ҒТАМР 14.35.09 https://doi.org/10.63597/UTO3105-4161.2026.5.1.005

Ауезханова А.Ж.[1*], Ерсултанова Т.С.[2], Қуандық А.Қ.[3], Әлен С.Ә.[4]
[1,2,3,4]ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Астана қ., Қазақстан Республикасы
*email: auezkhanova72@gmail.com

[1]ORCID 0009-0005-9335-9593, [2]ORCID 0000-0003-3286-9096,
[3]ORCID 0009-0001-4370-4636, [4]ORCID 0009-0009-2079-4331

2025 ЖЫЛЫ ПББ АПРОБАЦИЯСЫ АЯСЫНДА ӨТКІЗІЛГЕН ТЕСТІЛЕУ НӘТИЖЕЛЕРІНІҢ ПСИХОМЕТРИЯЛЫҚ ТАЛДАУЫ

Бұл зерттеу Қазақстан Республикасында педагогтердің білімін бағалауда қолданылатын физика пәні тест тапсырмаларының психометриялық сипаттамаларына талдауға арналған. Зерттеу 2025 жылы өткізілген педагогтердің білімін бағалау (ПББ) апробациясының нәтижелеріне негізделген. Талдау барысында Раш өлшеу моделі қолданып, тесттің сенімділігі, валидтілігі және дифференциациялық қабілеті бағаланды. Жалпы апробацияға 6000-нан астам педагог қатысқанымен, талдау үшін 118 қатысушыдан тұратын репрезентативті іріктеме пайдаланылды. Зерттеу нәтижелері тесттің жоғары сенімділігін көрсетті (Item Reliability = 0,99), тапсырмалардың 92%-ы валидтілік критерийлеріне сәйкес келді (Infit/Outfit MNSQ 0,5–1,5 аралығында). Сонымен қатар, тапсырмалардың қиындық деңгейі логит шкаласында бағаланып, дистракторлық талдау жүргізілді. Кейбір тапсырмаларда сәйкессіздік байқалып, қайта қарау ұсынылады. Тестіленушілер мен тапсырмалардың қабілеті мен күрделілігі арасындағы сәйкестік тесттің білім деңгейін кең ауқымда бағалауға жарамды екенін көрсетті. 2022 жылдан бастап енгізілген тақырыптық талдау педагогтерге өз нәтижелерін терең түсінуге мүмкіндік береді. Зерттеу нәтижелері қолданылған тесттің педагогтердің пәндік білім деңгейін объективті бағалауға жарамды екенін дәлелдейді және тест тапсырмаларының сапасын одан әрі жетілдіруге бағытталған ұсыныстар ұсынады.

Түйін сөздер: педагогтердің білімін бағалау, психометриялық талдау, тест, Раш моделі, сенімділік, валидтілік.

Кіріспе

ПББ тарихы және зерттеудің өзектілігі. Қазақстан Республикасында педагогтердің білімін бағалау жүйесі кезең-кезеңімен жетілдіріліп, нормативтік-құқықтық және технологиялық тұрғыдан жүйелі жаңғыртудан өтті. Бұл өзгерістер бағалау рәсімдерінің ашықтығын, объективтілігін және сенімділігін қамтамасыз етуге бағытталды.

2010 жылдан бастап педагогтердің білімін бағалау мазмұндық тұрғыдан жаңартылған кезеңі басталды. Осы кезеңде біліктілікті растаудың негізгі құралы ретінде педагог қызметкерлерді кешенді тестілеу енгізілді. Бұл жүйе Қазақстан Республикасы Білім және ғылым министрінің 2010 жылғы 22 қаңтардағы №16 «Педагог қызметкерлерді аттестаттау қағидаларын бекіту туралы» бұйрығы [1] негізінде жүзеге асырылды. Аталған құжат мектепке дейінгі тәрбие мен оқыту, бастауыш, негізгі орта және жалпы орта білім беру, техникалық және кәсіптік, орта білімнен кейінгі, қосымша білім беру бағдарламаларын іске асыратын білім беру ұйымдарының педагогтеріне қатысты болды.

2018 жылдан бастап педагог қызметкерлердің білімін бағалау мазмұны мен құрылымы қайта қарастырылып, тестілеудің жаңа форматы – Ұлттық біліктілік тестілеуі енгізілді. Осы жылы тестілеу қағаз форматында өткізілсе, 2019 жылдан бастан тестілеу электрондық форматқа көшірілді. Аталған өзгеріс бағалау үдерісінің ашықтығын, жеделдігін және объективтілігін арттыруға мүмкіндік берді, сондай-ақ адами фактордың ықпалын төмендетуге ықпал етті.

2020 жылдан бастап бүгінгі күнге дейін Ұлттық біліктілік тестілеуді арнайы бағдарламалық қамтамасыз ету арқылы жүргізіліп келеді. Жаңа жүйе тестілеу рәсімін толық автоматтандыруға, нәтижелердің дәлдігі мен сенімділігін қамтамасыз етуге бағытталған.

2022 жылдан бастап Педагогтердің білімін бағалау (ПББ) форматы енгізілді [2]. Тестілеу педагогтердің пәндік білімдерін бағалауға бағытталған. Максималды балл саны - 50. Тапсырмалар қиындық деңгейі бойынша үш деңгейге бөлінеді: базалық (26%), орташа (60%) және жоғары (14%). Тест 50 тапсырмадан тұрады. Тестілеуден өту үшін педагог кемінде барлық тапсырманың 50 %-на дұрыс жауап беруі қажет.

Педагогтердің біліктілік санатына сәйкес шекті балл деңгейлері белгіленген [2]:

«педагог» / «педагог-тағылымдамашы» – кемінде 50%;

«педагог-модератор» біліктілік санаты – кемінде 60%;

«педагог-сарапшы» біліктілік санаты – кемінде 70%;

«педагог-зерттеуші» біліктілік санаты – кемінде 80%;

«педагог-шебер» біліктілік санаты – кемінде 90%.

Осыған байланысты тест тапсырмаларының сапасын ғылыми негізде бағалау өзекті болып табылады. Зерттеудің мақсаты – физика пәні бойынша тест тапсырмаларының сапасын Раш моделі негізінде психометриялық талдау.

Халықаралық тәжірибеде педагогтердің білімін бағалаудың білім сапасына тікелей әсер ететінін көрсетеді [3]. АҚШ-та педагогтердің білімін бағалау үшін Praxis емтихандары пәндік білім мен педагогикалық құзыреттілікті кешенді түрде өлшейді [4], ал edTPA моделі мұғалімнің практикалық оқыту дағдыларын бағалауға бағытталған [5].

Ұлыбританияда мұғалімдер Qualified Teacher Status (QTS) алу үшін кәсіби стандарттарға сәйкестігін дәлелдеуі тиіс [5]. Оңтүстік Корея мен Қытайда педагогтерді іріктеу орталықтандырылған емтихандар арқылы жүзеге асырылады [3].

Финляндияда формалды тестілеу жүйесі қолданылмайды, алайда мұғалімдерді даярлау үдерісі жоғары талаптары мен қатаң іріктеуге негізделген [5]. TALIS халықаралық зерттеуі педагогтердің кәсіби дамуы мен оқыту тәжірибесін салыстырмалы түрде талдауға мүмкіндік береді [3].

Бұл тәжірибелер педагогтерді бағалау жүйесі ғылыми негізделген, сенімді және көпдеңгейлі болуы қажет екенін көрсетеді [6, 7].

Зерттеу материалдары мен әдістері

Зерттеу дизайны. Бұл зерттеу сандық әдіснамаға негізделген және тест тапсырмаларының психометриялық сипаттамаларын талдауға бағытталған. Негізгі талдау құралы ретінде Раш өлшеу моделі (кесте 1) қолданылды [6, 8]. Бұл модель тест тапсырмаларының қиындық деңгейін және тестіленушілердің қабілетін бір логиттік шкалада бағалауға мүмкіндік береді [8]. Тест тапсырмаларының базасын қалыптастыру үдерісі бірнеше кезеңнен тұрады: әзірлеу, бірінші және екінші сараптама, сондай-ақ бірінші және екінші түзету кезеңдері [9]. Тапсырмаларды әзірлеушілер мен сарапшылар ретінде колледждер, орта мектептер және жоғары және (немесе) жоғары оқу орнынан кейінгі білім беру ұйымдары (ЖЖОКБҰ) ұсынған тәжірибелі педагогтері мен оқытушылары іріктелді. Бұл құрылым тесттің ғылыми негізделген және әділ болуын қамтамасыз етеді [9, 10].

Деректер және құралдар. Зерттеу 2025 жылғы ПББ апробациясы нәтижелеріне негізделген [9]. Жалпы қатысушылар саны 6000-нан асқанымен, осы зерттеуде 188 педагогтің деректері талданды.

Тест 50 тапсырмадан тұрады:

- 40 – бір дұрыс жауапты таңдау;
- 10 – контекстке негізделген тапсырмалар.

Максималды балл – 50. Тапсырмалар дихотомиялық форматта ұсынылған, бұл классикалық тест теориясы мен Item Response Theory талаптарына сәйкес келеді [7].

Деректерді талдау. Деректер Winsteps бағдарламасы арқылы өңделді. Келесі көрсеткіштер талданды:

- сенімділік;
- валидтілік (Infit, Outfit MNSQ);
- қиындық деңгейі (логит);
- корреляция;
- дифференциациялық қабілет.

Валидтілік Infit және Outfit MNSQ статистикалары арқылы бағаланды (қалыпты диапазон: 0,5–1,5) [6, 11]. Сонымен қатар, дистракторлық талдау жүргізіліп, жауап нұсқаларының тиімділігі бағаланды [10].

Әр тапсырма бір дұрыс жауаптан және бірнеше дистрактордан тұрады, тапсырмалар күрделілік деңгейі мен когнитивтік дағдыларға сәйкес әзірленген [10]. Сонымен қатар, тапсырмалар мен жалпы балл арасындағы корреляция есептелді, бұл тесттің педагогтердің қабілетін объективті бағалау қабілетін арттырды [12].

Нәтижелер және талқылау

Психометриялық талдау нәтижелері тесттің жоғары сапасын көрсетті. Тапсырмалардың орташа стандартты қатесі 0,23 логитті құрады, бұл өлшеу дәлдігінің жоғары екенін көрсетеді [6]. Тесттің ішкі сенімділігі Cronbach α = 0,83, стандартталған сенімділік 0,84, SEM = 2,98, яғни тест педагогтердің қабілетін сенімді бағалауға қабілетті [10, 12]. Тапсырмалардың 92%-ы валидті, қалғандарында сәйкессіздік байқалды. Тапсырмалардың қиындық деңгейі кең диапазонда таралған, бұл тесттің әртүрлі деңгейдегі педагогтерді ажыратуға қабілетті екенін көрсетеді [7, 8].

Дистракторлық талдау нәтижесінде жауап нұсқаларының тиімділігін көрсетті.

Кесте 1 - Раш моделіндегі қорытынды статистика

Көрсеткіш	Орташа мән (Mean)	Стандарттық қате (SEM)	Стандарттық ауытқу (SD)	Минимум	Максимум
Жалпы ұпай (Score)	23.0	0.7	7.2	9.0	47.0
Өлшем (Measure, логит)	-0.14	0.08	0.85	-1.98	3.32
Қате (S.E.)	0.34	0.00	0.04	0.32	0.61
Infit MNSQ	1.00	0.02	0.19	0.65	1.52
Outfit MNSQ	1.03	0.03	0.36	0.55	2.18
Infit ZSTD	-0.11	0.11	1.20	-2.95	2.85
Outfit ZSTD	0.00	0.11	1.16	-2.30	3.13

Кесте 1-де көрсетілгендей, 118 педагогтің орташа ұпайы 23,0 (максималды 50 ұпайдан), қабілет деңгейі (Measure) -0,14 логит, яғни тест тапсырмаларының орташа күрделілік деңгейі тестіленушілердің қабілет деңгейінен сәл жоғары екенін білдіреді [8]. Тестіленушілердің қабілет деңгейінің стандартты ауытқуы 0,85 логитті құрады, бұл іріктеме ішінде білім деңгейінің жеткілікті деңгейде әртүрлі екенін көрсетеді. Модельге сәйкестік көрсеткіштері қалыпты аралықта (Infit MNSQ 1,00; Outfit MNSQ 1,03), ZSTD нөлге жақын, бұл деректердің Раш моделіне жақсы сәйкес келетінін көрсетеді [6, 11].

Сенімділік көрсеткіші (Person Reliability) 0,82–0,84 аралығында болды, бұл тесттің педагогтердің қабілет деңгейін сенімді түрде ажыратуға қабілетті екенін көрсетеді [11]. Жалпы алғанда, Кесте 1 нәтижелері тесттің педагогтердің білім деңгейін бағалауда тиімді құрал екенін көрсетеді [6, 8], алайда тест құрылымын одан әрі жетілдіру мақсатында жекелеген көрсеткіштерге назар аудару ұсынылады.

Кесте 2 - Тапсырмалар деңгейіндегі психометриялық көрсеткіштер (Раш моделі)

TABLE 14.1 118-1-420174 ZOU296WS.TXT Dec 01 2025 16:18
INPUT: 118 Person 50 Item REPORTED: 118 Person 50 Item 2 CATS WINSTEPS 5.8.3.0

Person: REAL SEP.: 2.15 REL.: .82 ... Item: REAL SEP.: 5.36 REL.: .97

Item STATISTICS: ENTRY ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	93	118	-1.64	.24	.85	-1.17	.73	-1.30	.45	.27	78.8	79.1	I0001
2	23	118	1.45	.25	.90	-.66	.98	-.04	.42	.34	84.7	82.0	I0002
3	44	118	.43	.20	.91	-1.20	.85	-1.41	.46	.35	72.0	68.1	I0003
4	80	118	-1.01	.21	.95	-.64	.87	-.88	.38	.31	72.0	70.0	I0004
5	52	118	.10	.20	.98	-.28	.98	-.21	.37	.35	66.1	65.0	I0005
6	34	118	.87	.22	1.28	2.57	1.49	3.07	.01	.35	65.3	74.0	I0006
7	36	118	.77	.21	1.11	1.10	1.15	1.13	.23	.35	72.0	72.6	I0007
8	109	118	-2.90	.35	1.04	.25	1.04	.24	.12	.18	92.4	92.4	I0008
9	40	118	.60	.21	1.07	.81	1.12	1.01	.27	.35	69.5	70.1	I0009
10	93	118	-1.64	.24	.93	-.54	1.45	1.87	.31	.27	78.8	79.1	I0010
11	81	118	-1.05	.21	.87	-1.53	.82	-1.18	.45	.31	78.0	70.6	I0011
12	98	118	-1.94	.25	.90	-.59	.77	-.87	.37	.25	84.7	83.1	I0012
13	26	118	1.28	.24	1.18	1.32	1.33	1.66	.12	.34	78.0	79.8	I0013
14	22	118	1.52	.25	.99	-.03	1.20	.95	.30	.33	83.9	82.8	I0014
15	35	118	.82	.22	1.15	1.45	1.16	1.17	.19	.35	71.2	73.3	I0015
16	58	118	-.13	.20	.79	-3.64	.75	-2.89	.58	.35	81.4	64.2	I0016
17	68	118	-.52	.20	.93	-1.02	.90	-.92	.41	.33	69.5	65.6	I0017
18	47	118	.30	.20	1.20	2.72	1.35	3.18	.10	.35	61.0	66.7	I0018
19	111	118	-3.18	.40	.98	.04	.65	-.66	.23	.16	94.1	94.1	I0019
20	89	118	-1.43	.22	.80	-1.92	.66	-1.92	.53	.28	76.3	75.9	I0020
21	16	118	1.95	.29	1.03	.22	1.12	.50	.28	.32	87.3	87.3	I0021
22	54	118	.03	.20	1.08	1.23	1.09	.93	.26	.35	64.4	64.6	I0022
23	17	118	1.87	.28	1.07	.43	1.56	1.89	.16	.32	86.4	86.5	I0023
24	98	118	-1.94	.25	.90	-.63	.72	-1.08	.39	.25	83.1	83.1	I0024
25	22	118	1.52	.25	.99	.00	1.00	.07	.34	.33	82.2	82.8	I0025
26	49	118	.22	.20	1.10	1.49	1.17	1.71	.23	.35	59.3	66.0	I0026
27	46	118	.34	.20	.93	-.97	1.03	.30	.41	.35	70.3	67.1	I0027
28	24	118	1.39	.24	1.00	.03	1.29	1.41	.28	.34	84.7	81.3	I0028
29	55	118	-.01	.20	.94	-1.03	.90	-1.13	.43	.35	63.6	64.4	I0029
30	49	118	.22	.20	1.01	.17	1.03	.36	.34	.35	69.5	66.0	I0030
31	64	118	-.36	.20	.85	-2.57	.81	-1.97	.51	.34	74.6	64.8	I0031
32	79	118	-.97	.21	.93	-.87	.91	-.57	.39	.31	71.2	69.5	I0032
33	36	118	.77	.21	.98	-.23	.93	-.48	.39	.35	70.3	72.6	I0033
34	30	118	1.06	.23	.95	-.41	.86	-.86	.42	.35	77.1	76.9	I0034
35	56	118	-.05	.20	1.03	.53	1.00	.08	.32	.35	59.3	64.3	I0035
36	43	118	.47	.20	1.01	.12	1.01	.14	.34	.35	69.5	68.6	I0036
37	32	118	.96	.22	1.07	.69	1.30	1.85	.23	.35	75.4	75.5	I0037
38	112	118	-3.34	.42	.95	-.01	.65	-.59	.24	.15	94.9	94.9	I0038
39	18	118	1.79	.27	1.08	.50	1.36	1.37	.19	.32	84.7	85.7	I0039
40	54	118	.03	.20	1.05	.85	1.07	.77	.29	.35	64.4	64.6	I0040
41	101	118	-2.14	.27	.99	.02	.96	-.05	.24	.23	85.6	85.6	I0041
42	60	118	-.21	.20	.89	-1.92	.86	-1.56	.47	.34	72.0	64.3	I0042
43	51	118	.14	.20	.95	-.79	.92	-.79	.41	.35	68.6	65.3	I0043
44	41	118	.55	.21	1.04	.50	1.02	.24	.31	.35	68.6	69.6	I0044
45	12	118	2.31	.32	.90	-.36	1.19	.65	.34	.30	91.5	90.4	I0045
46	48	118	.26	.20	1.15	2.13	1.35	3.19	.15	.35	58.5	66.3	I0046
47	78	118	-.92	.21	.99	-.08	.96	-.22	.33	.32	66.1	68.9	I0047
48	50	118	.18	.20	.94	-.83	.90	-1.02	.42	.35	67.8	65.6	I0048
49	39	118	.64	.21	.99	-.13	1.00	.00	.36	.35	72.9	70.7	I0049
50	42	118	.51	.21	1.04	.57	1.06	.52	.30	.35	66.1	69.1	I0050
MEAN	54.3	118.0	.00	.23	.99	-.09	1.03	.15			74.8	74.2	
P.SD	27.3	.0	1.32	.05	.10	1.18	.22	1.33			9.4	9.0	

TABLE 14.3 118-1-420174 ZOU296WS.TXT Dec 01 2025 16:18
INPUT: 118 Person 50 Item REPORTED: 118 Person 50 Item 2 CATS WINSTEPS 5.8.3.0

Раш моделі негізінде жүргізілген талдау тест тапсырмаларының сапасын жан-жақты бағалауға мүмкіндік берді [6, 8]. Кесте 2-де көрсетілгендей, тапсырмалардың орташа қиындық деңгейі 0.00 логит, диапозон -3.34-тен 2.31-ге дейін. Бұл диапазон тест құрамында өте жеңіл және өте күрделі тапсырмалардың бар екенін және оның тестіленушілердің әртүрлі қабілет деңгейлерін ажыратуға мүмкіндік беретінін дәлелдейді. № 6, 8, 13, 15, 18, 46, 23, 39

тапсырмалар корреляция көрсеткіштері төмен (<0,20), қайта қарауды қажет етеді. Ал, № 11, 17, 27, 43, 29, 48, 34 тапсырмалар модельге жақсы сәйкес келеді.

Корреляция коэффициенті төмен (<0,20) болған тапсырмаларға сәйкес келетін тақырыптар педагогтердің білімін бағалау сапасына әсер ететін маңызды мазмұндық аймақтарды көрсетеді. Жоғарыда баяндалғандай, 2022 жылдан бастап енгізілген тақырыптық талдау педагогтерге өз нәтижелерін терең түсінуге мүмкіндік береді. Талдау нәтижелері бойынша төмен корреляция байқалған негізгі тақырыптар төмендегідей:

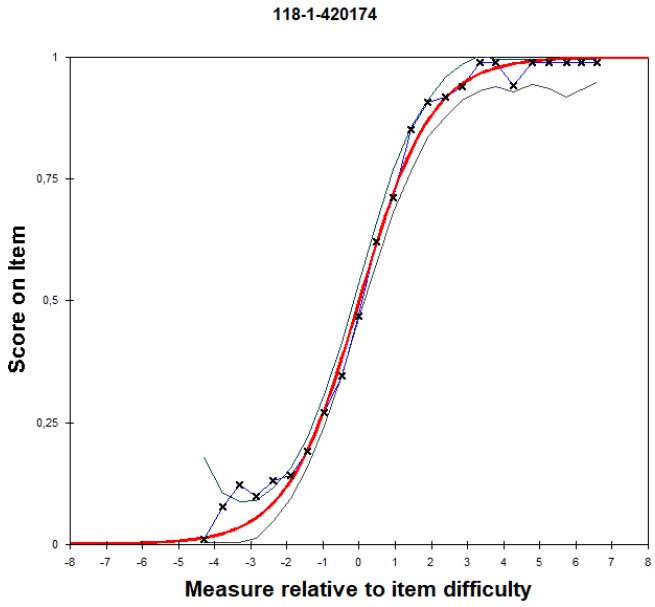
- Динамика (№6)
- Қысым (№8)
- Сақталу заңдары (№13)
- Тербелістер мен толқындар (№15)
- Жылу құбылыстары және заттың агрегаттық күйлері (№18)
- Термодинамика негіздері (№23)
- Атом ядросының физикасы (№39)

Тапсырмалардың жалпы тест нәтижесімен әлсіз байланысы олардың күрделілік деңгейінің сәйкес келмеуінен (тым жеңіл немесе тым күрделі), мазмұнның кешенділігінен және бірнеше дағдыны қатар талап етуінен туындауы мүмкін. Жоғары білім деңгейіндегі педагогтердің өзі тым жеңіл тапсырмаларға дұрыс жауап бере алмауы мүмкін, бұл олардың дискриминациялық қабілетін төмендетеді. Сондықтан аталған тақырыптардағы тапсырмалар мазмұндық, құрылымдық және психометриялық тұрғыдан қайта қарауды қажет етеді.

Ал, №11 (Статика), №17 (Сұйықтар мен газдар механикасы), №27 (Электростатика негіздері), №29 (Тұрақты электр тогы), №34 (Толқындық оптика), №43 және №48 (Мәнмәтін) тапсырмалар Раш моделіне жақсы сәйкес келеді. Олар нақты дағдыны өлшейді, мазмұны айқын, дистракторлары сапалы және тесттің сенімділігі мен валидтілігін арттырады.

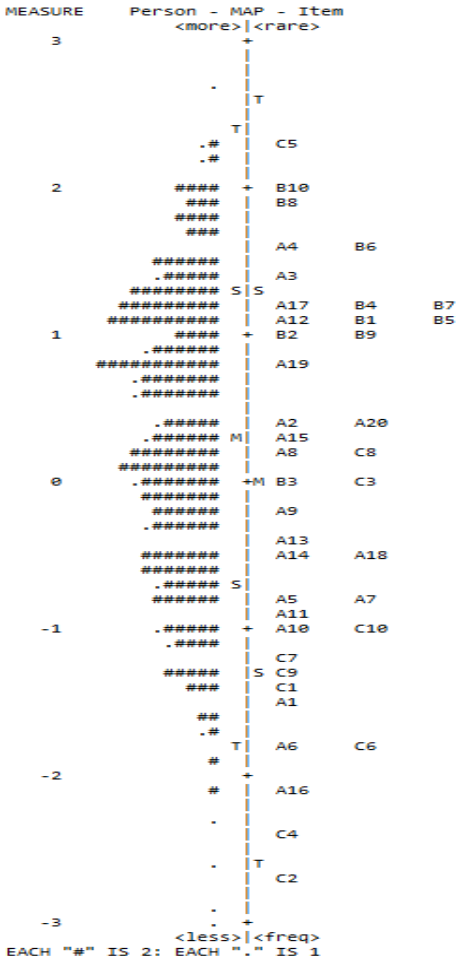
1-суретте тест тапсырмасының (Item) күрделілік деңгейіне (difficulty) байланысты тестіленушінің дұрыс жауап беру ықтималдығын көрсетеді.

Дегенмен, 1-сурет тест тапсырмасының күрделілік деңгейіне байланысты тестіленушінің дұрыс жауап беру ықтималдығы сызығы тесттің педагогтердің білім деңгейіне сәйкес келетінін дәлелдейді [8, 13].



1 сурет - Тестіленушілер мен тапсырмалардың қабілеті және күрделілік деңгейлерінің таралу сызығы

2-суретте көрсетілген тестіленушілердің қабілет деңгейі мен тапсырмалардың күрделілік деңгейі арасындағы сәйкестікті береді. Картада тестіленушілердің қабілет деңгейі (сол жақта) және тапсырмалардың күрделілік деңгейі (оң жақта) шамамен –3-тен +3 логит аралығында таралғаны байқалады.



2 сурет - Қатысушылар мен тапсырмалар картасын талдау

Егер тестіленушілердің қабілеті мен тапсырмалардың күрделілігі шкалада шамамен бірдей таралса, онда тест педагогтердің барлық деңгейін тиімді бағалайды деп айтуға болады. Мысалы, тапсырма С5 шкалада жоғарырақ орналасқан, бұл қиындық деңгейі «А» тапсырма, дегенмен ол педагогтер үшін күрделі тапсырма болғанын көрсетеді. Ал тестіленушілердің басым бөлігі шкаланың ортаңғы бөлігінде шоғырланған, яғни орташа қабілет деңгейі бар [14, 15].

Бұл карта тесттің деңгейлік балансын бағалауға және қажет болған жағдайда тапсырмаларды қайта қарауға мүмкіндік береді.

Талқылау. Раш моделі тапсырмаларды дәл бағалауға мүмкіндік береді. Кейбір тапсырмалар талапқа сай емес, сондықтан тест базасын жетілдіру қажет. 2022 жылдан бастап енгізілген тақырыптық талдау педагогтердің кері байланысын жақсартып, кәсіби дамуға септігін тигізеді.

Қорытынды

Физика пәні бойынша тест педагогтердің білімін бағалауда сенімді және валидті құрал екенін көрсетті. Зерттеу нәтижелері тесттердің жоғары сенімділігін (Item Reliability = 0,99)

және тапсырмалардың 92%-ының валидтілік критерийлеріне сәйкес келетінін растады (Infit/Outfit MNSQ 0,5–1,5 аралығында). Сондай-ақ тапсырмалардың қиындық деңгейі логит шкаласында бағаланып, дистракторлық талдау жүргізілді. Бұл нәтижелер педагогтердің білімін объективті бағалауға мүмкіндік беретінін және тесттің білім деңгейін кең ауқымда бағалауға жарамдылығын дәлелдейді.

Әдебиеттер тізімі

1. Педагог қызметкерлерді аттестаттау қағидаларын бекіту туралы : Қазақстан Республикасы Білім және ғылым министрінің 2010 жылғы 22 қаңтардағы № 16 бұйрығы. – Астана, 2010.

2. Педагогтерді аттестаттаудан өткізу қағидалары мен шарттарын бекіту туралы : Қазақстан Республикасы Оқу-ағарту министрінің 2016 жылғы 27 қаңтардағы № 83 бұйрығы. – Астана, 2016.

3. OECD. TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. – Paris: OECD Publishing, 2019. – P. 55–110. - https://doi.org/10.1787/1d0bc92a-en

4. ETS. Praxis Test Technical Manual. – Princeton : Educational Testing Service, 2022. – P. 15–60. - DOI: 10.1080/02619768.2017.1315399

5. Darling-Hammond L. Teacher Education around the World: What Can We Learn? / L. Darling-Hammond // European Journal of Teacher Education. – 2017. – Vol. 40, № 3. – P. 291–309. - https://doi.org/10.1080/02619768.2017.1315399

6. 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. – 230 p. - DOI:10.4324/9781315814698

7. Hambleton R. K. Fundamentals of Item Response Theory / R. K. Hambleton, H. Swaminathan, H. J. Rogers. – Newbury Park : Sage, 1991. – 174 p.

8. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests / G. Rasch. – Chicago : University of Chicago Press, 1980. – 199 p.

9. Тест тапсырмаларының базасын қалыптастыру қағидалары. – Астана : Ұлттық тестілеу орталығы, 2026. – 12–20 б.

10. DeVellis R. F. Scale Development: Theory and Applications / R. F. DeVellis. – 4th ed. – Los Angeles : Sage, 2017. – 280 p.

11. Linacre J. M. A User’s Guide to Winsteps Rasch-Model Computer Programs / J. M. Linacre. – Chicago : Winsteps.com, 2023. – 250 p.

12. Field A. Discovering Statistics Using IBM SPSS Statistics / A. Field. – 5th ed. – London : Sage, 2018. – 1104 p.

13. Boone W. J. Rasch Analysis in the Human Sciences / W. J. Boone, J. R. Staver, M. S. Yale. – Dordrecht : Springer, 2014. – 482 p.

14. Wright B. D. Rating Scale Analysis / B. D. Wright, G. N. Masters. – Chicago : MESA Press, 1982. – 206 p.

15. Brookhart S. M. How to Create and Use Rubrics for Formative Assessment and Grading / S. M. Brookhart. – Alexandria : ASCD, 2013. – 160 p.

Ауезханова А.Ж., Ерсултанова Т.С., Куандык А.К., Ален С.А.

ПСИХОМЕТРИЧЕСКИЙ АНАЛИЗ АПРОБАЦИИ ТЕСТИРОВАНИЯ ДЛЯ ОЦЕНКИ ЗНАНИЙ ПЕДАГОГОВ ПО ФИЗИКЕ В 2025 ГОДУ

Данное исследование посвящено психометрическому анализу тестовых заданий по физике, применяемых для оценки уровня знаний педагогов в Республике Казахстан. Исследование базируется на результатах апробации педагогического тестирования (ПББ) 2025 года. В анализе использовалась модель Раша для оценки надежности, валидности и дифференциационной способности теста. Несмотря на то, что в апробации приняли участие более 6000 педагогов, для анализа была использована репрезентативная выборка из 118 участников. Результаты исследования показали высокую надежность теста (Item Reliability = 0,99), 92% заданий соответствовали критериям валидности (Infit/Outfit MNSQ в диапазоне 0,5–1,5). Кроме того, оценивался уровень сложности

заданий в логитной шкале и проводился дистракторный анализ. Некоторые задания выявили несоответствия и требуют пересмотра. Сопоставление уровня знаний участников и сложности заданий показало, что тест эффективно оценивает уровень знаний педагогов в широком диапазоне. С 2022 года введенный тематический анализ позволяет педагогам глубже понять свои результаты. Полученные данные подтверждают, что используемый тест является надежным и валидным инструментом для объективной оценки предметных знаний педагогов и формируют рекомендации по дальнейшему совершенствованию заданий.

Ключевые слова: оценка знаний педагогов, психометрический анализ, тест, модель Раша, надежность, валидность.

A. Auezhanova, T. Ersultanova, A. Kuandyk, S. Alen

PSYCHOMETRIC ANALYSIS OF THE 2025 PILOT TESTING FOR ASSESSING THE KNOWLEDGE OF PHYSICS TEACHERS

This study focuses on the psychometric analysis of physics test items used to assess teachers' knowledge in the Republic of Kazakhstan. The research is based on the results of the 2025 pilot of the Teacher Knowledge Assessment (TKA). The Rasch measurement model was applied to evaluate the reliability, validity, and discriminatory power of the test. Although over 6,000 teachers participated in the pilot, a representative sample of 118 participants was analyzed. The results demonstrated high test reliability (Item Reliability = 0.99), with 92% of items meeting validity criteria (Infit/Outfit MNSQ within 0.5–1.5). Additionally, item difficulty was measured on the logit scale, and a distractor analysis was conducted. Some items revealed inconsistencies and require revision. The alignment between participants' ability levels and item difficulty indicated that the test effectively differentiates teachers across a broad range of knowledge levels. Since 2022, the introduction of thematic analysis has allowed teachers to gain deeper insights into their results. Overall, the findings confirm that the test is a reliable and valid instrument for the objective assessment of teachers' subject knowledge and provide recommendations for further item refinement.

Keywords: teacher knowledge assessment, psychometric analysis, test, Rasch model, reliability, validity.

References

1. Ministry of Education and Science of the Republic of Kazakhstan. (2010). On approval of the rules for certification of pedagogical workers (Order No. 16). Astana.

2. Ministry of Education of the Republic of Kazakhstan. (2016). On approval of rules and conditions for certification of teachers (Order No. 83). Astana.

3. OECD. (2019). *TALIS 2018 results (Volume I): Teachers and school leaders as lifelong learners*. OECD Publishing. https://doi.org/10.1787/1d0bc92a-en

4. Educational Testing Service. (2022). *Praxis test technical manual*. https://doi.org/10.1080/02619768.2017.1315399

5. Darling-Hammond, L. (2017). Teacher education around the world: What can we learn? *European Journal of Teacher Education, 40*(3), 291–309. https://doi.org/10.1080/02619768.2017.1315399

6. 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

7. Hambleton, R. K., Swaminathan, H., Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage.

8. Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.

9. National Testing Center. (2026). *Rules for the formation of the database of test items* (pp. 12–20). Astana.

10. DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.

11. Linacre, J. M. (2023). *A user's guide to Winsteps Rasch-model computer programs*. Winsteps.com.

12. Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.

13. Boone, W. J., Staver, J. R., Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.

14. Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

15. Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Автор туралы мәліметтер:

Ауезханова Анар Жумагалиевна (автор-корреспондент) – ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Орта білімдегі педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру ғылыми-әдістемелік зертханасының меңгерушісі, Астана қ., Қазақстан Республикасы, email: auezkhanova72@gmail.com

Ерсултанова Турсынкуль Серикбаевна - ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру басқармасының басшысы, Астана қ., Қазақстан Республикасы, email: Tursynkul-07@mail.ru

Қуандық Аманжол Қасқырбайұлы - ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Орта білімдегі педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру ғылыми-әдістемелік зертханасының аға сарапшысы, Астана қ., Қазақстан Республикасы, email: Amanzhol.uandy@bk.ru

Әлен Сабира Әділқызы - ҚР ҒЖБМ «Ұлттық тестілеу орталығы» ШЖҚ РМК, Орта білімдегі педагогтердің білімін бағалау тест тапсырмаларын қалыптастыру ғылыми-әдістемелік зертханасының жетекші сарапшысы, Астана қ., Қазақстан Республикасы, email: Sabira.alen@bk.ru

Сведения об авторах:

Ауезханова Анар Жумагалиевна (автор-корреспондент) - РГП на ПХВ «Национальный центр тестирования» МНВО РК, заведующий научно-методической лабораторией формирования тестовых заданий оценки знаний педагогов среднего образования, Астана, Республика Казахстан, email: auezkhanova72@gmail.com

Ерсултанова Турсынкуль Серикбаевна - РГП на ПХВ «Национальный центр тестирования» МНВО РК, руководитель управления формирования тестовых заданий оценки знаний педагогов, г. Астана, Республика Казахстан, email: Tursynkul-07@mail.ru

Қуандық Аманжол Қасқырбайұлы - РГП на ПХВ «Национальный центр тестирования» МНВО РК, старший эксперт научно-методической лаборатории формирования тестовых заданий оценки знаний педагогов среднего образования, г. Астана, Республика Казахстан, email: Amanzhol.uandy@bk.ru

Әлен Сабира Әділқызы - РГП на ПХВ «Национальный центр тестирования» МНВО РК, ведущий эксперт научно-методической лаборатории формирования тестовых заданий оценки знаний педагогов среднего образования, г. Астана, Республика Казахстан, email: Sabira.alen@bk.ru

Information about the authors:

Auezhanova Anar (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, Head of the Scientific-Methodological Laboratory for Developing Test Items for Assessing the Knowledge of Secondary Education Teachers, Astana, Republic of Kazakhstan, email: auezkhanova72@gmail.com

Ersultanova Tursynkul - 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, Head of the Department for Developing Test Items for Assessing Teacher Knowledge, Astana, Republic of Kazakhstan, email: Tursynkul-07@mail.ru

Kuandyk Amanzhol - 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, Senior Expert of the Scientific-Methodological Laboratory for Developing Test Items for Assessing the Knowledge of Secondary Education Teachers, Astana, Republic of Kazakhstan, email: Amanzhol.uandy@bk.ru

Alen Sabira - 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, Leading Expert of the Scientific-Methodological Laboratory for Developing Test Items for Assessing the Knowledge of Secondary Education Teachers, Astana, Republic of Kazakhstan, email: Sabira.alen@bk.ru