

**Information about the authors:**

*Abdikayumova Nodira Abdirakhimovna\* – PhD student, O.Zhanibekov South Kazakhstan Pedagogical University, Republic of Kazakhstan, 160012, Shymkent, 13 A. Baitursynov Str., tel.: +77024555030; e-mail: nodira.086@mail.ru.*

*Madybekova Galiya Mutaliyevna – Candidate of Chemical Sciences, Professor, O.Zhanibekov South Kazakhstan Pedagogical University, Republic of Kazakhstan, 160012, Shymkent, 13 A. Baitursynov Str., tel.: +77477014368; e-mail: galiya56@list.ru.*

*Абдикаюмова Нодира Абдирахимовна\* – PhD докторанты, Ө.Жәнібеков атындағы Оңтүстік Қазақстан педагогикалық университеті, Қазақстан Республикасы, 160012 Шымкент қ., А.Байтұрсынов көш. 13, тел.: +77024555030; e-mail: nodira.086@mail.ru.*

*Мадыбекова Галия Муталиевна – химия ғылымдарының кандидаты, профессор, Ө.Жәнібеков атындағы Оңтүстік Қазақстан педагогикалық университеті, Қазақстан Республикасы, 160012 Шымкент қ., А.Байтұрсынов көш. 13, тел.: +77477014368; e-mail: galiya56@list.ru.*

*Абдикаюмова Нодира Абдирахимовна\* – PhD докторант, Южно-Казахстанский педагогический университет имени О. Жанибекова, Республика Казахстан, 160012 г. Шымкент, ул. А. Байтұрсынова 13, тел.: +77024555030, e-mail: nodira.086@mail.ru.*

*Мадыбекова Галия Муталиевна – кандидат химических наук, профессор, Южно-Казахстанский педагогический университет имени О. Жанибекова, Республика Казахстан, 160012 г. Шымкент, ул. А. Байтұрсынова, 13, тел.: +77477014368, e-mail: galiya56@list.ru.*

XFTAP 14.35.01.

ӨОЖ 001.891

<https://doi.org/10.52269/SRDG2612014>

**ҚАЗТЕСТ ЖҮЙЕСІНІҢ ТЕСТ ТАПСЫРМАЛАРЫНА  
ПСИХОМЕТРИЯЛЫҚ ТАЛДАУ ЖҮРГІЗУ НӘТИЖЕЛЕРІ**

*Абдрасилов Б.С. – ҚР ҰҒА академигі, Ұлттық тестілеу орталығының директоры, Астана қ., Қазақстан Республикасы.*

*Ерсултанова Г.С.\* – PhD докторант, ҚАЗТЕСТ жүйесін дамыту жөніндегі басқарма басшысы, Ұлттық тестілеу орталығы, Астана қ., Қазақстан Республикасы.*

*Махмұтов Б.Б. – химия ғылымының кандидаты, доцент, директор кеңесшісі, Ұлттық тестілеу орталығы, Астана қ., Қазақстан Республикасы.*

*Тоғызбай Н.Б. – магистр, ғылыми-әдістемелік зертхана меңгерушісі, Ұлттық тестілеу орталығы, Астана қ., Қазақстан Республикасы.*

*Бұл мақалада ҚАЗТЕСТ жүйесінің тыңдалым және оқылым блоктары бойынша тест тапсырмаларының психометриялық көрсеткіштері Раш (Rasch) моделіне негізделіп, Winsteps 5.8.3.0 бағдарламасы арқылы талданды. Мұндай тәсіл тапсырмалардың сапасын, тест құрылымының теңгерімділігін және тестіленушілердің тілдік қабілеттерін объективті бағалауға мүмкіндік береді. Талдау барысында тест элементтерінің қиындық деңгейі, сенімділігі, ажыратушылық қабілеті және тестіленушілердің нәтижелері қарастырылды. Нәтижелер ҚАЗТЕСТ тест тапсырмаларының сапасының жоғары екенін көрсетті: Item Reliability – 0.99, Person Reliability – 0.76, Cronbach's Alpha – 0.84. Бұл көрсеткіштер тесттің ішкі біртұтастығы мен психометриялық тұрақтылығын дәлелдейді. Тыңдалым блогы бойынша мәліметтер тестіленушілердің мәтінді түсіну деңгейлері айқын ажыратылатынын көрсетті. Дегенмен төмен деңгейлі қатысушылардың нәтижелерін дәлірек анықтау үшін тыңдалым мәтіндерін екі рет тыңдату тәжірибесін енгізу ұсынылады. Оқылым блогы бойынша тапсырмалардың қиындық деңгейі –5.12 мен +2.63 logits аралығында бөлінді, бұл тесттің әртүрлі деңгейдегі тілдік құзыреттілікті сенімді өлшей алатынын дәлелдейді. Тапсырмалар мазмұнының және пішінінің үйлесімділігі тесттің валидтілігін қамтамасыз етеді. Жалпы зерттеу нәтижелері ҚАЗТЕСТ тесттерін психометриялық тұрғыдан жетілдіруге және халықаралық тілдік тестілермен үйлесімділігін арттыруға мүмкіндік береді. Тапсырмаларды күрделілік деңгейіне қарай тең бөлу, когнитивтік дағдыларды қамтитын элементтер үлесін арттыру және Раш моделіне сүйеніп тест спецификацияларын қайта қарау ұсынылады. Бұл зерттеу ҚАЗТЕСТ жүйесінің сапасын арттыруға және оны халықаралық стандарттарға жақындатуға бағытталған маңызды эмпирикалық негіз қалыптастырады.*

**Түйінді сөздер:** ҚАЗТЕСТ, психометриялық талдау, Раш моделі, тыңдалым, оқылым, сенімділік, тілдік құзыреттілік.

**РЕЗУЛЬТАТЫ ПСИХОМЕТРИЧЕСКОГО АНАЛИЗА ТЕСТОВЫХ ЗАДАНИЙ СИСТЕМЫ КАЗТЕСТ**

Абдрасилов Б.С. – академик НАН РК, директор Национального центра тестирования, г. Астана, Республика Казахстан.

Ерсултанова Г.С.\* – PhD докторант, руководитель управления по развитию системы КАЗТЕСТ, Национальный центр тестирования, г. Астана, Республика Казахстан.

Махмутов Б.Б. – кандидат химических наук, доцент, советник директора, Национальный центр тестирования, г. Астана, Республика Казахстан.

Тогызбай Н.Б. – магистр, заведующая научно-методической лабораторией, Национальный центр тестирования, г. Астана, Республика Казахстан.

В статье представлен комплексный психометрический анализ заданий блоков аудирования и чтения системы КАЗТЕСТ, проведённый на основе модели Раша (Rasch) с использованием программы Winsteps 5.8.3.0. Такой подход обеспечивает объективную оценку качества заданий, сбалансированности теста и языковых способностей тестируемых. Анализ охватил показатели трудности заданий, их надёжности, способности к дифференциации и результаты участников тестирования. Результаты показали высокий уровень качества заданий КАЗТЕСТ: надёжность заданий – 0.99, надёжность тестируемых – 0.76, коэффициент  $\alpha$  Кронбаха – 0.84, что подтверждает внутреннюю согласованность и психометрическую стабильность теста. Данные по блоку аудирования выявили чёткое разграничение уровней понимания текста. Для более точной оценки результатов участников низкого уровня предлагается внедрить практику двукратного прослушивания текстов, что соответствует международным стандартам (TOEFL, IELTS, TestDaF). В блоке чтения уровень трудности заданий варьируется от  $-5.12$  до  $+2.63$  логитов, что подтверждает способность теста надёжно измерять языковую компетенцию на базовом, среднем и высоком уровнях. Содержание и структура текстов обеспечивают валидность и репрезентативность теста. Полученные результаты создают эмпирическую основу для совершенствования психометрического качества КАЗТЕСТ и его гармонизации с международными тестовыми системами. Рекомендуется оптимизировать структуру по уровням сложности, увеличить долю заданий, требующих когнитивных операций, и пересмотреть спецификации теста на основе данных модели Раша. Исследование способствует повышению надёжности, объективности и интерпретационной точности национальной системы языкового тестирования.

**Ключевые слова:** КАЗТЕСТ, психометрический анализ, модель Раша, аудирование, чтение, надёжность, языковая компетенция.

**RESULTS OF PSYCHOMETRIC ANALYSIS OF TEST ITEMS IN THE KAZTEST SYSTEM**

Abdrassilov B.S. – Academician of the National Academy of Sciences of the Republic of Kazakhstan, Director of the National Testing Center, Astana, Republic of Kazakhstan.

Yersultanova G.S.\* – PhD student, Head of the Department for the Development of the KAZTEST System, National Testing Center, Astana, Republic of Kazakhstan.

Makhmutov B.B. – PhD, Associate Professor, Advisor to Director, National Testing Center, Astana, Republic of Kazakhstan.

Togyzbay N.B. – Master, Head of the Scientific and Methodological Laboratory, National Testing Center, Astana, Republic of Kazakhstan.

This article presents a comprehensive psychometric analysis of the listening and reading sections of the KAZTEST system based on the Rasch model using the Winsteps 5.8.3.0 software. This approach ensures objective evaluation of task quality, test balance, and test-takers' language abilities. The analysis examined test item difficulty, reliability, separation indices, and test-taker performance indicators. The results confirmed the high quality of KAZTEST test items: Item Reliability – 0.99, Person Reliability – 0.76, and Cronbach's Alpha – 0.84, demonstrating strong internal consistency and psychometric stability. Data from the listening section showed a clear differentiation of comprehension levels among test-takers. However, to enhance the accuracy of low-level performance assessment, it is recommended to allow two listenings of the audio texts, a practice widely used in international tests such as TOEFL, IELTS, and TestDaF. In the reading section, item difficulty ranged from  $-5.12$  to  $+2.63$  logits, confirming the test's ability to reliably measure language proficiency across basic, intermediate, and advanced levels. The balance of text content, length, and linguistic complexity supports test validity and content representativeness. Overall, the findings provide a solid empirical basis for improving the psychometric quality of KAZTEST and enhancing its alignment with international language testing standards. Recommendations include optimizing the structure by difficulty levels, increasing cognitively demanding tasks, and revising test specifications according to Rasch model data. The study contributes to strengthening the reliability, fairness, and interpretive accuracy of the national language testing system.

**Key words:** KAZTEST, psychometric analysis, Rasch model, listening, reading, reliability, language proficiency.

**Кіріспе.** Қазіргі таңда тілдік құзыреттілікті бағалаудың сапалы жүйелерін қалыптастыру мемлекеттік тіл саясатының маңызды құрамдас бөлігі болып табылады. Мемлекеттік тілдің қоғамдық және кәсіби салаларда тиімді қолданылуы оны меңгеру деңгейін объективті бағалау тетіктерін талап етеді [1]. ҚАЗТЕСТ жүйесі – қазақ тілін меңгеру деңгейін халықаралық стандарттарға сай бағалауға бағытталған ұлттық құрал. Жүйенің деңгейлік құрылымы мен бағалау қағидаттары халықаралық тілдік тестілеу тәжірибесіне негізделген. Қазақстан Республикасы Үкіметінің 2023 жылғы 16 қазандағы №914 Қаулысымен бекітілген «Қазақстан Республикасында тіл саясатын дамытудың 2023 – 2029 жылдарға арналған тұжырымдамасына» сәйкес ҚАЗТЕСТ жүйесі бойынша тестілеуге мемлекеттік қызметшілер, бюджеттік ұйымдар қызметкерлері, ұлттық холдингтер мен компаниялар өкілдері, сондай-ақ түрлі сала мамандары қатысады [2]. Осыған байланысты тілдік білім мен дағдыларды тәуелсіз және әділ бағалауға бағытталған тест тапсырмалары объективті және валидті сипатқа ие болуы қажет. Бұл өз кезегінде тестілеу нәтижелерінің сенімділігі мен салыстырмалылығын қамтамасыз етіп, бағалау үдерісінің ашықтығы мен әділдігін арттырады.

ҚАЗТЕСТ жүйесі 2006 жылы Ұлттық тестілеу орталығының базасында Қазақстан Республикасы азаматтарының және Қазақстан Республикасының аумағында әртүрлі қызметпен айналысатын шетел азаматтарының қазақ тілін білу деңгейін бағалау жүйесі ретінде құрылып, Қазақстан Республикасы Үкіметінің 2008 жылғы 11 қыркүйектегі № 835 Қаулысымен «ҚАЗТЕСТ жүйесін енгізу мен дамыту тұжырымдамасы» бекітілді. Ол TOEFL, IELTS (ағылшын тілін меңгеру деңгейін бағалау жүйесі), TÖMER (түрік тілін меңгеру деңгейін бағалау жүйесі), ТРКИ (орыс тілін меңгеру деңгейін бағалау жүйесі) сияқты тілдік білімді бағалаудың халықаралық жүйелерінің қағидаттары мен талаптарына негізделген, сондай-ақ қазақ тілін білу деңгейін бағалаудың отандық жүйесі болып табылады.

ҚАЗТЕСТ жүйесінің негізгі бағыттарының бірі – қазақ тілін меңгеру деңгейін бағалау мақсатында тест тапсырмаларын әзірлеу. Қазіргі таңда, Қазақстан Республикасы Ғылым және жоғары білім министрлігінің 2024 жылғы 24 қыркүйектегі № 458 бұйрығымен «ҚАЗТЕСТ қазақ тілін меңгеру деңгейін бағалау жүйесі бойынша сертификаттық тестілеуді өткізу қағидалары» бекітіліп, ҚАЗТЕСТ тестілеуі электронды және/немесе қағаз форматта өткізіледі [3].

ҚАЗТЕСТ жүйесі 4 блоктан тұрады және жабық, ашық тест тапсырмалары қолданылады. Бұл тілдік құзыреттіліктің рецептивті және продуктивті дағдыларын кешенді түрде бағалауға мүмкіндік береді [4].

– Тыңдалым блогы – диалогтар мен монологтарды тыңдау арқылы мәтінді түсіну дағдысын анықтауға арналған тест тапсырмалары.

– Оқылым блогы – әртүрлі мазмұндағы және стильдегі мәтіндерді оқып, түсіну дағдыларын анықтауға арналған тест тапсырмалары.

– Жазылым блогы – берілген сурет бойынша өз ойын жазбаша баяндау және берілген тақырыпқа эссе жазу.

– Айтылым блогы – әртүрлі тақырыптар бойынша қойылған сұрақтар/тақырыптар бойынша өз пікірін сауатты жеткізу ұсынылады [5, 365 б].

Бұл жүйеде қолданылатын тест тапсырмаларының сапасы мен олардың диагностикалық дәлдігі тесттің әділдігін және сенімділігін қамтамасыз етеді. Осы тұрғыда психометриялық талдау әдістерін қолдану тест сапасын бағалау мен жетілдірудің тиімді жолы болып табылады.

Зерттеудің **мақсаты** – ҚАЗТЕСТ жүйесінде қолданылатын тест тапсырмаларының сапасын психометриялық талдау әдістері негізінде ғылыми тұрғыда бағалау, олардың объективтілігі, валидтілігі мен сенімділік деңгейін айқындау және тестілеу жүйесін жетілдіруге бағытталған ғылыми-әдістемелік ұсыныстар әзірлеу.

Зерттеудің **міндеттері**:

1. ҚАЗТЕСТ жүйесіндегі тест тапсырмаларын әзірлеудің қолданыстағы әдістемелік қағидаттарын талдау.

2. Тыңдалым, оқылым, блоктары бойынша тест тапсырмаларының құрылымдық және мазмұндық ерекшеліктерін сипаттау.

3. Тест тапсырмаларының психометриялық көрсеткіштерін (күрделілік деңгейі, ажырату қабілеті, сенімділік коэффициенттері және т.б.) анықтау және талдау.

4. Тестілеу нәтижелерінің объективтілігі мен валидтілігін қамтамасыз етуге әсер ететін факторларды айқындау.

5. ҚАЗТЕСТ жүйесі бойынша сертификаттық тестілеудің сапасын арттыруға бағытталған ғылыми негізделген ұсыныстар ұсыну.

Бұл мақсат пен міндеттер зерттеудің теориялық және қолданбалы маңызын айқындап, ҚАЗТЕСТ жүйесінің тілдік тестілеудің халықаралық тәжірибесіне сүйене отырып, ұлттық тілдік бағалау құралы ретіндегі тиімділігін арттыруға бағытталады.

**Материалдар мен әдістері.** Қазіргі уақытта білім мен құзыреттілікті бағалаудың заманауи жүйелері стандартты тесттерді қолдануға негізделеді. Алайда тестілеудің тиімділігі мен объективтілігі көбіне тест тапсырмаларының сапасына және нәтижелердің дұрыс түсіндірілуіне байланысты. Осы міндеттерді шешу үшін психометриялық талдау қолданылады. Ол статистикалық және математикалық әдістердің кең ауқымын қамтиды. Психометриялық талдау әрбір тапсырманың негізгі параметрлерін, тапсырмалардың қиындық деңгейін, дискриминативтік қабілетін және сенімділігін анықтауға мүмкіндік береді [6], ол:

– Қиындық деңгейі – дұрыс жауап берген тестіленушілер үлесімен өлшенеді, бұл деңгейде тым жеңіл немесе аса күрделі тапсырмалар тесттің диагностикалық құндылығын төмендетеді;  
 – Дискриминативтік қабілет – тапсырманың күшті және әлсіз жақтарын ажырату қабілетін көрсетеді;  
 – Сенімділік және ішкі келісімділік – тест ішіндегі тапсырмалардың өзара байланысы (мысалы, Кронбах альфа коэффициенті арқылы).  
 – Дистракторлардың жұмыс істеуі – балама жауаптардың шынайылығы мен тиімділігі бағаланады.  
 Психометриялық талдау білім мен дағдылар деңгейін неғұрлым дәл анықтауға жағдай жасайды [7, 9 б].

Зерттеу барысында ҚАЗТЕСТ жүйесінің «Тыңдалым» және «Оқылым» блоктарына психометриялық талдау жүргізілді.

Бұл зерттеудің мақсаты – ҚАЗТЕСТ жүйесінің тыңдалым, оқылым блоктары бойынша тест тапсырмаларының сапасын психометриялық тұрғыда бағалау, олардың сенімділігі мен валидтілігін айқындау және қиындық деңгейінің сәйкестігін талдау.

Психометриялық бағалау Раш моделі негізінде «Winsteps 5.8.3.0» бағдарламасы арқылы жүзеге асырылады. Талдау барысында тапсырмалардың қиындық деңгейі, сенімділік коэффициенттері, ішкі үйлесімділігі және дистракторлардың жұмыс істеу тиімділігі анықталады [8, 42 б.].

Қосымша ретінде дәстүрлі тест теориясы (СТТ) көрсеткіштері – тапсырма қиындық индексі ( $\rho$ ), дискриминативтік қабілет ( $r_{pbis}$ ) және Кронбах альфа коэффициенті қолданылды. Сонымен қатар, Differential Item Functioning (DIF) талдауы тапсырмалардың әділдігін аймақтық және әлеуметтік топтар тұрғысынан бағалау үшін қарастырылды. Confirmatory Factor Analysis (CFA) әдісі арқылы тесттің құрылымдық валидтілігі тексерілді.

**Нәтижелер және талқылау.** Қазақ тілін меңгеру деңгейін бағалаудың маңызды аспектілерінің бірі – тыңдалым және оқылым. Тыңдалым және оқылым дағдыларын бағалау тілдік құзыреттілікті өлшеудің негізгі компоненттері болып табылады және халықаралық тестілеу тәжірибесінде кеңінен қолданылады. Тыңдалым блогы арқылы тестіленушінің қазақ тіліндегі мәтіннің мазмұнын түсіну, ақпаратты іріктеу және түсіндіру қабілеті өлшенеді. Оқылым блогы тестіленушінің қазақ тіліндегі жазбаша мәтіндерді түсіну, негізгі және жанама ақпаратты анықтау, мәтін құрылымын және логикалық байланыстарды тану қабілетін өлшейді [9, 23.б].

Тест тапсырмаларының сапасы мен сенімділігін қамтамасыз ету мақсатында психометриялық талдау жүргізу – ҚАЗТЕСТ жүйесінің ғылыми негізін күшейтудің негізгі тетіктерінің бірі болып табылады. Алынған психометриялық көрсеткіштер тапсырмалардың сенімділігі мен валидтілігінің жоғары екенін көрсетті, бұл халықаралық зерттеулер нәтижелерімен сәйкес келеді.

Cronbach L.J. «Coefficient alpha and the internal structure of tests» еңбегінде Кронбах альфа тестің немесе шкаланың ішкі келісімділігін (internal consistency) сипаттайтын көрсеткіш ретінде анықталады. Ол тест тапсырмаларының бір өлшенетін психологиялық қасиетке қаншалықты біртұтас және өзара байланысты екенін көрсетеді. Кронбах бойынша, альфа коэффициенті шкаладағы тапсырмалар арасындағы ковариацияға және олардың жалпы дисперсиясына негізделіп есептеледі және өлшеудің кездейсоқ қатеден қаншалықты еркін екенін айқындайды.

Ғылыми тұрғыда бұл көрсеткіш психологиялық өлшеулердің сенімділігін бағалаудың әмбебап индексі ретінде қарастырылады және тест нәтижелерін интерпретациялаудың, сондай-ақ оларды ғылыми және практикалық мақсатта қолданудың негізділігін қамтамасыз етеді [10, 297–334 бб].

Осыған орай ҚАЗТЕСТ жүйесі бойынша «Тыңдалым» блогының психометриялық талдау есебін қарастырайық. Талдау «Тыңдалым» блогынан №364370 тест нұсқасы негізінде орындалған нәтижелер бойынша жүргізілді. Зерттеу деректері 1199 тестіленушінің 20 тест тапсырмасына берген жауаптарына негізделді.

1-кесте. – Тест тапсырмаларының көрсеткіштері (Item)

Көрсеткіш	Мәні
Орташа көрсеткіш (Item Mean Measure)	0.00 logits
Item Separation	10.03
Item Reliability	0.99
Қиындық диапазоны	-3.12 → +3.65 logits
Infit/Outfit орташа мәні	≈1.0
Max Outfit	9.90 (I0015)

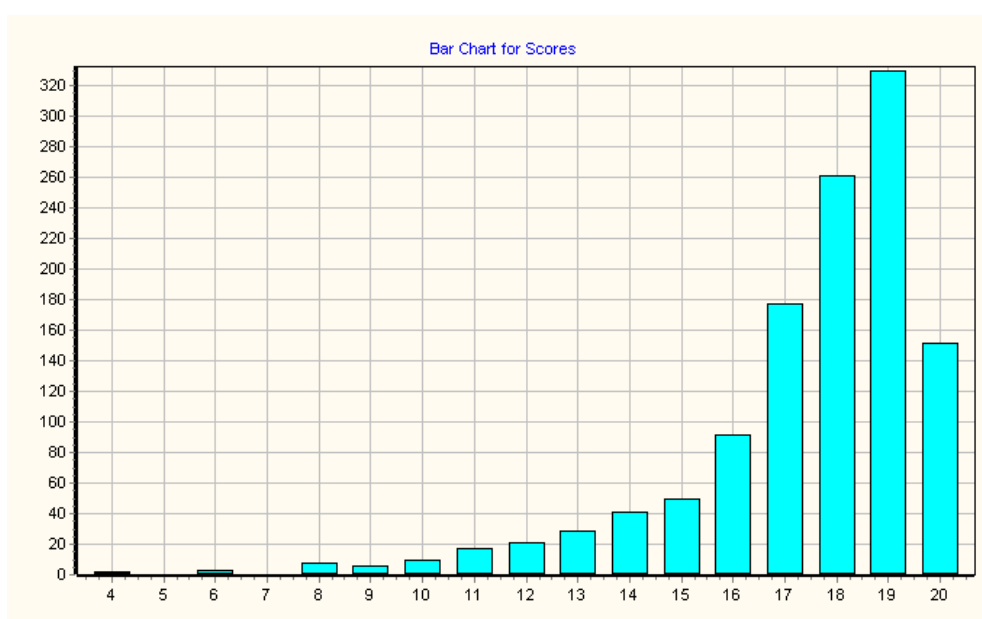
Тапсырмалардың көрсеткіштеріне (Item) тоқталып өтсек, 0.00 logits – бұл Раш моделінде қалыпты нәтиже, өйткені тапсырмалар шкаланың ортасында өлшенген. Item Separation 10.03 өте жоғары мән. Бұл тапсырмалардың қиындық деңгейі өте жақсы ажыратылғанын, яғни тестте әртүрлі деңгейдегі тапсырмалар бар екенін көрсетеді. Item Reliability 0.99 бұл көрсеткіштен тест тапсырмаларының сенімділігінің өте жоғары екенін көруге болады. Сенімділік психодиагностикада валидтіліктің

алғышарты екені және ол зерттеу нәтижелерінің дәлдігі мен практикалық құндылығын айқындайтын басты өлшемдердің бірі [11, 102.6] Бұл көрсеткіш тест тапсырмаларының қиындық параметрлерін тұрақты және дәл өлшейтінін білдіреді [12, 301.6]. Егер бұл тест басқа топтарға қолданылса да, тапсырмалар қиындық ретімен ұқсас орналасады. *Қиындық диапазоны*  $-3.12 \rightarrow +3.65 \text{ logits}$  – бұл өте кең диапазон. Яғни тестте өте жеңілден өте қиынға дейінгі тапсырмалар бар. Бұл сапалы тест үшін оң сипат, себебі ол әртүрлі деңгейдегі қатысушыларға бейімделе алады. *Infit/Outfit орташа мәні*  $\approx 1.0$  – бұл тапсырмалардың көпшілігі Раш моделінің болжамына жақсы сәйкес келетінін білдіреді. Infit және Outfit мәндерінің 0.7–1.3 аралығында болуы – жақсы сәйкестік белгісі. *Max Outfit = 9.90 (Item I0015)* – бұл бір тапсырмада (I0015) айқын сәйкессіздік (misfit) бар екенін көрсетеді. Мұндай тапсырма қатысушылардың күтілген үлгісінен қатты ауытқиды – яғни кейбір төмен деңгейлі қатысушылар оны дұрыс орындаған, ал жоғары деңгейлілер қате орындаған болуы мүмкін. Бұл жағдайда тапсырманы қайта қарап, тілдік немесе мазмұндық тұрғыдан сараптау қажет.

2-кесте. – Қатысушылардың көрсеткіштері (Person)

Көрсеткіш	Мәні
Орташа көрсеткіш (Mean measure)	<b>3.16 logits</b>
Person Separation	<b>0.83</b>
Person Reliability	<b>0.41</b>
Cronbach Alpha (KR-20)	<b>0.73</b>
Raw score-to-measure correlation	<b>0.94</b>

Ал, тестке қатысушылар бойынша көрсеткіштерді талдау барысында, *орташа көрсеткіш (Mean Measure)*  $3.16 \text{ logits}$  – бұл қатысушылардың орташа қабілет деңгейінің тапсырмалардың орташа қиындық деңгейінен жоғары екенін білдіреді. Яғни, тест тапсырмалары респонденттердің жалпы деңгейіне қарағанда салыстырмалы түрде жеңіл болған. Мұндай нәтиже тесттің жоғары қабілетті топқа сәйкес еместігін, тесттің қиындық параметрін біршама арттыру қажет екенін көрсетеді. *Person Separation = 0.83* – қатысушыларды шамамен екі деңгейге (төмен және жоғары) бөлуге болатынын көрсетеді. Separation мәні 2.0-ден жоғары болғанда деңгейлер нақтырақ ажырайды. Сондықтан бұл тест қатысушылардың қабілет деңгейін нақты саралауда шектеулі ажыратушылыққа ие екенін білдіреді. *Person Reliability = 0.41* – қатысушылардың сенімділік коэффициенті төмен. Бұл топтың гомогенді (біртектес) екенін және тесттің қатысушыларды сенімді түрде деңгей бойынша ажырата алмайтынын көрсетеді. Себептері ретінде тесттің тым жеңіл болуы немесе үлгінің (sample) біртектілігі (мысалы, барлығы бір деңгейдегі тыңдаушылар) болуы мүмкін. *Cronbach Alpha (KR-20) = 0.73* – бұл тесттің ішкі біртектілігінің орташа деңгейде екенін көрсетеді. Тест тапсырмаларының басым бөлігі бір ғана дағдыны (мысалы, тыңдалым қабілетін) өлшейтінін білдіреді, бірақ кейбір тапсырмаларда қосымша факторлар (лексика, контексті түсіну және т.б.) әсер етуі ықтимал. *Raw score-to-measure correlation = 0.94* – эмпирикалық деректер мен модель арасындағы сәйкестік өте жақсы. Бұл модельдің деректерді дәл сипаттайтынын және Раш талдауы нәтижелерінің интерпретациясы сенімді екенін дәлелдейді.



1-сурет. – Тестіленушілердің Тыңдалым блогы бойынша жинаған ұпайларының таралу диаграммасы (Bar Chart for Scores)

Диаграммада тестіленушілердің басым бөлігі жоғары нәтижелер көрсеткені анық байқалады. 17-20 ұпай аралығында шоғырланған қатысушылар саны айтарлықтай көп, ал төмен ұпай жинағандар үлесі өте аз. Орташа балл – 17,41 (87%), медиана – 18 (90%), стандартты ауытқу – 2,46.

Бұл көрсеткіштер тыңдалым тапсырмаларының жалпы алғанда жеңіл деңгейде болғанын көрсетеді.

Таратудың асимметрия коэффициенті (-1,83) теріс мәнге ие, бұл нәтижелердің солға қарай қисайғанын, яғни жоғары балл көрсеткендердің басымдығын білдіреді. Куртозис (4,16) мәні таралудың өткір әрі пішіні бойынша шоғырланған екенін көрсетеді.

Тесттің ішкі сенімділік көрсеткіші (KR-20) 0,732 деңгейінде, яғни қанағаттанарлық сенімді деп бағаланады. Алайда орташа тапсырма қиындығының (0,87) және алты тапсырманың тым жеңіл (difficulty  $\geq 0,95$ ) болуы фактісі тесттің ажыратушылық деңгейінің төмендеуіне әсер еткен.

Мұндай нәтиже тесттің жоғары деңгейлі тестіленушілер үшін оңай болғанын және ажыратушылық деңгейінің жеткіліксіздігін аңғартады. Сондықтан болашақта орта және жоғары қиындықтағы тыңдалым тапсырмаларының үлесін арттыру ұсынылады.

3-кесте. – Тапсырмалардың қиындық деңгейлері бойынша бөлінуі (A1-C1)

№	Деңгей	Қиындық диапазоны (logits)	Тапсырмалар саны
1	A1 (қарапайым)	< -2.0	2
2	A2 (базалық)	-2.0 → -0.5	5
3	B1 (орта)	-0.5 → +1.0	6
4	B2 (ортадан жоғары)	+1.0 → +2.5	5
5	C1 (жоғары)	> +2.5	2

Тест тапсырмаларының қиындық деңгейі Раш моделі бойынша логит шкаласында бағаланып, CEFR деңгейлеріне (A1-C1) сәйкес келтірілді. Талдау нәтижелері тесттің қиындық шкаласының логикалық және бірізді өсуін көрсетеді. A1 және A2 деңгейлеріндегі тапсырмалар бастапқы және төмен деңгейдегі қабылдау дағдыларын бағалауға бағытталған, ал B1 және B2 деңгейлеріндегі тапсырмалар мәтінді түсіну, негізгі мазмұн мен логикалық байланыстарды анықтау қабілетін тексереді. C1 деңгейіндегі тапсырмалар күрделі мәтіндерді талдау және түсіндіру дағдыларын талап етеді [13].

Тесттің деңгейлік бөлінуі пирамидалық сипатқа ие, яғни орталық деңгейлер (B1-B2) көп тапсырмалармен ұсынылған. Бұл тесттің орта деңгейдегі қатысушылардың қабілетін сенімді бағалауға арналғанын көрсетеді. Алайда A1 және C1 деңгейлеріндегі тапсырмалар саны аз болғандықтан, төменгі және жоғарғы деңгейдегі қатысушылар үшін тесттің сезімталдығы шектеулі.

Психометриялық тұрғыдан, бұл деңгейлік құрылым тесттің әртүрлі қабілет деңгейлерін қамтитынын, бірақ орта деңгейге бейімделгенін дәлелдейді. Тесттің толық диагностикалық әлеуетін арттыру үшін A1 және C1 деңгейлеріндегі тапсырмалардың санын көбейту ұсынылады. Сонымен қатар, әр деңгейдегі тапсырмалардың мазмұндық және тілдік сипаттамаларын CEFR дескрипторларымен нақты сәйкестендіру тесттің сенімділігін арттырады.

Талдау нәтижелері көрсеткендей, тыңдалым бөлімі бойынша тест тапсырмалары жалпы алғанда модель талаптарына сай келеді және өлшем біртектілігі қамтамасыз етілген. Item Separation мәнінің 10.03 болуы тест тапсырмаларының қиындық деңгейі бойынша жақсы дифференциацияланғанын білдіреді. Сонымен бірге Person Reliability = 0.41 көрсеткіші тестіленушілер тобының қабілет деңгейі бойынша айырмашылықтың айқын еместігін байқатады, яғни тест бұл іріктеме үшін біршама жеңіл болған.

Кейбір тапсырмаларда (I0015, I0018) Outfit мәндерінің жоғары болуы олардың модельге толық сәйкес келмейтінін және мазмұндық тұрғыда қайта қарауды қажет ететінін көрсетеді. Бұл тапсырмаларда аудиомәтіннің ұзындығы немесе жауап нұсқаларының бірімәнді болмауы ықтимал.

Жалпы алғанда, тыңдалым бөлімі бойынша тест тапсырмаларының сенімділігі жоғары, валидтілігі жеткілікті, ал қиындық деңгейлері әртүрлі деңгейдегі қатысушыларды қамтуға мүмкіндік береді. Дегенмен, қатысушылардың дифференциациясы төмен болғандықтан, келесі кезеңдерде тест құрылымын кеңейту, аудиомәтінді бірнеше емес тек екі рет тыңдату әсерін қосымша эмпирикалық зерттеу, аудиоматериалдарды тілдік деңгейлерге нақты сәйкестендіру және төмен және жоғары деңгейдегі қатысушыларға арналған қосымша тапсырмалар енгізу ұсынылады.

Біздің келесі зерттеуімізде «Оқылым» блогына психометриялық талдау жасалынды. Зерттеу деректері 1197 тестіленушінің 40 тест тапсырмасына берген жауаптарына негізделді.

4-кесте. – Тест тапсырмаларының көрсеткіштері (Item)

Көрсеткіш	Мәні
Орташа көрсеткіш (Item Mean Measure)	<b>0.00 logits</b>
Item Separation	<b>10.33</b>
Item Reliability	<b>0.99</b>
Қиындық диапазоны	<b>-5.12 → +2.63 logits</b>
Infit/Outfit орташа мәні	<b>≈1.0</b>
Max Outfit	<b>9.62 (бір тапсырма)</b>

Талдау нәтижелері оқылым блогының тест тапсырмаларының психометриялық сапасының жоғары деңгейде екенін көрсетті. Оқылымға арналған тапсырмалардың орташа көрсеткіші *0.00 logits* деңгейінде, яғни логит шкаласының орталық нүктесіне сәйкес келеді. Бұл көрсеткіш тесттің жалпы қиындық деңгейі теңгерімді екенін және тапсырмалардың орташа алғанда тестіленушілердің қабілет деңгейіне сәйкес келетінін білдіреді. Демек, оқылым тапсырмалары *neither* тым жеңіл, *neither* тым қиын емес, орта деңгейлі қатысушыларға бағытталған.

*Item Separation* коэффициентінің мәні *10.33* болып анықталды. Бұл өте жоғары көрсеткіш оқылым блогындағы тапсырмалардың қиындық деңгейлері бір-бірінен айқын ажыратылғанын және тесттің түрлі деңгейдегі оқырмандық қабілеттерді тиімді бағалай алатынын дәлелдейді.

*Item Reliability* мәні *0.99* болып, тест тапсырмаларының қиындық параметрлерін өлшеу сенімділігінің өте жоғары екенін дәлелдеді. Бұл көрсеткіш тапсырмалардың тұрақтылығын және олардың түрлі үлгілер мен қатысушылар топтарында қолданылған жағдайда да қиындық ретінің сақталатынын білдіреді.

Оқылым тапсырмаларының *қиындық диапазоны*  $-5.12 \rightarrow +2.63 logits$  аралығында орналасқан. Бұл көрсеткіш тесттің өте жеңілден (A1 деңгейіне жуық) бастап, өте күрделіге (C1 деңгейіне жуық) дейінгі мәтіндерді қамтитынын білдіреді. Яғни, тест мазмұны тестіленушінің түсіну дағдыларының кең спектрін – нақты ақпаратты табудан бастап, астарлы мағынаны түсінуге дейінгі қабілеттерді қамтиды.

*Infit және Outfit* статистикалық көрсеткіштерінің орташа мәндері шамамен *1.0* деңгейінде болды, бұл тапсырмалардың Раш моделінің болжамдарына жоғары дәрежеде сәйкес келетінін айғақтайды. Яғни, тестіленушілердің жауап беру үлгілері күтілетін модельге сай келеді және тест құрылымы ішкі тұрғыдан үйлесімді. Мұндай үйлесімділік тесттің мазмұндық және когнитивтік талаптарының дұрыс теңестірілгенін дәлелдейді.

Сонымен бірге, бір тапсырмада *Max Outfit = 9.62* шамасында сәйкессіздік (*misfit*) байқалды. Бұл тапсырманың Раш моделіне сәйкес келмейтінін, яғни күтілетін жауап үлгісінен айтарлықтай ауытқығанын көрсетеді. Мұндай сәйкессіздік мазмұндық тұрғыдан күрделі тұжырым, мәтіннің семантикалық ерекшеліктері немесе дистракторлардың сапасына байланысты болуы мүмкін. Сондықтан бұл тапсырмаға қосымша сараптама жүргізіп, қажет болған жағдайда оның тұжырымдамасын немесе жауап нұсқаларын қайта қарау ұсынылады.

Жалпы алғанда, талдау нәтижелері оқылым блогының тест тапсырмаларының психометриялық сапасының өте жоғары екенін көрсетті. Тапсырмалар логит шкаласы бойынша теңгерімді орналасқан, модельмен жақсы үйлеседі және сенімділігі жоғары (*Item Reliability = 0.99*). Бұл деректер КАЗТЕСТ жүйесінің оқылым блогы халықаралық психометриялық талаптарға сәйкес әзірленгенін және тестіленушілердің оқу дағдыларын тиімді бағалай алатынын дәлелдейді.

5-кесте. – Қатысушылардың көрсеткіштері (Person)

Көрсеткіш	Мәні
Орташа көрсеткіш (Mean measure)	<b>1.92 logits</b>
Person Separation	<b>1.80</b>
Person Reliability	<b>0.76</b>
Cronbach Alpha (KR-20)	<b>0.84</b>
Raw score-to-measure correlation	<b>0.98</b>

Бұл кестедегі талдау нәтижелері тестіленушілердің жалпы оқу дағдылары мен тест тапсырмаларына жауап беру қабілеттерінің психометриялық сипаттамасын анықтауға мүмкіндік берді.

Қатысушылардың орташа көрсеткіші (*Mean Measure*) *1.92 logits* деңгейінде тіркелді. Бұл мән нөлдік логит нүктесінен жоғары болғандықтан, тестіленушілердің орташа қабілет деңгейі тест тапсырмаларының орташа қиындық деңгейінен біршама жоғары екенін көрсетеді. Яғни, тест қатысушыларының басым бөлігі берілген оқылым тапсырмаларын сәтті орындап, олардың деңгейі тесттің жалпы қиындығынан жоғары болғаны байқалады.

*Person Separation* коэффициентінің мәні *1.80* құрады. Бұл көрсеткіш тестіленушілердің қабілет деңгейіне қарай шамамен үш деңгейге (жоғары, орта, төмен) саралай алатынын білдіреді. *Separation* коэффициентінің шамасы 2.0-ден төмен болғанымен, *1.80* мәні тесттің саралау қуатының орташа-жоғары деңгейде екенін және ол тестіленушілер арасындағы айырмашылықтарды белгілі бір дәрежеде сенімді түрде ажырата алатынын көрсетеді.

*Person Reliability* мәні *0.76* деңгейінде анықталды. Бұл көрсеткіш тестіленушілерді қабілет деңгейіне қарай тұрақты түрде жіктеу қабілетінің жеткілікті екенін көрсетеді. Раш моделіндегі *Person Reliability* шамамен *0.70–0.80* аралығында болған жағдайда тестіленушілердің қабілеттерін ажырату сенімділігі қанағаттанарлық деңгейде деп саналады. Демек, оқылым блогы бойынша алынған нәтижелер тестіленушілердің оқу, түсіну қабілеттерін жалпы деңгейде сенімді өлшей алатынын дәлелдейді.

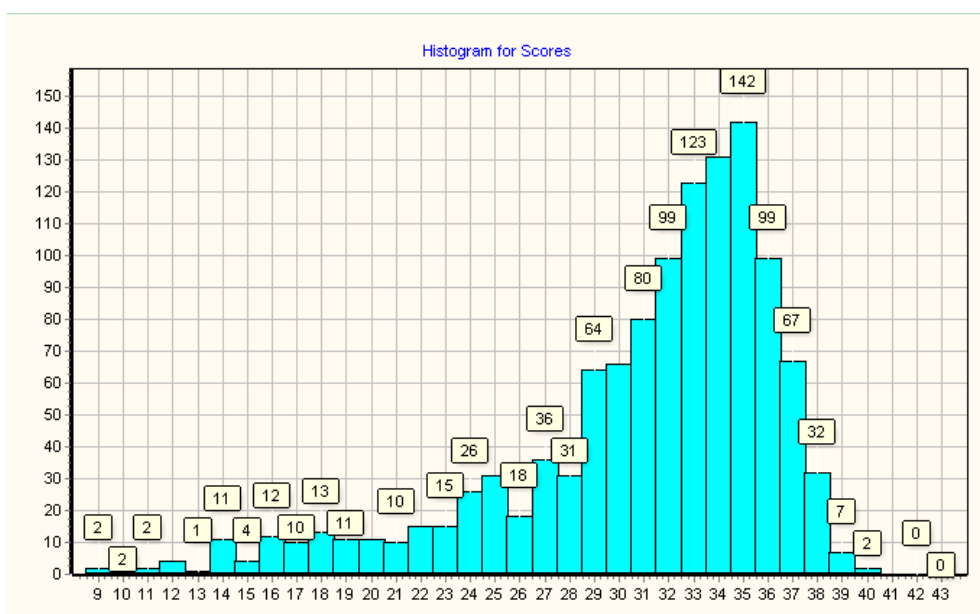
Сонымен қатар, *Cronbach's Alpha (KR-20)* мәні *0.84* болып анықталды. Бұл көрсеткіш ішкі біртектіліктің (*internal consistency*) жоғары екенін, яғни тест тапсырмаларының мазмұндық және

статистикалық тұрғыдан өзара үйлесімді екенін көрсетеді. KR-20 өрнегі бойынша 0.80 және одан жоғары мән тесттің сенімділігінің жоғары деңгейін білдіреді. Осыған сәйкес, оқылым блогы тапсырмаларының мазмұны тестіленушілердің оқу дағдыларын жүйелі және бірізді өлшей алатынын айтуға болады.

*Raw score-to-measure correlation* шамасы 0.98 болды, бұл өте жоғары корреляциялық байланыс деңгейін көрсетеді. Мұндай жоғары мән тест бойынша алынған бастапқы баллдар мен Раш моделінде есептелген логит көрсеткіштері арасындағы өзара байланыстың күшті екенін, яғни өлшеу нәтижелерінің сәйкестігі мен тұрақтылығы жоғары екенін дәлелдейді. Басқаша айтқанда, тесттің бағалау жүйесі мен модельдік есептеулер бір-біріне сәйкес келеді, сондықтан өлшеу нәтижелерін түсіндіру сенімді.

Жалпы алғанда, қатысушылардың көрсеткіштеріне жүргізілген психометриялық талдау КАЗТЕСТ жүйесінің оқылым блогының сенімділігі мен өлшеу дәлдігінің жеткілікті жоғары деңгейде екенін айқындайды.

Келесі кестеден «Оқылым» болгы бойынша тапсырмалардың деңгейге сәйкес бөлінуі бойынша талдау нәтижесін көруге болады.



2-сурет. – Тестіленушілердің Оқылым блогы бойынша жинаған ұпайларының таралу диаграммасы (Bar Chart for Scores)

Бұл диаграммада таралу пішіні қоңырау тәрізді (bell-shaped) сипатқа ие, яғни деректердің басым бөлігі орта деңгей маңында шоғырланған. Ең жоғары жиілік 33-35 ұпай аралығында байқалады (шамамен 142 тестіленуші).

Төменгі ұпайлар (20-ға дейінгі нәтижелер) сирек кездеседі, ал жоғары ұпайлар (38 және одан жоғары) да аз, яғни үлестірімнің ұштарында қатысушылар саны төмен. Бұл жағдай тест тапсырмаларының қиындығы мен мазмұны теңгерімді екенін және тестіленушілердің деңгейлік айырмашылығын жақсы ажырата алғанын көрсетеді.

Төменгі және жоғары ұпай аймақтарында қатысушылар саны аз, бұл тесттің диагностикалық функциясының тиімділігін және бағалау дәлдігінің жеткіліктілігін дәлелдейді.

6-кесте. – Тапсырмалардың қиындық деңгейлері бойынша бөлінуі (A1-C1)

Деңгей	Диапазон (logit)	Тапсырмалар саны
A1 (қарапайым)	< -3.0	3
A2 (базалық)	-3.0 → -1.0	9
B1 (орта)	-1.0 → +0.5	13
B2 (ортадан жоғары)	+0.5 → +2.0	10
C1 (жоғары)	> +2.0	5

Раш моделінің нәтижелері негізінде оқылым блогындағы тапсырмалар қиындық деңгейлері бойынша бес деңгейге бөлінді: A1, A2, B1, B2 және C1. Мұндай деңгейлік саралау тест мазмұнының Еуропалық тілдік құзыреттердің ортақ шеңбері (CEFR) талаптарына сәйкес құрылымданғанын және тесттің диагностикалық әлеуетінің кең екенін көрсетеді.

Жалпы алғанда, тапсырмалардың қиындық деңгейлері  $-3.0$ -мен  $+2.0$  *logits* аралығында теңгерімді бөлінген. Мұндай құрылым тесттің диагностикалық тиімділігін арттырып, қатысушылардың тілдік қабілеттерін кең диапазонда өлшеуге мүмкіндік береді. Тапсырмалар жиынтығы А1-ден С1-ге дейінгі барлық деңгейлерді қамтитындықтан, тест оқу дағдыларын кешенді бағалауға бағытталған деп қорытынды жасауға болады.

Сонымен қатар, деңгейлердің үлестік арақатынасы (А1-А2 – 12 тапсырма, В1-В2 – 23 тапсырма, С1 – 5 тапсырма) тест мазмұнының орта деңгейге (В1-В2) бағытталғанын, яғни тестіленушілердің басым көпшілігіне арналған теңгерімді құрылымды білдіреді. Бұл үлестік модель КАЗТЕСТ жүйесінің функционалды сауаттылыққа және нақты оқу жағдайларында тілдік құралдарды тиімді қолдану қабілетін бағалауға бағытталғанын айғақтайды.

Талдау нәтижесінен оқылым блогындағы тапсырмалардың қиындық деңгейлері бойынша бөлінуі тесттің валидтілігі мен мазмұндық үйлесімділігін дәлелдейді. Тапсырмалар жиынтығының кең логит диапазоны (А1–С1) тесттің әртүрлі деңгейдегі қатысушылардың оқу дағдыларын нақты және сенімді өлшеуге мүмкіндік беретінін көрсетеді.

Жалпы алғанда, оқылым бөлімі бойынша тест тапсырмаларының сенімділігі мен валидтілігі жоғары. Тапсырмалардың қиындық деңгейі тілдік құзыреттіліктің А1-С1 деңгейлерін толық қамтиды, бұл КАЗТЕСТ жүйесінің коммуникативтік бағыттылығына сәйкес келеді. Психометриялық тұрғыда тесттің өлшеу дәлдігі, ішкі біртектілігі және құрылымдық үйлесімділігі қанағаттанарлық деңгейде.

Белгілі бір өлшемдерден кейінгі келесі міндет – белгілі бір бағалау жүргізу. Бұл бағалар оқушыны немесе оқытуды бағалау түрінде болады. Оқушыны бағалау барысында бір жағынан оқушының біліміндегі кемшілігі мен қиыншылықтарын, ал екінші жағынан соңғы қол жеткізген білім деңгейін анықтауға арналады [14. 67.б].

Зерттеу нәтижесінде КАЗТЕСТ жүйесіне алғаш рет кешенді психометриялық және статистикалық талдау жүргізілді. Бұл зерттеу қазақ тілін меңгеру деңгейін бағалаудың ұлттық жүйесін халықаралық стандарттарға сәйкестендіруге мүмкіндік береді. Практикалық тұрғыдан алынған нәтижелер тест әзірлеушілер мен сарапшылар үшін тапсырмалардың сапасын бағалау мен жетілдірудің ғылыми негізін қалыптастырады. Алынған психометриялық көрсеткіштер тапсырмалардың сенімділігі мен валидтілігінің жоғары екенін көрсетті, бұл халықаралық зерттеулер нәтижелерімен сәйкес келеді. Сонымен қатар, тапсырмалардың CEFR деңгейлеріне сәйкес бөлінуі тест мазмұнының деңгейлік үйлесімділігін дәлелдейді.

**Қорытынды.** Жүргізілген зерттеу нәтижелері КАЗТЕСТ жүйесінің тыңдалым және оқылым блоктарындағы тест тапсырмаларының психометриялық сапасы мен өлшеу дәлдігін жан-жақты сипаттады. Раш моделіне негізделген талдау нәтижелері екі блоктың да жоғары сенімділік пен ішкі үйлесімділікке ие екенін, тест мазмұнының тестіленушілердің қабілет деңгейіне сәйкес қалыптасқанын көрсетті. Бұл КАЗТЕСТ жүйесінің тілдік құзыреттілікті бағалау құралы ретінде тиімділігін және халықаралық өлшеу стандарттарына сәйкес келетінін дәлелдейді.

Тыңдалым блогы бойынша талдау нәтижелері тест тапсырмаларының сапалы құрылымдалғанын және өлшеу дәлдігінің жоғары екенін айқындады. Орташа көрсеткіштің (Mean Measure = 3.16 *logits*) оң бағытта орналасуы қатысушылардың орташа қабілет деңгейі тесттің орташа қиындық деңгейінен жоғары екенін білдіреді. Бұл тыңдалым тапсырмаларының жалпы алғанда салыстырмалы түрде жеңіл болғанын көрсетеді. Сонымен қатар, Item Reliability = 0.99 және Item Separation = 10.03 көрсеткіштері тест мазмұнының тұрақтылығы мен қиындық деңгейлерінің айқын ажыратылғанын дәлелдейді. Алайда бір тапсырмада анықталған misfit (Outfit = 9.90) мәні тест мазмұнын жетілдіру қажеттілігін көрсетеді.

Оқылым блогы бойынша талдау нәтижелері де оң сипатқа ие болды. Қатысушылардың орташа қабілет көрсеткіші 1.92 *logits* деңгейінде болып, тест тапсырмаларының орташа қиындығынан жоғары екенін көрсетті. Person Reliability = 0.76 және Cronbach Alpha (KR-20) = 0.84 мәндері тесттің ішкі біртектілігі мен сенімділігінің жоғары деңгейде екенін дәлелдейді. Сонымен қатар, Raw score-to-measure correlation = 0.98 шамасында тіркеліп, эмпирикалық деректер мен модель арасындағы сәйкестіктің өте жоғары екенін көрсетті. Бұл КАЗТЕСТ жүйесінің оқылым блогы тілдік қабілеттерді дәл және тұрақты өлшей алатынын айқындайды.

Тапсырмалардың қиындық деңгейлері бойынша бөлінуі екі блогта да А1–С1 аралығын қамтитынын көрсетті. Тыңдалым блогындағы қиындық диапазоны  $-3.12 \rightarrow +3.65$  *logits*, ал оқылым блогында  $-5.12 \rightarrow +2.63$  *logits* аралығында анықталды. Бұл көрсеткіштер тесттің диагностикалық әлеуеті жоғары екенін, яғни төменнен жоғары деңгейге дейінгі барлық қатысушылардың тілдік қабілеттерін нақты саралай алатынын көрсетеді.

TOEFL iBT Test, IELTS, DELF/DALF тыңдалым бөлімі академиялық сөйлеуді түсінуді бағалауға бағытталған және аудиомәтіндер бір рет тыңдатылады. Бұл формат тест тапсырушылардың ақпаратты қайта тыңдамай, нақты уақыт режимінде қабылдау қабілетін бағалауды көздейді және шынайы академиялық орта талаптарына сәйкес келеді [15,16,17,18].

Бұл ретте зерттеу нәтижелері негізінде КАЗТЕСТ жүйесінің әрі қарай жетілдіру бағыттарына қатысты төмендегі ұсыныстар беріледі:

1. Тыңдалым блогы бойынша:

- Халықаралық тілдік білімді бағалау тестілерінің тәжірибесіне сәйкес (мысалы, IELTS, TOEFL, TestDaF және т.б.) аудиомәтінді 1 немесе 2 рет тыңдату тәжірибесін эксперименттік түрде зерттеп, оның тест нәтижелеріне, тестіленушілердің түсіну деңгейіне және сенімділік көрсеткіштеріне әсерін бағалау;
- Тыңдалым тапсырмаларының күрделілігін арттыру мақсатында түрлі жанрлар мен сөйлесім дағдыларын (репортаж, лекция, пікірталас) енгізу;
- 2. Оқылым блогы бойынша:
  - Жоғары деңгейдегі (B2-C1) мәтіндердің үлесін көбейтіп, күрделі грамматикалық және мағыналық құрылымдарды қамтитын материалдарды енгізу;
  - Мәтіндердің тақырыптық ауқымын кеңейту (ғылыми-көпшілік, кәсіби, мәдени тақырыптар);
- 3. Жалпы тест құрылымы бойынша:
  - Психометриялық талдауды тұрақты негізде жүргізу және әр сынақтан кейін тапсырмаларды саралау;
  - Тест мазмұнының теңгерімділігі мен валидтілігін сақтау мақсатында деректер базасын жүйелі түрде жаңарту;
  - Тестілеу нәтижелерін халықаралық жүйелермен (CEFR, ALTE) сәйкестендіру жұмыстарын жалғастыру.

Қорытындылай келе, тыңдалым және оқылым блоктары бойынша алынған психометриялық деректер ҚАЗТЕСТ жүйесінің тілдік құзыреттілікті бағалауда сенімді, ғылыми негізделген және халықаралық талаптарға сәйкес құрал екенін дәлелдейді. Талдау нәтижелері тест мазмұнын жетілдірудің, тыңдалым блогындағы аудиомәтінді тыңдату жиілігін оңтайландырудың және тапсырмалардың когнитивтік күрделілігін арттырудың ғылыми-тәжірибелік негізін айқындайды. Алынған деректер тест тапсырмаларын жетілдіру және оларды халықаралық жүйелермен үйлестіру үшін ғылыми негіз қалыптастырады. Бұл өз кезегінде тест тапсырмаларын әрі қарай жетілдіру арқылы оларды халықаралық тілдік бағалау практикасына үйлестіруге әдіснамалық және мазмұндық негіз қалыптастыратынын дәлелдейді.

#### ӘДЕБИЕТТЕР:

- 1 **Қазақстан Республикасының «Тіл туралы» Заңы.** – Астана, 1997. [Электронды ресурс] URL: <https://adilet.zan.kz/kaz/docs/Z970000151> (жүгінген күні: 8.09.2025).
- 2 **Қазақстан Республикасында тіл саясатын дамытудың 2023-2029 жылдарға арналған тұжырымдамасы: Қазақстан Республикасы Үкіметінің 2023 жылғы 29 желтоқсандағы №1219 қаулысымен бекітілген.** [Электронды ресурс] URL: <https://www.gov.kz/memleket/entities/language/press/news/details/717131> (жүгінген күні: 8.09.2025).
- 3 **Қазақстан Республикасы Ғылым және жоғары білім министрінің 2024 жылғы 23 қыркүйектегі №458 бұйрығы «ҚАЗТЕСТ қазақ тілін меңгеру деңгейін бағалау жүйесі бойынша қағидаларды бекіту туралы».** [Электронды ресурс]. URL: <https://adilet.zan.kz/kaz/docs/V2400035109> (жүгінген күні: 9.09.2025).
- 4 **Қазақ тілін меңгерудің коммуникативтік тілдік құзыреттіліктері: Қарапайым, Базалық, Орта, Ортадан жоғары және Жоғары деңгейлер. Жалпы меңгеру** [Мәтін]: ҚР СТ 1926-2023, ҚР СТ 1928-2023, ҚР СТ 1929-2023, ҚР СТ 1925-2023, ҚР СТ 1927-2023 – Қазақстан Республикасы Сауда және интеграция министрлігі Техникалық реттеу және метрология комитетінің «Қазақстан стандарттау және метрология институты». – Астана: 2023. – 20 б.
- 5 **Ерсултанова Г.С. Тілді меңгеру деңгейін бағалаудағы ҚАЗТЕСТ жүйесінің рөлі: жаңа форматтағы тиімділік** [Мәтін] / Г.С. Ерсултанова // Қазақстан Республикасында мемлекеттік тілді оқытудың жүйесі: деңгейлері мен сатылары: конф. материалдары. – Алматы, 2024. – Б. 362–367.
- 6 **Дубина И.Н. Математико-статистические методы в эмпирических социально-экономических исследованиях** [Текст] учеб. пособие. / И.Н.Дубина – М.: Финансы и статистика; ИНФРА-М, 2010. – 416 с.
- 7 **Карданова Е.Ю., Иванова А.Е. Психометрические исследования: современные методы и новые возможности для образования** [Текст] / Е.Ю. Карданова, А.Е. Иванова // Вопросы образования. – 2023. -№3. – С. 8–19.
- 8 **Бейсенова Ж., Шинетова Л., Жабаева Г. Психометрический анализ тестовых заданий по оценке знаний педагогов в Республике Казахстан (по результатам апробации)** [Мәтін] / Ж.Бейсенова, Л. Шинетова, Г. Жабаева // Педагогикалық өлшемдер. – 2025. – №3. – Б. 41–52. DOI: <https://doi.org/10.63597/UTO3105-4161.2025.3.3.002>.
- 9 **Садық Ұ.Р., Макыбаева А.К., Нурымбетов Н.Б. ҚАЗТЕСТ жүйесіндегі диагностикалық тестілеудің өткізілу барысы мен тәжірибесі** [Мәтін] / Ұ.Р. Садық, А.К. Макыбаева, Н.Б. Нурымбетов // Педагогикалық өлшемдер. – 2025. – №3. – Б. 30-39. DOI: <https://doi.org/10.63597/UTO3105-4161.2025.3.3.003>.
- 10 **Cronbach L.J. Coefficient alpha and the internal structure of tests** [Text] / L.J. Cronbach // Psychometrika. – 1951. – Vol. 16. – P. 297–334.

- 11 Корнилов С.А. Проблема надежности психологических шкал и ее значение для психодиагностики в научных и практических целях [Текст] / С.А. Корнилов // Экспериментальная психология в России: традиции и перспективы. – 2010. – С. 101–105.
- 12 Crocker L., Algina J. **Introduction to Classical and Modern Test Theory** [Text] / L. Crocker, J. Algina. – Belmont, CA: Cengage Learning, 2008. – P. 482–498.
- 13 **Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)**. – Strasbourg: Council of Europe, 2001. [Electronic resource] URL: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4> (accessed 1.10. 2025).
- 14 Сагиндиқов И.У., Динаева Б.Б., Убайдуллаева Г.Ж. **Өлшеу және бағалау** [Мәтін]: монография / Сагиндиқов И.У., Динаева Б.Б., Убайдуллаева Г.Ж. – Өлшеу: «Imprime» ЖШС баспасы, - Нұр-Сұлтан, 2021. – 284 б.
- 15 **ETS. TOEFL iBT Test** [Electronic resource] URL: <https://www.ets.org/toefl> (accessed 3.11.2025).
- 16 **IELTS. International English Language Testing System** [Electronic resource] URL: <https://www.ielts.org> (accessed 1.10. 2025).
- 17 **TORFL. Test of Russian as a Foreign Language** [Electronic resource] URL: <https://russian-test.com> (accessed 1.10. 2025).
- 18 **DELFDALF. Diplôme d'Études en Langue Française** [Electronic resource] URL: <https://www.france-education-international.fr> (accessed 1.10 2025).

## REFERENCES:

- 1 **Kazakstan Respublikasynyn «Til turaly» Zany** [Law of the Republic of Kazakhstan «On Languages in the Republic of Kazakhstan»]. Astana, 1997. Available at: <https://adilet.zan.kz/kaz/docs/Z970000151> (accessed 08 September 2025). (In Kazakh)
- 2 **Kazakstan Respublikasynda til sayasatyn damytudyn 2023-2029 zhyldarga arналган tuzhyrymdamasy: Kazakstan Respublikasy Ukimetinin 2023 zhylygy 29 zheltokсандagy №1219 kaulysymen bekитilgen** [Concept for the Development of Language Policy in the Republic of Kazakhstan for 2023–2029: Approved by the Decree of the Government of the Republic of Kazakhstan №1219 dated December 29, 2023]. Available at: <https://www.gov.kz/memleket/entities/language/press/news/details/717131> (accessed 08 September 2025). (In Kazakh)
- 3 **Kazakstan Respublikasy Gylym zhane zhogary bilim ministrinin 2024 zhylygy 23 kyrkuyektegi №458 buyrygy «KAZTEST kazak tilin mengeru dengeyin bagalau zhuyesi bojnsha kagidalardy bekitu turaly»** [Order of the Minister of Science and Higher Education of the Republic of Kazakhstan №458 dated September 23, 2024 «On Approval of the Rules of the KAZTEST System for Assessing Proficiency in the Kazakh Language»]. Available at: <https://adilet.zan.kz/kaz/docs/V2400035109> (accessed 09 September 2025). (In Kazakh)
- 4 **Kazak tilin mengerudin kommunikativtik tildik kuzyrettilikteri: Karapajym, Bazalyk, Orta, Ortadan zhogary zhane Zhogary dengeiler. Zhalpy mengeru: KR ST 1926-2023, KR ST 1928-2023, KR ST 1929-2023, KR ST 1925-2023, KR ST 1927-2023** [Communicative language competences of the Kazakh language proficiency: elementary, basic, intermediate, upper-intermediate and advanced levels. General proficiency]. Developed and introduced by the Republican State Enterprise on the Right of Economic Management “Kazakhstan Institute of Standardization and Metrology”, Astana, 2023, 20 p. (In Kazakh)
- 5 **Ersultanova G.S. Tildi mengeru dengeyin bagalaudagy KAZTEST zhujesinin roli: zhana format tagy tiimdilik** [The role of the KAZTEST system in assessing language proficiency: effectiveness of a new format]. *Kazakstan Respublikasynda memlekettik tildi okytudyn zhujesi: dengejleri men satylary: konferenciya materialdary*, Almaty, 2024, 362 p. (In Kazakh)
- 6 **Dubina I.N. Matematiko-statisticheskie metody' v e'mpiricheskikh social'no-e'konomicheskikh issledovaniyah** [Mathematical and statistical methods in empirical socio-economic research: study guide], Moscow, Finansy' i statistika, INFRA-M, 2010, 416 p. (In Russian)
- 7 **Kardanova E.Yu., Ivanova A.E. Psihometricheskije issledovaniya: sovremenny'e metody' i novy'e vozmozhnosti dlya obrazovaniya** [Psychometric research: modern methods and new opportunities for education]. *Voprosy' obrazovaniya*, 2023, no. 3, pp.8–19. (In Russian)
- 8 **Beisenova Zh., Shinetova L., Zhabayeva G. Psychometric analysis of test items used to assess teachers' knowledge in the Republic of Kazakhstan (Based on Approbation Results)**. *Pedagogikalyyk olshemder*, 2025, no. 3, pp. 41-52. <https://doi.org/10.63597/UTO3105-4161.2025.3.3.002>.
- 9 **Sadyk U.R., Makybayeva A.K., Nurimbetov N.B. KAZTEST zhujesindegi diagnostikalyyk testileudin otkizilu barysy men tazhiribesi** [Organization and practice of diagnostic testing within the KAZTEST system]. *Pedagogikalyyk olshemder*, 2025, no. 3, pp.30-39. <https://doi.org/10.63597/UTO3105-4161.2025.3.3.003>.
- 10 **Cronbach L.J. Coefficient alpha and the internal structure of tests**. *Psychometrika*, 1951, no. 16, pp. 297–334.
- 11 **Kornilov S.A. Problema nadezhnosti psihologicheskikh shkal i ee znachenie dlya psihodiagnostiki v nauchny'h i prakticheskikh celyah** [The problem of reliability of psychological scales and its

significance for psychodiagnostics in scientific and practical purposes]. *E'ksperimental'naya psihologiya v Rossii: tradicii i perspektivy'*, 2010, pp. 101-105. (In Russian)

12 **Crocker L., Algina J. Introduction to Classical and Modern Test Theory.** Belmont, CA: Cengage Learning, 2008, pp. 482–498.

13 **Council of Europe. Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR).** Strasbourg, Council of Europe, 2001, available at: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4> (accessed 1 October 2025).

14 **Sagindikov I.U., Dinaeva B.B., Ubaydullaeva G.Zh. Olsheu zhane bagalau: monografiya.** [Measurement and assessment: a monograph]. Nur-Sultan, 2021, 284 p. (In Kazakh)

15 **ETS. TOEFL iBT Test.** Available at: <https://www.ets.org/toefl> (accessed 3 November 2025).

16 **IELTS. International English Language Testing System.** Available at: <https://www.ielts.org> (accessed 1 October 2025).

17 **TORFL. Test of Russian as a Foreign Language.** Available at: <https://russian-test.com> (accessed 1 October 2025).

18 **DELFDALF. Diplôme d'Études en Langue Française.** Available at: <https://www.france-education-international.fr> (accessed 1 October 2025). (In French)

#### Авторлар туралы мәліметтер:

*Абдрасилов Болатбек Серикбаевич – ҚР ҰҒА академигі, директор, Ұлттық тестілеу орталығы, Қазақстан Республикасы, 010000 Астана қ., Жеңіс даңғ., 60 ғимарат, e-mail: Bolatbek\_s@mail.ru.*

*Ерсұлтанова Гульнур Серикбаевна\* – PhD докторант, ҚАЗТЕСТ жүйесін дамыту жөніндегі басқарма басшысы, Ұлттық тестілеу орталығы, Қазақстан Республикасы, 010000 Астана қ., Жеңіс даңғ., 60 ғимарат, тел.: 87477936297, e-mail: ersultanovagulnur@gmail.com.*

*Махмұтов Болат Бижанұлы – химия ғылымының кандидаты, доцент, директор кеңесшісі, Ұлттық тестілеу орталығы, Қазақстан Республикасы, 010000 Астана қ., Жеңіс даңғ., 60 ғимарат, e-mail: bolat200@mail.ru.*

*Тоғызбай Нұрғұл Батырбекқызы – магистр, ғылыми-әдістемелік зертхана меңгерушісі, Ұлттық тестілеу орталығы, Қазақстан Республикасы, 010000 Астана қ., Жеңіс даңғ., 60 ғимарат, e-mail: nurgul.9bai@bk.ru.*

*Абдрасилов Болатбек Серикбаевич – академик НАН РК, директор, Национальный центр тестирования, Республика Казахстан, 010000 г. Астана, ул. Победа, здание 60, e-mail: Bolatbek\_s@mail.ru.*

*Ерсұлтанова Гульнур Серикбаевна\* – PhD докторант, руководитель управления по развитию системы КАЗТЕСТ, Национальный центр тестирования, Республика Казахстан, 010000 г. Астана, ул. Победа 60, здание, тел.: 87477936297, e-mail: ersultanovagulnur@gmail.com.*

*Махмұтов Болат Бижанұлы – кандидат химических наук, доцент, советник директора, Национальный центр тестирования, Республика Казахстан, 010000 г. Астана, ул. Победа, здание 60, e-mail: bolat200@mail.ru.*

*Тоғызбай Нұрғұл Батырбекқызы – магистр, заведующая научно-методической лабораторией, Национальный центр тестирования, Республика Казахстан, 010000 г. Астана, ул. Победа, здание 60, e-mail: nurgul.9bai@bk.ru.*

*Abdrasilov Bolatbek Serikbayevich – Academician of the National Academy of Sciences of the Republic of Kazakhstan, Director of the National Testing Center, Republic of Kazakhstan, 010000, Astana, 60 Pobeda Str., e-mail: Bolatbek\_s@mail.ru.*

*Yersultanova Gulnur Serikbayevna\* – PhD student, Head of the Department for the Development of the KAZTEST System, National Testing Center, Republic of Kazakhstan, 010000, Astana, 60 Pobeda Str., tel.: 87477936297, e-mail: ersultanovagulnur@gmail.com.*

*Makhmutov Bolat Bizhanuly – Candidate of Chemical Sciences, Associate Professor, Advisor to Director, National Testing Center, Republic of Kazakhstan, 010000, Astana, 60 Pobeda Str., e-mail: bolat200@mail.ru.*

*Togyzбай Nurgul Batyrbekkyzy – Master, Head of the Scientific and Methodological Laboratory, National Testing Center, Republic of Kazakhstan, 010000, Astana, 60 Pobeda Str., e-mail: nurgul.9bai@bk.ru.*