

Ш.Б. Алтыбаева¹, М.Ш.Тасбулатова^{2*}, Г.Е. Нұрмұханова³, Р.Б.Аубакирова⁴

^{1,2,3}ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Астана қ, Қазақстан Республикасы

⁴Астана қаласы әкімдігінің «№82 «Дарын» мамандандырылған мектеп-лицейі» КММ,

Астана қ, Қазақстан Республикасы

*e-mail: tasbulatova.makhabbat@mail.ru

¹ORCID 0000-0003-0306-861X, ²ORCID 0009-0009-2938-5307,

³ORCID 0009-0007-3665-0602, ⁴ORCID 0009-0008-7823-4946

ҰБТ-ДАҒЫ ҚАЗАҚ ТІЛІН ӨЛШЕУ МОДЕЛІНІҢ ВАЛИДТІЛІГІ: КОГНИТИВТІ ДЕҢГЕЙЛЕР МЕН СТАНДАРТТАРДЫ ВЕРИФИКАЦИЯЛАУ

Бұл зерттеу Ұлттық бірыңғай тестілеудің (ҰБТ) «Қазақ тілі» бейіндік пәні бойынша қолданылатын өлшеу моделінің валидтілігін талдауға бағытталған. Зерттеудің өзектілігі елімізде жоғары оқу орындарына түсу емтиханының білім сапасын басқарудағы стратегиялық рөлімен айқындалады. Зерттеу мақсаты тест мазмұнының оқу бағдарламасына сәйкестігін, когнитивтік деңгейлердің теңгерімін және психометриялық сапасын бағалау болып табылады. Мазмұндық салыстыру мен Блум таксономиясы бойынша когнитивтік верификация, сондай-ақ Раш моделіне негізделген психометриялық талдау (N=257) әдіснамалық негіз ретінде қолданылды. Нәтижелер тесттің мазмұндық тұрғыдан оқу бағдарламасына жалпы сәйкес келетінін және ішкі сенімділік көрсеткішінің жоғары екенін көрсетті (Reliability = 0,81; Item Separation = 7,58). Тапсырмалардың 70%-ы төменгі ретті ойлау дағдыларына (білу, түсіну) тиесілі болса, жоғары деңгейдегі талдау, бағалау және жинақтау дағдылары 10%-дан аз үлесті құрайды. Тапсырмалар қиындығы -2,89-дан +1,77 логитке дейінгі кең ауқымды қамтиды, алайда айнымалылар картасы жоғары қабілетті талапкерлер (+1,5 логиттен жоғары) үшін тапсырмалар тапшылығын айқын көрсетті. Қорытындылай келе тесттің конструкттық валидтілігін арттыру үшін жоғары когнитивтік деңгейдегі тапсырмалар үлесін ұлғайту, QTI стандарты негізінде интерактивті форматтарды енгізу және компьютерлік бейімделген тестілеу технологиясына кезең-кезеңімен көшу ұсынылады.

Түйін сөздер: ҰБТ, қазақ тілі, валидтілік, Раш моделі, Блум таксономиясы, QTI стандарты.

Кіріспе

Қазіргі білім беруде жоғары ставкалы емтихандар тек оқу жетістіктерін тіркеп қоймайды. Олар білім сапасын басқарудың, академиялық іріктеудің, әлеуметтік әділеттілікті қамтамасыз етудің маңызды құралына айналды. Осы тұрғыдан Қазақстандағы Ұлттық бірыңғай тестілеу (ҰБТ) стратегиялық мәнге ие – ол талапкерлердің жоғары оқу орнына түсуге дайындық деңгейін айқындайды.

Қазіргі психометрия ғылымы өлшеу құралының сапасы оның мақсатына, қолданыс аясына және өлшенетін конструкттың теориялық негізіне сәйкестігімен тікелей байланысты екенін көрсетіп отыр [1, 2]. Яғни, тест нәтижелерінің валидтілігі – бұл балдардың қандай когнитивтік қабілеттерді білдіретініне және сол нәтижелерге сүйеніп қабылданатын шешімдердің қаншалықты негізді екеніне тәуелді.

С. Мессик [1] валидтілікті тек статистикалық сипаттама ретінде емес, тестті қолдану мақсаты, нәтижелерді түсіндіру және олардың әлеуметтік салдары бірлігінде қарастыру керек деп есептейді. Ол мазмұндық, құрылымдық, сыртқы және салдарлық аспектілерді өлшеу құралын бағалаудың ажырамас стандарттары деп санайды. Ал Р. Мислеви [2] ұсынған Evidence-Centered Design тұжырымдамасы бағалау жүйесін білім алушының жасырын қабілеттері, бақыланатын әрекеттері және олардың арасындағы эмпирикалық дәлелдермен байланыстырады.

Экономикалық ынтымақтастық және даму ұйымының зерттеулерінде [3] ұлттық емтихандардың екі негізгі қызметі атап көрсетіледі: біріншісі – білім алушының білім

деңгейін растау, екіншісі – одан әрі академиялық траекториясын анықтау үшін іріктеу. Бұл екі қызметті тиімді орындау үшін өлшеу моделі тест мазмұнына, құрылымына және бағалау шешімдерінің салдарына сай болуы керек.

ҰБТ нәтижелері талапкерлердің келесі оқу жолына да, мемлекеттік гранттардың бөлінуіне де тікелей әсер етеді. Сондықтан тесттің қалай құрастырылғаны – оның оқу бағдарламасындағы білім мен дағдыларды шынымен өлшейтіні, нәтижелердің әділдігі, қабылданатын шешімдердің орындылығы – басты назарда болуы тиіс. Бұл ҰБТ-дағы әр пән бойынша өлшеу моделін жүйелі түрде қайта қарастыру қажеттігін көрсетеді.

ҰБТ құрылымы үш міндетті және екі бейіндік пәнді қамтиды [4]. Тестілеуді үш тілде, оқу жылы ішінде бірнеше рет тапсыруға болады – бұл жүйеге белгілі бір икемділік береді.

ҰБТ пәндері бойынша өлшеу модельдерінің валидтілігін зерттеу – білім сапасын басқарудың маңызды бағыттарының бірі. Ұлттық тестілеу орталығы осы бағытта жұмыс істеп келеді. Мәселен, Алтыбаева [5] информатика пәні бойынша тест спецификациясында Блум таксономиясының «қолдану» және «жасау» деңгейлеріне басымдық берілгенін атап өтті. Бұл өзге де бейіндік пәндер бойынша валидтілік зерттеулерін жүргізу қажеттігін көрсетеді.

Біз бұл зерттеуде ҰБТ-дағы қазақ тілін бағалау моделінің мазмұндық және құрылымдық валидтілігін, оның заманауи технологиялық стандарттар негізінде жаңғырту мүмкіндіктерін қарастырдық. Тест тапсырмаларының оқу бағдарламасына сәйкестігін, әсіресе жоғары ретті ойлау дағдыларын (талдау, бағалау, шығармашылық) қаншалықты өлшейтінін анықтауды мақсат еттік [6, 7].

Сонымен қатар тест тапсырмаларының Блум таксономиясы бойынша когнитивтік деңгейлерде қалай таралғанын, психометриялық сапа көрсеткіштерін (сенімділік, дискриминативтілік, күрделілік) және QTI сияқты халықаралық стандарттарға сәйкес технологиялық трансформация мүмкіндіктерін талдадық [8].

Осы мәселелерді Мессиктің валидтілік тұжырымдамасы аясында зерделеп, ҰБТ-дағы қазақ тілін бағалауды теориялық және практикалық тұрғыдан жетілдіруге ғылыми негіз дайындауға тырыстық.

Зерттеу материалдары мен әдістері

Зерттеу барысында ҰБТ қазақ тілі бойынша өлшеу моделінің валидтілігі мен технологиялық трансформациясын бағалауды көздедік. Қойылған міндеттерді шешу үшін сандық психометриялық талдау мен тест мазмұнын сараптауды біріктіретін аралас әдіснаманы қолдандық [9, 10].

Алдымен нормативтік құжаттарға сапалық талдау жасадық. «Қазақ тілі» пәнінің типтік оқу бағдарламаларын, ондағы күтілетін нәтижелерді, олардың когнитивтік күрделілігін қарастырдық. Сонымен қатар 2026 жылға арналған тест спецификациясындағы тақырыптық ауқым, тапсырмалар пішімі және күрделілік деңгейлерінің үлесі (А – 50%, В – 30%, С – 20%) салыстырмалы түрде талданды.

Тест тапсырмаларының когнитивтік талаптарын бағалау үшін Блум таксономиясының классикалық нұсқасын қолдандық [11, 12]. Әр тапсырманы білу, түсіну, қолдану, талдау, синтез, бағалау деңгейлерінің қайсысына сәйкес келетінін анықтап, олардың таралу құрылымын бағаладық.

Негізгі материал ретінде ҰТО-ның ресми базасынан алынған «Қазақ тілі» тапсырмаларын пайдаландық. Үш форматтағы тапсырманы талдадық:

- бір немесе бірнеше дұрыс жауапты таңдау (MC) – грамматика, лексика, орфография;
- контекстік тапсырмалар (CONT) – мәтінді оқу, түсіну, контекстпен жұмыс;
- сәйкестікті анықтау (SOOTV) – логикалық, мағыналық байланыстар.

Деректер статистикалық өндеуге жарамды форматта шығарылды, әр тапсырма мен тестіленушіге бірегей идентификатор берілді.

Психометриялық талдау үшін Раштың бірпараметрлік моделі мен Winsteps 5.2.3 бағдарламасын қолдандық [13]. Талдау 2025 жылғы ҰБТ апробациясының анонимделген деректеріне негізделді. Іріктеме кездейсоқ стратификациялық әдіспен құрастырылды, барлық өңірден 257 мектеп бітірушіні қамтыды. Іріктеуде гендерлік арақатынас (шамамен 55% әйелдер, 45% ерлер) және елді мекен түрі (қала/ауыл) ескерілді. Толық жауап берген бітірушілердің деректері ғана енгізілді.

Зерттеу барысында тесттің ішкі сәйкестік сенімділігін Кронбах альфа коэффициентімен (α) [14], тапсырмалардың күрделілік индексі (р-мәні) және дискриминациялық қабілетімен бағаладық. Раш моделіне сәйкестікті Infit/Outfit Mean-Square (MNSQ) статистикасымен тексердік. Қолайлы диапазон ретінде 0,5–1,5 аралығын алдық [15].

Психометриялық талдау ҰТО-дан алынған анонимді деректерге негізделген. Құжаттық талдау ашық дереккөздерге (ресми сайттар, жарияланған жинақтар) сүйенді. Нәтижелер тек жиынтық түрде ұсынылды, жеке талапкерлер немесе мекемелер туралы мәлімет ашылған жоқ.

Біздің қолданған әдістер тест мазмұнының оқу бағдарламасына сәйкестігін, когнитивтік деңгейлердің таралуын және психометриялық сапаны кешенді бағалауға, сондай-ақ қазақ тілін бағалау моделін технологиялық жаңғырту әлеуетін талдауға мүмкіндік берді.

Нәтижелер және талқылау

ҰБТ қазақ тілі пәнінің спецификациясы мен оқу бағдарламасын салыстырдық. Талдау көрсеткендей, тест мазмұны бағдарламаға жалпы сәйкес келеді. Спецификациядағы бес негізгі блок (орфография, лексика, грамматика, пунктуация, оқылым) бағдарламадағы тақырыптарды толық қамтиды. Оқылым бөлімінің функционалдық сауаттылыққа бағдарлануы да бағдарламадағы коммуникативтік және талдамалық құзыреттерді бағалауға мүмкіндік береді.

Дегенмен, бірқатар алшақтықтар да бар. Бағдарламада басымдық берілген жоғары деңгейлі когнитивтік дағдылар (мысалы, графикалық мәліметтерді салыстырмалы талдау, академиялық жанрларды талдау) тест спецификациясында міндетті тапсырма түрлері ретінде нақты көрсетілмеген. Бұл дағдылар бағалау практикасында шектеулі қамтылады.

Зерттеу барысында қазіргі тест моделінің күрделі коммуникативтік және талдамалық құзыреттерді бағалаудағы шектеулерін анықтадық. ҰБТ электронды форматта өткізілетіндіктен, QTI стандартын енгізу арқылы тапсырмаларды интерактивті етуге мүмкіндік бар.

Қазақ тілі пәні 40 тест тапсырмасынан тұрады: 25 бір дұрыс жауапты таңдау, 5 бір дұрыс жауапты таңдаудан тұратын 1 контекст, 5 бір немесе бірнеше дұрыс жауапты таңдау, 5 сәйкестікті анықтау [7]. Тапсырмаларды орындауға орта есеппен 2 минут беріледі.

Дәстүрлі көп таңдаулы тапсырмалардың (MCQ) мүмкіндігі шектеулі – олар негізінен жаттанды білімді тану немесе қарапайым қолдану деңгейінде тексереді. Ал спецификацияда жоғары когнитивтік деңгейдегі тапсырмалар күрделі білім мен дағдыларды, мәселелерді шешу үшін оларды біріктіруді, деректерді талдауды, пайымдау мен тұжырымдарды негіздеуді талап етеді.

QTI сияқты ашық стандарттар интерактивті тапсырмаларды енгізуге жол ашады. Мәтіндік фрагменттерді ретке келтіру, графикалық деректермен жұмыс, элементтерді сәйкестендіру – мұндай тапсырмалар жоғары когнитивтік дағдыларды (талдау, синтез, пайымдау) бағалаудың валидтілігін арттыра алады. Ал компьютерлік бейімделген тестілеу әр талапкерге өз деңгейіне сәйкес тест ұсынып, бағалау дәлдігін арттырады.

Біздің байқауымызша, ҰБТ-ның қазіргі моделі кешенді құзыреттерді бағалауда шектеулі. Валидтілікті арттыру үшін тестті халықаралық цифрлық стандарттармен (QTI) үйлестіріп, адаптивтік технологияларды енгізу қажет.

Тест тапсырмаларын Блумның жаңартылған таксономиясы (Remembering, Understanding, Applying, Analyzing, Evaluating, Creating) негізінде талдадық. Мақсат – спецификацияда жарияланған бағалау мақсаттары мен нақты тапсырмалар арқылы өлшенетін когнитивтік үдерістердің сәйкестігін анықтау, сондай-ақ жоғары деңгейлі ойлау дағдыларының (HOTS) тестте қаншалықты қамтылғанын бағамдау.

Талдау нәтижелері бойынша, спецификацияда көрсетілген «пайдалану», «дәлелдер мен ұғымдарды пайдалану» сияқты әрекеттер жабық форматтағы тапсырмаларда негізінен төменгі деңгейлі когнитивтік әрекеттерге сәйкес келеді. Тестіленуші көбіне тілдік ережелерді тану, еске түсіру, қарапайым интерпретациялаумен шектеледі. Бұл тапсырмалар Remembering және Understanding деңгейлеріне жатады.

Спецификациядағы «талдау», «салыстыру», «жалпылау» сияқты етістіктер теориялық тұрғыдан Applying және Analyzing деңгейлеріне сәйкес келеді. Бірақ жабық типтегі тапсырмаларда бұл әрекеттер тестіленушінің өзіндік аналитикалық ойлауын емес, дайын нәтижені тануын талап етеді. Сондықтан мұндай тапсырмалардың нақты когнитивтік сұранысы Understanding немесе стандартты Applying деңгейінен аспайды.

Спецификацияда көрсетілген «біріктіру», «пайымдау», «негіздеу» сияқты күрделі әрекеттер Analyzing, Evaluating, Creating деңгейлеріне жатады. Алайда практикада олар көпнұсқалы немесе контексті тапсырмалар арқылы жүзеге асырылады, мұнда «пайымдау» дайын қорытындыны таңдауға, «негіздеу» дұрыс аргументті нұсқалар ішінен анықтауға сүйенеді.

Салыстырмалы талдау нәтижесінде оқыту мақсаттарында көзделген когнитивтік дағдылар мен тапсырмалар арқылы өлшенетін когнитивтік деңгейлер арасында айқын алшақтық бар екенін көрдік. Тест спецификациясында мәтінді талдау, ақпаратты синтездеу, аргументациялау, бағалау сияқты жоғары деңгейлі дағдылар көзделгенімен, тест құрылымы негізінен төменгі деңгейдегі ойлау дағдыларын (LOTS) тексеруге бағытталған.

Бұл теңгерімсіздік 1-кестеде анық көрінеді. Тапсырмалардың шамамен 70%-ы Remembering және Understanding деңгейлеріне тиесілі – бұл жабық форматтағы тапсырмалардың (40-тың 35-і) басым болуымен байланысты. Applying және Analyzing деңгейлері шектеулі көлемде қамтылған. Ал Evaluating және Creating деңгейлері тестте фрагментарлы, жанама түрде ғана көрініс табады.

1-кесте – Блум таксономиясы негізінде «Қазақ тілі» пәні бойынша тест тапсырмаларының когнитивтік деңгейлерінің бағалауыштық үлестірімі

Когнитивтік деңгей (Блум таксономиясы)	Негізгі тексерілетін әрекеттер	ҰБТ тестіндегі шамамен үлесі	Спецификациядағы күрделілік деңгейлеріне сәйкестігі
Remembering (Білу, қайта жаңғырту)	Терминдерді, ережелерді, фактілерді тану және есте сақтау	~50–60%	Негізінен базалық деңгей (А), ішінара В деңгейі
Understanding (Түсіну)	Интерпретациялау, қайта тұжырымдау, жіктеу, түсіндіру	~20–30%	Базалық (А) және орташа (В) деңгейлер
Applying (Қолдану)	Ережелерді, үлгілерді стандартты жағдаятта пайдалану	~15–20%	Көбіне орташа деңгей (В)
Analyzing (Талдау)	Байланыстарды анықтау, құрылымдық бөліктерді ажырату, аргументациялау	~5–10%	Орташа (В) және жоғары (С) деңгейдің жекелеген элементтері
Evaluating (Бағалау) Creating (Жасау)	Тексеру, сыни бағалау, көзқарасты негіздеу; Идеяларды генерациялау, жоспарлау, жаңа мәтін құрастыру	< 5%	Жоғары деңгейде (С) фрагментарлы түрде

Қазіргі тест моделі тілдік білімді репродуктивті түрде меңгеру деңгейін өлшеуге бағдарланған және тілді коммуникативтік әрі аналитикалық құрал ретінде қолдану қабілетін толыққанды бағаламайды. Анықталған когнитивтік дисбаланс тесттің конструктивтік валидтілігінің шектеулі екенін көрсетеді және жоғары деңгейдегі ойлау дағдыларын өлшеу тұрғысынан тест мазмұнын қайта қарау қажеттігін айқындайды. Тест спецификациясындағы тапсырмалардың когнитивтік күрделілік деңгейлері Блум таксономиясының бес деңгейі бойынша салыстырылды. Оқу бағдарламасында мақсаттардың басым бөлігі «қолдану» (Applying), «талдау» (Analyzing) және «бағалау» (Evaluating) сияқты жоғары когнитивтік сатыларға бағытталған. Бұл оқушылардан тілдік білімдерін күрделі тапсырмаларды шешуге, сыни тұрғыдан талдау және бағалау дағдыларын пайдалануға талап етеді.

Психометриялық талдау талапкерлердің апробация бойынша алынған нәтижелілікке қатысты статистикалық деректер негізінде жүргізілді (N = 257).

Зерттеу аясында әрбір тапсырма элементінің өлшеу функциясын қаншалықты тиімді орындайтыны талданды. Талдау Winsteps бағдарламалық қамтамасыз етуінің (5.8.3.0 нұсқасы) көмегімен жүзеге асырылды. Қорытынды өңдеуге барлық қатысушылар мен тапсырмалар енгізілді, деректердің жетіспеушілігі немесе статистикалық шектен тыс мәндер анықталған жоқ.

Тесттің жинақталған психометриялық сипаттамаларына сәйкес (2-кесте), тапсырмалардың қиындығын бағалаудағы орташа стандарттық қате 0,32 логитті құрайды. Бұл көрсеткіш берілген іріктеме шегінде өлшеудің жеткілікті дәлдігін және алынған бағалардың сенімділігін айғақтайды.

2-кестеде ұсынылған деректер тесттің жинақталған психометриялық сипаттамаларының жалпы алғанда Раш моделінің талаптарына сәйкес келетінін көрсетеді. Нормаланған қалдықтардың (ZSTD) орташа мәндері Infit үшін 0,00, ал Outfit үшін 0,22 болып, рұқсат етілген -2,0 мен +2,0 аралығында орналасқан. ZSTD көрсеткіштері, әдетте, модельге сәйкестік индекстері (MNSQ) қабылданған шектерден асқан жағдайларда диагностикалық маңызға ие болатынын ескерсек, орташа ZSTD мәндерінің нөлге жуық болуы деректердің модельге жалпы сәйкестігін көрсетеді.

2-кесте - Тесттің жинақталған психометриялық сипаттамалары

Көрсеткіш	SCORE	COUNT	MEASURE	S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
MEAN	161,6	257,0	0,00	0,14	0,99	0,00	1,02	0,22
SEM	11,1	0,0	0,18	0,01	0,01	0,21	0,03	0,26
P.SD	69,1	0,0	1,10	0,03	0,08	1,32	0,18	1,60
S.SD	70,0	0,0	1,11	0,04	0,09	1,34	0,18	1,62
MAX	389,0	257,0	1,77	0,27	1,24	4,05	1,76	5,16
MIN	49,0	257,0	-2,89	0,09	0,85	-2,78	0,69	-3,04
Көрсеткіш	REAL RMSE	TRUE SD	Separation					
Item (Real)	0,14	1,09	7,58					
Item (Model)	0,14	1,09	7,66					

Нормаланған қалдықтар (ZSTD) Infit үшін 0,00, Outfit үшін 0,22 – рұқсат етілген -2,0 мен +2,0 аралығында. Бұл деректердің Раш моделіне жалпы сәйкестігін көрсетеді. Infit MNSQ (0,99) және Outfit MNSQ (1,02) рұқсат етілген 0,5–1,5 аралығында. Яғни, эмпирикалық деректер модельге жақсы үйлеседі.

Дегенмен, жеке тапсырмалар үшін максималды Infit ZSTD 4,05-ке, Outfit ZSTD 5,16-ға жетеді. Бұл кейбір тапсырмалардың модельден айтарлықтай ауытқуын көрсетеді. Оларды

жеке сараптап, мазмұндық немесе құрылымдық түзету қажет. Бірақ бұл ауытқулар жалпы тест деңгейіндегі психометриялық сапаға елеулі әсер етпейді.

Бөлу коэффициенттері (Item Separation = 7,58–7,66) тесттің тапсырмаларды күрделілігі бойынша айқын ажырата алатынын көрсетеді. Нақты және модельдік орташа квадраттық қате (RMSE = 0,14) төмен, шынайы стандартты ауытқу 1,09 логит. Бұл өлшеу дәлдігінің жоғары екенін білдіреді.

3-кесте - Тапсырмалар параметрлерінің психометриялық сипаттамалары

№	Item	саны	Measure	S.E.	Infit MNS Q	Infit ZSTD	Outfit MNS Q	Outfit ZSTD	PTM EA Corr	Obs %	Exp %
1	MC-1	116	0,38	0,13	1,01	0,23	1,03	0,5	0,32	60,7	64,1
2	MC-2	105	0,58	0,13	1	0,1	1,02	0,34	0,32	68,1	65,3
3	MC-3	162	-0,44	0,14	0,99	-0,27	1,01	0,16	0,34	71,6	67,7
4	MC-4	185	-0,9	0,15	0,94	-0,88	0,88	-1,3	0,4	73,9	74
5	MC-5	223	-1,93	0,19	0,94	-0,4	0,99	0,01	0,32	86,8	86,8
6	MC-6	201	-1,27	0,16	0,94	-0,64	0,95	-0,4	0,37	80,2	79
7	MC-7	209	-1,49	0,17	0,9	-1,05	0,8	-1,46	0,43	82,5	81,8
8	MC-8	108	0,52	0,13	1,09	2,06	1,14	1,24	0,21	59,9	64,9
9	MC-9	80	1,05	0,14	1,09	1,45	1,12	1,38	0,33	68,9	71
10	MC-10	125	0,22	0,13	1,17	4,05	1,19	3,35	0,13	53,7	63,8
11	MC-11	173	-0,66	0,14	0,85	-2,78	0,78	-3,04	0,53	73,5	70,4
12	MC-12	124	0,24	0,13	1,02	0,6	1,05	0,91	0,3	64,2	63,8
13	MC-13	49	1,77	0,17	1,08	0,88	1,29	1,7	0,34	79	81,2
14	MC-14	133	0,08	0,13	1,04	0,98	1,07	1,36	0,28	62,3	63,9
15	MC-15	189	-0,99	0,15	0,93	-1,02	0,88	-1,17	0,41	78,6	75,2
16	MC-16	131	0,12	0,13	0,99	-0,19	0,99	-0,25	0,35	63,8	63,8
17	MC-17	211	-1,55	0,17	0,98	-0,17	0,87	-0,88	0,33	82,5	82,5
18	MC-18	79	1,07	0,14	0,98	-0,31	1,01	0,18	0,32	71,6	71,3
19	MC-19	185	-0,9	0,15	0,93	-0,97	0,88	-1,29	0,41	75,5	74
20	MC-20	165	-0,5	0,14	0,99	-0,09	0,97	-0,41	0,34	68,9	68,4
21	MC-21	163	-0,46	0,14	0,96	-0,84	1	0,06	0,37	72,8	67,9
22	MC-22	142	-0,08	0,13	1,04	1	1,07	1,32	0,28	61,9	64,5
23	MC-23	66	1,35	0,15	1,02	0,27	1,08	0,76	0,36	75,1	75,3
24	MC-24	64	1,39	0,15	0,93	-0,93	1	0,01	0,36	77,8	76
25	MC-25	97	0,72	0,14	0,91	-1,94	0,9	-1,55	0,43	75,1	66,6
26	CONT-26	228	-2,13	0,2	0,92	-0,5	0,87	-0,56	0,33	88,7	88,7
27	CONT-27	242	-2,89	0,27	0,92	-0,28	0,69	-1	0,31	94,2	94,2
28	CONT-28	187	-0,95	0,15	0,95	-0,77	0,93	-0,72	0,39	75,5	74,6
29	CONT-29	75	1,15	0,14	1,08	1,31	1,27	1,74	0,16	74,7	72,5
30	CONT-30	51	1,71	0,16	0,97	-0,29	1,06	0,48	0,29	81,3	80,4
31	SOOTV-31	187	0,69	0,09	1,09	1,25	1,06	0,69	0,4	44	48,2

32	SOOTV-32	389	-0,98	0,1	1	0,04	1,18	1,56	0,42	66,5	60,8
33	SOOTV-33	124	1,25	0,1	1,24	1,76	5,16	2,16	0,36	51,4	59,7
34	SOOTV-34	288	-0,12	0,09	0,97	-0,34	0,97	-0,44	0,47	50,2	50,7
35	SOOTV-35	219	0,44	0,09	1,19	1,77	1,28	2,15	0,32	38,9	44,2
36	MC-36	237	0,32	0,09	0,95	-0,71	0,96	-0,51	0,5	44,7	45,9
37	MC-37	132	1,28	0,1	0,96	-0,41	0,95	-0,42	0,43	57,2	58,5
38	MC-38	216	0,53	0,1	0,9	-1,45	0,89	-1,48	0,53	54,9	51,7
39	MC-39	209	0,58	0,09	0,86	-2,07	0,84	-2,21	0,57	57,2	50,1
40	MC-40	193	0,8	0,1	0,96	-0,5	0,97	-0,42	0,45	56,4	55,1
Орташа мән (MEAN)		161,6	257	0	0,14	0,99	0	1,02	0,22	68,1	67,9
Стандартты ауытқу (P.SD)		69,1	0	1,1	0,03	0,08	1,32	0,18	1,6	12,7	11,7

3-кестеде көрсетілгендей, тапсырмалардың көпшілігі Раш моделінің талаптарына қанағаттанарлық сәйкес келеді. Күрделілік деңгейі кең ауқымды қамтиды (-2,89-дан +1,77 логитке дейін). Ең оңай тапсырма – CONT-27 (-2,89 логит), оны 94,2% қатысушы дұрыс орындаған. Ең қиыны – MC-13 (1,77 логит), дұрыс орындағандар 79%.

Модельге сәйкессіздік көрсеткен тапсырмалар да бар. SOOTV-33 (Outfit MNSQ = 1,76; ZSTD = 5,16) және MC-10 (Outfit MNSQ = 1,19; ZSTD = 3,35) ерекшеленеді. Бұл тапсырмалардың жауаптарында күтілмеген үлгілер кездесетінін білдіреді – мұндай тапсырмалардың құрылымын немесе жауап нұсқаларын қайта қарау қажет.

Тапсырма түрлері бойынша:

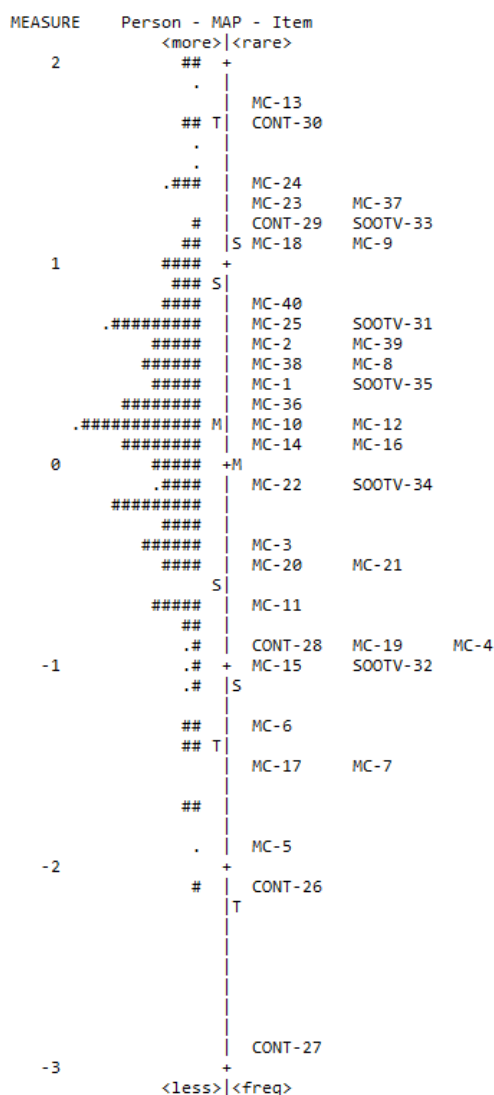
- MC тапсырмалары тұрақты психометриялық сипаттамаларға ие. Жоғары дұрыс жауап пайызы (75–87%) олардың салыстырмалы жеңілдігін көрсетеді.
- CONT тапсырмалары жоғары орындалу пайызымен (74–94%) және тұрақты Infit/Outfit мәндерімен ерекшеленеді.
- SOOTV тапсырмаларының кейбірінде проблемалар бар. SOOTV-33 сәйкессіздігі жоғары, SOOTV-35 (38,9%) және SOOTV-31 (44%) төмен орындалу пайызы олардың қиындығын көрсетеді.

Жалпы тест сапасы:

- Item Separation = 7,58 – тапсырмалар күрделілігі бойынша жақсы ажыратылған.
- Reliability = 0,81 – ішкі сәйкестік жоғары.
- RMSE = 0,14 – өлшеу дәлдігі жоғары.

Айнымалылар картасы (Сурет 1) бойынша респонденттердің қабілеттері -2,5-тен +2,0 логит аралығында, негізгі массасы 0,0-ден +1,0 логит аралығында шоғырланған. Ең көп адам шоғыры +0,5 логит шамасында. Орташа қабілет 0,0 логит – бұл қатысушылардың орташа дайындық деңгейінің басым екенін және тапсырмалардың орташа күрделілікке сәйкес келетінін көрсетеді.

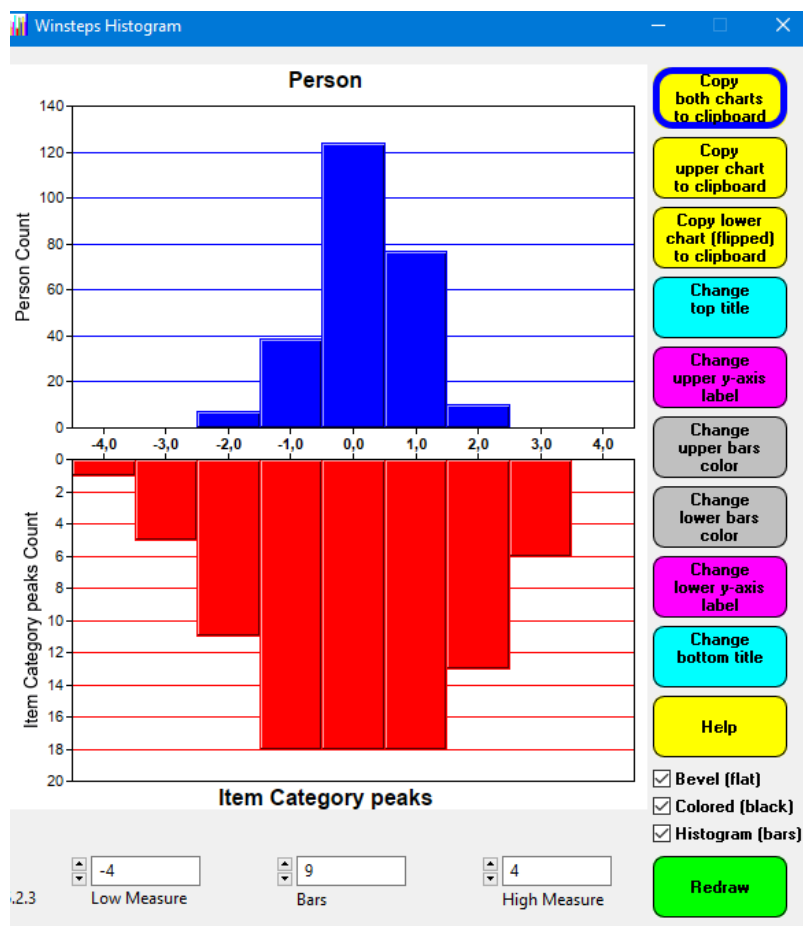
INPUT: 257 Person 40 Item REPORTED: 257 Person 40 Item 90 CATS WINSTEPS 5.8.3.0



Сурет 1 - Айнымалылар картасы: Респонденттер – Тапсырмалар

Тапсырмалар қиындығы $-2,89$ -дан $+1,77$ логитке дейінгі аралықты қамтиды. Ең сәйкес диапазон $-1,5$ -тен $+1,5$ логит аралығы. Бірақ $+1,5$ логиттен жоғары аймақта тапсырмалар саны азаяды (MC-13, MC-24, CONT-30, MC-23, MC-37). Бұл жоғары қабілетті қатысушылар үшін өлшеу дәлдігін төмендетеді. Мұны түзету үшін «Талдау» және «Бағалау» деңгейлеріне сәйкес қосымша тапсырмалар әзірлеу керек.

Жалпы, тест қабілеттер континуумын жақсы қамтиды, төмен және орта деңгейдегі қатысушыларды сенімді бағалауға мүмкіндік береді. Негізгі тест нұсқасын жасауда тапсырмаларды қиындық деңгейіне сәйкес қалыпқа келтіру принципі қолданылды.



Сурет 2 - Вraith диаграммасы

Гистограмма көрсеткендей, қатысушылар қабілеті мен тапсырмалар қиындығы арасындағы сәйкестік негізінен қанағаттанарлық. Негізгі қатысушылар (0,0–1,0 логит) үшін тапсырмалар жеткілікті. Бірақ жоғары қабілеттілер (+1,5 логиттен жоғары) үшін тапсырмалар саны аз – бұл «төбе» эффектісіне әкеледі. Төмен қабілеттілер (–1,5 логиттен төмен) үшін де өте оңай тапсырмалар саны шектеулі

Талқылау. Зерттеу барысында 2026 жылдан бастап қолданылатын «Қазақ тілі» пәні бойынша ҰБТ спецификациясының мазмұны мен құрылымы талданды. Нәтижелер тест мазмұнының мемлекеттік жалпыға міндетті білім беру стандарты талаптарына жалпы сәйкес келетінін көрсетті. Спецификация тілдік нормалардың негізгі компоненттерін (фонетика, лексика, морфология, синтаксис, пунктуация) және оқылым дағдыларын қамтып, мазмұндық тұтастықты қамтамасыз етеді.

Тапсырмалар қиындығының үш деңгейде (А – 50%, В – 30%, С – 20%) ұйымдастырылуы қатысушыларды саралап бағалауға мүмкіндік береді. Алайда когнитивтік сұраныс тұрғысынан деңгейлер арасындағы айырмашылық жеткілікті дәрежеде айқын емес. Жоғары деңгейге (С) жатқызылған тапсырмалар көбінесе мазмұндық немесе форматтық күрделілікке (мысалы, бірнеше дұрыс жауапты таңдау, сәйкестікті анықтауға арналған тест тапсырмасы) негізделеді. Зерттеу нәтижелері бұл тапсырмаларда жоғары ретті ойлау дағдыларының (HOTS), атап айтқанда синтездеу және пайымдаудың шектеулі түрде көрініс табатынын көрсетті. Соның салдарынан тест Блум таксономиясының төменгі сатыларын – білу, түсіну және қолдануды (Remembering, Understanding, Applying) басым өлшейді, бұл оның конструкттық валидтілігін шектейді.

Психометриялық тұрғыдан тесттің сенімділік көрсеткіші (Reliability = 0,81) жақсы деңгейде және нәтижелердің тұрақтылығын қамтамасыз етеді. Алайда сенімділік пен

валидтілік ұғымдарының мазмұны әртүрлі: сенімділік өлшеудің тұрақтылығын сипаттаса, валидтілік тесттің нақты қандай конструкты өлшейтінін анықтайды. Жүргізілген талдау тесттің негізінен реттеуші білімді (ережелерді білу және қолдану) сенімді бағалайтынын, бірақ коммуникативтік және аналитикалық құзыреттіліктерді толыққанды қамтымайтынын көрсетті.

Раш моделінің көрсеткіштері (Item Separation = 7,58) тапсырмалардың күрделілік деңгейі бойынша жақсы ажыратылғанын дәлелдейді. Дегенмен, бұл ажырату көбіне төмен және орта деңгейдегі когнитивтік тапсырмалар шеңберінде жүзеге асады. Айнымалылар картасы (Wright Map) бұл қорытындыны растайды: тапсырмалар күрделілік континуумында кең таралғанымен, қатысушылардың басым бөлігі (шамамен +0,5 логит деңгейінде) негізінен орташа және салыстырмалы жеңіл тапсырмалар арқылы өлшенеді. +1,5 логиттен жоғары аймақта тапсырмалар санының шектеулі болуы жоғары қабілетті талапкерлерді дәл саралау мүмкіндігін төмендетеді.

Аталған нәтижелерді OECD-нің «Future of Education and Skills 2030» [16] тұжырымдамасы аясында қарастырсақ, белгілі бір дисбаланс байқалады. OECD құжаты XXI ғасыр оқушылары үшін тек пәндік білімді ғана емес, сонымен қатар сыни ойлау, проблеманы шешу, рефлексия, жауапкершілік, құндылықтарға негізделген шешім қабылдау сияқты кешенді құзыреттіліктерді дамыту қажеттігін айқындайды. Бұл парадигма білімді репродукциялаудан гөрі оны жаңа жағдайларда трансферлеу мен интеграциялауға басымдық береді.

Зерттеу шектеулері. Алынған нәтижелер 257 қатысушының деректеріне негізделген бір тест нұсқасын талдаумен шектеледі. Раш моделі бойынша стандартты қателік (0,32 логит) параметрлердің бағалану сенімділігін көрсеткенімен, іріктемені кеңейту мен тестің басқа нұсқаларын салыстыру тұжырымдардың жалпылану дәрежесін арттырады. Дегенмен, бұл шектеулер тест құрылымындағы когнитивтік дисбаланс пен оны жаңғырту бағыттарын айқындауға кедергі келтірмейді.

Осы жұмысты дайындау барысында авторлар мәтінді тілдік редакциялау және құрылымын жетілдіру үшін ChatGPT құралын пайдаланды. Аталған құралды қолданғаннан кейін мазмұнды қайта қарап, қажетті түзетулер енгізілді. Авторлар жарияланымның мазмұны үшін толық жауапкершілікті өз мойнына алады.

Әдебиеттер тізімі

1. Messick, S. Validity // Educational Measurement / ed. by R. L. Linn. - 3rd ed. - New York: American Council on Education ; Macmillan, 1989. - P. 13-103.
2. Mislevy, R. J. Evidence and Inference in Educational Assessment. - Princeton, NJ : Educational Testing Service, 1994. - 45 p.
3. Strengthening national examinations in Kazakhstan to achieve national goals / OECD // OECD Education Policy Perspectives. - 2020. - No. 24. - Paris : OECD Publishing. - <https://doi.org/10.1787/0bf8662b-en>
4. Ұлттық бірыңғай тестілеуді өткізу қағидалары : Қазақстан Республикасы Білім және ғылым министрінің 2017 жылғы 2 мамырдағы № 204 бұйрығы. - <https://adilet.zan.kz/kaz/docs/V1700015173>
5. Алтыбаева, Ш. Б. Валидация содержания теста по информатике в рамках ЕНТ: экспертная оценка и психометрический анализ / Ш. Б. Алтыбаева, И. Сагиндиқов, Г. Жабаева, Ш. Шақралиева // Қ. Алтынсарин атындағы Ұлттық білім академиясының «Білім-Образование» ғылыми-педагогикалық журналы. - 2025. - Т. 114, № 3. - С. 68-82. - <https://doi.org/10.59941/2960-0642-2025-3-68-82>
6. Мемлекеттік жалпыға міндетті білім беру стандарттары : Қазақстан Республикасы Оқу-ағарту министрінің 2022 жылғы 3 тамығындағы № 348 бұйрығы. - <https://adilet.zan.kz/kaz/docs/V2200029031>
7. «Қазақ тілі» пәні бойынша Ұлттық бірыңғай тестілеуге арналған тест спецификациясы / Ұлттық тестілеу орталығы. - 2025. - <https://testcenter.kz/wp->

content/uploads/2025/10/01_%D2%9A%D0%B0%D0%B7%D0%B0%D2%9B-%D1%82%D1%96%D0%BB%D1%96_%D0%BA%D0%B0%D0%B7.pdf

8. Question & Test Interoperability (QTI) Specification Documents / 1EdTech Consortium. - 2024. - <https://www.ledtech.org/standards/qti>

9. Yin, R. K. Case Study Research and Applications: Design and Methods. - 6th ed. - Los Angeles : SAGE Publications, 2018. - 352 p.

10. Creswell, J. W. Qualitative Inquiry and Research Design: Choosing Among Five Approaches / J. W. Creswell, C. N. Poth. - 4th ed. - Los Angeles : SAGE Publications, 2018. - 488p.

11. Bloom, B. S. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain / B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, D. R. Krathwohl. - New York : Longmans, Green, 1956. - 207 p.

12. Krathwohl, D. R. A Revision of Bloom's Taxonomy: An Overview // Theory Into Practice. - 2002. - Vol. 41, No. 4. - P. 212-218. - https://doi.org/10.1207/s15430421tip4104_2

13. Linacre, J. M. Winsteps® Rasch Measurement Computer Program. - Beaverton, Oregon : Winsteps.com, 2024. - <https://www.winsteps.com>

14. Cronbach, L. J. Coefficient Alpha and the Internal Structure of Tests // Psychometrika. - 1951. - Vol. 16, No. 3. - P. 297-334. - <https://doi.org/10.1007/BF02310555>

15. Bond, T. G. Applying the Rasch Model: Fundamental Measurement in the Human Sciences / T. G. Bond, C. M. Fox. - 3rd ed. - New York : Routledge, 2015. - 406p. - <https://doi.org/10.4324/9781315814698>

16. Future of Education and Skills 2030 / OECD. - 2022. - <https://www.oecd.org/en/about/projects/future-of-education-and-skills-2030.htm>

Ш.Б. Алтыбаева, М.Ш. Тасбулатова, Г.Е. Нұрмұханова, Р.Б. Аубакирова

ВАЛИДНОСТЬ ИЗМЕРИТЕЛЬНОЙ МОДЕЛИ ПО КАЗАХСКОМУ ЯЗЫКУ В ЕНТ: ВЕРИФИКАЦИЯ КОГНИТИВНЫХ УРОВНЕЙ И СТАНДАРТОВ

Данное исследование направлено на анализ валидности измерительной модели, применяемой в рамках Единого национального тестирования (ЕНТ) по профильному предмету «Казахский язык». Актуальность исследования определяется стратегической ролью вступительного экзамена в высшие учебные заведения в управлении качеством образования в стране. Целью исследования является оценка соответствия содержания теста учебной программе, баланса когнитивных уровней, а также его психометрического качества. В качестве методологической основы использованы содержательное сопоставление и когнитивная верификация на основе таксономии Блума, а также психометрический анализ, основанный на модели Раша (N=257). Результаты показали, что тест в целом соответствует учебной программе с содержательной точки зрения и характеризуется высоким уровнем внутренней надежности (Reliability = 0,81; Item Separation = 7,58). При этом 70% заданий относятся к навыкам низшего порядка (знание, понимание), тогда как задания, направленные на развитие навыков высокого уровня – анализа, оценки и синтеза – составляют менее 10%. Диапазон сложности заданий варьируется от -2,89 до +1,77 логита, однако карта переменных выявила дефицит заданий для абитуриентов с высоким уровнем подготовки (выше +1,5 логита). В заключение, для повышения конструктивной валидности теста рекомендуется увеличить долю заданий высокого когнитивного уровня, внедрить интерактивные форматы на основе стандарта QTI, а также поэтапно перейти к технологии компьютерного адаптивного тестирования.

Ключевые слова: ЕНТ, казахский язык, валидность, модель Раша, таксономия Блума, стандарт QTI.

Sh.B. Altybayeva, M.Sh. Tasbulatova, G.E. Nurmukhanova, R.B. Aubakirova

VALIDITY OF THE KAZAKH LANGUAGE MEASUREMENT MODEL IN THE UNIFIED NATIONAL TESTING: VERIFICATION OF COGNITIVE LEVELS AND STANDARDS

This study analyzes the validity of the measurement model for the «Kazakh Language» subject within the Unified National Testing (UNT) and explores its potential for modernization based on contemporary technological standards. The relevance of the research is determined by the strategic role of high-stakes exams in academic selection and educational quality management. The aim of the study is to assess the alignment of the test content with the curriculum, the balance of cognitive levels, and its psychometric quality. A mixed-method approach was employed as a methodological framework: content audit and cognitive verification using Bloom's Taxonomy, along with psychometric analysis based on the Rasch model (N=257). The results indicate that the test generally aligns with the curriculum content and demonstrates high internal reliability (Reliability = 0,81; Item Separation = 7,58). 70% of the items target Lower Order Thinking Skills (LOTS: remembering, understanding), while Higher Order Thinking Skills – analyzing, evaluating, and creating (HOTS) – account for less than 10%. Item difficulty covers a wide range (from -2,89 to +1,77 logits); however, the variable map clearly shows a shortage of items for high-ability test-takers (above +1,5 logits). In conclusion, to enhance the test's construct validity, it is recommended to increase the proportion of high-cognitive-level items, introduce interactive formats based on the QTI standard, and gradually transition to computerized adaptive testing technology.

Keywords: UNT, Kazakh language, validity, Rasch model, Bloom's Taxonomy, QTI standard.

References

1. Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). American Council on Education; Macmillan.
2. Mislevy, R. J. (1994). *Evidence and inference in educational assessment*. Educational Testing Service.
3. OECD. (2020). *Strengthening national examinations in Kazakhstan to achieve national goals*. OECD Education Policy Perspectives, (24). Paris: OECD Publishing. <https://doi.org/10.1787/0bf8662b-en>
4. Ultyq biryngai testileudi otkizu qagidalary: Kazakhstan Respublikasy Bilim jane gylym ministrinin 2017 jylgy 2 мамыrdagy No. 204 buirygy [Rules for conducting the Unified National Testing: Order of the Minister of Education and Science of the Republic of Kazakhstan dated May 2, 2017, No. 204]. (2017). <https://adilet.zan.kz/kaz/docs/V1700015173>
5. Altybayeva, S. B., Sagindikov, I., Zhabayeva, G., Shakraliyeva, S. (2025). Validatsiya sodержaniya testa po informatike v ramkakh ENT: ekspertnaya otsenka i psikhometricheskii analiz [Validation of the content of the ICT test within the framework of the UNT: Expert assessment and psychometric analysis]. *Q. Altynsarin atyndagy Ultyq bilim akademiyasynyn «Bilim-Obrazovanie» gylimi-pedagogikalyq jurnaly* [Scientific-Pedagogical Journal «Bilim-Obrazovanie» of the National Academy of Education named after Y. Altynsarin], 114(3), 68-82. <https://doi.org/10.59941/2960-0642-2025-3-68-82>
6. Memlekettyk jalpyga mindetty bilim beru standarttary: Kazakhstan Respublikasy Oqu-agartu ministrinin 2022 jylgy 3 tamyzyndagy No. 348 buirygy [State compulsory standards of education: Order of the Minister of Education of the Republic of Kazakhstan dated August 3, 2022, No. 348]. (2022). <https://adilet.zan.kz/kaz/docs/V2200029031>
7. «Kazakh tili» pani boiynsha Ultyq biryngai testileuge arналган test spetsifikatsiyasy [Test specification for the Unified National Testing in the subject «Kazakh Language»]. (2025). Ultyq testileu ortalygy. https://testcenter.kz/wp-content/uploads/2025/10/01_%D2%9A%D0%B0%D0%B7%D0%B0%D2%9B-%D1%82%D1%96%D0%BB%D1%96_%D0%BA%D0%B0%D0%B7.pdf
8. 1EdTech Consortium. (2024). *Question & test interoperability (QTI) specification documents*. <https://www.1edtech.org/standards/qti>
9. Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

10. Creswell, J. W., Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
11. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longmans, Green.
12. Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2
13. Linacre, J. M. (2024). *Winsteps® Rasch measurement computer program*. <https://www.winsteps.com>
14. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
15. Bond, T. G., Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315814698>
16. OECD. (2022). *Future of education and skills 2030*. <https://www.oecd.org/en/about/projects/future-of-education-and-skills-2030.htm>

Авторлар туралы мәлімет:

Алтыбаева Шугыла Болатовна - ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК Орта және жоғары білім берудегі тест тапсырмаларын қалыптастыру басқармасының басшысы, Астана, Қазақстан, e-mail: shugyla.altybayeva@gmail.com

Тасбулатова Махаббат Шалабаевна (автор-корреспондент) - ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, бас сарапшы, Астана, Қазақстан, e-mail: tasbulatova.makhabbat@mail.ru

Нұрмұханова Гүлнар Ерғалиқызы - ҚР ҒЖБМ Ұлттық тестілеу орталығы» ШЖҚ РМК, бас сарапшы, Астана, Қазақстан, e-mail: Gulnarae@bk.ru

Аубакирова Риза Бекмухамедовна - Астана қаласы әкімдігінің «№82 «Дарын» мамандандырылған мектеп-лицейі» КММ, қазақ тілі мен әдебиеті пәні мұғалімі, Астана, Қазақстан, e-mail: riza211989@mail.ru

Сведения об авторах:

Алтыбаева Шугыла Болатовна - руководитель Управления по формированию тестовых заданий для высшего и среднего образования РГП на ПХВ «Национальный центр тестирования» МНВО РК, Астана, Казахстан, e-mail: shugyla.altybayeva@gmail.com

Тасбулатова Махаббат Шалабаевна (автор-корреспондент) - РГП на ПХВ «Национальный центр тестирования» МНВО РК, заведующий лабораторией, Астана, Казахстан, e-mail: tasbulatova.makhabbat@mail.ru

Нұрмұханова Гүлнар Ерғалиқызы - РГП на ПХВ «Национальный центр тестирования» МНВО РК, главный эксперт, Астана, Казахстан, e-mail: Gulnarae@bk.ru

Аубакирова Риза Бекмухамедовна - учитель казахского языка и литературы КГУ «Специализированный школа-лицей №82 «Дарын» акимата города Астаны, Астана, Казахстан, e-mail: riza211989@mail.ru

Information about authors:

Altybayeva Shugyla Bolatovna - Head of the Office for the Formation of Test Items for Higher and Secondary Education «National Testing Center» Ministry of Science and Higher Education of the Republic of Kazakhstan, Astana, Kazakhstan, e-mail: shugyla.altybayeva@gmail.com

Tasbulatova Makhabbat Shalabaevna (corresponding author) - Republican State Enterprise on the Right of Economic Management «National Testing Center» of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Chief expert, Astana, Kazakhstan, e-mail: tasbulatova.makhabbat@mail.ru

Nurmukhanova Gulnar Ergalikyzy - Republican State Enterprise on the Right of Economic Management «National Testing Center» of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Chief expert, Astana, Kazakhstan, e-mail: Gulnarae@bk.ru

Aubakirova Riza Bekmukhamedovna - teacher of Kazakh language and literature KSU «Specialized school- lyceum №82 «Daryn» of the akimat of Astana city, Astana, Kazakhstan, e-mail: riza211989@mail.ru