

ПСИХОМЕТРИЧЕСКИЙ АНАЛИЗ ТЕСТА ПО ИНФОРМАТИКЕ ДЛЯ ПЕДАГОГОВ ТЕХНИЧЕСКОГО И ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

Ш.Ш. Альшурина* , Н.А. Дидарбекова , Д.С. Оспанов , А.А. Кулдеева 

РГП на ПХВ «Национальный центр тестирования» МНВО РК,

Республика Казахстан, г. Астана

*e-mail: shinara.a@mail.ru, nauzhan.didarbekova@mail.ru,
dimokospan@gmail.com, lunochka_90@mail.ru

Аннотация. В статье представлены результаты валидации теста по информатике, используемого для оценки знаний педагогов технического и профессионального образования (ОЗП ТиПО) в Республике Казахстан. Рассмотрены нормативно-правовые основы аттестации, структура квалификационных категорий и процедура тестирования. Эмпирическую базу составили данные 112 педагогов, выполнивших 50 дихотомических заданий. Психометрический анализ проводился с применением классической теории тестирования (КТТ) и модели Раиа (IRT). Установлена умеренная надежность теста ($KR-20 = 0,655$; $\alpha = 0,65$), средняя трудность заданий (0,434). Выявлены задания с избыточной сложностью (11 заданий, $p \leq 0,2$), чрезмерной легкостью (1 задание, $p = 0,89$), отрицательной (4 задания) и низкой дискриминацией (6 заданий). IRT-анализ подтвердил несоответствие ряда заданий модели: отрицательная корреляция с общей мерой способности (№ 1, 13, 37), отклонения Fit-статистик (Outfit MNSQ $> 1,3$ у № 10; $< 0,7$ у № 35). Wright Map показала дефицит заданий повышенной сложности. На основе анализа предложены рекомендации по переработке проблемных заданий, усилению дистракторов и повторной апробации. Сделан вывод о содержательной валидности теста и его пригодности для объективной оценки компетентности педагогов после внесения корректировок.

Ключевые слова: валидация теста, педагоги ТиПО, информатика, классическая теория тестирования, модель Раиа, надежность, дискриминативность.

Введение

Система технического и профессионального образования (ТиПО) в Республике Казахстан на современном этапе переживает период институциональной трансформации. В условиях глобальных вызовов «Индустрии 4.0» и цифровизации экономики, качество подготовки специалистов в колледжах напрямую коррелирует с профессиональной компетентностью педагогического состава. Оценка и аттестация педагогов в этом контексте перестают быть формальными административными процедурами, превращаясь в инструмент управления человеческим капиталом и обеспечения качества образовательного процесса.

Нормативно-правовое регулирование аттестации педагогов ТиПО основывается на ряде нормативных правовых актов, ключевое место среди которых занимает приказ Министра просвещения Республики Казахстан от 27 января 2016 года № 83 (с изменениями, внесёнными приказом Министра просвещения Республики Казахстан от 25 февраля 2025 года № 32), утвердивший Правила и условия проведения аттестации педагогов. Данный документ интегрирует положения Трудового кодекса РК, законов «Об образовании», «О статусе педагога» и «О государственных услугах».

Философия современной системы аттестации, согласно программным документам, основывается на принципах социальной справедливости, академической добросовестности и объективности. В 2025 году в указанный приказ были внесены существенные изменения приказом Министра просвещения Республики Казахстан от 25 февраля 2025 года № 32 «О внесении изменений в правила аттестации педагогических работников». Этот акт, зарегистрированный в Министерстве юстиции РК и приводящий Правила и условия проведения аттестации педагогов в новую редакцию, обновляет порядок и критерии аттестации.

Одним из наиболее значимых нововведений является внедрение Национальной платформы непрерывного профессионального развития педагога. Эта цифровая система,

классифицируемая как объект информатизации, предназначена для автоматизированного сбора, хранения и обработки данных о профессиональной деятельности учителя [1]. Платформа позволяет реализовать концепцию проактивного формирования портфолио, минимизируя субъективизм при оценке достижений. Согласно новым правилам, аттестационный период для педагогов, методистов и заместителей руководителей составляет 5 лет, в то время как первые руководители организаций образования проходят аттестацию каждые 3 года [1].

Процедура аттестации разделена на три четких этапа: квалификационная оценка, оценка знаний педагогов (далее – ОЗП) и комплексное аналитическое обобщение результатов деятельности. Этап ОЗП представляет собой стандартизированное тестирование, направленное на проверку предметных компетенций. Педагогу предоставляется право на две попытки сдачи ОЗП.

Современная система категорий в Казахстане выстроена таким образом, как показано в Таблице 1, чтобы стимулировать не только педагогический стаж, но и научно-методическую активность.

Таблица 1 – Квалификационные категории педагогов

Квалификационная категория	Блок	Квалификационные пороги	Минимальный стаж и требования
Педагог	Предметные знания	50%	Стаж более 1 года
Педагог-модератор		60%	Наличие соответствующих достижений
Педагог-эксперт		70%	Высокие показатели качества знаний обучающихся
Педагог-исследователь		80%	Наличие научных публикаций и авторских программ
Педагог-мастер		90%	Трансляция опыта на республиканском уровне

Таким образом, дифференциация квалификационных категорий и установление пороговых значений успешности аттестации формируют институциональную основу для объективной оценки профессиональных достижений педагогов. В данной системе ключевую роль играет этап ОЗП, выступающий универсальным инструментом сопоставимой и стандартизированной проверки предметных компетенций вне зависимости от квалификационной категории и уровня образовательной организации.

ОЗП проводится по основному предмету, преподаваемому педагогом, и включает 50 тестовых заданий с выбором одного правильного ответа. Для первых руководителей, заместителей руководителя организаций образования, методических кабинетов (центров) ОЗП проводится по направлению «Знание законодательства Республики Казахстан и нормативных правовых актов в области образования» и включает 60 (шестьдесят) тестовых заданий с выбором одного правильного ответа [1]. Все задания оцениваются дихотомически – по 1 баллу за каждый правильный ответ. Максимально возможный результат составляет: для педагогов – 50 баллов. Для обеспечения содержательной валидности ОЗП по каждому предмету разрабатываются спецификации тестов, определяющие структуру контрольных материалов, распределение заданий по разделам учебной программы, уровни сложности и критерии оценивания. Публикация спецификаций на официальном сайте Национального центра тестирования (НЦТ) обеспечивает прозрачность и открытость процедуры тестирования. Тестовые задания разрабатываются квалифицированными педагогами и проходят многоуровневую экспертную оценку на соответствие утвержденным учебным программам и учебникам.

ОЗП построено по единым принципам для всех предметов, однако специфика отдельных предметов требует учета когнитивной нагрузки и особенностей контента. Наиболее ярко это проявляется в предмете «Информатика», где задания связаны с алгоритмическим мышлением и решением задач повышенной сложности. Ввиду этого нормативно закреплены специфические условия проведения теста для учителей информатики: продолжительность экзамена составляет 125 минут [1].

Актуальность исследования обусловлена необходимостью перехода от дескриптивных методов оценки к объективным психометрическим измерениям, обеспечивающим валидность и надежность результатов аттестации в контексте международных стандартов.

Настоящее исследование направлено на обоснование валидности результатов ОЗП на основе теории валидности Samuel Messick [2]. Для ответа на главный исследовательский вопрос – какие доказательства подтверждают валидность интерпретации результатов ОЗП как тестирования, способного адекватно описывать текущую компетентность педагога и прогнозировать его профессиональные достижения, – проводились валидизационные исследования в соответствии с Объединёнными стандартами [3].

Обзор литературы

Анализ научной литературы свидетельствует о растущем интересе исследователей к проблемам ТиПО в Казахстане. Одной из центральных тем является развитие методической компетентности учителей ИТ-дисциплин. В условиях дефицита кадров – по данным на 2023 год в стране не хватает около 5 тысяч педагогов, а каждый пятый выпускник не трудоустраивается по профилю – вопрос качества подготовки становится критическим [4].

Другим важным вектором исследований является цифровая компетентность будущих педагогов ТиПО. Для оценки этого параметра в казахстанском контексте была успешно адаптирована шкала Digital Competence Scale for University Students (DC-US), изначально разработанная в Китае. Валидизация шкалы на выборке из 202 студентов показала, что инструмент является надежным (приемлемые значения альфа Кронбаха и Rasch-анализа) и способен дифференцировать уровни технической грамотности и операционных навыков. В среднем, будущие педагоги ТиПО в Казахстане демонстрируют «средний» уровень цифровой компетенции, что указывает на необходимость усиления технологической подготовки в вузах [5].

Таким образом, помимо оценки цифровой компетентности будущих педагогов, исследования подчеркивают необходимость системного подхода к их профессиональному развитию. Систематические обзоры литературы выявляют ключевые направления, требующие внимания при подготовке специалистов ТиПО: усиление непрерывного образования, адаптация к современным условиям и учет специфики преподавания технических дисциплин, где важен актуальный индустриальный опыт [6].

Исследования, посвящённые аттестации педагогических кадров в Казахстане, демонстрируют особое внимание к нормативно-правовой базе, изменениям в регулировании аттестации, а также к возможностям сопоставления с международным опытом. Так, Пенкина (2021) осуществляет сравнительный анализ действующих и отменённых нормативных актов, выявляя ключевые сложности прохождения аттестации и подчёркивая необходимость адаптации процедур к специфике казахстанской системы образования [7].

Булатбаева и Кожаяева (2022) фокусируются на возможностях использования международного опыта аттестации (Сингапур, Финляндия) и отмечают восприятие процедуры самими педагогами [8]. Кроме того, качественное исследование Серикбаевой (2025) анализирует отношение учителей к процессу аттестации и его влияние на профессиональное развитие и благополучие педагогов [9].

Анализ научной литературы по подготовке педагогов ТиПО, вопросам аттестации и оценке показывает, что внимание исследователей более сосредоточено на нормативно-правовой базе и профессиональном развитии, но при этом акцент смещён на

содержательные, организационные и квалиметрические аспекты, такие как дефицит кадров, методическая компетентность и возможности профессионального развития. Однако наблюдается исследовательский пробел: отсутствуют работы, в которых результаты ОЗП подвергаются строгому психометрическому анализу с использованием как классической теории тестирования (КТТ), так и современных моделей отклика на задания (IRT), с целью проверки их соответствия международным стандартам надежности [10]. Несмотря на критическую важность качества тестов для ЕНТ и ОЗП, конкретные данные о дискриминативности заданий, эффективности дистракторов и соответствии модели Раша для тестов по отдельным предметам (например, информатике) практически отсутствуют.

Данное исследование призвано восполнить этот пробел, применив методологию, общепринятую в международной практике, к анализу реальных данных тестирования педагогов ТиПО в Казахстане. В частности, классическая теория тестирования позволяет провести первичный анализ трудности и дискриминативности заданий, оценить надежность теста и эффективность дистракторов, тогда как современные методы, включая Раш-моделирование, обеспечивают более детальную оценку тестовых заданий, включая дихотомические задания и задания с частичным начислением баллов (Partial Credit Model) [11-13].

Таким образом, исследования в области педагогических измерений демонстрируют, что обеспечение высокого качества тестов является неотъемлемой частью системы аттестации и профессиональной оценки педагогов, напрямую влияя на валидность и надежность интерпретации результатов экзаменов.

Материалы и методы

Для проведения психометрического анализа теста по предмету «Информатика» (вариант 419358) был использован комплекс методов, основанный на двух парадигмах: классической теории тестирования (КТТ) и современной теории ответов на задания (IRT) в рамках модели Раша. Исследование включало следующие этапы:

Характеристика выборки и подготовка данных. Эмпирическую базу исследования составили результаты тестирования 112 педагогов технического и профессионального образования (ТиПО), прошедших процедуру оценки знаний педагогов (ОЗП) по предмету «Информатика». В выборку вошли педагоги различных квалификационных категорий («педагог», «педагог-модератор», «педагог-эксперт» и др.) со стажем профессиональной деятельности более одного года. Исследование охватывало представителей различных регионов Республики Казахстан, что позволило обеспечить вариативность выборки и повысить репрезентативность полученных результатов. Тест состоял из 50 заданий с дихотомической оценкой (1 – верно, 0 – неверно). Первичная обработка данных включала расчет описательных статистик суммарного балла (среднее значение, стандартное отклонение, минимум, максимум, асимметрия, эксцесс) для проверки распределения результатов на соответствие нормальному закону распределения.

Анализ в рамках Классической Теории Тестов (КТТ). Данный этап выполнен с использованием специализированных программ TAP (Test Analysis Program) и Jamovi. Целью было получение базовых психометрических характеристик теста и его заданий.

- **Оценка надежности теста.** Для определения внутренней согласованности использовался коэффициент KR-20 (Кьюдера-Ричардсона), так как он наиболее адекватен для дихотомических данных. Дополнительно рассчитывался коэффициент α -Кронбаха для сопоставимости с другими исследованиями. Значения коэффициентов интерпретировались следующим образом: $< 0,5$ – низкая надежность; $0,5-0,7$ – умеренная надежность; $> 0,7$ – высокая надежность; $> 0,8$ – уровень, достаточный для высокостратных экзаменов (high-stakes testing).

- **Анализ трудности заданий (Difficulty).** Для каждого задания рассчитывался индекс трудности (p -value) как доля испытуемых, ответивших правильно. В соответствии с рекомендациями [10], задания классифицировались как:

- Слишком сложные: $p \leq 0,2$.
- Оптимальные: $0,2 < p < 0,8$.

- Слишком легкие: $p \geq 0,8$.

• **Анализ дискриминативности (Discrimination).** Оценивалась способность задания дифференцировать сильных и слабых испытуемых. Использовались два показателя:

- Индекс дискриминации (Disc. Index), рассчитываемый методом крайних групп (27%).
- Скорректированный точечно-бисериальный коэффициент корреляции (AdjPtBiserial) между ответом на задание и общим баллом по тесту (без учета данного задания). Задания с AdjPtBiserial $< 0,1$ считались имеющими низкую дискриминативность, а с отрицательными значениями (< 0) – непригодными для использования, так как они "работают" против общей оценки.

• **Дистракторный анализ.** Проверялась эффективность неправильных вариантов ответов. Анализировалась частота выбора каждого дистрактора и корреляция выбора дистрактора с общим баллом. Дистрактор считался неработающим, если его выбирали менее 5% испытуемых, или "парадоксальным", если его предпочитали сильные участники (положительная корреляция с общим баллом).

Анализ в рамках Современной Теории Тестов (IRT). Для более глубокого изучения свойств заданий и проверки их соответствия конструкту теста был применен анализ с использованием однопараметрической логистической модели Г. Раша (Rasch model). Расчеты выполнялись в программе Winsteps.

• **Обоснование выбора модели.** Модель Раша была выбрана, так как она преобразует сырые баллы в интервальную шкалу (логиты) и обеспечивает инвариантность измерений, что позволяет объективно сравнивать уровень подготовленности педагогов и сложность заданий независимо от распределения выборки [11-16].

• **Оценка соответствия заданий модели (Fit Statistics).** Для каждого задания анализировались статистики:

– Infit MNSQ (взвешенный средний квадрат остатков) – чувствителен к неожиданным ответам на задания, близкие по уровню к способности испытуемого.

– Outfit MNSQ (невзвешенный средний квадрат остатков) – чувствителен к выбросам (неождаанным ответам на слишком легкие или слишком сложные задания). Следуя рекомендациям Б. Райта и М. Линакра, задания считались соответствующими модели при значениях MNSQ в диапазоне 0,7 – 1,3. Выход за эти пределы интерпретировался как «шум» в данных: значения $< 0,7$ указывают на избыточную предсказуемость (задание "скучно" для выборки), $> 1,3$ – на несоответствие модели (задание измеряет нечто иное или содержит ошибки).

• **Анализ корреляции «пункт-мера» (Pt-Measure Corr.).** Рассчитывалась корреляция ответов на задание с итоговой мерой способности в логитах. Задания с отрицательной корреляцией (Corr. < 0) считались грубо нарушающими конструкт теста.

• **Построение карты "Person-Item" (Wright Map).** Для визуальной оценки соответствия распределения уровня способностей испытуемых и уровня сложности заданий строилась совместная карта (Wright Map). Она позволяла выявить "потолок" и "пол" теста, а также наличие пробелов в покрытии шкалы способностей заданиями соответствующей трудности.

Интерпретационный этап. На заключительном этапе проводилась триангуляция данных, полученных методами КТТ и IRT. Совпадение проблемных заданий по разным критериям (низкая трудность p и высокий Outfit MNSQ; отрицательная дискриминация в КТТ и отрицательная корреляция в IRT) служило основанием для выработки итоговых рекомендаций по их переработке или исключению из тестового банка.

Результаты и обсуждение

Анализ в классической теории тестов (КТТ)

В ходе исследования проводилась оценка психометрических характеристик теста по информатике, предназначенного для оценки знаний педагогов ТиПО. Представленные в

таблице 2 результаты анализа варианта 419358, включавшего 50 заданий и апробированного на выборке из 112 педагогов с применением методов классической теории тестирования (программы TAP, Jamovi), свидетельствуют об умеренном уровне надежности и приемлемой сбалансированности теста в целом.

Таблица 2 – Основные результаты анализа (КТТ)

Показатель	Вариант 419358
Число участников	112
Средний балл	21,7
Стандартное отклонение	5,33
Минимум / Максимум	10/40
Асимметрия	0,68
Экспесс	1,18
Надежность (KR-20 и Cronbach's Alpha)	0.655
Средняя трудность	0.434
Число заданий трудности < 0,2	11 (№ 1,4,9,10,13,14,22,35,43,44,46)
Число заданий трудности > 0,8	1 (№ 29)
Проблемные задания с дискриминативностью < 0	4 (№ 1,10,13,37)
Задания с дискриминативностью < 0,1	9 (№ 7,8,10,24,25,37,44,46,50)
Число заданий с неработающими дистракторами	0

Тест демонстрирует умеренную надежность ($KR-20 = 0,655$), что указывает на относительную стабильность результатов. Средняя трудность заданий (0,434) свидетельствует о приемлемой сбалансированности теста. Однако выявлен дисбаланс: 11 заданий избыточно сложны ($*p* \leq 0,2$), одно – чрезмерно легкое ($*p* = 0,89$). Дискриминативная способность нарушена у 4 заданий (отрицательные значения) и еще у 6 заданий низкая ($Disc. Index \leq 0$ или $AdjPtBiserial \leq 0,1$). На основании анализа, для повышения надежности рекомендуется переработка проблемных заданий и добавление новых элементов аналогичного качества.

Для выявления причин низкой дискриминативной способности заданий был проведен дистракторный анализ, результаты которого представлены в таблице 3. Анализ выполнялся на основе данных, полученных в программах TAP и Jamovi.

Таблица 3 – Результаты дистракторного анализа проблемных заданий

№ задания	Diff (трудность)	Disc. Index / AdjPtBiserial	Тип проблемы	Рекомендация
1	0,2	-0,08	Слишком сложное, отрицательная дискриминация	Переработать формулировку или исключить
4	$\leq 0,2$	0,1–0,15	Слишком сложное	Смягчить формулировку
7	0,15–0,2	0,15	Низкая дискриминация	Переработать ключ или дистракторы
8	0,2–0,3	0,12	Низкая дискриминация	Усилить дистракторы
9	$\leq 0,2$	0,2	Слишком сложное	Переработать
10	$\leq 0,2$	-0,02	Слишком сложное, отрицательная дискриминация	Переработать формулировку

№ задания	Diff (трудность)	Disc. Index / AdjPtBiserial	Тип проблемы	Рекомендация
13	$\leq 0,2$	-0,11	Слишком сложное, отрицательная дискриминация	Переработать формулировку
14	$\leq 0,2$	0,1	Слишком сложное	Смягчить формулировку
22	$\leq 0,2$	0,2	Слишком сложное	Переработать
24	0,3–0,4	0,08	Низкая дискриминация	Переработать ключ или дистракторы
25	0,3–0,4	0,02	Низкая дискриминация	Переработать ключ или дистракторы
29	0,893	0,3	Слишком лёгкое	Проверить варианты ответа
35	$\leq 0,2$	0,15	Слишком сложное	Переработать
37	0,2–0,3	0,00	Низкая/отрицательная дискриминация	Переработать ключ или дистракторы
43	$\leq 0,2$	0,1	Слишком сложное	Смягчить формулировку
44	$\leq 0,2$	0,04	Слишком сложное, низкая дискриминация	Переработать формулировку/дистракторы
46	$\leq 0,2$	-0,05	Слишком сложное, низкая дискриминация	Переработать формулировку/дистракторы
50	0,3–0,4	0,04	Низкая дискриминация	Переработать ключ или дистракторы

Примечания: задания с $Diff \leq 0,2$ классифицируются как слишком сложные; с $Diff \geq 0,8$ – как слишком лёгкие; $Disc. Index < 0$ или $AdjPtBiserial \leq 0$ указывают на отсутствие различительной способности.

В таблице 3 проблемные задания можно разделить на несколько категорий. У заданий № 1, 10, 13, 37 и 46 зафиксирована отрицательная или нулевая дискриминация ($Disc. Index \leq 0$), что свидетельствует об их полной непригодности для дифференциации педагогов по уровню подготовки – сильные участники отвечали на них хуже слабых. Задания № 7, 8, 24, 25, 44 и 50 имеют крайне низкую дискриминативность ($AdjPtBiserial \leq 0,1$), что указывает на неэффективность ключей или дистракторов.

Отдельную группу составляют задания с неоптимальной трудностью. Одиннадцать заданий (№ 1, 4, 9, 10, 13, 14, 22, 35, 43, 44, 46) оказались избыточно сложными ($Diff \leq 0,2$) – с ними справились менее 20% участников. Задание №29, напротив, оказалось чрезмерно легким ($Diff = 0,893$), что снижает его диагностическую ценность.

На основе проведенного анализа рекомендуется провести экспертизу заданий с отрицательной дискриминацией, переработать задания с низкой дискриминацией с усилением дистракторов, скорректировать формулировки слишком сложных заданий, а также исключить или заменить чрезмерно легкое задание №29.

Результаты анализа КТТ подтверждают умеренную надежность теста и приемлемое качество большинства заданий. Однако наличие отдельных заданий с низкими психометрическими характеристиками требует доработки заданий и последующей повторной проверки для повышения надежности и информативности теста.

Анализ в современной теории тестов (IRT)

Для более точной оценки качества заданий и проверки их соответствия единой шкале способностей был проведен анализ в рамках современной теории тестирования (IRT) с применением модели Раша (программа Winsteps). Основные результаты представлены в таблице 4.

Таблица 4 – Основные результаты анализа по модели Раша

Показатель	Вариант 419358
Количество респондентов	112
Надежность (альфа-Кронбаха)	0,65
Надежность Person Reliability	0,64
Разделяющая способность Person Separation	1,34
Реальная стандартная ошибка (Real RMSE)	0,33
Средняя способность участников (Mean Measure Persons)	-0,34
Стандартное отклонение SD Measure (Persons)	0,56
Число заданий и номера с $\text{Corr} < 0$	3 (№1,13,37)
Число заданий и номера с $\text{Corr} < 0.1$	4 (№18,10,25,37)
Число заданий и номера с $\text{Infit MNSQ} < 0.7$	0
Число заданий и номера с $\text{Infit MNSQ} > 1.3$	0
Число заданий и номера $\text{Outfit MNSQ} < 0.7$	1 (№ 35)
Число заданий и номера $\text{Outfit MNSQ} > 1.3$	5 (№ 10,18,25,37,44)

Как следует из таблицы 4, тест обладает умеренной надежностью ($\alpha = 0,65$; Person Reliability = 0,64) и средней разделяющей способностью (Person Separation = 1,34), что позволяет дифференцировать испытуемых, но лишь в ограниченном числе уровней. Средняя мера способности участников (-0,34 логита) при стандартном отклонении 0,56 указывает на относительно однородную по подготовленности выборку.

Несмотря на приемлемые общие показатели, выявлены существенные отклонения на уровне отдельных заданий. Три задания (№ 1, 13, 37) демонстрируют отрицательную корреляцию с общей мерой способности ($\text{Corr} < 0$), а четыре задания (№ 10, 18, 25, 37) – крайне низкую ($\text{Corr} < 0,1$), что свидетельствует об их несоответствии измеряемому конструкту. Анализ Fit-статистик показал, что пять заданий имеют завышенный Outfit MNSQ ($> 1,3$), что указывает на наличие непредсказуемых ответов (№ 10, 18, 25, 37, 44), а одно задание (№ 35) – заниженный Outfit MNSQ ($< 0,7$), что указывает на избыточную предсказуемость и возможную тривиальность. Полученные данные подтверждают необходимость содержательной ревизии указанных заданий.

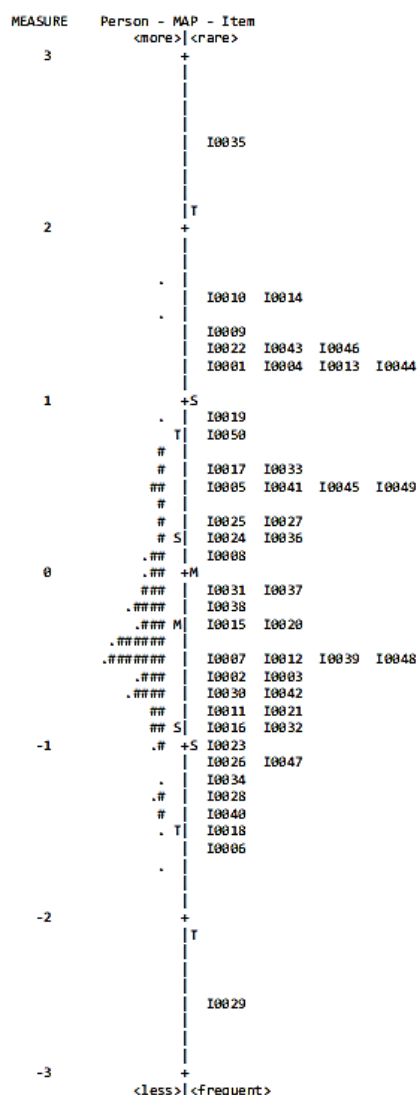


Рисунок 1 – Карта (Wright Map) «Person-Item»
Построено с применением модели Раша в программе Winsteps.

Анализ карты Wright Map показывает, что распределение большинства тестовых заданий соответствует уровню подготовки основной группы педагогов (диапазон –1 до 0 логитов) как показано на Рисунке 1. При этом выявлены экстремальные задания: чрезмерно сложное задание №35 (I0035) и чрезмерно лёгкое задание №29 (I0029), которые находятся за пределами типичного диапазона подготовки участников. Отмечается также недостаток заданий повышенной трудности в верхней части шкалы, что ограничивает возможность дифференциации наиболее подготовленных педагогов. В целом тест наиболее эффективно различает участников со средним уровнем подготовки. Для повышения диагностической ценности рекомендуется расширить банк сложных заданий и пересмотреть чрезмерно лёгкие задания для улучшения информативности в нижней части шкалы [17].

К ограничениям исследования следует отнести то, что анализ проводился на основе одного варианта теста и выборки, включающей 112 педагогов. Несмотря на то, что данный объем является достаточным для первичной валидации инструмента и выявления ключевых направлений его доработки, полученные результаты требуют дальнейшего подтверждения на более широкой выборке. Включение дополнительных вариантов теста и расширение числа участников исследования позволит уточнить психометрические характеристики инструмента и повысить надежность и обобщаемость сделанных выводов.

Рекомендуется:

- Провести экспертную проверку всех заданий с отклонениями от модели Раша (Infit/Outfit MNSQ за пределами 0.7–1.3), а также заданий с низкой или отрицательной корреляцией. При необходимости переформулировать или исключить такие задания. Особое внимание следует уделить заданиям с отрицательной корреляцией, поскольку они работают против общей оценки теста.

- Усилить контроль за формулировкой заданий на этапе разработки и совершенствовать процедуру содержательной экспертизы, что позволит предотвратить появление заданий низкого качества в дальнейшем.

- Ввести механизм обязательной апробации всех заданий перед их основным использованием в аттестации для предварительного выявления психометрических характеристик и внесения необходимых корректировок до официального применения.

- После корректировки или исключения проблемных заданий провести повторный Rasch-анализ для оценки достигнутых улучшений и подтверждения соответствия теста требованиям качества, надежности и справедливости.

Практические рекомендации также включают переработку заданий с отрицательной и низкой дискриминацией, усиление качества дистракторов, корректировку формулировок избыточно сложных заданий, а также добавление элементов повышенной трудности. Реализация данных мер позволит повысить надежность теста до требуемого уровня и обеспечить объективность аттестации педагогов.

Практическая значимость исследования заключается в возможности использования полученных результатов для совершенствования системы аттестации педагогов технического и профессионального образования Республики Казахстан. Выявленные психометрические характеристики тестовых заданий могут быть использованы НЦТ при разработке, экспертизе и апробации тестовых материалов, а также при повышении объективности и надежности процедур оценки знаний педагогов. Результаты исследования способствуют совершенствованию качества тестового банка и обеспечению справедливости аттестационных процедур.

Заключение

Проведенное исследование было направлено на психометрическое обоснование валидности теста по информатике, используемого в рамках оценки знаний педагогов технического и профессионального образования Республики Казахстан. Применение комплекса методов классической теории тестирования и современной теории ответов на задания (модель Раша) позволило получить многомерную характеристику качества тестового инструмента.

Полученные данные позволили сформулировать конкретные направления доработки теста: переработка заданий с отрицательной и низкой дискриминацией, усиление дистракторов, корректировка формулировок избыточно сложных элементов, добавление заданий повышенной трудности для улучшения дифференциации сильных педагогов, а также введение обязательного этапа апробации всех заданий перед основным использованием.

Несмотря на выявленные недостатки, тест демонстрирует содержательную валидность и принципиальную пригодность для оценки профессиональной компетентности педагогов ТиПО после внесения предложенных корректировок. Реализация рекомендаций и последующий повторный Rasch-анализ позволят повысить надежность инструмента до уровня, требуемого для высокостратных аттестационных процедур ($KR-20 \geq 0,8$), и обеспечить объективность и справедливость оценки. Дальнейшие исследования предполагают расширение выборки, включение других вариантов теста и углубленный анализ дифференциального функционирования заданий для различных групп педагогов.

В заключении следует отметить, что Перспективы дальнейших исследований связаны с расширением банка тестовых заданий, анализом дифференциального функционирования заданий (DIF) в различных группах педагогов, а также последующей психометрической валидацией обновленного инструментария на основе Rasch-модели, включая проверку параметрической стабильности и инвариантности измерений.

Вклад авторов

Альшурина Ш.Ш. – руководила исследованием, разработала концепцию статьи, выполнила анализ и интерпретацию результатов, подготовила окончательный текст статьи.

Дидарбекова Н.А. – участвовала в сборе и обработке данных, проведении психометрического анализа, интерпретации результатов исследования.

Оспанов Д.С. – осуществил психометрический анализ данных, включая расчёты по классической теории тестирования (КТТ) и модели Раша (IRT), а также подготовку результатов статистической обработки.

Кулдеева А.А. – участвовала в сборе данных, оформлении результатов исследования и подготовке рукописи.

ЛИТЕРАТУРА:

1 Приказ Министра просвещения Республики Казахстан от 25 февраля 2025 г. № 32 «О внесении изменений в приказ Министра образования и науки Республики Казахстан от 27 января 2016 г. № 83 “Об утверждении Правил и условий проведения аттестации педагогов”» // Эділет: правовой интернет-ресурс. URL: https://online.zakon.kz/Document/?doc_id=39067173

2 Мессик С. Валидность психологической оценки: валидация выводов на основе ответов и результатов испытуемых как научное исследование значения тестовых баллов // *American Psychologist*. – 1995. – Т. 50, № 9. – С. 741–749.

3 American Educational Research Association; American Psychological Association; National Council on Measurement in Education. Standards for Educational and Psychological Testing. – Washington, D.C.: American Educational Research Association, 2014. – 230 p.

4 Methodology for Assessing the Level of Methodological Competence of IT Teachers // *Scientific Journal of Astana IT University*. – 2024. – Vol. 18. – P. 120–130. URL: https://sj.astanait.edu.kz/wp-content/uploads/2024/07/Journal_AITU_18vol_june2402-120-130.pdf

5 Bitemirova S., Zholdasbekova S., Mussakulov K., Anesova A., Zhanbirshiyev S. Pre-service TVET Teachers' Digital Competence: Evidence from Survey Data // *TEM Journal*. – 2023. – Vol. 12, No. 2. – P. 1182–1189.

6 Abd Wahab N., Ali D. F., Ahmad A.R. Professional Development for Novice TVET Teachers: A Systematic Review // *International Journal of Academic Research in Progressive Education and Development*. – 2024. – Vol. 13, No. 4. – P. 2669–2692. URL: https://hrmars.com/papers_submitted/23654/professional-development-for-novice-tvet-teachers-a-systematic-review.pdf

7 Пенькина С.А. Сравнительный анализ документов, регламентирующих аттестацию педагогов в Республике Казахстан // *Вестник Карагандинского университета. Серия Педагогика*. – 2021. – Т. 104, № 4. – С. 68–74. URL: <https://pedagogy-vestnik.buketov.edu.kz/index.php/pedagogy-vestnik/article/view/197>

8 Булатбаева А.А., Кожаева А.А. Трансформация в Казахстане положительного опыта системы сертификации педагогических кадров в Сингапуре и Финляндии // *Вестник образовательных наук*. – 2022. – Т. 70, № 1. – С. 60–70. URL: <https://bulletin-pedagogic-sc.kaznu.kz/index.php/1-ped/article/view/1196> (дата обращения: 24.12.2025).

9 Серикбаева Ж. Изучение отношения учителей к аттестации педагогов в одной отборочной и одной массовой школе Казахстана : магистерская диссертация. – Астана : Высшая школа образования Назарбаев Университета, 2025. URL: <https://nur.nu.edu.kz/handle/123456789/8547>

10 Алгина Дж., Пенфилд Р.Д. Классическая теория тестирования // Руководство СЕЙДЖ по количественным методам в психологии / под ред. Р. Э. Милсап, А. Мадейу-Оливарес. – Лондон : SAGE Publications, 2009. – С. 93–122. URL: <https://doi.org/10.4135/9780857020994.n5>

11 Райт Б.Д., Мастерс Г.Н. Анализ шкал оценивания. – Чикаго : MESA Press, 1982. – URL: <https://www.rasch.org/rmt/rmt73g.htm> (дата обращения: 16.03.2026).

12 Райт Б.Д., Стоун М.Х. Лучший дизайн теста. – Чикаго : MESA Press, 1979. URL: <https://www.rasch.org/btd.htm>

13 Айяла Р.Х. Теория ответов на задания и моделирование Раша // Справочник рецензента по количественным методам в социальных науках / под ред. Г. Р. Хэнкока, Л. М. Стэплтон, Р. О. Мюллера. – 2-е изд. – Нью-Йорк : Routledge, 2018. – С. 145–163. URL: <https://doi.org/10.4324/9781315755649>

14 Бонд Т.Г., Фокс К.М. Применение модели Раша: фундаментальные измерения в гуманитарных науках. – Routledge, 2015.

15 Бун В.Дж., Ставер Дж.Р., Йейл М.С. Rasch-анализ в гуманитарных науках. – Springer, 2014.

16 Эмбертсон С.Е., Райз С.П. Теория ответов на задания для психологов. – Psychology Press, 2000.

17 Линеик Дж.М. Winsteps®: руководство пользователя программы Rasch-измерений. – Winsteps.com, 2022.

REFERENCES:

- 1 Prikaz Ministra prosvešenia Respubliki Kazakhstan ot 25 fevralä 2025 g. № 32 «O vnesenii izmeneni v prikaz Ministra obrazovania i nauki Respubliki Kazakhstan ot 27 ianvarä 2016 g. № 83 “Ob utverjdenii Pravil i uslovi provedenia atestatii pedagogov”» [Order of the Minister of Education of the Republic of Kazakhstan No. 32 dated February 25, 2025 “On Amendments to Order No. 83 dated January 27, 2016 ‘On Approval of the Rules and Conditions for Teacher Certification’”]. Ädilet: pravovoi internet-resurs. URL: https://online.zakon.kz/Document/?doc_id=39067173 (in Russian).
- 2 Messick S. (1995). Validnost psihologicheskoi osenki: validasia vyvodov na osnove otvetov i rezülatov ispytuemykh kak nauchnoe issledovanie znacheniia testovykh balov [Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning]. American Psychologist, vol. 50, no. 9, pp. 741–749. (in Russian).
- 3 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). Standards for Educational and Psychological Testing. Washington, D.C.: American Educational Research Association, 230 p.
- 4 Methodology for Assessing the Level of Methodological Competence of IT Teachers. (2024). Scientific Journal of Astana IT University, vol. 18, pp. 120–130. URL: https://sj.astanait.edu.kz/wp-content/uploads/2024/07/Journal_AITU_18vol_june2402-120-130.pdf5
- 5 Bitemirova S., Zholdasbekova S., Mussakulov K., Anesova A., Zhanbirshiyev S. (2023). Pre-service TVET Teachers' Digital Competence: Evidence from Survey Data. TEM Journal, vol. 12, no. 2, pp. 1182–1189.
- 6 Abd Wahab N., Ali D.F., Ahmad A.R. (2024). Professional Development for Novice TVET Teachers: A Systematic Review. International Journal of Academic Research in Progressive Education and Development, vol. 13, no. 4, pp. 2669–2692. URL: https://hrmars.com/papers_submitted/23654/professional-development-for-novice-tvet-teachers-a-systematic-review.pdf
- 7 Penkina S.A. (2021). Sravnitelnyi analiz dokumentov, reglamentiruiushih atestasiu pedagogov v Respublike Kazakhstan [Comparative analysis of documents regulating teacher certification in the Republic of Kazakhstan]. Vestnik Karagandinskogo universiteta. Seriya Pedagogika, vol. 104, no. 4, pp. 68–74. URL: <https://pedagogy-vestnik.buketov.edu.kz/index.php/pedagogy-vestnik/article/view/197> (in Russian).
- 8 Bulatbaeva A.A., Kozhaeva A.A. (2022). Transformasiia v Kazahstane polojitel'nogo opyta sistemy sertifikatsii pedagogicheskikh kadrov v Singapore i Finlãndii // Vestnik obrazovatel'nykh nauk [Transformation in Kazakhstan of the positive experience of teacher certification systems in Singapore and Finland]. Vestnik obrazovatel'nykh nauk, vol. 70, no. 1, pp. 60–70. URL: <https://bulletin-pedagogic-sc.kaznu.kz/index.php/1-ped/article/view/1196> (in Russian).
- 9 Serikbaeva Zh. (2025). İzuchenie otnošenii uchitelei k atestatsii pedagogov v odnoi otborochnoi i odnoi masovoi škole Kazahstana : magisterskaia disertatsia [Exploring teachers' attitudes toward teacher certification in a selective and a mainstream school in Kazakhstan]. Master's thesis. Astana: Nazarbayev University Graduate School of Education. URL: <https://nur.nu.edu.kz/handle/123456789/8547> (in Russian).
- 10 Algina J., Penfield R.D. (2009). Klasicheskaiia teoriia testirovaniia [Classical Test Theory]. In: Millsap R.E., Maydeu-Olivares A. (eds.). Sage Handbook of Quantitative Methods in Psychology. London: SAGE Publications, pp. 93–122. Available at: <https://doi.org/10.4135/9780857020994.n5> (in Russian).
- 11 Wright B.D., Masters G.N. (1982). Analiz škala osenivaniia [Rating Scale Analysis]. Chicago: MESA Press. URL: <https://www.rasch.org/rmt/rmt73g.htm> (in Russian).
- 12 Wright B.D., Stone M.H. (1979). Luchši dizain testa [Best Test Design]. Chicago: MESA Press. URL: <https://www.rasch.org/btd.htm> (in Russian).
- 13 Ayala R.J. (2018). Teoria otvetov na zadania i modelirovanie Raša [Item Response Theory and Rasch Modeling]. In: Hancock G.R., Stapleton L.M., Mueller R.O. (eds.). Reviewer's Guide to Quantitative Methods in the Social Sciences. 2nd ed. New York: Routledge, pp. 145–163. DOI: <https://doi.org/10.4324/9781315755649> (in Russian).
- 14 Bond T.G., Fox C.M. (2015). Primenenie modeli Raša: fundamentãlnye izmereniia v gumanitarnykh naukakh [Applying the Rasch Model: Fundamental Measurement in the Human Sciences]. Routledge. (in Russian).
- 15 Boone W.J., Staver J.R., Yale M.S. (2014). Rasch-analiz v gumanitarnykh naukakh [Rasch analysis in the humanities]. Springer.
- 16 Embretson S.E., Reise S.P. (2000). Teoria otvetov na zadania dlã psihologov [Item Response Theory for Psychologists]. Psychology Press.
- 17 Linacre J.M. (2022). Winsteps®: rukovodstvo pölzovatelã programy Rasch-izmereni [Winsteps®: Rasch Measurement Software User Guide]. Winsteps.com.

ИНФОРМАТИКА БОЙЫНША ТЕХНИКАЛЫҚ ЖӘНЕ КӘСІПТІК БІЛІМ БЕРУ ПЕДАГОГТЕРІНЕ АРНАЛҒАН ТЕСТТІҢ ПСИХОМЕТРИЯЛЫҚ ТАЛДАУЫ

Ш.Ш. Альшурина*, Н.А. Дидарбекова, Д.С. Оспанов, А.А. Кулдеева

ҚР ҒЖБМ «Ұлттық тестілеу орталығы», Қазақстан Республикасы, Астана қ.

*e-mail: shinara.a@mail.ru, nauzhan.didarbekova@mail.ru,
dimokospan@gmail.com, lunochka_90@mail.ru

Аңдатпа. Мақалада Қазақстан Республикасында техникалық және кәсіптік білім беру педагогтерінің білімін бағалауда (ТЖКББ ПББ) қолданылатын информатика тестін валидациялау нәтижелері ұсынылған. Аттестаттаудың нормативтік-құқықтық негіздері, біліктілік санаттарының құрылымы және тестілеу процедурасы қарастырылған. Эмпирикалық базаны 50 дихотомиялық тапсырманы орындаған 112 педагогтің деректері құрады. Психометриялық талдау классикалық тестілеу теориясы (КТТ) және Раш моделі (IRT) қолданылып жүргізілді. Тесттің орташа сенімділігі ($KR-20 = 0,655$; $\alpha = 0,65$) және тапсырмалардың орташа қиындығы (0,434) анықталды. Шамадан тыс қиын (11 тапсырма, $p \leq 0,2$), шамадан тыс жеңіл (1 тапсырма, $p = 0,89$), теріс дискриминациясы (4 тапсырма) және төмен дискриминациясы (6 тапсырма) бар тапсырмалар айқындалды. IRT талдауы бірқатар тапсырмалардың модельге сәйкес келмейтінін растады: жалпы қабілет өлшемімен теріс корреляция (№ 1, 13, 37), Fit-статистика ауытқулары (Outfit MNSQ $>1,3$ № 10 үшін; $<0,7$ № 35 үшін). Райт картасы жоғары қиындықтағы тапсырмалардың жетіспеушілігін көрсетті. Талдау негізінде проблемалық тапсырмаларды қайта өңдеу, дистракторларды күшейту және қайта апробациялау бойынша ұсыныстар берілген. Тест мазмұндық валидтілікті көрсетеді және ұсынылған түзетулер енгізілгеннен кейін педагогтердің кәсіби құзыреттілігін объективті бағалауға жарамды деген қорытынды жасалды.

Кілт сөздер: тестілік валидация, ТЖКБ педагогтері, информатика, классикалық тестілеу теориясы, Раш моделі, сенімділік, дискриминативтілік.

SYCHOMETRIC ANALYSIS OF INFORMATICS TEST FOR TEACHERS OF TECHNICAL AND VOCATIONAL EDUCATION

Sh. Alshurina*, N. Didarbekov, D. Ospanov, A. Kuldeeva

National Testing Center of the MSHE, Republic of Kazakhstan, Astana

*e-mail: shinara.a@mail.ru, nauzhan.didarbekova@mail.ru,
dimokospan@gmail.com, lunochka_90@mail.ru

Abstract. This article presents the results of the validation of an informatics test used for the knowledge assessment of Technical and Vocational Education and Training (TVET) teachers in the Republic of Kazakhstan. The legal and regulatory framework for teacher attestation, the structure of qualification categories, and the testing procedure are examined. The empirical base consisted of data from 112 teachers who completed 50 dichotomous items. Psychometric analysis was conducted using Classical Test Theory (CTT) and the Rasch model (IRT). Moderate test reliability ($KR-20 = 0.655$; $\alpha = 0.65$) and average item difficulty (0.434) were established. Items with excessive difficulty (11 items, $p \leq 0.2$), excessive ease (1 item, $p = 0.89$), negative discrimination (4 items), and low discrimination (6 items) were identified. IRT analysis confirmed the misfit of several items to the model: negative correlation with the overall ability measure (No. 1, 13, 37), deviations in fit statistics (Outfit MNSQ >1.3 for No. 10; <0.7 for No. 35). The Wright Map revealed a lack of items with high difficulty. Based on the analysis, recommendations are proposed for revising problematic items, strengthening distractors, and conducting re-approbation. It is concluded that the test demonstrates content validity and is suitable for objectively assessing teacher competence after implementing the proposed corrections.

Keywords: test validation, TVET teachers, informatics, Classical Test Theory, Rasch model, reliability, discrimination.

Сведения об авторах:

Альшурина Шынар Шименбаевна* – РГП на ПХВ «Национальный центр тестирования» МНВО РК, главный эксперт Научно-методической лаборатории по формированию тестовых заданий для оценки квалификации Педагогов технического и профессионального образования, Астана, Казахстан, e-mail: shinara.a@mail.ru, ORCID: <https://orcid.org/0009-0009-8254-3271>).

Дидарбекова Наужан Абдисатаровна – кандидат филологических наук, заместитель директора Национального центра тестирования МНВО РК (Казахстан Астана г., e-mail: nauzhan.didarbekova@mail.ru, ORCID: <https://orcid.org/0000-0002-7085-324X>).

Оспанов Динмухамед Сакенович – РГП на ПХВ «Национальный центр тестирования» МНВО РК, старший научный сотрудник лаборатории научных исследований и психометрики (Казахстан Астана г., e-mail: dimokospan@gmail.com, ORCID: <https://orcid.org/0009-0006-4418-0770>).

Кулдеева Айнур Алихановна – магистр, РГП на ПХВ «Национальный центр тестирования» МНВО РК, ведущий эксперт научно-методической лаборатории по формированию тестовых заданий для оценки квалификации Педагогов технического и профессионального образования (Казахстан Астана г., e-mail: lunochka_90@mail.ru, ORCID: <https://orcid.org/0009-0000-4970-3549>).

Авторлар туралы мәліметтер:

Альшурина Шынар Шименбаевна* – ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Техникалық және кәсіптік білім берудегі Педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру ҒӨЗ бас сарапшысы (Қазақстан, Астана қ., e-mail: shinara.a@mail.ru, ORCID: <https://orcid.org/0009-0009-8254-3271>).

Дидарбекова Наужан Абдисатаровна – филология ғылымдарының кандидаты, ҚР ҒЖБМ Ұлттық тестілеу орталығы директорының орынбасары (Қазақстан, Астана қ., e-mail: nauzhan.didarbekova@mail.ru, ORCID: <https://orcid.org/0000-0002-7085-324X>).

Оспанов Динмухамед Сакенович – ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Ғылыми зерттеулер мен психометрика зертханасының аға ғылыми қызметкері (Қазақстан, Астана қ., e-mail: dimokospan@gmail.com, ORCID: <https://orcid.org/0009-0006-4418-0770>).

Кулдеева Айнур Алихановна – магистр, ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Техникалық және кәсіптік білім берудегі Педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру ҒӨЗ жетекші сарапшысы (Қазақстан, Астана қ., e-mail: lunochka_90@mail.ru, ORCID: <https://orcid.org/0009-0000-4970-3549>).

Information about authors:

Alshurina Shynar* – State-Owned Enterprise “National Testing Center” under the Ministry of Education and Science of the Republic of Kazakhstan, Chief Expert at the Scientific and Methodological Laboratory for the Development of Test Items to Assess the Qualifications of Technical and Vocational Education Teachers (Kazakhstan, Astana, e-mail: shinara.a@mail.ru, ORCID: <https://orcid.org/0009-0009-8254-3271>).

Didarbekova Nauzhan – Candidate of Philological Sciences, Deputy Director of the National Testing Center of the Ministry of National Education and Science of the Republic of Kazakhstan (Kazakhstan, Astana, e-mail: nauzhan.didarbekova@mail.ru, ORCID: <https://orcid.org/0000-0002-7085-324X>).

Ospanov Dinmukhamed – State Enterprise “National Testing Center” under the Ministry of National Education and Science of the Republic of Kazakhstan, Senior Researcher at the Laboratory of Scientific Research and Psychometrics (Kazakhstan, Astana, e-mail: dimokospan@gmail.com, ORCID: <https://orcid.org/0009-0006-4418-0770>).

Kuldeeva Ainur – Master’s, State Joint-Stock Company “National Testing Center” under the Ministry of Education and Science of the Republic of Kazakhstan, Lead Expert at the Scientific and Methodological Laboratory for the Development of Test Items to Assess the Qualifications of Technical and Vocational Education Teachers (Kazakhstan, Astana, e-mail: lunochka_90@mail.ru, ORCID: <https://orcid.org/0009-0000-4970-3549>).