruhlarida ham o‘zbek guruhlari singari 1- va 2-smenalarda foydalanilgan test topshiriqlarining distraktorlarini 95 foiz va undan yuqori tanlagan abituriyentlar kuzatilmadi.

*4-jadval*

**2025-2026- o‘quv yilida huquqshunoslik fanidan bakalavriatga kirish test sinovlari natijalari haqida statistik ma’lumotlar (*rus guruhi, 1-kun, 1-smena*)**

| Test topshiriqlari raqami | Kalit | Distraktorlar | | | NA, % |
|---|---|---|---|---|---|
| | A, % | B, % | C, % | D, % | |
| **1-variant** | | | | | |
| 1 | 49,41 | 14,33 | 15,51 | 20,55 | 0,20 |
| 2 | 40,91 | 19,37 | 23,81 | 15,51 | 0,40 |
| 3 | 43,38 | 25,59 | 19,57 | 10,97 | 0,49 |
| 4 | 60,08 | 8,10 | 14,23 | 17,29 | 0,30 |
| 5 | 49,01 | 18,58 | 19,17 | 12,85 | 0,40 |
| 6 | 52,96 | 17,49 | 13,14 | 16,11 | 0,30 |
| 7 | 58,60 | 25,59 | 10,18 | 5,24 | 0,40 |
| 8 | 38,93 | 22,92 | 12,85 | 25,20 | 0,10 |
| 9 | 60,67 | 15,51 | 11,56 | 11,76 | 0,49 |
| 10 | 41,01 | 24,70 | 13,64 | 20,26 | 0,40 |
| 11 | 42,98 | 11,26 | 11,17 | 34,39 | 0,20 |
| 12 | 50,20 | 18,68 | 18,18 | 12,85 | 0,10 |
| 13 | 39,53 | 34,29 | 13,04 | 12,85 | 0,30 |

| | | | | | |
|---|---|---|---|---|---|
| 14 | 65,91 | 13,93 | 10,47 | 9,49 | 0,20 |
| 15 | 39,23 | 14,62 | 9,29 | 36,66 | 0,20 |
| 16 | 22,04 | 26,68 | 23,22 | 27,57 | 0,49 |
| 17 | 59,19 | 17,09 | 12,75 | 10,87 | 0,10 |
| 18 | 53,36 | 17,09 | 19,37 | 9,78 | 0,40 |
| 19 | 36,56 | 18,58 | 21,44 | 22,92 | 0,49 |
| 20 | 31,32 | 16,21 | 28,36 | 24,01 | 0,10 |
| 21 | 37,65 | 22,43 | 8,79 | 31,03 | 0,10 |
| 22 | 45,75 | 18,58 | 21,74 | 13,54 | 0,40 |
| 23 | 76,58 | 9,98 | 8,30 | 5,04 | 0,10 |
| 24 | 65,81 | 6,23 | 7,21 | 20,45 | 0,30 |
| 25 | 53,36 | 15,51 | 20,06 | 10,77 | 0,30 |
| 26 | 17,39 | 44,07 | 7,51 | 30,73 | 0,30 |
| 27 | 48,52 | 24,80 | 15,12 | 11,17 | 0,40 |
| 28 | 56,62 | 9,98 | 21,84 | 11,17 | 0,40 |
| 29 | 43,38 | 26,68 | 14,03 | 15,51 | 0,40 |
| 30 | 38,54 | 18,48 | 20,36 | 22,33 | 0,30 |
| **2-variant** | | | | | |
| 1 | 49,79 | 25,15 | 15,05 | 9,79 | 0,21 |
| 2 | 44,12 | 18,66 | 17,01 | 19,90 | 0,31 |
| 3 | 79,59 | 6,60 | 8,45 | 5,26 | 0,10 |
| 4 | 32,89 | 11,75 | 35,57 | 19,48 | 0,31 |
| 5 | 47,94 | 22,99 | 14,33 | 14,74 | 0,00 |
| 6 | 47,53 | 16,91 | 10,10 | 25,46 | 0,00 |
| 7 | 51,24 | 11,34 | 23,40 | 13,81 | 0,21 |
| 8 | 49,90 | 18,76 | 18,76 | 12,47 | 0,10 |
| 9 | 67,53 | 8,25 | 13,81 | 10,21 | 0,21 |

| | | | | | |
|---|---|---|---|---|---|
| 10 | 49,48 | 20,31 | 15,98 | 13,92 | 0,31 |
| 11 | 59,69 | 12,68 | 13,40 | 13,71 | 0,52 |
| 12 | 46,60 | 18,56 | 18,45 | 16,08 | 0,31 |
| 13 | 41,65 | 11,65 | 34,33 | 12,37 | 0,00 |
| 14 | 58,14 | 8,66 | **4,33** | 28,56 | 0,31 |
| 15 | 48,45 | 21,13 | 19,79 | 10,52 | 0,10 |
| 16 | 67,73 | 5,88 | 14,95 | 11,24 | 0,21 |
| 17 | 58,76 | 15,15 | 7,63 | 18,25 | 0,21 |
| 18 | 44,33 | 19,90 | 23,30 | 12,27 | 0,21 |
| 19 | 40,10 | 24,02 | 17,01 | 18,66 | 0,21 |
| 20 | 29,79 | 19,79 | 30,72 | 19,59 | 0,10 |
| 21 | 34,12 | 47,42 | 7,22 | 11,13 | 0,10 |
| 22 | 56,19 | 14,23 | 19,38 | 10,21 | 0,00 |
| 23 | 64,95 | 9,18 | 8,56 | 17,22 | 0,10 |
| 24 | 50,62 | 24,74 | 15,57 | 8,97 | 0,10 |
| 25 | 53,40 | 20,62 | 14,02 | 11,75 | 0,21 |
| 26 | 32,47 | 24,02 | 16,29 | 27,01 | 0,21 |
| 27 | 60,82 | 11,34 | 10,82 | 16,91 | 0,10 |
| 28 | 69,59 | 11,65 | 6,80 | 11,86 | 0,10 |
| 29 | 54,33 | 16,70 | 9,48 | 19,18 | 0,31 |
| 30 | 46,49 | 14,02 | 16,70 | 22,68 | 0,10 |
| **3- variant** | | | | | |
| 1 | 52,31 | 17,11 | 13,86 | 16,52 | 0,20 |
| 2 | 41,40 | 15,44 | 15,93 | 26,84 | 0,39 |
| 3 | 32,84 | 25,17 | 26,06 | 15,63 | 0,29 |
| 4 | 67,26 | 14,26 | 11,11 | 6,98 | 0,39 |
| 5 | 49,26 | 24,09 | 14,85 | 11,70 | 0,10 |

| | | | | | |
|---|---|---|---|---|---|
| 6 | 49,16 | 12,78 | 25,86 | 12,00 | 0,20 |
| 7 | 54,67 | 9,44 | 29,70 | 6,10 | 0,10 |
| 8 | 29,30 | 13,37 | 30,29 | 26,75 | 0,29 |
| 9 | 62,34 | 9,34 | 20,45 | 7,28 | 0,59 |
| 10 | 51,82 | 24,09 | 9,93 | 14,06 | 0,10 |
| 11 | 57,92 | 12,49 | 14,75 | 14,45 | 0,39 |
| 12 | 48,18 | 18,68 | 19,17 | 13,47 | 0,49 |
| 13 | 43,95 | 6,49 | 7,67 | 41,79 | 0,10 |
| 14 | 70,60 | 9,24 | 8,85 | 11,21 | 0,10 |
| 15 | 45,43 | 15,04 | 13,27 | 26,06 | 0,20 |
| 16 | 63,52 | 8,16 | 14,26 | 14,06 | 0,00 |
| 17 | 60,47 | 14,85 | 11,60 | 12,78 | 0,29 |
| 18 | 63,52 | 14,85 | 14,16 | 7,28 | 0,20 |
| 19 | 33,92 | 14,36 | 27,24 | 23,89 | 0,59 |
| 20 | 24,29 | 22,42 | 24,48 | 28,42 | 0,39 |
| 21 | 45,82 | 26,94 | 14,06 | 12,78 | 0,39 |
| 22 | 44,94 | 29,89 | 10,62 | 14,16 | 0,39 |
| 23 | 67,26 | 12,78 | 9,05 | 10,32 | 0,59 |
| 24 | 39,13 | 10,42 | 25,07 | 24,98 | 0,39 |
| 25 | 54,87 | 14,55 | 13,96 | 16,22 | 0,39 |
| 26 | 38,74 | 17,50 | 24,58 | 18,68 | 0,49 |
| 27 | 63,62 | 10,03 | 11,11 | 15,04 | 0,20 |
| 28 | 50,84 | 11,80 | 8,65 | 28,61 | 0,10 |
| 29 | 40,61 | 18,78 | 25,37 | 15,24 | 0,00 |
| 30 | 31,96 | 24,19 | 23,30 | 20,26 | 0,29 |

*5-jadval*

**2025-2026- o‘quv yilida huquqshunoslik fanidan bakalavriatga kirish test sinovlari natijalari haqida statistik ma’lumotlar (*rus guruhi, 1-kun, 2-smena*)**

| Test topshiriqlari raqami | Kalit | Distraktorlar | | | NA, % |
|---|---|---|---|---|---|
| | A, % | B, % | C, % | D, % | |
| **1-variant** | | | | | |
| 1 | 55,42 | 16,67 | 16,25 | 10,42 | 1,25 |
| 2 | 39,17 | 21,67 | 20,00 | 18,75 | 0,42 |
| 3 | 46,67 | 10,00 | 21,67 | 20,83 | 0,83 |
| 4 | 34,17 | 9,17 | 27,92 | 28,33 | 0,42 |
| 5 | 44,58 | 16,25 | 20,42 | 18,33 | 0,42 |
| 6 | 48,33 | 19,17 | 12,92 | 19,17 | 0,42 |
| 7 | 42,50 | 16,67 | 21,67 | 18,75 | 0,42 |
| 8 | 41,67 | 12,08 | 21,67 | 24,17 | 0,42 |
| 9 | 58,75 | 9,58 | 16,67 | 14,17 | 0,83 |
| 10 | 30,83 | 20,42 | 26,25 | 22,08 | 0,42 |
| 11 | 49,17 | 15,83 | 20,42 | 14,17 | 0,42 |
| 12 | 39,58 | 19,58 | 18,33 | 22,08 | 0,42 |
| 13 | 64,17 | 14,58 | 12,08 | 8,75 | 0,42 |
| 14 | 44,17 | 15,83 | 9,58 | 30,00 | 0,42 |
| 15 | 31,67 | 24,17 | 17,50 | 25,42 | 1,25 |
| 16 | 48,75 | 17,08 | 17,92 | 15,83 | 0,42 |
| 17 | 53,75 | 12,50 | 12,50 | 20,83 | 0,42 |
| 18 | 76,25 | **4,58** | 10,42 | 7,50 | 1,25 |
| 19 | 19,17 | 38,75 | 16,67 | 24,58 | 0,83 |

| | | | | | |
|---|---|---|---|---|---|
| 20 | 43,75 | 21,67 | 19,58 | 14,58 | 0,42 |
| 21 | 49,58 | 13,75 | 10,42 | 25,42 | 0,83 |
| 22 | 35,00 | 19,58 | 24,17 | 20,83 | 0,42 |
| 23 | 63,33 | 16,67 | 9,58 | 10,00 | 0,42 |
| 24 | 64,17 | 12,92 | 14,17 | 8,33 | 0,42 |
| 25 | 38,75 | 23,33 | 17,50 | 20,00 | 0,42 |
| 26 | 32,92 | 12,92 | 10,83 | 42,50 | 0,83 |
| 27 | 53,75 | 22,92 | 7,50 | 15,00 | 0,83 |
| 28 | 50,00 | 11,67 | 25,83 | 11,67 | 0,83 |
| 29 | 49,58 | 21,67 | 15,42 | 12,92 | 0,42 |
| 30 | 52,08 | 23,33 | 11,25 | 12,08 | 1,25 |
| **2-variant** | | | | | |
| 1 | 38,04 | 30,59 | 12,55 | 18,43 | 0,39 |
| 2 | 49,41 | 21,96 | 14,90 | 13,73 | 0,00 |
| 3 | 27,45 | 21,57 | 25,49 | 25,49 | 0,00 |
| 4 | 42,75 | 19,22 | 7,06 | 30,59 | 0,39 |
| 5 | 44,31 | 14,51 | 19,22 | 21,96 | 0,00 |
| 6 | 52,94 | 18,04 | 13,33 | 15,69 | 0,00 |
| 7 | 24,71 | 14,51 | 39,61 | 20,78 | 0,39 |
| 8 | 46,67 | 12,55 | 5,10 | 35,29 | 0,39 |
| 9 | 66,67 | 10,98 | 12,16 | 9,80 | 0,39 |
| 10 | 36,86 | 16,08 | 23,14 | 23,92 | 0,00 |
| 11 | 50,59 | 16,47 | 16,47 | 16,47 | 0,00 |
| 12 | 46,27 | 18,04 | 20,39 | 15,29 | 0,00 |
| 13 | 51,76 | 15,29 | 16,47 | 16,47 | 0,00 |
| 14 | 66,27 | 9,41 | 8,63 | 15,29 | 0,39 |
| 15 | 45,10 | 19,22 | 15,29 | 20,00 | 0,39 |

| | | | | | |
|---|---|---|---|---|---|
| 16 | 32,94 | 34,12 | 17,65 | 15,29 | 0,00 |
| 17 | 39,61 | 24,31 | 12,16 | 23,92 | 0,00 |
| 18 | 72,55 | 6,67 | 5,88 | 14,90 | 0,00 |
| 19 | 30,98 | 30,20 | 24,31 | 14,12 | 0,39 |
| 20 | 47,06 | 18,43 | 17,65 | 16,47 | 0,39 |
| 21 | 50,98 | 16,86 | 13,73 | 17,65 | 0,78 |
| 22 | 46,27 | 10,98 | 15,29 | 26,67 | 0,78 |
| 23 | 57,65 | 12,55 | 14,90 | 14,51 | 0,39 |
| 24 | 65,49 | 16,08 | 14,51 | **3,92** | 0,00 |
| 25 | 52,16 | 14,12 | 11,37 | 22,35 | 0,00 |
| 26 | 25,10 | 19,22 | 19,22 | 36,47 | 0,00 |
| 27 | 50,20 | 7,45 | 31,37 | 10,59 | 0,39 |
| 28 | 57,25 | 7,84 | 23,92 | 10,98 | 0,00 |
| 29 | 54,12 | 14,51 | 11,37 | 19,61 | 0,39 |
| 30 | 52,55 | 16,08 | 17,25 | 14,12 | 0,00 |
| **3-variant** | | | | | |
| 1 | 44,81 | 12,03 | 22,82 | 20,33 | 0,00 |
| 2 | 37,76 | 21,58 | 23,24 | 17,43 | 0,00 |
| 3 | 45,23 | 15,35 | 16,18 | 22,82 | 0,41 |
| 4 | 53,11 | 10,79 | 16,60 | 19,09 | 0,41 |
| 5 | 42,32 | 27,39 | 19,92 | 9,96 | 0,41 |
| 6 | 45,23 | 21,58 | 16,18 | 17,01 | 0,00 |
| 7 | 44,81 | 12,03 | 19,50 | 23,65 | 0,00 |
| 8 | 53,11 | 23,65 | 18,26 | **4,98** | 0,00 |
| 9 | 56,85 | 16,60 | 10,37 | 15,77 | 0,41 |
| 10 | 51,87 | 14,11 | 16,60 | 17,43 | 0,00 |
| 11 | 49,79 | 12,45 | 26,14 | 11,62 | 0,00 |

| | | | | | |
|---|---|---|---|---|---|
| 12 | 52,70 | 20,75 | 15,77 | 10,37 | 0,41 |
| 13 | 64,73 | 8,30 | 11,62 | 14,94 | 0,41 |
| 14 | 58,51 | 18,26 | 9,96 | 13,28 | 0,00 |
| 15 | 42,32 | 10,79 | 34,85 | 11,62 | 0,41 |
| 16 | 39,83 | 19,50 | 17,01 | 23,65 | 0,00 |
| 17 | 72,61 | 10,79 | 8,30 | 8,30 | 0,00 |
| 18 | 65,56 | 8,30 | 14,94 | 11,20 | 0,00 |
| 19 | 29,05 | 26,56 | 13,69 | 30,71 | 0,00 |
| 20 | 36,93 | 19,50 | 25,73 | 17,84 | 0,00 |
| 21 | 20,33 | 55,19 | 14,94 | 9,13 | 0,41 |
| 22 | 35,68 | 14,94 | 21,99 | 26,97 | 0,41 |
| 23 | 64,32 | 11,62 | 16,60 | 6,64 | 0,83 |
| 24 | 44,81 | 13,69 | 10,79 | 30,29 | 0,41 |
| 25 | 59,75 | 14,11 | 11,20 | 14,52 | 0,41 |
| 26 | 24,07 | 14,52 | 40,66 | 20,33 | 0,41 |
| 27 | 61,83 | 16,18 | 15,77 | 6,22 | 0,00 |
| 28 | 58,09 | 12,03 | 18,26 | 11,62 | 0,00 |
| 29 | 55,60 | 20,75 | 12,45 | 11,20 | 0,00 |
| 30 | 48,13 | 9,54 | 33,20 | 8,30 | 0,83 |

Huquqshunoslik fanidan (o‘zbek va rus guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining ishlamayotgan (2-5-jadvallarda ajratib ko‘rsatilgan) distraktorlarini kengroq ko‘rib chiqamiz. Abituriyentlarning tanlagan muqobil javoblariga asoslanib, 5 foizdan kam va 95 foizdan ko‘p tanlangan distraktorlarning har birini abituriyentlarning tanlagan muqobil javoblari bilan birga ularning qobiliyat darajalari bo‘yicha tanlagan javoblarini quyida keltiriladigan jadvallar va chizmalar (o‘zbek guruhlari uchun 6-10-jadvallar va 1-5-rasmlar, rus guruhlari uchun 11-14-jadvallar va 6-9-rasmlar) yordamida batafsil tahlil qilib chiqamiz.

O‘zbek (6-10-jadvallar) va rus (11-14-jadvallar) guruhlari uchun keltirilgan ushbu jadvallarda har bir test topshirig‘iga test topshiruvchilarning tanlagan javoblari foizi va ushbu javoblar ichidan quyi, o‘rtacha va yuqori qobiliyatli test topshiruvchilarning tanlangan javoblari foizi keltirilgan bo‘lib, bu yerda: TTS – test topshiruvchilar soni; TMJ – tanlangan muqobil javoblar foizi; 0÷25% - quyi qobiliyatli – variantdagi test topshiriqlarining 0÷25 foizi oralig‘ida tog‘ri javob bergan test topshiruvchilarning tanlagan muqobil javoblari foizi; 25÷75 foizi – o‘rtacha qobiliyatli – variantdagi test topshiriqlarining 25÷75 foizi oralig‘ida tog‘ri javob bergan test topshiruvchilarning tanlagan muqobil javoblari foizi; 75÷100 foizi - yuqori qobiliyatli – variantdagi test topshiriqlarining 75÷100 foizi oralig‘ida to‘g‘ri javob bergan test topshiruvchilarning tanlagan muqobil javoblari foizi; A*- kalit; B, C, va D distraktorlar; NA – belgilanmagan javoblardir. O‘zbek (1-5-rasmlar) va rus (6-9-rasmlar) guruhlari uchun keltirilgan ushbu rasmlarda esa keltirilgan jadvallardagi distraktorlarning chizmalari keltirilgan. Bu yerda ham yuqorida aytilganidek, javob bermagan abituriyentlarning javobi NA bilan belgilangan bo‘lib, chizmalarning pastki qismida to‘g‘ri javoblar ulushi (Pval), element va umumiy ball korrelyatsiyasi (Rit), element chiqarilgandagi umumiy ball bilan korrelyatsiyasi (Rir) keltirilgan. Rasmning chap burchagida distraktorlarga berilgan javoblarga mos keluvchi chiziqlar mos raqamlar (ranglar) bilan ko‘rsatilgan. Qavs ichidagi 1 raqami to‘g‘ri javobni, (0) esa noto‘g‘ri javoblarni bildiradi.

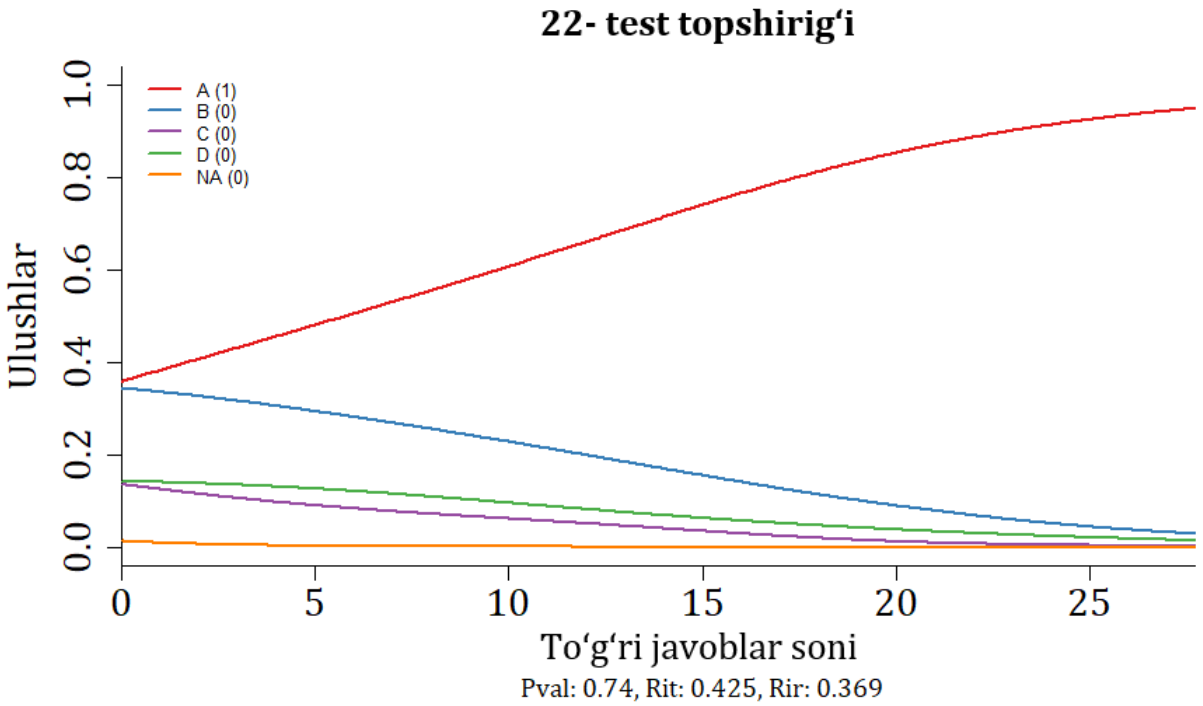
Demak, yuqorida huquqshunoslik fanidan o‘zbek guruhlari uchun (3-jadvaldagi) adabiyotlarda [7-8] mezon sifatida 5 foizdan kam va 95 foizdan ko‘p tanlangan distraktorlarni, ishlamaydigan distraktorlar deb hisoblangan distraktorlarni batafsil ko‘rib chiqamiz.

Quyidagi 6-jadvaldan 22-test topshirig‘ining C distraktoriga tanlangan muqobil javoblar (TMJ) foizi 3,83 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 9,00 foiz, o‘rtacha va yuqori qobiliyatli abituriyentlar mos ravishda 2,26 va 0,34 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 0,23 foizi ushbu test topshirig‘ini umuman belgilamaganlar.

Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir bo‘ladi.

*6-jadval*

| | 22- test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 1983 | 74,46 | 48,54 | 80,56 | 95,92 |
| B | 401 | 15,06 | 29,28 | 12,05 | 2,55 |
| C | 102 | 3,83 | 9,00 | 2,26 | 0,34 |
| D | 171 | 6,42 | 12,67 | 4,98 | 1,19 |
| NA | 6 | 0,23 | 0,51 | 0,16 | 0,00 |




1-rasm. To‘g‘ri javoblar sonining va 22-test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

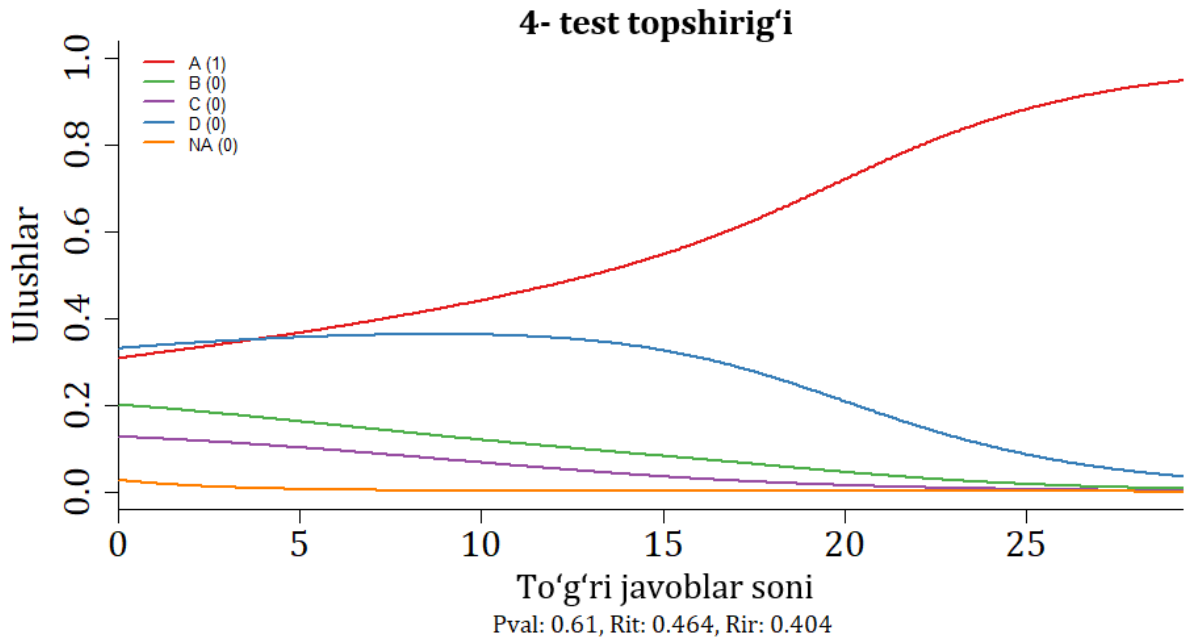
1-rasmdan 22-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,74 ekanligi, ushbu test topshirig‘i 1-qiyinlik darajasida juda yaqin ekanligini bildiradi. 1-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,425, u umumiy balldan

chiqarilgandagi korelyatsiyaning (Rir) 0,369 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korelyatsiyasi yaxshi ekanligini anglatadi. 22- test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,3÷0,4 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test top-shiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distrak-torlariga javob berganlar ulushi 0,1÷0,4 oraliqdan boshlanib, 0,1 gacha kamayib borishini 1-rasmdan ko‘rish mumkin.

7-jadvaldan 4-test topshirig‘ining C distraktoriga tanlangan muqobil javoblar (TMJ) foizi 4,36 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 11,59 foiz, o‘rtacha va yuqori qobiliyatli abituriyentlar mos ravishda 2,61 va 0,44 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyent-larning 0,47 foizi ushbu test topshirig‘ini umuman belgilamaganlar. Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklarini yoki javobni belgilashga ulgurma-ganliklari tufayli sodir bo‘ladi.

***7-jadval***

| | 4- test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 1677 | 60,87 | 34,76 | 57,97 | 93,79 |
| B | 228 | 8,28 | 17,02 | 7,39 | 1,04 |
| C | 120 | 4,36 | 11,59 | 2,61 | 0,44 |
| D | 717 | 26,03 | 35,62 | 31,67 | 4,59 |
| NA | 13 | 0,47 | 1,00 | 0,36 | 0,15 |




2-rasm. To‘g‘ri javoblar sonining va 4- test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

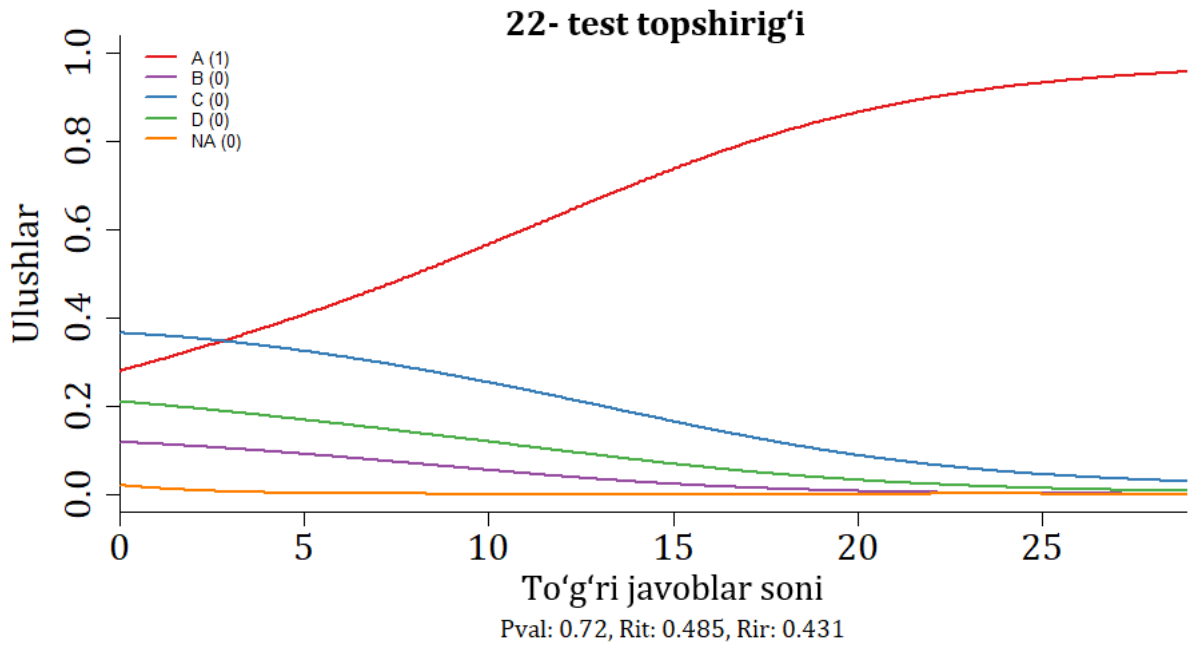
2-rasmdan 4-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,61 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korelyatsiyaning (Rit) 0,464, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,404 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 4-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,3÷0,4 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,15÷0,4 oraliqdan boshlanib, 0,1 gacha kamayib borishini 2-rasmdan ko‘rish mumkin.

8-jadvaldan 22-test topshirig‘ining B distraktoriga tanlangan muqobil javoblar (TMJ) foizi 3,41 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 10,01 foiz, o‘rtacha va yuqori qobiliyatli abituriyentlar mos ravishda 1,67 va 0,15 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib

borayotganligini ko‘rish mumkin. Bu esa, umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 0,25 foizi ushbu test topshirig‘ini umuman belgilamaganlar. Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir bo‘ladi.

*8-jadval*

| | 22-test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 1978 | 71,80 | 37,20 | 77,68 | 95,56 |
| B | 94 | 3,41 | 10,01 | 1,67 | 0,15 |
| C | 463 | 16,81 | 33,62 | 14,93 | 3,25 |
| D | 213 | 7,73 | 18,45 | 5,72 | 0,74 |
| NA | 7 | 0,25 | 0,72 | 0,00 | 0,30 |




3-rasm. To‘g‘ri javoblar sonining va 22-test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

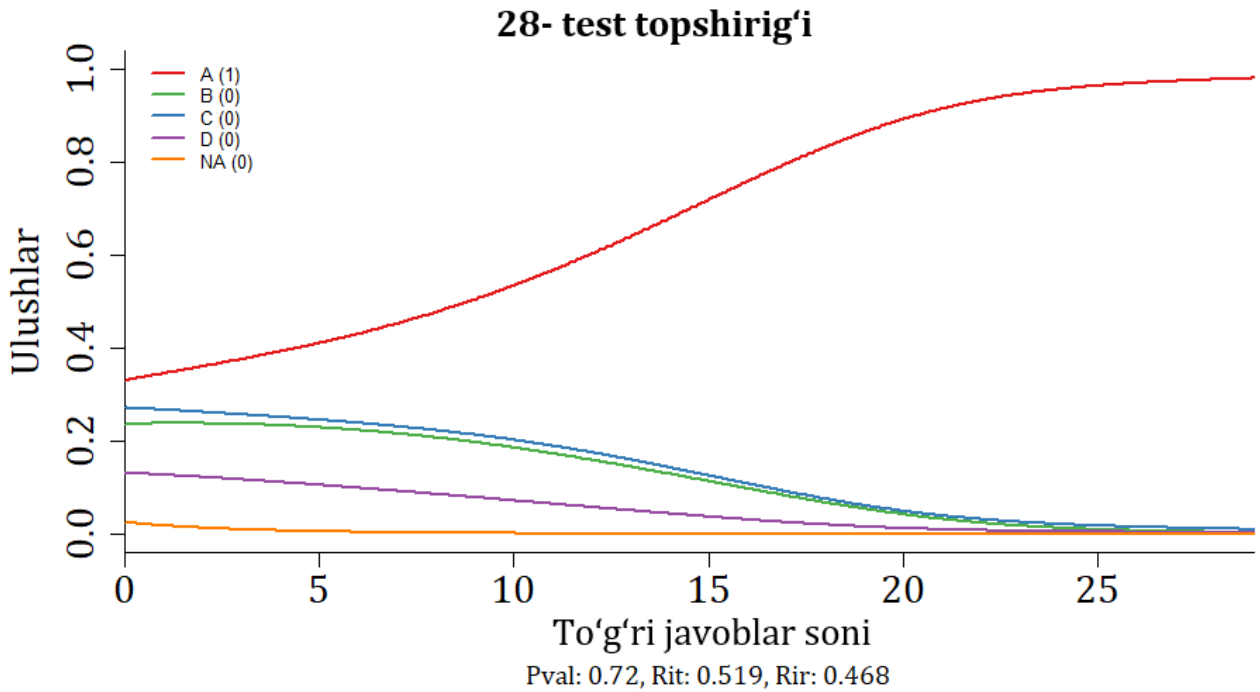
3-rasmdan 22-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,72 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,485, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,431 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 22-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,3÷0,4 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,1÷0,4 oraliqdan boshlanib, 0,1 gacha kamayib borishini 3-rasmdan ko‘rish mumkin.

9-jadvaldan 28-test topshirig‘ining D distraktoriga tanlangan muqobil javoblar (TMJ) foizi 4,43 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 11,30 foiz, o‘rtacha va yuqori qobiliyatli abituriyentlar mos ravishda 2,90 va 0,44 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 0,25 foizi ushbu test topshirig‘ini umuman belgilamaganlar. Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir bo‘ladi.

***9-jadval***

| | 28- test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 1982 | 71,94 | 38,05 | 76,30 | 98,08 |
| B | 305 | 11,07 | 24,46 | 9,57 | 0,30 |

| | | | | | |
|---|---|---|---|---|---|
| C | 339 | 12,30 | 25,32 | 11,16 | 1,18 |
| D | 122 | 4,43 | 11,30 | 2,90 | 0,44 |
| NA | 7 | 0,25 | 0,86 | 0,07 | 0,00 |




4-rasm. To‘g‘ri javoblar sonining va 28- test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

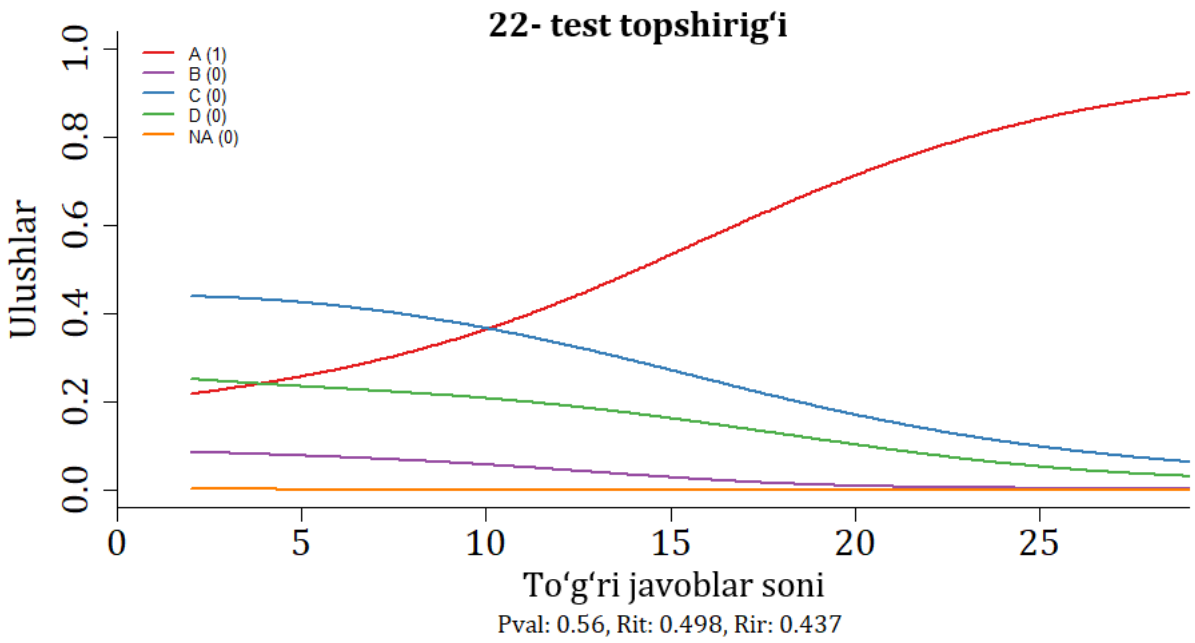
4-rasmdan 28- test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,72 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,519, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,468 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 28- test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,3÷0,4 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distrak-torlariga javob berganlar ulushi 0,15÷0,4 oraliqdan boshlanib, 0,1

gacha kamayib borishini 4-rasmdan ko‘rish mumkin.

10-jadvaldan 22-test topshirig‘ining B distraktoriga tanlangan muqobil javoblar (TMJ) foizi 3,35 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 8,02 foiz, o‘rtacha va yuqori qobiliyatli abituriyentlar mos ravishda 1,76 va 0,31 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 0,11 foizi ushbu test topshirig‘ini umuman belgilamaganlar. Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir bo‘ladi.

***10-jadval***

| | 22-test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 1578 | 56,22 | 25,81 | 60,20 | 89,03 |
| B | 94 | 3,35 | 8,02 | 1,76 | 0,31 |
| C | 718 | 25,58 | 42,21 | 23,53 | 7,37 |
| D | 414 | 14,75 | 23,84 | 14,36 | 3,29 |
| NA | 3 | 0,11 | 0,12 | 0,15 | 0,00 |




5-rasm. To‘g‘ri javoblar sonining va 22- test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

5-rasmdan 22-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,56 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,498, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,437 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 22-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,2÷0,3 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining dis-traktorlariga javob berganlar ulushi 0,1÷0,5 oraliqdan boshlanib, 0,1 gacha kamayib borishini 5-rasmdan ko‘rish mumkin.

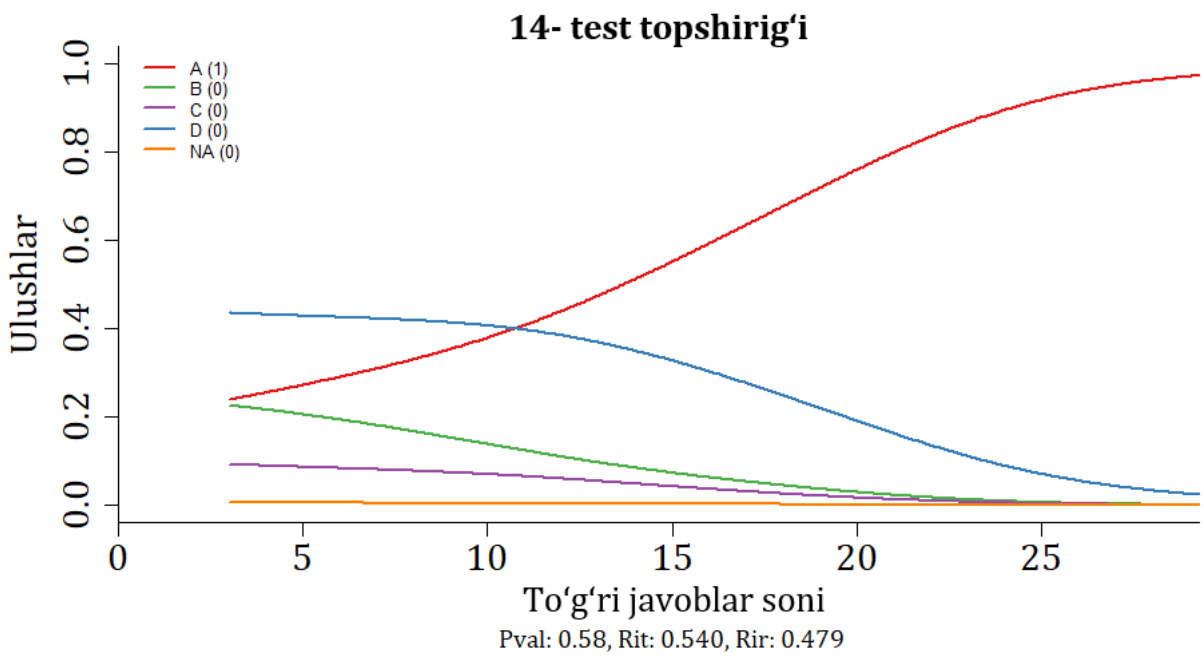
Endi yuqorida huquqshunoslik fanidan rus guruhlari uchun (4-5-jad-vallardagi, bu yerda ham adabiyotlarda [7-8] mezon sifatida 5 foizdan kam va 95 foizdan ko‘p tanlangan dis-traktorlarni ishlamaydigan distrak-torlar deb hisoblangan) distraktorlar-ning tahliliga o‘tamiz.

11-jadvaldan 14-test topshi-rig‘ining C distraktoriga tanlangan muqobil javoblar (TMJ) foizi 4,33 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 9,13 foiz va o‘rtacha qobiliyatli abituriyentlar 3,85 foiz

javob berganligini ko‘rish mumkin. Qobiliyat darajasi yuqori bo‘lgan abituriyentlar tomonidan mazkur distraktor umuman tanlanmagan, sababi yuqori qobiliyatli abituriyentlar mazkur distraktorning noto‘g‘ri, ya’ni to‘g‘ri javob emasligini aniq bilganliklaridan dalolat beradi. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 0,31 foizi ushbu test topshirig‘ini umuman belgilamaganlar. Bu ko‘pincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir bo‘ladi.

***11-jadval***

| | 14- test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 564 | 58,14 | 27,78 | 56,07 | 96,88 |
| B | 84 | 8,66 | 21,83 | 5,87 | 0,00 |
| C | 42 | 4,33 | 9,13 | 3,85 | 0,00 |
| D | 277 | 28,56 | 40,87 | 33,81 | 3,13 |
| NA | 3 | 0,31 | 0,40 | 0,40 | 0,00 |




6-rasm. To‘g‘ri javoblar sonining va 14- test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

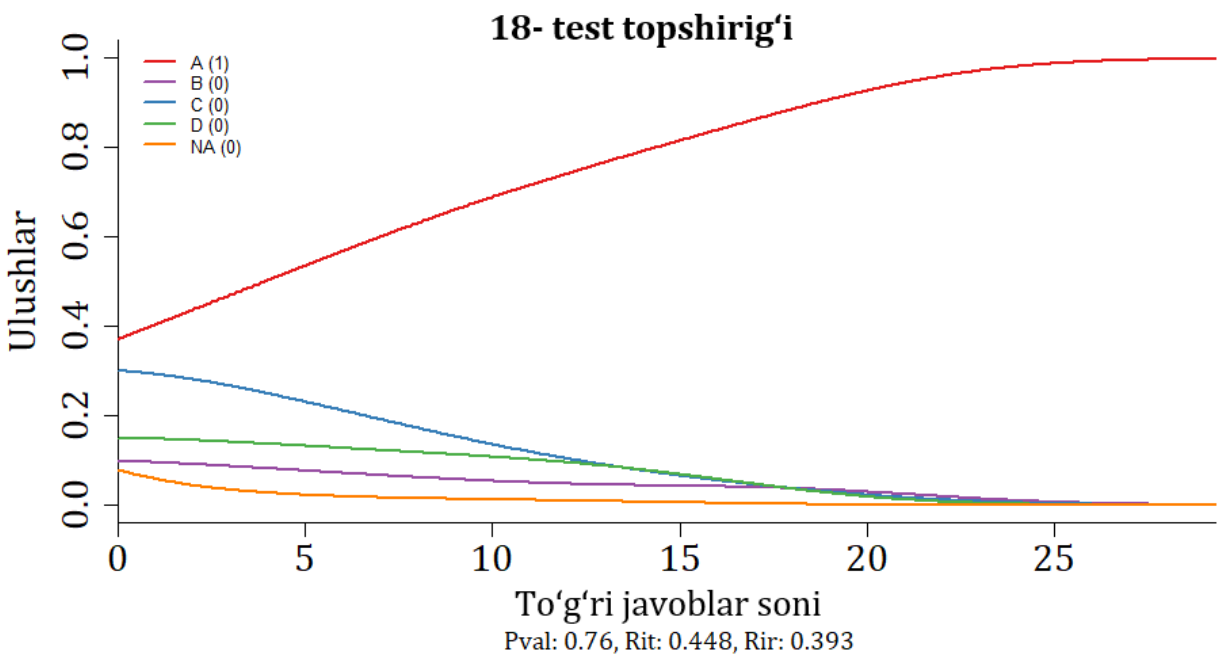
6- rasmdan 14- test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,58 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,540, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,479 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 14- test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,2÷0,3 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,1÷0,5 oraliqdan boshlanib, 0,1 gacha kamayib borishini 6-rasmdan ko‘rish mumkin.

12-jadvaldan 8-test topshirig‘ining B distraktoriga tanlangan muqobil javoblar (TMJ) foizi 4,58 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 8,45 foiz va o‘rtacha qobiliyatli abituriyentlar 4,55 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajasi yuqori bo‘lgan abituriyentlar tomonidan nafaqat mazkur distraktor, balki boshqa distraktorlar ham umuman tanlanmagan, sababi yuqori qobiliyatli abituriyentlar mazkur test topshirig‘idagi distraktorlarning

notoʻgʻri, yaʼni toʻgʻri javob emasligini aniq bilganliklaridan dalolat beradi. Qobiliyat darajalari boʻyicha quyidan yuqoriga qarab yurganimizda toʻgʻri javob beruvchi abituriyentlar foizi oshib borganligi koʻrinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori boʻlib borgan sari kamayib borayotganligini koʻrish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi. Abituriyentlarning 1,25 foizi ushbu test topshirigʻini umuman belgilamaganlar. Bu koʻpincha abituriyentlar ushbu test topshiriqlarini yecha olmaganliklari yoki javobni belgilashga ulgurmaganliklari tufayli sodir boʻladi.

***12-jadval***

| | 18- test topshirigʻi | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 183 | 76,25 | 49,30 | 80,91 | 100,00 |
| B | 11 | 4,58 | 8,45 | 4,55 | 0,00 |
| C | 25 | 10,42 | 26,76 | 5,45 | 0,00 |
| D | 18 | 7,50 | 12,68 | 8,18 | 0,00 |
| NA | 3 | 1,25 | 2,82 | 0,91 | 0,00 |




7-rasm. To‘g‘ri javoblar sonining va 18-test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

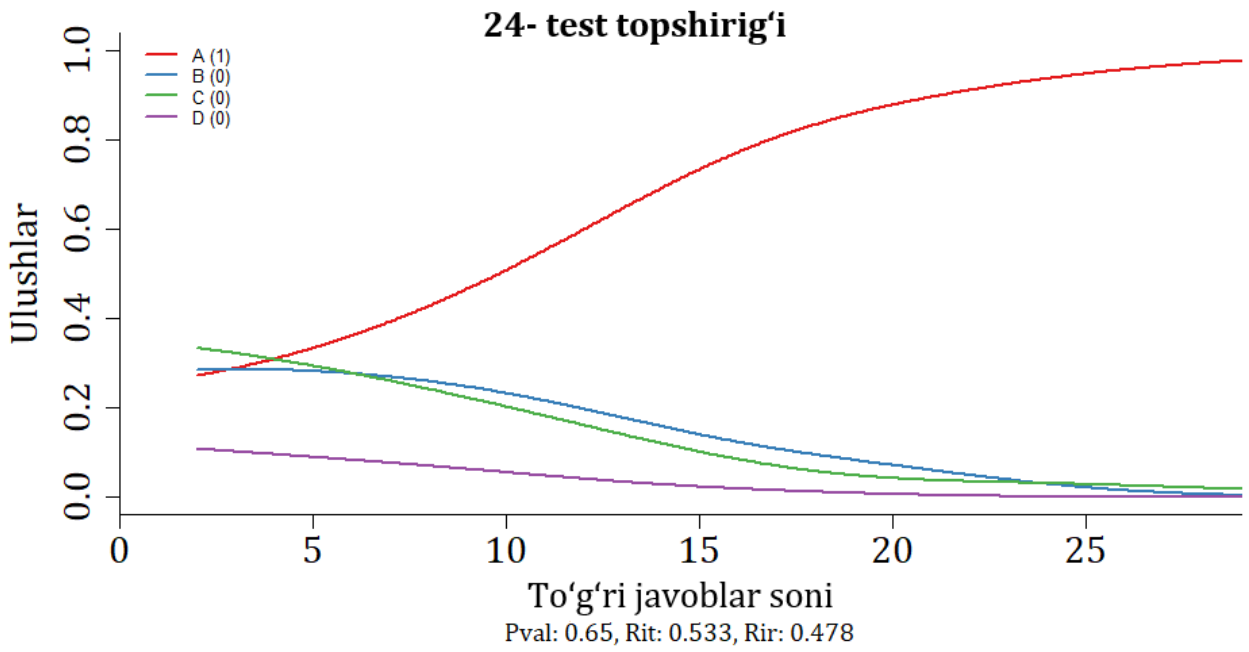
7-rasmdan 18-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,76 ekanligi, ushbu test topshirig‘i 1-qiyinlik darajasida ekanligini bildiradi. 1-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,448, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,393 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 18-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,3÷0,4 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,1÷0,3 oraliqdan boshlanib, 0,1 gacha kamayib borishini 7-rasmdan ko‘rish mumkin.

13-jadvaldan 24-test topshirig‘ining D distraktoriga tanlangan muqobil javoblar (TMJ) foizi 3,92 bo‘lsa-da, aslida bu disrtaktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 11,11 foiz va o‘rtacha qobiliyatli abituriyentlar 1,59 foiz javob berganligidan ko‘rish mumkin. Qobiliyat darajasi yuqori bo‘lgan abituriyentlar tomonidan mazkur distraktor umuman tanlanmagan, sababi yuqori qobiliyatli abituriyentlar mazkur distraktorning noto‘g‘ri, ya’ni

to‘g‘ri javob emasligini aniq bilganliklaridan dalolat beradi. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Dis-traktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi.

*13-jadval*

| | 24- test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 167 | 65,49 | 26,39 | 74,60 | 94,74 |
| B | 41 | 16,08 | 30,56 | 14,29 | 1,75 |
| C | 37 | 14,51 | 31,94 | 9,52 | 3,51 |
| D | 10 | 3,92 | 11,11 | 1,59 | 0,00 |
| NA | 0 | 0,00 | 0,00 | 0,00 | 0,00 |




8-rasm. To‘g‘ri javoblar sonining va 24- test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

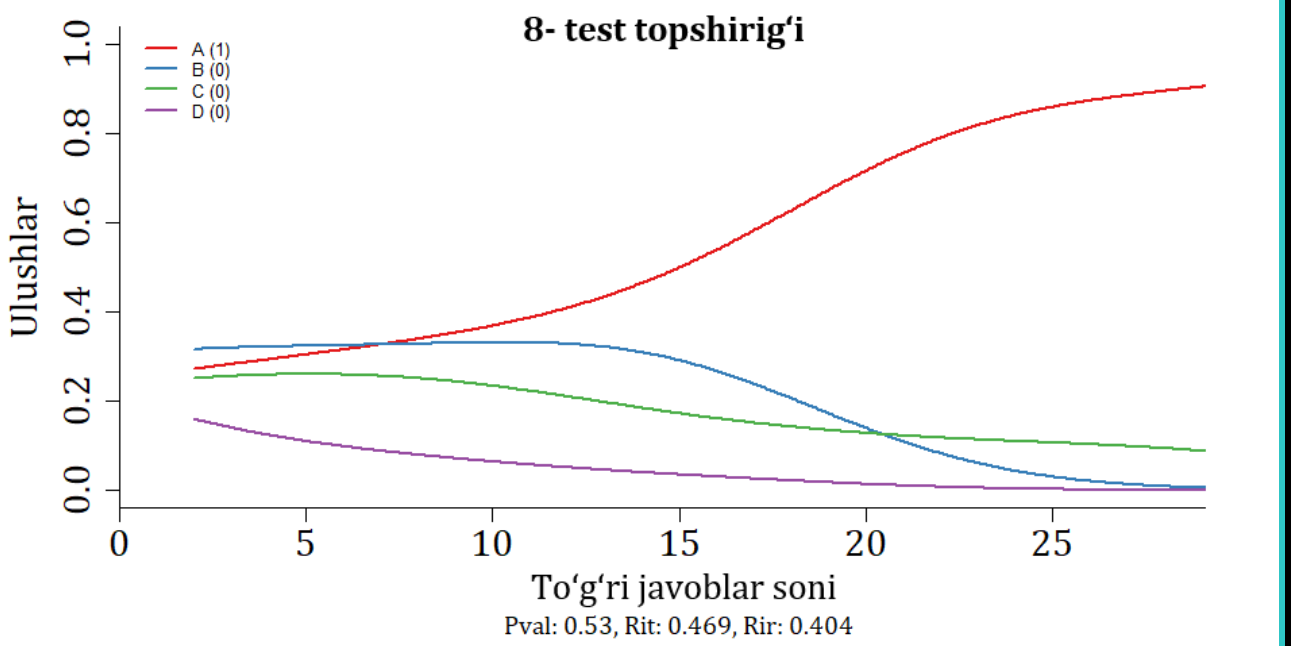
8-rasmdan 24-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,65 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,533, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,478 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 24-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,2÷0,3 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,1÷0,4 oraliqdan boshlanib, 0,1 gacha kamayib borishini 8-rasmdan ko‘rish mumkin.

14-jadvaldan 8-test topshirig‘ining D distraktoriga tanlangan muqobil javoblar (TMJ) foizi 4,98 bo‘lsa-da, aslida bu distraktorga qobiliyat darajasi quyi bo‘lgan abituriyentlar 14,52 foiz va o‘rtacha qobiliyatli abituriyentlar 2,31 foiz javob berganligini ko‘rish mumkin. Qobiliyat darajasi yuqori bo‘lgan abituriyentlar tomonidan mazkur distraktor umuman tanlanmagan, sababi yuqori qobiliyatli abituriyentlar mazkur distraktorning noto‘g‘ri, ya’ni to‘g‘ri javob emasligini aniq bilganliklaridan dalolat beradi. Qobiliyat darajalari bo‘yicha quyidan yuqoriga qarab yurganimizda to‘g‘ri javob beruvchi abituriyentlar foizi oshib borganligi ko‘rinadi. Distraktorlarni tanlaganlar foizi esa qobiliyat darajasi yuqori bo‘lib borgan sari kamayib borayotganligini ko‘rish mumkin. Bu esa umuman olganda ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligini bildiradi.

***14-jadval***

| | 8-test topshirig‘i | | | | |
|---|---|---|---|---|---|
| | *TTS* | *TMJ, %* | *0÷25, %* | *25÷75, %* | *75÷100, %* |
| A* | 128 | 53,11 | 27,42 | 50,77 | 91,84 |
| B | 57 | 23,65 | 30,65 | 29,23 | 0,00 |
| C | 44 | 18,26 | 27,42 | 17,69 | 8,16 |

| | | | | | |
|---|---|---|---|---|---|
| D | 12 | 4,98 | 14,52 | 2,31 | 0,00 |
| NA | 0 | 0,00 | 0,00 | 0,00 | 0,00 |




9-rasm. To‘g‘ri javoblar sonining va 8-test topshirig‘i muqobil javoblarini tanlagan abituriyentlar ulushlarining bog‘liqligi

9-rasmdan 8-test topshirig‘iga to‘g‘ri javob berganlar ulushi (Pval) 0,53 ekanligi, ushbu test topshirig‘i 2-qiyinlik darajasida ekanligini bildiradi. 2-qiyinlik darajasidagi test topshiriqlari uchun umumiy ball bilan korrelyatsiyaning (Rit) 0,469, u umumiy balldan chiqarilgandagi korrelyatsiyaning (Rir) 0,404 ga tengligi ushbu test topshirig‘ining umumiy ball bilan korrelyatsiyasi juda yaxshi ekanligini anglatadi. 8-test test topshirig‘ining to‘g‘ri javobini (A-chiziq) tanlaganlar ulushi 0,2÷0,3 oralig‘idan boshlanib oshib borayotganini ko‘rish mumkin. To‘g‘ri javoblar soni oshib borishi bilan absissa o‘qidagi o‘zgaruvchining barcha o‘zgarish sohasida oshib borganini ko‘rish mumkin. Bu tayyorgarlik darajasi oshib borishi bilan abituriyentlarning test topshiriqlariga to‘g‘ri javob berish ehtimoli oshib borayotganini bildiradi. Ushbu test topshirig‘ining distraktorlariga javob berganlar ulushi 0,1÷0,4 oraliqdan boshlanib, 0,1 gacha kamayib borishini 9-rasmdan ko‘rish mumkin.

Huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining distraktorlari o‘rganib chiqildi. O‘rganish natijasida o‘zbek guruhlari uchun 1- smenada 1-, 2- va

3-variantlarda foydalanilgan test topshiriqlarining distraktorlari o‘z vasifasini yaxshi bajarayotganligi kuzatildi. 2-smenada 1-variantda foydalanilgan 1 ta test topshirig‘ining C distraktorining 3,86 foizi, 2-variantda foydalanilgan 3 ta test topshiriqlarining mos ravishda C, B va D distraktorlarining mos ravishda 4,36, 3,41 va 4,43 foizi va 3-variantda foydalanilgan test topshiriqlaridan esa, 1 ta test topshirig‘ining B distraktorining 3,35 foizi abituriyentlar tomonidan tanlanganligi aniqlandi. O‘zbek guruhlari uchun 1- va 2-smenalarda foydalanilgan test topshiriqlarining distraktorlarini 95 foiz va undan yuqori tanlagan abituriyentlar aniqlanmadi.

Rus guruhlari uchun 1-smenada 1- va 3-variantlarda foydalanilgan test topshiriqlarining distraktorlari o‘z vasifasini yaxshi bajarayotganligi kuzatildi. 2-variantda foydalanilgan 1 ta test topshirig‘ining C distraktorining 4,33 foizi abituriyentlar tomonidan tanlanganligi, 2-smenada esa 1-variantda foydalanilgan 1 ta test topshirig‘ining B distraktorining 4,58 foizi, 2-variantda foydalanilgan 1 ta test topshirig‘ining D distraktorining 3,92 foizi va 3-variantda foydalanilgan test topshiriqlaridan esa 1 ta test topshirig‘ining D distraktorining 4,98 foizi abituriyentlar tomonidan tanlanganligi aniqlandi. Rus guruhlarida ham o‘zbek guruhlari singari 1- va 2-smenalarda foydalanilgan test topshiriqlarining distraktorlarini 95 foiz va undan yuqori tanlagan abituriyentlar aniqlanmadi.

Huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining distraktorlari abituriyentlar tomonidan tanlangan muqobil javoblar (TMJ) foizi bo‘yicha o‘rganib chiqildi. Mazkur distraktorlarni batafsil tahlil qilish orqali ularning barchasi abituriyentlarning qobiliyat darajalariga mos holda tanlagan javoblari va distraktorlarning chizmalari asosida tahlil qilindi va ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligi aniqlandi.

## 3. Xulosa

Huquqshunoslik fanidan (o‘zbek va rus guruhlari uchun) 2025-2026-o‘quv yilida bakalavriatga kirish test sinovlari natijalarining klassik test nazariyasi asosidagi tahlili natijasida olingan ma’lumotlar solishtirish orqali o‘rganildi. 2025-2026-o‘quv yilida huquqshunoslik fanidan (o‘zbek va rus guruhlari uchun) test sinovlari natijalarining statistik xarakteris-

tikalari - o‘rtacha qiymati, mediana, moda, o‘lchashning standart xatoligi aniqlandi.

Huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining distraktorlari abituriyentlar tomonidan tanlangan muqobil javoblar (TMJ) foizi bo‘yicha o‘rganib chiqildi.

O‘rganish natijasida o‘zbek guruhlari uchun 1-smenada 1-, 2- va 3-variantlarda foydalanilgan test topshiriqlarining distraktorlari o‘z vasifasini yaxshi bajarayotganligi kuzatildi. 2-smenada 1-variantda foydalanilgan 1 ta test topshirig‘ining C distraktorining 3,86 foizi, 2-variantda foydalanilgan 3 ta test topshiriqlarining mos ravishda C, B va D distraktorlarining mos ravishda 4,36, 3,41 va 4,43 foizi va 3-variantda foydalanilgan test topshiriqlaridan esa, 1 ta test topshirig‘ining B distraktorining 3,35 foizi abituriyentlar tomonidan tanlanganligi aniqlandi.

Rus guruhlari uchun 1-smenada 1- va 3-variantlarda foydalanilgan test topshiriqlarining distraktorlari o‘z vasifasini yaxshi bajarayotganligi kuzatildi. 2-variantda foydalanilgan 1 ta test topshirig‘ining C distraktorining 4,33 foizi abituriyentlar tomonidan tanlanganligi, 2-smenada esa 1-variantda foydalanilgan 1 ta test topshirig‘ining B distraktorining 4,58 foizi, 2-variantda foydalanilgan 1 ta test topshirig‘ining D distraktorini 3,92 foizi va 3-variantda foydalanilgan test topshiriqlaridan esa 1 ta test topshirig‘ining D distraktorining 4,98 foizi abituriyentlar tomonidan tanlanganligi aniqlandi.

O‘zbek va rus guruhlarida 1- va 2-smenalarda foydalanilgan test topshiriqlarining distraktorlarini 95 foiz va undan yuqori tanlagan abituriyentlar aniqlanmadi.

Distraktorlar quyi, o‘rtacha va yuqori qobiliyatli test topshiruvchilarning tanlagan javoblari foizi va shunga mos holda chizmalari asosida tahlil qilindi va ushbu distraktorlar quyi va yuqori qobiliyatlarni ajratishga xizmat qilayotganligi aniqlandi.

Umuman olganda huquqshunoslik fanidan (o‘zbek va rus tili guruhlari uchun) 2025-2026 - o‘quv yilida bakalavriatga kirish test sinovlarida foydalanilgan test topshiriqlarining distraktorlari juda yaxshi tanlanganligi ko‘rsatib berildi.

## ADABIYOTLAR


1. H. Tyrone Black, David L. Duhon. Evaluating and Improving Student Achievement in Business Programs: The Effective Use of Standardized Assessment Test. Journal of Education for Business, 2014, 90-98.
2. A.B. Normurodov, M.Dj. Ermamatov, A.A. Baratov, I.A. Boyxonov. Umumiy o‘rta ta’lim maktablarining 9–sinf bitiruvchilari uchun biologiya fanidan bilimlarni baholashda standart testlardan foydalanish. Axborotnoma, №1, 63-78, 2023.
3. A.B. Normurodov, M.Dj. Ermamatov. Parallel testlar yordamida standartlikni ta’minlash: 2022-2023-yillarda biologiya fanidan o‘tkazilgan test sinovi natijalari. Axborotnoma, №4, 36-66, 2023.
4. A.R. Sattiyev, M. Dj. Ermamatov. Matematika fanidan milliy sertifikat uchun o‘tkazilgan test sinovlari natijalari tahlili. Axborotnoma, №2, 35-55, 2023.
5. A.R. Sattiyev, M.Dj. Ermamatov. Tarix fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalarining klassik test nazariyasi va rash modeli asosida tahlili. Axborotnoma, №3, 51-72, 2023.
6. Sattiyev A.R., Ermamatov M.Dj., Alimov D.M. Ona tili va adabiyot fanidan milliy sertifikat uchun o‘tkazilgan test sinovida foydalanilgan test topshiriqlarining distraktorlari va korrelyatsiyalari// BMBA “Axborotnoma” ilmiy uslubiy jurnali, 2024-yil, 2-son, 4-96 betlar.
7. В.С. Аванесов Дистракторный анализ//Журнал Педагогические измерение, Москва, №1, 2013, стр.70-78.
8. Marie Tarrant, James Ware and Ahmed M. Mohammed, An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis, BMC Medical Education 2009, 9:40.
9. M.Dj. Ermamatov, A. Abbosov, A.A. Baratov. Test topshiriqlarini kalibrovkalash va qobiliyatlarni tenglashtirish, Axborotnoma, № 3-4, 4-15, 2022.

# ANALYSIS OF TEST RESULTS IN THE FIELD OF JURISPRUDENCE

**A.R. Sattiev, M.Dj. Ermamatov**

*Scientific-Educational Practical Center under the Agency for Assessment of Knowledge and Skills, 100084, Tashkent, Bog'ishamol Street, 12*

**Abstract.** In this paper analyzes the results of entrance test examinations in the subject of Law for undergraduate admission conducted in the 2025–2026 academic year for Uzbek- and Russian-language groups. The analysis is based on the framework of Classical Test Theory and includes key statistical indicators such as the mean, median, mode, standard error of measurement, and distractors. The study of distractor performance revealed that the distractors used in the test items functioned effectively, indicating a high quality of the test items employed in the examination process.

**Keywords:** test items, classical test theory, distractors.

# TEST MA'LUMOTI VA XARAKTERISTIKASI CHIZIQLARINI RASH VA IKKI PARAMETRLI MODELLAR ORQALI TAQQOSLASH

**A.B. Normurodov, M.Dj. Ermamatov**

*Bilim va malakalarni baholash agentligi huzuridagi Ilmiy-o'quv amaliy markazi, 100084, Toshkent sh., Bog'ishamol k., 12, a.normurodov@uzbmb.uz*

**Qisqacha mazmuni.** Ushbu maqolada ilmiy tadqiqot uchun biologiya fanidan o'tkazilgan test sinovi natijalari asosida test ma'lumoti va xarateristikasi chiziqlari Rash modeli va ikki parametrli modellar bilan tahlil qilindi. Tadqiqotda har ikki model bo'yicha testning qobiliyat darajasiga ko'ra ma'lumot funksiyalari taqqoslandi. Natijalar orqali testning validligi va ishonchliligini oshirishda model tanlashning ahamiyati ko'rsatildi. Ikki parametrli model bo'yicha hisoblangan test ma'lumot chiziqlari yuqori qobiliyatlar uchun ko'proq ma'lumot berishi ko'rsatildi.

**Kalit so'zlar:** test ma'lumoti chizig'i, test xarakteristikasi chizig'i, Rash modeli, ikki parametrli model

## 1. Kirish

Pedagogik o'lchovlarda asosiy maqsad – sinaluvchilarning bilimlarini baholashning ishonchli va valid bo'lishini ta'minlashdir. Shu maqsadda o'lchanayotgan xususiyatlarni aniqroq ko'rsatib bera oladigan modellar va nazariyalar ishlab chiqilmoqda. Baholashda hamma o'zgaruvchilar yashirin (latent) holda bo'lib, ularni minimal xatolar doirasida o'lchay oladigan o'lchov vositalari tanlanishi lozim. Hozirgi paytda keng qo'llaniladigan asosiy nazariyalar quyidagilardan iborat: Klassik test nazariyasi (ing. Classic Test Theory) va zamonaviy test nazariyasi (ing. Item Response Theory). An'anaviy klassik test nazariyasi (KTN) ko'plab amaliy vaziyatlarda o'z samaradorligini ko'rsatgan bo'lsa-da, test topshiriqlari va talabgor qobiliyatlari o'rtasidagi murakkab o'zaro bog'liqliklarni to'liq ochib berishda ba'zi cheklovlarga ega. Shu sababli klassik test nazariyasi zamonaviy test nazariyasi darajasidagi ma'lumotlarni taqdim eta olmaydi.

Klassik test nazariyasida shaxsning qobiliyat darajasi unga berilgan topshiriq xususiyatlariga va topshiriq xususiyatlarining sinaluvchilar guruhiga bog'liq bo'lib qoladi. Bundan tashqari har xil qiyinlikdagi test variantlarini yechgan sinaluvchilar qobiliyatini taqqoslash qiyin, ishonchlilikni aniqlash uchun esa parallel testlar kerak bo'ladi [1].

Zamonaviy test nazariyasi modellarida bu muammolar yechilgan bo'lib, qobiliyat darajalaridagi xatolikni hisobga olib, aniqroq natijalar olish imkoniyati mavjud. Element parametrlari guruhlar o'zgarishidan

mustaqil bo‘ladi (ya’ni invariant) va shaxsning qobiliyat darajasi test topshiriqlardan mustaqil ravishda baholanadi [2-4].

Zamonaviy test nazariyasi – sinaluvchining har bir elementga bergan javobi, uning qobiliyat darajasi va element xususiyatlari bilan qanday bog‘liqligini tahlil qiladi. Bu tahlil orqali shaxsning qobiliyati va element xususiyatlari (masalan, qiyinlik darajasi, ajratish darajasi)ni baholash imkonini beradi [5]. Chunki shaxsning o‘lchanayotgan xususiyatlari yoki qobiliyat darajalarini to‘g‘ridan-to‘g‘ri o‘lchab bo‘lmaydi. Zamonaviy test nazariyasi sinaluvchilarning test topshiriqlari orqali olingan natijalari bilan ularning yashirin qobiliyatlari o‘rtasidagi munosabatni aniqlashga harakat qiladi [6-9]. Shu bilan birga zamonaviy test nazariyasining ham bir nechta modellari mavjud bo‘lib, ularning har birini maqsadga mos holda qulayligi va kamchilliklari mavjuddir. Biz test topshiriqlarining qiyinlik va ajratish darajalariga ko‘ra zamonaviy test nazariyasinig Rash va ikki parametrli modellarini solishtirib ularning kamchilliklari va afzalliklarini o‘rgangan edik. [10].

Ushbu maqolada tadqiqot obyekti sifatida 9-sinflar uchun biologiya fanidan tadqiqot maqsadida o‘tkazilgan test sinovlari natijalaridan foydalanilgan. Test sinovlarida umumiy elementli 4 ta variant ishlatilgan va turli xil hududlardan jami 841 nafar litsey va maktab o‘quvchilari qatnashgan.

Ilmiy tadqiqotda umumiy elementli test variantlarining ma’lumot va xarakteristika chiziqlarini Rash [6] va ikki parametrli modellar [11-12] bilan hisoblash orqali ular o‘rtasidagi farq va modellarning afzalliklari o‘rganildi. Ushbu tadqiqot biologiya fanidan o‘tkazilgan test sinovi natijalarini baholashda zamonaviy test nazariyasi modellaridan foydalanish bo‘yicha amaliy misol sifatida taqdim etilishi jihatidan muhimdir. Bundan tashqari, test sinovi natijalarining validligi va ishonchliligi ham muhim ahamiyatga ega bo‘lib, bilimni aniqroq baholash uchun turli statistik modellar va nazariyalar ishlab chiqilgan. Tadqiqotda esa aynan shu nazariya va modellar orqali hisoblashlar o‘tkazildi va hisoblash natijalari taqdim etildi.

Ushbu yondashuv nafaqat har bir modelning afzallik va kamchiliklarini aniqlash, balki testning ishonchliligini oshirish hamda talabgorlarning bilim darajasini yanada aniqroq baholash imkonini beradi.

Rash modeli doirasida hisoblashlarda R dasturining dexter to‘plamidan [13], ikki parametrli model doirasida hisoblashlarda esa kengaytirilgan dexter MML to‘plamidan foydalanildi.

Garchi manbalarda [14-15] ko’p o’lchovli modellar ham rivojlantirilgan bo’lsa-da, eng ko’p ishlatiladigan modellar bir o’lchovli modellardir. Shuning uchun ushbu ilmiy tadqiqotda bir o‘lchovli modellardan foydalanildi.

## 2. Tadqiqot usullari

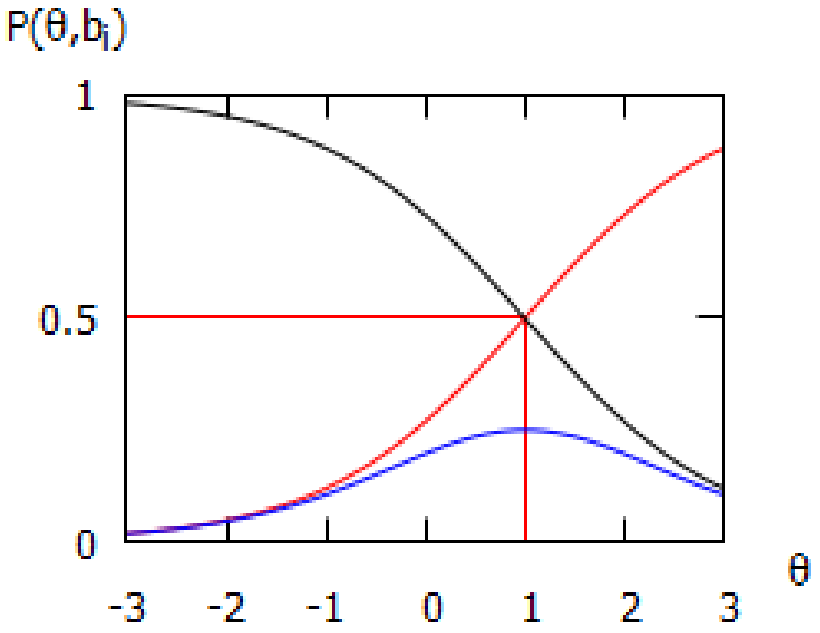
Test ma’lumot funksiyalari zamonaviy test nazariyasida muhim rol o‘ynaydi. Testdagi ixtiyoriy element test topshiruvchi haqida qandaydir ma’lumot beradi. Lekin bu ma’lumotning miqdori elementning qiyinlik darajasining shaxsning qobiliyat darajasiga qanchalik mos kelishiga bog‘liq. Bir parametrli logistik modelida bu element ma’lumoti funksiyasiga ta’sir qiladigan yagona omildir. Bir parametrli modelda ma’lumot funksiyasining formulasi juda oddiy [9]:

$$\boldsymbol{I_i(\theta) = P_i(\theta, b_i)Q_i(\theta, b_i)} \qquad \mathbf{(1)}$$

Zamonaviy test nazariyasining barcha modellarida ma’lumot funksiyasi talabgorning elementlarga (test topshiriqlariga) to‘g‘ri javob berish ehtimoli orqali ifodalanadi [5-9]. Har bir talabgor o‘ziga xos yashirin (latent) qobiliyat darajasiga ($\theta_j$) ega bo‘lib, elementlar esa o‘z parametrlari bilan ajralib turadi.

1-rasmdagi birinchi holatga, ya’ni $\boldsymbol{b_i = 1}$ ga mos keladigan bir parametrli logistik model bilan olingan ma’lumot funksiyasi ham keltirilgan (pastga qaragan parabola shaklidagi chiziq). 1-rasmda bitta elementning ma’lumot funksiyasi to‘g‘ri va noto‘g‘ri javob berish ehtimolligi chiziqlari bilan birgalikda berilgan.

Bir parametrli logistik modelda ma’lumot funksiyasining maksimal qiymati 0,25 ga teng. Bu - to‘g‘ri va noto‘g‘ri javoblar ehtimolligi 0,5 ga teng bo‘lgan holatga to‘g‘ri keladi.



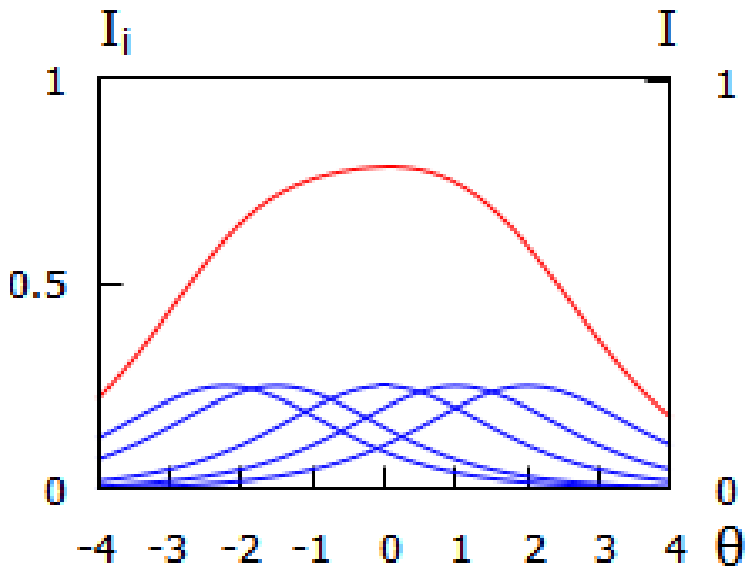

1-rasm. To‘g‘ri va noto‘g‘ri javob berish ehtimolligi va ma’lumot funksiyasi

Boshqacha qilib aytganda, bir parametrli logistik modelda ixtiyoriy element, qobiliyat darajasi element qiyinligi darajasiga teng bo‘lgan test topshiruvchilar uchun eng ko‘p ma’lumot bera oladi. Test top-shiruvchining qobiliyat darajasi element qiyinligi darajasidan katta yoki kichik bo‘lishi bilanoq element ma’lumoti kamayadi. Bundan esa har xil qobiliyatlarni aniqlash uchun turli xil qiyinlikdagi elementlarga ega test variant shakllantirilishi lozimligi kelib chiqadi.

Xuddi testga javob funksiyasi elementga javob funksiyalarining yig‘indisi bo‘lgani kabi test ma’lumot funksiyasi ham quyidagi formula bilan aniqlanadi [9]:

$$I_i(\theta_j) = \sum_i I_{ij}(\theta_j, b_i) \qquad (2)$$

2-rasmda element ma’lumoti funksiyalari va test ma’lumoti funksiyasi bir xil shkalada berilgan. Test ma’lumoti funksiyasi alohida element ma’lumoti funksiyalariga nisbatan ko‘proq ma’lumotni o‘z ichiga olishiga va kengroq qobiliyat sohasiga to‘g‘ri kelishiga e’tibor qaratamiz. Element ma’lumoti funksiyalari esa ularning qiyinliklari atrofida mujas-samlashgan. Yashirin qobi-liyatning aniqligini bashorat qila olishi - ma’lumot funksiyasining eng muhim xususiyatidir. Rasmdan har bir element berayotgan ma’lumot miqdoriga (ularning har biri 1 ga teng ma’lumot beradi) va test berayotgan ma’lumot miqdoriga (5 ga teng) e’tibor qaratmoqchimiz.



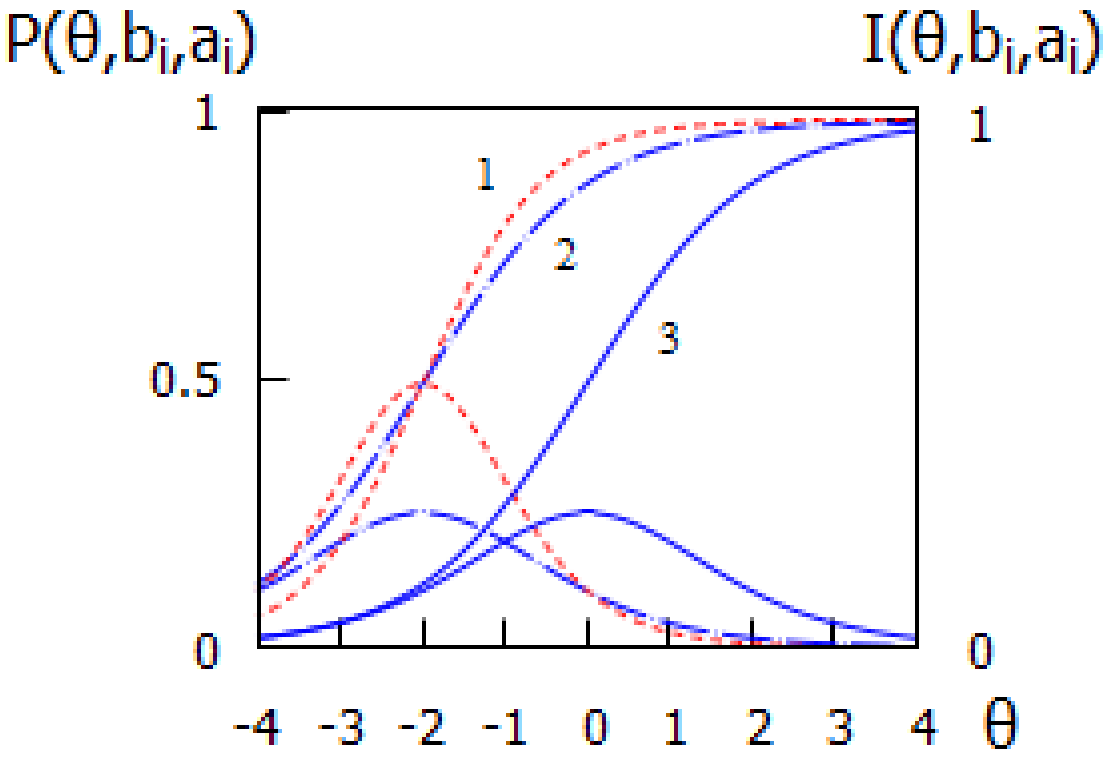

2-rasm. Bir parametrli modelning element va test ma’lumoti funksiyalari

Bunda elementlar soni qanchalik ko‘p bo‘lsa, sinaluvchilarning qobiliyati haqida shunchalik ko‘p ma’lumot olish mumkin ekanligi ko‘rinadi. Ikki parametrli modelning ma’lumot funksiyasi formulasining ko‘rinishi quyidagicha [9]:

$$\mathbf{I}(\boldsymbol{\theta}_j, \mathbf{b}_i, \mathbf{a}_i) = \mathbf{a}_i^2 \mathbf{P}(\boldsymbol{\theta}_j, \mathbf{b}_i, \mathbf{a}_i)\mathbf{Q}(\boldsymbol{\theta}_j, \mathbf{b}_i, \mathbf{a}_i) \quad (3)$$

Bu yerda ham elementning ajratish darajasi parametri muhim rol o‘ynaydi, chunki uning kvadrati ma’lumot funksiyasini hisoblashda ishlatiladi. Agar ajratish darajasi parametri birdan ancha kichik bo‘lsa, u ma’lumot funksiyasini juda sezilarli darajada kamaytiradi, aksincha birdan katta qiymatlari esa ma’lumot funksiyasi qiymatini oshirib yuboradi.

3-rasmda ikki parametrli modelning ma’lumot funksiyalari elementga javob funksiyalari bilan birga chizilgan. Elementning ajratish darajasi qiymati *a=1,4* bo‘lganda, 1-chiziqqa mos keluvchi ma’lumot funksiyasining balandligi ajratish darajasi qiymati *a*=1 bo‘lgan 2-ma’lumot funksiyasiga nisbatan sezilarli darajada oshganligini ko‘rish mumkin.




3-rasm. Ikki parametrli modelning elementga javob va element ma’lumoti funksiyalari

Ko‘rib o‘tganimizdek, bir parametrli modelda ma’lumot funksiyalarining balandligi bir xil va ularning cho‘qqisi ko‘pi bilan 0,25 ga teng bo‘lishi mumkin. Ular qiyinlik darajalariga qarab o‘ngga yoki chapga suriladi. Bunda qiyinlik darajasi va qobiliyat darajasi teng bo‘lgan nuqta qiyinlikni belgilaydi va aynan shu nuqtada eng ko‘p ma’lumot olish mumkin.

Ikki parametrli modelda ham ma’lumot funksiyalari qiyinlik darajasida o‘zining maksimumiga erishadi, lekin ularning shakli va cho‘qqidagi qiymati ajratish darajasiga kuchli tarzda bog‘liq. Ajratish darajasi katta bo‘lganda (shunga mos ravishda ma’lumot funksiyasi ham tik bo‘lganda), element qobiliyat haqida ko‘proq ma’lumot beradi va bu ma’lumot qiyinlik atrofida mujassamlashgan bo‘ladi. Kichik ajratish darajasiga ega elementlar esa qobiliyat haqida kamroq ma’lumot beradi, lekin kengroq qobiliyat sohasini qamrab oladi.

Ikki parametrli modelning test ma’lumoti funksiyasi ham bir parametrli modelning test ma’lumoti funksiyasiga o‘xshab element ma’lumoti funksiyalarining yig‘in-disidan hosil qilinadi. Lekin ikki parametrli modelning elementga javob funksiyasi bir parametrli modeldan farq qiladi. Shunday qilib, test ma’lumoti funksiyasi uchun ifoda quyidagicha ko‘rinishga ega [9]:

$$\boldsymbol{I_j(\theta_j) = \sum_i I_{ij}\left(b_i, a_i, \theta_j\right)} \qquad \textbf{(10)}$$

Elementlar ma’lumoti funksiyalari elementning ajratish darajasiga juda kuchli bog‘liqligi uchun, test ma’lumoti funksiyasi kutilmagan ko‘rinishga ega bo‘lishi mumkin. Bunday holat, ayniqsa, juda kam sondagi elementlari bo‘lgan testlarda uchrashi mumkin. Amaliyotda biz (–3:3) oraliqdagi qobiliyatlarni qamrab oladigan bir tekis taqsimlangan test ma’lumoti funksiyasini olishni xohlaymiz. Bu talab yetarlicha ko‘p sondagi ajratish darajasi yaxshi bo‘lgan va qiyinlik darajasi qobiliyat sohasini qamrab oladigan elementlarga ega bo‘lishni ta’minlash bilan bajarilishi mumkin.

Ajratish darajasi juda past bo‘lgan elementlar odatda testdan chiqarib tashlanadi. Lekin chiziq qanchalik tik bo‘lsa, shunchalik yaxshi deb qarash ham to‘g‘ri emas, chunki kompyuter adaptiv testlarida ajratish darajasi past bo‘lgan test topshiriqlaridan ham unumli foydalaniladi [9].

## 3. Natijalar va muhokamalar

Biologiya fanidan tadqiqot uchun o‘tkazilgan test sinovlarida foydalanilgan umumiy elementli 4 ta variantlar uchun test ma’lumoti va

xarakteristikasi chiziqlari Rash va ikki parametrli modellar bilan hisoblandi.

4-rasmda Rash (4a-rasm) va ikki parametrli (4b-rasm) modellar asosida hisoblangan umumiy elementli 4 ta variant uchun test ma'lumoti chiziqlari keltirilgan. Rash modeli bilan hisoblangan test ma'lumoti chiziqlari cho'qqilarining barchasi deyarli teng ma'lumot berayotganligini ko'rish mumkin (4a-rasm). Bu modelda barcha elementlar uchun ajratish darajasi parametri bir xil deb qabul qilinadi (odatda *a=1*), ya'ni barcha elementlar bir xil ajratish darajasiga ega deb hisoblanadi. Natijada test ma'lumot chiziqlari shakl va balandlik jihatidan deyarli bir xil bo'lib, test variantlari o'rtasida aniqlik farqlari sezilarli emas.

Ikki parametrli modelda har bir elementning ajratish darajasi parametrini hisobga olingani uchun test variantlari o'rtasidagi bu farqlar elementlar ajratish darajasiga bog'liq. Ya'ni ko'proq yuqori ajratish darajasiga ega elementlar mavjud bo'lgan test varianti ko'proq ma'lumot beradi. Bu holat test variantlari orasidagi ma'lumot farqlari tafovutlarni kuchaytiradi.

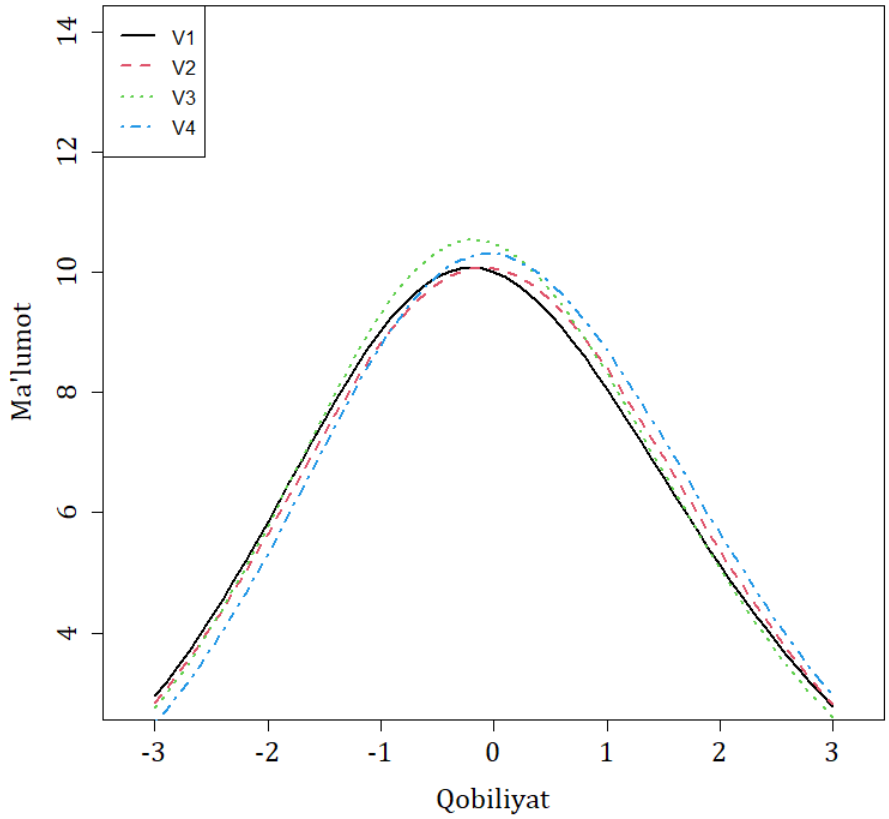
4b-rasmda ham V1 va V3 variantlar test ma'lumot chiziqlari ko'proq ma'lumot berayotganligini ko'rish mumkin. Bunga sabab ushbu ikkala variantda ajratish darajasi yuqori bo'lgan test topshiriqlar soni ko'pligidir. Shuning uchun bu chiziqlar eng baland cho'qqiga ega, ya'ni maksimal ma'lumot beradi. Demak, ikki parametrli modelda elementlar ajratish darajasi oshgan sari test ma'lumoti yuqoriroq bo'ladi, bu esa baholash aniqligi yaxshiroq ekanligini bildiradi.

Rash modelida ajratish darajasi o'zgarmas, test ma'lumot cho'qqisi doimo bir xil darajada bo'ladi, ya'ni test ma'lum bir qobiliyat darajasi atrofida maksimal ma'lumot beradi.

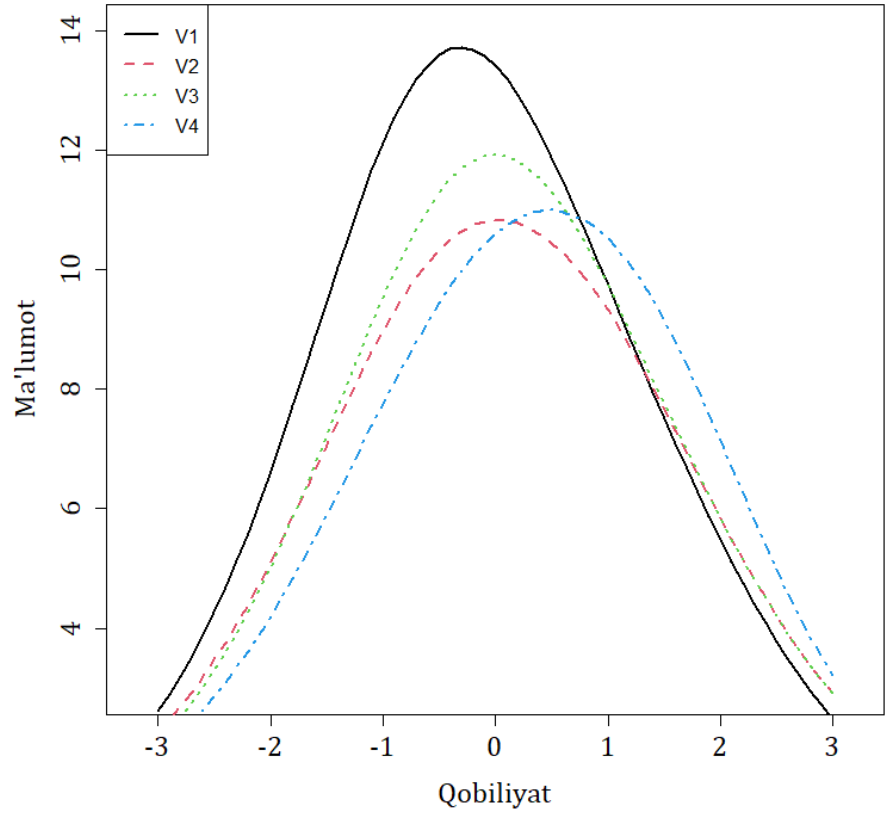
Ikki parametrli model ajratish darajasi yuqori bo'lgan topshiriqlar orqali baholashda ko'proq aniqlik va moslashuvchanlik olib keladi.

Rash modeli oddiy baholashlar uchun yetarli, biroq test turlicha ajratish darajasiga ega topshiriqlardan tashkil topgan bo'lsa, ikki parametrli model ustunlik qiladi.

4a va 4b-rasmlardagi ikki parametrli modeli bilan hisoblangan test ma'lumoti chiziqlari yuqori qobiliyatga ega guruh uchun Rash modeli bilan solishtirilganda ko'proq ma'lumot berishini ko'rish mumkin.




a)




b)

4-rasm. Rash (a) va ikki parametrli (b) model bo‘yicha test ma’lumoti chiziqlari

Ushbu ikki modelni solishtirganda, ikki parametrli modeli test variantlari orasidagi real tafovutlarni aniq ko‘rsatib bera oladi, chunki u elementalar ajratish darajasi farqini hisobga oladi. Bu, ayniqsa, test natijalarini chuqurroq tahlil qilish va test variantlarini farqlash zarur bo‘lgan holatlarda muhim ahamiyatga ega.

Rash modeli test variantlari orasida aniqlikni barqaror saqlaydi va umumiy qobiliyatni baholashda qulaylik yaratadi. Bu model testlarni soddalashtirilgan shaklda tenglashtirish (equating) va diagnostika maqsadlarida samarali qo‘llaniladi. Rash modeli eng keng qamrovli qobiliyat darajasi uchun ma’lumot taqdim etuvchi model bo‘lib, u elementlar ajratish darajasiga ko‘ra kam miqdordagi ma’lumotni beradi.

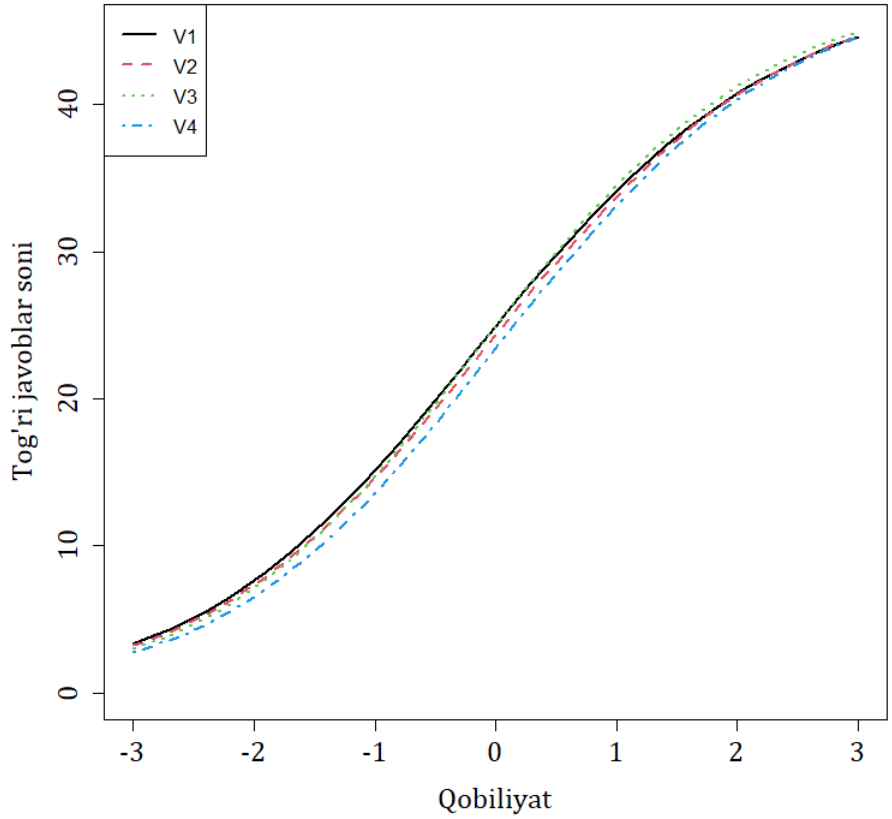
Xulosa qilib aytganda, agar test variantlari o‘rtasidagi farqlarni aniqlash va baholash muhim bo‘lsa, ikki parametrli model afzal. Agar test variantlari orasida farq ma’lumotlari haqida barqarorlik talab qilinsa va oddiyroq tushuntirishlar zarur bo‘lsa, Rash modeli maqsadga muvofiqdir.

5-rasmda Rash (5a-rasm) va ikki parametrli (5b-rasm) modellar asosida hisoblangan umumiy elementli 4 ta variant uchun test xarakteristikasi chiziqlari keltirilgan.

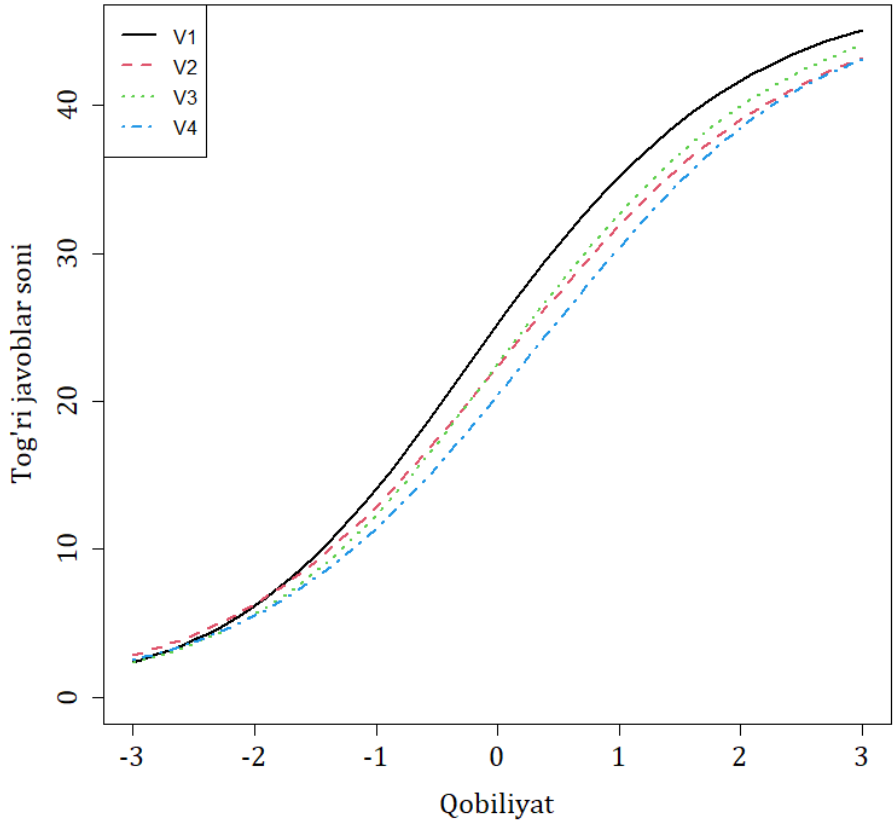
Rasmlarda har ikkala modelda ham sinaluvchilarning qobiliyat darajasi ortgani sari kutilayotgan to‘g‘ri javoblar soni qanday o‘zgarishi tasvirlangan. Agar biz sinaluvchilar qobiliyat darajalarini shartli ravishda quyi, o‘rtacha va yuqori qobiliyat darajalariga ajratib olsak, umumiy elemenli 4 ta variant uchun ham Rash modeli bilan hisoblangan test xarakteristikasi chiziqlari qobiliyat darajalarinini to‘g‘ri javoblar soniga nisbatan mos ravishda o‘zgarayotganligini ko‘rish mumkin (5a-rasm).

Ikki parametrli modelda esa quyi qobiliyatlar uchun to‘g‘ri javoblar soni barcha variantlarda deyarli bir xil o‘zgaradi. O‘rtacha va yuqori darajadagi qobiliyatlar uchun esa variantlarda farqlar kuzatiladi. V1 – variant bu modelda test topshiriqlariga to‘g‘ri javoblar soni barcha qobiliyatlarda kattaroq va bu variant qolgan variantlarga nisbatan eng oson. V4 – variant ikki parametrli modelda qolgan variantlarga nisbatan eng qiyin, ya’ni test topshiriqlariga to‘g‘ri javoblar soni qobiliyat oshgan sari natija ko‘rsatkichining o‘sishi sekin bo‘ladi.

Bu ikki parametrli modelda elementlar ajratish darajasi hisobga olinishi bilan izohlanadi. Ya’ni ba’zi variantlarda ajratish darajasi yuqori elementlar ko‘proq bo‘lib, bu test natijasi ma’lum qobiliyat sohalarida boshqalaridan ajralib turadi. Ushbu umumiy 4 ta umumiy elementli variantlardagi test topshiriqlarning qiyinlik va ajratish darajasi qiymatlariga asoslanib quyidagilarni




a)




b)

5-rasm. Rash (a) va ikki parametrli (b) model bo‘yicha test xarakteristikasi chiziqlari

aytish mumkin. Ya'ni test variantlari quyi qobiliyat darajalari uchun qiyin bo'lib, ularni o'z miqdorida ajratadi. O'rta darajadagi qobiliyatlar uchun o'rtacha darajada qiyin bo'lib, ularni yaxshi ajratadi. Yuqori qobiliyat darajalari uchun esa test topshiriqlari ancha oson va ajratish darajasi kichik bo'ladi.

Umuman olganda, har ikki modelda ham test variantlari o'r-tasidagi test xarakteristikasi chiziqlari yaqin bo'lsa-da, ikki parametrli modelda farqlar biroz ko'proq. Bu esa ikki parametrli modelining test strukturasi va ajratish darajasini aniqroq aks ettirilishini ko'rsatadi. Ikki parametrli model ajratish darajasi yuqori bo'lgan topshiriqlar soni ko'p bo'lsa, qobiliyat darajasi yuqori bo'lgan sinaluvchilarni aniqroq ajrata oladi. Rash modeli esa o'lchashlarda barqaror, biroq individual farqlarni aniqlashda unchalik sezgir emas.

Shu boisdan tanlangan model maqsadga muvofiq holda belgilanmog'i lozim. Agar test variantlari o'rtasida maksimal tenglik va moslik kerak bo'lsa, unda Rash modelidan, elementlarining ajratish darajasi hisobga olinishi kerak bo'lsa, ikki parametrli model afzaldir.

## Xulosa

Ushbu tadqiqotda biologiya fanidan o'tkazilgan test sinovlarida foydalanilgan umumiy elementli 4 ta variantlar uchun test ma'lumoti va xarakteristikasi chiziqlari Rash va ikki parametrli modellar bilan tahlil qilindi.

Ikki parametrli modelda ele-mentlarning qiyinlik darajasi para-metridan tashqari ajratish darajasi parametri bilan ham ma'lumotlar moslashtirilganligi sababli mazkur model bilan hisoblangan element xarakteristikasi chiziqlari bilan taj-ribada olingan ma'lumotlar mosligi yaxshi bo'ladi. Biroq ular Rash mo-delidagi parametr kabi barqaror bo'lmasligi mumkin.

Rash modeli oddiy baholashlar uchun yetarli bo'ladi, lekin variantlar turlicha ajratish darajasiga ega topshiriqlardan tashkil topgan bo'lsa, ikki parametrli model ustunlik qiladi.

Ikki parametrli modelida test ma'lumoti chiziqlari yuqori qobiliyatga ega guruh uchun Rash modeliga qaraganda ko'proq ma'lumot berishi aniqlandi.

Agar tadqiqot maqsadi turli test variantlarini ishonchli taqqoslash yoki barqaror o'lchashga asoslanish zarur bo'lsa, ikki parametrli model ustunlik qilishi ko'rsatib berildi.

Agar test elementlari oldindan muvozanatli tuzilgan bo'lsa va tu-shuntirishlar soddaligi ustuvor bo'lsa, Rash modeli samarali hisoblanadi.

## ADABIYOTLAR


1. R.K. Hambleton, H. Swaminathan, H.J. Rogers. Fundamentals of item response theory. Newbury Park, CA: Sage, 1991.
2. R.J. De Ayala, S.Y. Santiago. An introduction to mixture item response theory models. Journal of School Psychology, 60, 25-40, 2017.
3. C. DeMars. Item response theory. New York: Oxford University Press. 2010.
4. I.J. Egberink, R.R. Meijer, B. P. Veldkamp. Conscientiousness in the workplace: Applying mixture IRT to investigate scalability and predictive validity. Journal of Research in Personality, 44, 232–244, 2010.
5. S.E. Embretson, & S.P. Reise. Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
6. G. Rash. Probabilistic models for some intelligence and attainment tests, Copenhagen, Danish Institute for Educational research, 1960.
7. R.K. Hambleton and W.J. Van der Linden. Advances in item response theory and applications: An introduction, Advanced Psycological Measurement, 6, 373-378, 1982.
8. Baker, Frank. The Basics of Item Response Theory, ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD, 2001.
9. Ivailo Partchev. A visual guide to item response theory, FriedrichSchiller-Universitat¨ Jena, 2004.
10. M.Dj. Ermamatov, A.B. Normurodov, A.R. Sattiyev. Klassik va zamonaviy test nazariyalarida element xususiyatlari, Axborotnoma №2, 4-26, 2025.
11. F.M. Lord. A theory of test scores, Psychometric Monograph. Psychometric Society, No.7, 1952.
12. A. Birnbaum. Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord, M.R. Novick, Statistical theories of mental test scores (chapter 17-20), Reading, M.A.: Addison-Wisley.
13. Gunter Maris, Timo Bechger, Jesse Koops and Ivailo Parchev, Data Management and Analysis of Tests, p. 1-49, 2022.

14. R.P. McDonald. The dimensionality of tests and items, British Journal of Mathematical and Statistical Psychology, 34, pp. 100–117, 1981.

15. R.P. McDonald. Linear versus nonlinear models in item response theory, Applied Psychological Measurement, 6(4), pp. 379–396, 1982.

# COMPARISON OF TEST INFORMATION AND TEST CHARACTERISTIC CURVES USING RASH AND 2PL MODELS

**A.B. Normurodov, M.J. Ermamatov**

*Scientific and Educational Practical Center Under the Agency for Assessment of Knowledge and Competences*, a.normurodov@uzbmb.uz

**Abstract.** This paper presents a study on the results of a biology test conducted for scientific research, test information and characteristic curves were analyzed using the Rash model and two-parameter models of modern test theory. In the study, the information functions of the test with respect to ability were compared under both models. The results demonstrated the importance of model selection in improving the validity and reliability of the test. It was found that the test information curves calculated using the two-parameter model provide more information for higher ability levels.

**Keywords:** test information curve , test characteristic curve, Rash model, 2PL model.

# FIZIKA FANINING OPTIKA BO‘LIMINI O‘QITISHDA TALABALARNING BILIM VA MALAKALARINI BAHOLASH MEXANIZMI

**X.N. Qodirova**
*Andijon davlat pedagogika instituti*
*“Fizifa va texnoliogik ta’lim” kafedrasi dotsenti v.b., p.f.f.d. (PhD)*
khamidaqodirova1983@gmail.com

**Qisqacha mazmuni.** Ushbu maqolada fizika fanining optika bo‘limini o‘qitish jarayonida talabalarning bilim va ko‘nikmalarini baholashning zamonaviy usullari va mezonlari tahlil qilingan. Test, laboratoriya ishi va portfolio asosida baholash usullari yoritilgan. Shuningdek, ta’lim sifatini oshirishda diagnostika va refleksiyaning o‘rni ham ko‘rsatilgan. Talabalarda optika bo‘limini innovatsion yondashuv asosida o‘zlashtirish darajasini baholash mezonlari va shakllanganlik darajalari ko‘rib chiqilgan.

**Kalit so‘zlar:** optika, baholash mexanizmi, diagnostika, kompetensiya, innovatsion yondashuv, refleksiya, baholash mezonlari va o‘zlashtirish, shakllanganlik darajalari.

## 1. Kirish

Yurtimizda fizika fanini chuqur va sifatli o‘qitishga alohida e’tibor qaratilib, ilm-fan taraqqiyoti hamda zamonaviy va uzluksiz ta’lim tizimini yanada takomillashtirishga qaratilgan keng ko‘lamli islohotlar izchil amalga oshirilmoqda. Davlat tomonidan “Fizika sohasidagi ta’lim sifatini oshirish va ilmiy-tadqiqotlarni rivojlantirish chora-tadbirlari to‘g‘risida”gi qarorining qabul qilinishi ham fizika sohasi oldida bir qancha vazifalarni belgilab berdi.

Bugungi kunda ta’lim muassasalarida fizika fanini o‘qitish sifatini oshirish, ta’lim jarayoniga zamonaviy pedagogik texnologiyalarni joriy etish, iqtidorli talabalarni saralash, yuqori malakali kadrlarni tayyorlash, ilmiy-tadqiqot va innovatsiyalarni rivojlantirish hamda amaliy samaradorlikka yo‘naltirishga katta e’tibor qaratilmoqda. Shu bilan birga, sohada yechimini topmagan qator masalalar, jumladan, fizika sohasidagi raqobatbardosh mutaxassislarni tayyorlashga, bilim va malakalarini baholashga, ta’lim sifati va ilmiy-tadqiqot samaradorligini oshirishga qaratilgan chora-tadbirlarni amalga oshirish zaruratini ko‘rsatmoqda [1].

Olimlar baholash faoliyatining asosiy funksiyalari sifatida ta’lim-tarbiyaviy, diagnostik va nazorat funksiyalarini alohida ajratib ko‘rsatadilar. Ularning fikricha baholash jarayoni ta’lim-tarbiya jarayonida talabalardan mustaqil faoliyatni talab

qilmaydi, balki u asosan o‘qituvchilarga bog‘liq bo‘lgan jarayon bo‘lib, u o‘qituvchilarning diagnostik faoliyatiga tengdir [2].

O‘qituvchilarning baholash faoliyati talabalarning egallagan bilimlarini tizimli nazorat qilish jarayonidan iborat bo‘lib, ushbu jarayon natijasida ularning bilim, ko‘nikma va munosabatlar darajasi aniqlanadi va baholanadi.

Sh.A. Amonashvili tomonidan ta’kidlanishicha, baho – baholash jarayonining natijasi va u inson tomonidan amalga oshiriladi. Amonashvili bahoni talabalarda o‘zlarining ta’lim jarayonidagi o‘sishlarini mustaqil ravishda baholashni o‘rgatish deb beradi. Baholash funksiyalari ichida olim rag‘batlantiruvchi funksiyaning ahamiyatini alohida ta’kidlaydi va uni ta’lim jarayonining muhim tarkibiy qismi sifatida ko‘rsatadi. Uning ta’kidlashicha, baholash ta’lim-tarbiya jarayonida talabalarning o‘qishga bo‘lgan ichki motivatsiyasini kuchaytiruvchi eng muhim omillardan biridir.

Baholash jarayonida o‘qituvchi talabalarning shakllangan va rivojlangan bilim, ko‘nikma va malakalarini sifatli baholashi, ularni takomillashtirish yo‘llarini ko‘rsatishi, nima va qanday qilish lozimligini hamda qaysi namunaga amal qilishni ko‘rsatishi kerak [3].

G.Yu. Ksenzova o‘zining tadqiqotida o‘qituvchining baholash faoliyati atamasini qo‘lladi. Uning fikricha, o‘qituvchining baholash faoliyatining maqsadi talabalarda o‘z-o‘zini mustaqil baholash ko‘nikmasini rivojlantirishdir. Olim baholash faoliyatining obyekti haqida gapirganda, u talabaning individual o‘sish jarayonida mustaqil rivojlanishini ko‘zda tutadi. Shunday qilib, o‘qituvchining baholash faoliyatining natijasi talabaning o‘z-o‘zini adekvat baholashi hisoblanadi [4]. Bryan va Clegglarning fikriga ko‘ra, aksariyat hollarda baholash ta’lim jarayoni tugagandan so‘ng amalga oshiriladi, vaholanki, u ta’lim berish va ta’lim olish jarayonining bir bo‘lagi sifatida amalga oshirilishi kerak deb ta'kidlaydi [5].

Fizika ta’limi yo‘nalishida tahsil olayotgan talabalarning ta’lim jarayonida shakllanayotgan bilimlari, ko‘nikmalari va malakalarini aniq hamda xolisona baholash ta’lim tizimining samaradorligini oshirishda muhim ahamiyat kasb etadi. Bo‘lajak fizika o‘qituvchilarining ta’lim sohasida mavjud zamonaviy moddiy-texnika bazasidan samarali foydalanish imkoniyatlarini kengaytirish, oliy ta’lim muassasalari malaka talablarini, o‘quv dasturlari va metodik ta’minotni takomillashtirish hamda ularni zamon talablari asosida qayta ko‘rib chiqish masalasiga katta ahamiyat berilmoqda. Shu sababli baholashning ilg‘or texnologiyalaridan foydalangan holda oliy ta’lim mazmuni va sifatini xalqaro darajaga ko‘tarishda

pedagoglarning baholash kompetensiyasini takomillashtirish dolzarb ahamiyat kasb etadi.

Ushbu maqolada fizika fanining optika bo‘limini innovatsion yondashuv asosida o‘qitishda talabalarning bilimlarini baholashning mexanizmlari va talabalarning o‘zlashtirish darajasiga asoslanib baholash mezonlari o‘rganib chiqilgan.

## 2. Natijalar va muhokamalar

Bugungi kunda oliy ta’lim tizimida talabalar bilimini baholashdagi asosiy maqsadlardan biri baholash tizimida xalqaro innovatsion yondashuvlardan keng foydalanishdan iborat. Yuqorida keltirilgan misollar bugungi kunda o‘qituvchining baholash faoliyatiga nisbatan turli yondashuvlar mavjudligini ko‘rsatadi. Biroq ushbu yondashuvlardan qat’i nazar, baholash faoliyatining mohiyatini belgilovchi asosiy omillar mavjud bo‘lib, xususan, baholash faoliyati [6]:

kasbiy-pedagogik faoliyatning ajralmas tarkibiy qismi bo‘lib, o‘qituvchining kasbiy vazifalaridan biridir;

talabani shaxs sifatida shakllantirish va uning o‘zini anglash jarayoniga xizmat qiladi;

ta’lim jarayonini rivojlantiruvchi faoliyat turi sifatida namoyon bo‘ladi.

Bugungi kunda baholashning asosiy maqsadi nafaqat talabalarda o‘zlashtirilgan bilimlarni baholash, balki ularda o‘z-o‘zini baholash ko‘nikmasini ham rivojlantirishdan iborat. Baholash jarayonini samarali olib borish uchun o‘qituvchilar ham baholash kompetensiyasiga ega bo‘lishlari lozim.

Fizika fanining optika bo‘limini innovatsion yondashuv asosida o‘qitishda talabalarning nafaqat nazariy bilimlarini, balki amaliy ko‘nikmalarini, ijodiy fikrlash salohiyatini hamda muammolarga ilmiy yechim topa olish qobiliyatini hisobga olish muhim ahamiyat kasb etadi [7]. Bunda quyidagi kompetensiyalarni shakllantirishga alohida e’tibor qaratiladi:

**Ilmiy tafakkur** – optik hodisalarni ilmiy asosda tahlil qilish va izohlash qobiliyatini bildiradi;

**Amaliy ko‘nikma** – laboratoriya sharoitida optik asbob-uskunalar bilan samarali ishlay olish malakasini namoyon etadi;

**Raqamli savodxonlik -** raqamli texnologiyalardan faol, ma’suliyatli, tanqidiy va maqsadli foydalana olish qobiliyati va bu qobiliyat shaxsiy, ta’limiy, kasbiy va ijtimoiy sohalardagi keng kompetensiyalar bilan uzviy bog‘langan bo‘ladi. Raqamli savodxonlik – insonning raqamli vositalardan (kompyuter, internet, mobil qurilmalar, ilovalar, platformalar va

h.k.) axborot izlash, baholash va tahlil qilish, yaratish va tarqatish, onlayn muloqot qilish, shaxsiy ma'lumotlarni himoya qilish, va raqamli muhitda xulqatvor kabi sohalarda ongli va samarali foydalanish qobiliyatini o'z ichiga oladi.

**Ijodiy fikrlash** – muammoli vaziyatlarda mustaqil va samarali yechim ishlab chiqish qobiliyatini namoyon qiladi.

Fizika fanining optika bo'limini o'qitish jarayonida talabalar nurlanish, interferensiya, difraksiya va qutb-lanish kabi fundamental optik hodisalar mohiyatini chuqur anglab, ularning fizik asoslarini tahlil qilish, yorug'lik manbalari, optik asboblar va ularning ishlash jarayoni kabi ko'nikmalarini shakllantiradi. Labo-ratoriya sharoitida yorug'lik hodi-salarini tadqiq qilish ko'nikmasiga ega bo'ladi. Bu maqsadlar talabalarda fizik tafakkur, tahlil qilish qobiliyati va ilmiy tadqiqot ko'nikmasini rivoj-lantiradi.

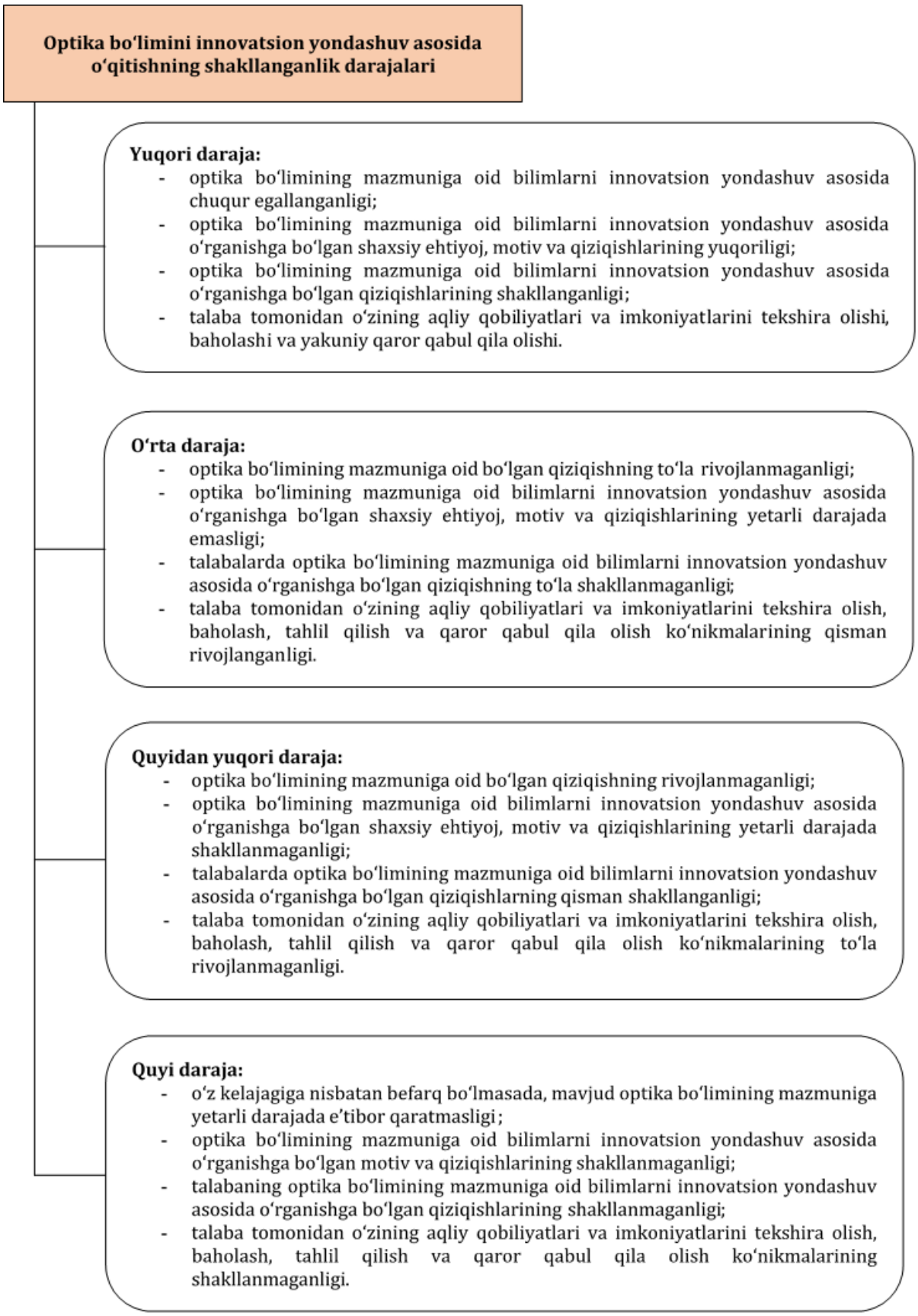
Yuqoridagi fikrlarni inobatga olgan holda, talabalarning fizika fanining optika bo'limini o'zlash-tirishda, mashg'ulotlarda egallagan bilim, malaka, ko'nikma va kom-petensiyalarini baholashda, baholash mexanizmi, baholash mezonlari va darajalariga asoslanish kerak.

Bilimlarni baholash mexanizimi quyidagi yo'nalishlarda amalga oshiriladi:

**Diagnostik baholash** – bu baholash turi ta'lim jarayonining boshlang'ich bosqichida o'tkaziladi. U orqali talabalarning avvalgi bilimlari, bazaviy tayyorgarlik darajasi hamda individual qobiliyatlari aniqlanadi.

**Joriy baholash** – bu test topshiriqlari, amaliy va laboratoriya ishlarining natijalari, guruhli muhokamalar, kichik loyihalar hamda elektron ta'lim platformalari orqali avtomatik baholash shaklida amalga oshiriladi.

**Yakuniy baholash** – bu baholash turi semestr yoki fan yakunida talabalarning umumiy bilim darajasini aniqlashga qaratilgan bo'lib, unda yozma imtihonlar, laboratoriya hisobotlari, elektron testlar, kurs ishlari yoki kichik tadqiqot loyihalari natijalari inobatga olinadi. Bundan tashqari yakuniy baholashda portfolio usuli ham mavjud bo'lib, bunda har bir talaba o'zining o'quv natijalari, laboratoriya ishlari, taqdimotlar va mustaqil ishlarini elektron shaklda to'plab, o'qituvchiga taqdim etadi. Ushbu usul talabalarning fan bo'yicha individual rivojlanish dinamikasini tizimli ravishda kuzatish imkonini beradi.




1-rasm. Talabalarda optika bo'limini innovatsion yondashuv asosida o'zlashtirishning shakllanganlik darajalari

Talabalarda optika bo‘limi innovatsion yondashuv asosida o‘zlashtirishning shakllanganlik dara-jasini aniqlashda onlayn test topshiriqlari, nazorat savollariga yozma javoblar, shuningdek, taqdimotlar tayyorlash orqali amalga oshirish mumkin. Biz talabalar bilimini baholashda o‘zlashtirishning to‘rt bosqichdan iborat shakllanganlik darajalarini ishlab chiqdik (1-rasm). Talabalarning optika bo‘limini o‘zlashtirish darajasiga asoslanib, quyidagi mezonlar bo‘yicha baholash mumkin (2-rasm):

1. **Motivatsion** - optika bo‘limini innovatsion yondashuv asosida o‘zlashtirishga oid bilimlarni o‘rganishga bo‘lgan shaxsiy ehtiyoj, motiv va qiziqish.

2. **Kognitiv** - talabalarda optika bo‘limini innovatsion yondashuv asosida o‘zlashtirishga oid bilimlarni o‘rganishga qo‘yiladigan shartlarning to‘la anglanganligi.

3. **Reproduktiv** – talabalarda optika bo‘limini innovatsion yondashuv asosida o‘zlashtirishga oid bilimlarni chuqur egallangani.

4. **Produktiv** – talaba tomonidan o‘zining aqliy qobiliyatlari va jismoniy imkoniyatlarini tekshirish, baholash, qayta aloqa va yakuniy qaror qabul qilish.


2-rasm. Talabalarda optika bo‘limini innovatsion yondashuv asosida o‘zlashtirish darajasini baholash mezonlari

Bo‘lajak fizika fani o‘qituvchilarini bakalavr ta’lim bosqichida optika bo‘limini o‘qitishda dars mash-g‘ulotlarini innovatsion-integratsion zamonaviy shaklda maxsus tashkil etilgan, tashkiliy va amalga oshiruvchi, qadriyatga yo‘naltirilgan, mazmunli-texnologik, maqsadga qaratilgan, rag‘batlantiruvchi-motivatsion, kogni-tiv-akslantiruvchi, hissiyirodali, nazo-rat qiluvchi, amalga oshiruvchi, tartibga soluvchi, natija va baholash kabi pedagogik-psixologik, ergonomik shart-sharoitlar optika bo‘limini o‘qitishni takomillashtirishning o‘ziga xos bo‘lgan eng ustivor ko‘rinishlari sifatida talqin etish zarur [8].

## Xulosa

Xulosa o‘rnida shuni aytish mumkinki, baholashning ta’lim berish va ta’lim olish jarayonida o‘z o‘rni bor, bu nafaqat fizika fanining “Optika” bo‘limini o‘qitish jarayonida, balki ta’limning barcha ko‘rinishlarida mavjud. Baholash o‘qituvchi, talaba, materiallar, kontekst va ta’lim jarayonining sifatini ko‘rsatuvchi beshta komponent bilan o‘zaro bog‘liqdir. Samarali qo‘llangan nazorat turi kontentning to‘g‘riligi va dars jarayoni qanday kechishi haqida ma'lumot olish uchun talabalarning o‘zlashtirish ko‘rsatgichlarini tahlil qilishni o‘zi yetarlidir. Mamlakat taraqqiyot yo’lida malakali kadrlar, yuqori saviyadagi bilimli muta-xassislarni tayyorlab berishning xalqaro tan olingan baholash tizimlari, usul va vositalaridan unumli foydalanishni, izchil ravishda doimiy oshirib borish zarur. Natijaga esa ta’lim tizimining asosiy bo‘g‘ini bo‘lgan professor - o‘qituvchilarga mehnat faoliyatlarida o‘sish imkoniyatlarini yaratib berish orqali erishish mumkin.

## ADABIYOTLAR


1. O‘zbekiston Respublikasi Prezidentining 2021- yil 19- martdagi “Fizika sohasidagi ta’lim sifatini oshirish va ilmiy-tadqiqotlarni rivojlantirish chora-tadbirlari to‘g‘risida”gi PQ-5032-sonli qarori 2021. www.lex.uz.
2. Muradkasimova K. Sh. Chet tillarni baholash kompetensiyasi: amaliyotdagi holat va ehtiyojlar. Scientific-methodological electronic journal “Foreign Languages in Uzbekistan”, 2022, No 3 (44),72-85
3. Ш. А. Амонашвили. “Педагогические притчи”-Москва.: -2010, ст-65.
4. Ксензова Г.Ю. Оценочная деятельность учителя // Педагогическое общество России. – М.; 1999. – 46 с.
5. Bryan C. & Clegg K. Innovative Assessment in Higher education. New York. Routledge. 2006. – P 82-83.
6. Кузьмина.Н.В. Профессионализм личности преподавателя и мастера производственного обучения. – М.: «Высшая школа», 1990.
7. Х.Н. Қодирова Инновацион ёндашув асосида оптика бўлимини ўқитишнинг методик таъминоти// Янгиланаётган Ўзбекистонда хотин-қизларнинг илм-фан тараққиётидаги роли ва гендер тенглиги мавзусидаги I Халқаро илмий-амалий анжуман материаллари тўплами, 2-3-март, Тошкент, 2021-йил, 78-81-бетлар.
8. Qodirova.X.N. Bo‘lajak fizika fani o‘qituvchilarini tayyorlashda mustaqil ta’lim jarayonini tashkil etish. Zamonaviy mikroelektronikaning rivojlanishida fan, ta’lim va innovatsiya integratsiyasi. Respublika ilmiy-uslubiy anjumani materiallari. Andijon - 2020 yil. 24-25 dekabr. -B 434-435.

# ASSESSMENT MECHANISM OF STUDENTS' KNOWLEDGE AND COMPETENCIES IN THE TEACHING OF THE OPTICS SECTION OF PHYSICS

**Kh.N. Kodirova**

*Andijan State Pedagogical Institute Acting Associate Professor of the Department of "Physics and Technological Education", PhD in Pedagogical Sciences*

khamidaqodirova1983@gmail.com

**Abstract.** In this paper, analyzes modern methods and criteria for assessing students' knowledge and skills in the process of teaching the optics section of physics. It highlights assessment approaches based on tests, laboratory work, and portfolios. The role of diagnostics and reflection in improving the quality of education is also discussed. Furthermore, the article examines the criteria for evaluating the level of students' mastery of optics based on an innovative approach, along with the levels of competency formation.

**Keywords:** optics, assessment mechanism, diagnostics, competence, innovative approach, reflection, assessment criteria, mastery, levels of competency formation.