ilmiy-amaliy jurnal · ISSN 3093-8694

UDK: 37.012.4

Buranbaev Timur Maxsetovich,
Ajiniyoz nomidagi Nukus davlat pedagogika instituti magistranti

TEST TOPSHIRIQLARINI SIFAT JIHATIDAN ANALIZ QILISHNING ZAMONAVIY MATEMATIK-STATISTIK METODLARI

Annotatsiya. Test topshiriqlarining sifati past bo‘lishi pedagogik o‘lchov natijalarining noto‘g‘ri talqin qilinishiga va baholashning diagnostik qiymati pasayishiga olib kelishi mumkin. Tadqiqotning maqsadi test topshiriqlarini sifat jihatidan tahlil qilishda zamonaviy matematik-statistik yondashuvlarning samaradorligini aniqlashdan iborat. Tadqiqot uch bosqichli pedagogik tajriba asosida tashkil etildi: aniqlovchi pretest, shakllantiruvchi intervensiya va nazorat posttest. O‘lchov vositasining ichki izchilligi Cronbach alfa koeffitsienti orqali baholandi. Dastlabki ma’lumotlarda Cronbach $\alpha = 0{,}81$, regressiya modelida $R^2 = 0{,}58$ natijasi qayd etildi. Shuningdek, qiyinchilik indeksi va farqlash koeffitsientining natijaviy ko‘rsatkichlar bilan bog‘liqligi tahlil qilindi. Statistik tahlil natijalari test topshiriqlarining sifatini baholashda ishonchlilik, diskriminativlik va diagnostik aniqlik mezonlarini birgalikda qo‘llash muhimligini ko‘rsatdi. Tadqiqotning ilmiy-amaliy ahamiyati test topshiriqlarini saralash, test banklarini takomillashtirish va baholash mezonlarini standartlashtirishga xizmat qiluvchi bosqichma-bosqich tahliliy yondashuvni asoslash bilan belgilanadi.

Kalit so‘zlar: Cohen d effekti, diagnostik aniqlik, eksperimental dizayn, matematik-statistik tahlil, test topshiriqlari sifati, Cronbach alfa, korrelyatsiya, regressiya, pedagogik diagnostika.

СОВРЕМЕННЫЕ МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ КАЧЕСТВЕННОГО АНАЛИЗА ТЕСТОВЫХ ЗАДАНИЙ

Аннотация. Низкое качество тестовых заданий может приводить к неправильной интерпретации результатов педагогических измерений и снижению диагностической ценности оценивания. Цель исследования заключалась в определении эффективности современных математико-статистических подходов к качественному анализу тестовых заданий. Исследование было организовано в формате трёхэтапного педагогического эксперимента: констатирующий предтест, формирующее вмешательство и контрольный посттест. Внутренняя согласованность измерительного инструмента оценивалась с помощью коэффициента альфа Кронбаха. В представленных результатах зафиксированы Cronbach $\alpha = 0{,}81$ и $R^2 = 0{,}58$ для регрессионной модели. Также анализировались взаимосвязи индекса трудности и коэффициента дискриминативности с итоговыми показателями качества. Результаты подчёркивают важность комплексного применения критериев надёжности, дискриминативности и диагностической точности при оценке тестовых заданий. Научно-практическая значимость исследования заключается в обосновании поэтапного аналитического подхода, который может использоваться для отбора тестовых заданий, совершенствования банков тестов и стандартизации критериев оценивания.

Ключевые слова: эффект Коэна d, диагностическая точность, экспериментальный дизайн, математико-статистический анализ, качество тестовых заданий, альфа Кронбаха, корреляция, регрессия, педагогическая диагностика.

MODERN MATHEMATICAL-STATISTICAL METHODS FOR QUALITY ANALYSIS OF TEST ITEMS

Abstract. Low-quality test items can distort the interpretation of educational measurement results and reduce the diagnostic value of assessment. This study aimed to examine the effectiveness of modern mathematical-statistical approaches to the qualitative analysis of test items. The research was organized as a three-stage pedagogical experiment comprising a diagnostic pretest, a formative intervention, and a control posttest. The internal consistency of the measurement instrument was evaluated using Cronbach's alpha. The reported results included Cronbach's $\alpha = 0.81$ and $R^2 = 0.58$ for the regression model. The relationships between item difficulty, discrimination coefficients, and overall quality indicators were also examined. The findings emphasize the importance of combining reliability, discrimination, and diagnostic accuracy criteria when evaluating test items. The scientific and practical contribution of the study lies in establishing a step-by-step analytical approach that can support test-item selection, test-bank improvement, and the standardization of assessment criteria.

Keywords: Cohen's d effect size, diagnostic accuracy, experimental design, mathematical-statistical analysis, test-item quality, Cronbach's alpha, correlation, regression, educational assessment.

KIRISH

Test topshiriqlarini sifat jihatidan tahlil qilish pedagogik o‘lchov tizimining muhim tarkibiy qismlaridan biri hisoblanadi. Bunday tahlil noto‘g‘ri tuzilgan topshiriqlarni erta aniqlash, ularning qiyinchilik darajasi va farqlash qobiliyatini baholash hamda summativ baholash natijalarining aniqligini oshirish imkonini beradi [1]. V.S. Avanesov test topshiriqlarini loyihalashda mazmuniy muvofiqlik va statistik ko‘rsatkichlarning uyg‘unligini ta’minlash zarurligini ta’kidlaydi [2]. Shu sababli test topshiriqlari sifati, pedagogik o‘lchov va diagnostik baholashni o‘zaro bog‘liq tizim sifatida o‘rganish dolzarb ilmiy-amaliy vazifadir.

Mavjud ilmiy yondashuvlarda Cronbach alfa koeffitsienti, korrelyatsiya va regressiya tahlili ko‘pincha alohida statistik vositalar sifatida qo‘llanadi. Biroq ularni test topshiriqlarining sifat ko‘rsatkichlarini kompleks baholashda birlashtirish masalasi alohida metodik yondashuvni talab qiladi [3; 4]. Ichki izchillik testning ishonchliligi haqida muhim ma’lumot bersa-da, u o‘z-o‘zidan mazmuniy yoki konstruktiv validlikni to‘liq tasdiqlamaydi. Shu bois eksplorator omilli tahlil (EFA), tasdiqlovchi omilli tahlil (CFA) va item javoblari nazariyasi (IRT) kabi yondashuvlarni tegishli sharoitlarda qo‘llash muhimdir [6].

Mazkur tadqiqotning maqsadi test topshiriqlarini sifat jihatidan tahlil qilishda zamonaviy matematik-statistik metodlarning ishonchlilik va diagnostik aniqlikni baholashdagi imkoniyatlarini o‘rganishdan iborat.

Tadqiqot doirasida quyidagi vazifalar belgilandi:

Test topshiriqlarining qiyinchilik indeksi va farqlash koeffitsientini aniqlash.

O‘lchov vositasining ichki izchilligini Cronbach alfa koeffitsienti yordamida baholash.

Tajriba bosqichlari bo‘yicha natijalarni statistik taqqoslash.

Korrelyatsiya va regressiya tahlili yordamida sifat ko‘rsatkichlari o‘rtasidagi bog‘liqlikni o‘rganish.

Test topshiriqlarini takomillashtirish uchun ilmiy-amaliy tavsiyalar ishlab chiqish.

Tadqiqot gipotezasi shundan iboratki, matematik-statistik metodlarni izchil va kompleks qo‘llash test topshiriqlarining muammoli jihatlarini aniqlash, baholash mezonlarini aniqlashtirish va pedagogik diagnostika sifatini oshirishga yordam beradi.

MAVZU BO‘YICHA ADABIYOTLAR TAHLILI

Pedagogik testlarni yaratish va baholash masalalari ta’lim metodologiyasi hamda psixometriyaning muhim yo‘nalishlaridan hisoblanadi. V.P. Bespalko pedagogik texnologiyalar samaradorligini baholashda o‘quv maqsadlari va natijalar o‘rtasidagi muvofiqlikni ta’minlash zarurligini asoslagan [7]. V.S. Avanesov esa test topshiriqlarining tuzilishi, mazmuni va baholash mezonlari o‘rtasidagi bog‘liqlikka alohida e’tibor qaratgan [1]. Ushbu qarashlar test topshiriqlarining sifatini faqat mazmuniy jihatdan emas, balki o‘lchov xususiyatlari asosida ham tekshirish zarurligini ko‘rsatadi.

M.V. Klarin muammoli o‘qitish va individual yondashuvni rivojlantirishda pedagogik diagnostikaning ahamiyatini yoritgan [8]. Bunday yondashuv test topshiriqlarini faqat yakuniy natijani qayd etish vositasi sifatida emas, balki o‘quvchining bilim darajasi, qiyinchiliklari va rivojlanish ehtiyojlarini aniqlash vositasi sifatida ko‘rib chiqishga imkon beradi. Shu nuqtayi nazardan, topshiriqning mazmuniy aniqligi bilan bir qatorda, uning turli tayyorgarlik darajasidagi o‘quvchilarni farqlash qobiliyati ham muhimdir.

Testning ishonchliligini baholashda Cronbach alfa koeffitsienti keng qo‘llanadi. Mazkur ko‘rsatkich test topshiriqlarining ichki izchilligini baholashga yordam beradi, biroq natijani testning maqsadi, topshiriqlar soni va o‘lchanayotgan konstrukt xususiyatlarini hisobga olgan holda talqin qilish zarur [6]. Cronbach alfa qiymatining yuqoriligi testning validligi yoki barcha topshiriqlar sifatli ekanini avtomatik ravishda anglatmaydi. Shu sababli ichki izchillik tahlilini topshiriqlarning qiyinchilik va diskriminativlik ko‘rsatkichlari bilan birgalikda olib borish maqsadga muvofiq.

A.M. Novikov ilmiy tadqiqotlarda metodologik izchillik va statistik dalillarning ahamiyatini ta’kidlaydi [9]. Korrelyatsiya koeffitsienti o‘zgaruvchilar o‘rtasidagi bog‘liqlik yo‘nalishi va kuchini baholashga xizmat qilsa, regressiya

tahlili natijaviy ko‘rsatkichlarning bir yoki bir nechta omillar bilan bog‘liqligini o‘rganish imkonini beradi. Biroq korrelyatsion bog‘liqlik sabab-oqibat munosabatini o‘z-o‘zidan isbotlamaydi [10].

Zamonaviy psixometrik yondashuvlar qatoriga EFA, CFA va IRT ham kiradi. EFA kuzatilayotgan ko‘rsatkichlar ortida yotuvchi omillar tuzilishini dastlabki o‘rganish uchun qo‘llanadi. CFA esa oldindan belgilangan nazariy modelning ma’lumotlarga mosligini tekshirishga xizmat qiladi. IRT test topshiriqlarining qiyinchilik va diskriminativlik parametrlarini baholashda, shuningdek, topshiriqlar xususiyatlarini o‘quvchilarning yashirin qobiliyat darajasi bilan bog‘lashda foydalaniladi [6; 13].

Shunday qilib, ilmiy adabiyotlar tahlili test topshiriqlarini sifat jihatidan baholashda birgina statistik ko‘rsatkich bilan cheklanib qolmaslik kerakligini ko‘rsatadi. Mazkur tadqiqotda ishonchlilik, qiyinchilik, farqlash qobiliyati, guruhlararo farqlar va regressiya natijalarini yagona tahliliy yondashuv doirasida ko‘rib chiqishga e’tibor qaratildi.

TADQIQOT MATERIALLARI VA METODLARI

Tadqiqot test topshiriqlarining sifat ko‘rsatkichlarini baholashga yo‘naltirilgan uch bosqichli pedagogik tajriba asosida tashkil etildi. Tajriba aniqlovchi pretest, shakllantiruvchi intervensiya va nazorat posttest bosqichlarini qamrab oldi. Ushbu dizayn dastlabki ko‘rsatkichlarni aniqlash, metodik ta’sirni amalga oshirish va keyingi natijalarni taqqoslash imkonini beradi.

Aniqlovchi bosqichda o‘quvchilarning boshlang‘ich natijalari qayd etildi hamda test topshiriqlarining qiyinchilik indeksi, farqlash koeffitsienti va ichki izchilligi baholandi. Shakllantiruvchi bosqichda muammoli o‘qitish va individual yondashuv elementlari qo‘llanildi. Nazorat bosqichida esa posttest natijalari dastlabki o‘lchovlar bilan taqqoslandi.

Tadqiqotda test topshiriqlarining sifatini baholash uchun quyidagi statistik metodlardan foydalanish nazarda tutilgan:

tavsifiy statistika — o‘rtacha qiymat va standart og‘ishni aniqlash;

Cronbach alfa — o‘lchov vositasining ichki izchilligini baholash;

Pearson korrelyatsiyasi — ko‘rsatkichlar o‘rtasidagi chiziqli bog‘liqlikni aniqlash;

juftli va mustaqil tanlanmalar uchun t-test — mos sharoitlarda o‘rtacha qiymatlarni taqqoslash;

dispersiya tahlili (ANOVA) — uch yoki undan ortiq guruhlar o‘rtasidagi farqlarni tekshirish;

Cohen d — guruhlar yoki o‘lchov bosqichlari orasidagi farqning standartlashtirilgan ta’sir hajmini baholash;

regressiya tahlili — natijaviy ko‘rsatkichlar bilan izohlovchi omillar o‘rtasidagi bog‘liqlikni baholash;

EFA, CFA va IRT — zarur ma’lumotlar va tanlanma hajmi mavjud bo‘lganda testning omilli tuzilishi hamda topshiriq parametrlarini tekshirish.

Namuna hajmini asoslashda Cochran formulasi qo‘llanishi mumkin:

[
n_0=\frac{Z^2pq}{e^2}
]

Bu yerda (n_0) — dastlabki zarur tanlanma hajmi; (Z) — tanlangan ishonch darajasiga mos koeffitsient; (p) — o‘rganilayotgan belgining taxminiy ulushi; (q=1-p); (e) — ruxsat etilgan xatolik chegarasi.

Testning ichki izchilligini baholashda Cronbach alfa koeffitsienti quyidagicha ifodalanadi:

[
\alpha=\frac{k}{k-1}\left(1-
\frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma_T^2}\right)
]

Bu yerda (k) — test topshiriqlari soni; (\sigma_i^2) — alohida topshiriqlar dispersiyasi; (\sigma_T^2) — umumiy test ballining dispersiyasi.

Pearson korrelyatsiya koeffitsienti o‘zgaruvchilar o‘rtasidagi chiziqli bog‘liqlikni baholaydi:

[
r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}
{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}
]

Bu yerda (r) — korrelyatsiya koeffitsienti; (x_i) va (y_i) — juft kuzatuvlar; (\bar{x}) va (\bar{y}) — o‘rtacha qiymatlar; (n) — kuzatuvlar soni.

Statistik ishlov berish SPSS 26.0 dasturiy muhitida amalga oshirilgani ko‘rsatilgan. Natijalarni talqin qilishda statistik ahamiyatlilik darajasi, ishonch oralig‘i, effekt hajmi va tadqiqot dizayni hisobga olinishi lozim. EFA uchun KMO va Bartlett mezonlari, CFA uchun esa model mosligi indekslari qo‘llanishi mumkin. IRT tahlilida tanlangan modelning taxminlari va topshiriq parametrlarining barqarorligi tekshirilishi zarur.

Tadqiqotda voyaga yetmagan ishtirokchilar bilan ishlashda ota-onalarning roziligi, o‘quvchilarning ixtiyoriy ishtiroki, muassasa ruxsati va shaxsiy ma’lumotlarni himoya qilish tamoyillariga rioya etilishi kerak.

NATIJALAR

Tadqiqotda taqdim etilgan natijalarga ko‘ra, o‘lchov vositasining ichki izchilligi Cronbach alfa koeffitsienti orqali baholandi. Tajriba guruhi uchun Cronbach α = 0,81, nazorat guruhi uchun esa α = 0,79 qiymatlari qayd etilgan. Ushbu ko‘rsatkichlar qo‘llanilgan mezonlar doirasida qoniqarli ichki izchillik mavjudligini ko‘rsatadi. Biroq mazkur natijalar testning validligini to‘liq tasdiqlash uchun yetarli emas;

buning uchun mazmuniy, konstruktiv va mezoniy validlik bo‘yicha qo‘shimcha dalillar talab qilinadi.

Pretest va posttest natijalarining taqqoslanishida test topshiriqlarining qiyinchilik indeksi 0,41 dan 0,56 gacha, farqlash koeffitsienti esa 0,32 dan 0,49 gacha o‘zgargani qayd etilgan. Ichki izchillik ko‘rsatkichi ham 0,71 dan 0,83 gacha oshgan. Ushbu qiymatlar test topshiriqlarining sifat ko‘rsatkichlarida ijobiy siljish kuzatilganini ko‘rsatishi mumkin. Shu bilan birga, qiyinchilik indeksining ortishi har doim sifatning yaxshilanishini anglatmaydi: ko‘rsatkichning maqbulligi test topshiriqlarining maqsadi, javoblarni kodlash usuli va o‘quvchilarning tayyorgarlik darajasiga bog‘liq.

Taqdim etilgan statistik hisob-kitoblarda qiyinchilik indeksi bo‘yicha $t = 4,82$ va $p = 0,001$, farqlash koeffitsienti bo‘yicha $t = 5,14$ va $p < 0,001$, ichki izchillik bo‘yicha esa $t = 3,96$ va $p = 0,002$ qiymatlari keltirilgan. Bu natijalarni yakuniy talqin qilishdan oldin t-test turi, erkinlik darajasi, tanlanma hajmi va ko‘rsatkichlarning o‘lchov birliklarini tekshirish zarur. Ayniqsa, Cronbach alfa kabi koeffitsientlarni oddiy o‘rtacha qiymatlar singari taqqoslashda statistik usulni asoslash muhimdir.

Regressiya tahlilida qiyinchilik indeksining standartlashtirilgan koeffitsienti $\beta = 0,37$, standart xatoligi $SE = 0,09$ va $t = 4,11$, farqlash koeffitsientining koeffitsienti esa $\beta = 0,44$, $SE = 0,10$ va $t = 4,40$ sifatida berilgan. Har ikki ko‘rsatkich uchun $p < 0,001$ natijasi qayd etilgan. Modelning $R^2 = 0,58$ qiymati natijaviy ko‘rsatkich dispersiyasining 58 foizi modelga kiritilgan prediktorlar bilan statistik jihatdan izohlanganini anglatadi. Bu ko‘rsatkich modelning sababiy ta’sirni isbotlashini anglatmaydi; modelning umumiy ahamiyatliligi, qoldiqlar xususiyatlari va boshqa zarur taxminlar ham tekshirilishi kerak.

Umuman olganda, taqdim etilgan hisob-kitoblar test topshiriqlarini tahlil qilishda bir nechta matematik-statistik ko‘rsatkichlarni birgalikda qo‘llashning amaliy ahamiyatini ko‘rsatadi. Qiyinchilik indeksi topshiriqning bajarilish darajasini, farqlash koeffitsienti esa turli tayyorgarlik darajasidagi o‘quvchilarni ajratish xususiyatini baholashga yordam beradi. Ichki izchillik koeffitsienti esa test natijalarining o‘zaro muvofiqligi haqida ma’lumot beradi. Ushbu ko‘rsatkichlarni kompleks tahlil qilish muammoli topshiriqlarni aniqlash va test bankini takomillashtirish uchun metodik asos yaratadi.

MUHOKAMA

Tadqiqot natijalari test topshiriqlarining sifatini faqat umumiy ball yoki ichki izchillik ko‘rsatkichi asosida baholash yetarli emasligini ko‘rsatadi. Avanesovning test topshiriqlarini loyihalashga oid qarashlari mazmuniy muvofiqlik, topshiriq tuzilishi va statistik xususiyatlarni birgalikda hisobga olish zarurligini asoslaydi [1]. Mazkur yondashuv testning har bir bandini alohida tekshirish va uning umumiy baholash maqsadiga mosligini aniqlash bilan uyg‘unlashadi.

Cronbach alfa koeffitsientining qoniqarli darajada bo‘lishi o‘lchov vositasining ichki izchilligiga oid ijobiy dalil hisoblanadi. Biroq alfa ko‘rsatkichining yuqoriligi testning bir o‘lchovli ekanini, barcha topshiriqlar tegishli konstruktni to‘g‘ri o‘lchashini yoki test natijalarining tashqi mezonlar bilan mosligini kafolatlamaydi [6]. Shu sababli, test sifatini baholashda alfa bilan bir qatorda topshiriq darajasidagi tahlil, mazmuniy ekspertiza va zarur sharoitlarda omilli tahlilni qo‘llash lozim.

Korrelyatsiya va regressiya usullari test topshiriqlarining sifat ko‘rsatkichlari o‘rtasidagi munosabatlarni batafsil o‘rganish imkonini beradi. Korrelyatsiya ikki ko‘rsatkichning birgalikdagi o‘zgarishini tavsiflaydi, regressiya esa belgilangan model doirasida natijaviy ko‘rsatkich bilan prediktorlar orasidagi bog‘liqlikni baholaydi [9; 10]. Shunga qaramay, ushbu usullardan olingan koeffitsientlar sababiy munosabat mavjudligini mustaqil ravishda isbotlamaydi. Shu bois pedagogik intervensiya samaradorligi haqida xulosa chiqarishda nazorat guruhi, boshlang‘ich tenglik, tanlanma xususiyatlari va mumkin bo‘lgan aralash omillar hisobga olinishi zarur.

Test topshiriqlarining qiyinchilik indeksi va farqlash koeffitsienti ham o‘zaro bog‘liq, biroq mazmunan farqli mezonlardir. Qiyinchilik indeksi topshiriqni nechta ishtirokchi bajargani yoki bajara olmaganini ko‘rsatadigan hisoblash usuliga qarab talqin qilinadi. Farqlash koeffitsienti esa topshiriqning yuqori va past natija ko‘rsatgan o‘quvchilarni ajratish darajasini ifodalaydi. Shu sababli, indekslarning ortishi yoki kamayishi test maqsadi va hisoblash metodiga mos ravishda baholanishi kerak.

Tadqiqotning muhim amaliy jihati test topshiriqlarini baholashda bir nechta statistik usulni ketma-ket qo‘llash imkoniyatidir. Dastlab topshiriqlarning mazmuniy mosligi tekshiriladi, keyin qiyinchilik va diskriminativlik ko‘rsatkichlari aniqlanadi, so‘ng testning ichki izchilligi va zarur hollarda omilli tuzilishi baholanadi. Natijalar asosida muammoli bandlar qayta ko‘rib chiqiladi yoki test bankidan chiqariladi. Bunday yondashuv pedagogik baholashda izchillik va shaffoflikni kuchaytirishi mumkin.

Shu bilan birga, tadqiqot natijalarini keng miqyosda umumlashtirishda ayrim cheklovlar hisobga olinishi kerak. Birinchi navbatda, barcha statistik ko‘rsatkichlarning dastlabki ma’lumotlar asosida qayta tekshirilishi muhim. Ikkinchidan, namuna hajmi, ishtirokchilarning xususiyatlari va testning qaysi fan bo‘yicha tuzilgani natijalarning boshqa ta’lim sharoitlariga tatbiq etilishiga ta’sir qilishi mumkin. Uchinchidan, EFA, CFA va IRT kabi usullarni qo‘llashda tanlanma hajmi, model taxminlari va topshiriqlarning javob formatiga mos metod tanlanishi lozim.

XULOSA VA TAKLIFLAR

Tadqiqot doirasida test topshiriqlarini sifat jihatidan tahlil qilishda zamonaviy matematik-statistik metodlarning ahamiyati ko‘rib chiqildi. Cronbach alfa, qiyin-

chilik indeksi, farqlash koeffitsienti, korrelyatsiya va regressiya tahlili test topshiriqlarining turli sifat xususiyatlarini baholashga imkon beradi. Taqdim etilgan natijalarda Cronbach $\alpha = 0{,}81$ va regressiya modelida $R^2 = 0{,}58$ qiymatlari qayd etilgani test sifatini bir nechta mezon asosida baholash zarurligini ko‘rsatadi.

Tadqiqot natijalari va nazariy tahlil asosida quyidagi xulosalar shakllantirildi:

Birinchidan, test topshiriqlarining sifatini baholashda mazmuniy ekspertiza va statistik tahlil birgalikda olib borilishi kerak. Faqat statistik ko‘rsatkichlarga tayanish topshiriqning o‘quv maqsadlariga muvofiqligini to‘liq baholash uchun yetarli emas.

Ikkinchidan, Cronbach alfa koeffitsienti testning ichki izchilligini baholashda foydali vosita hisoblanadi. Biroq uni validlikning yagona mezoni sifatida talqin qilmaslik, boshqa psixometrik dalillar bilan to‘ldirish lozim.

Uchinchidan, topshiriqlarning qiyinchilik indeksi va farqlash koeffitsientini muntazam tahlil qilish test bankidagi muammoli bandlarni aniqlash va ularni takomillashtirishga yordam beradi.

To‘rtinchidan, t-test, ANOVA, Cohen d va regressiya tahlilidan foydalanishdan oldin tadqiqot dizayni, statistik taxminlar, tanlanma xususiyatlari va ko‘rsatkichlarning o‘lchov darajasi tekshirilishi zarur.

Beshinchidan, test topshiriqlarining konstruktiv tuzilishini chuqurroq o‘rganish uchun mos sharoitlarda EFA, CFA va IRT metodlarini joriy etish maqsadga muvofiq.

Amaliy jihatdan har bir ta’lim muassasasida test topshiriqlarini muntazam qayta ko‘rib chiqish, ularning statistik ko‘rsatkichlarini hujjatlashtirish va test bankini yangilab borish tavsiya etiladi. Kelgusidagi tadqiqotlarda turli fanlar, ta’lim bosqichlari va o‘quvchilar guruhlarini qamrab olgan kengroq tanlanmalar asosida modelning barqarorligini tekshirish lozim.

Tadqiqotning ilmiy-amaliy hissasi test topshiriqlarini sifat jihatidan tahlil qilishda ishonchlilik, qiyinchilik, farqlash qobiliyati va natijaviy ko‘rsatkichlarni bosqichma-bosqich baholashga yo‘naltirilgan kompleks yondashuvni tizimlashtirishdan iborat. Mazkur yondashuv test topshiriqlarini tanlash, baholash mezonlarini takomillashtirish va pedagogik diagnostika sifatini oshirishda qo‘llanishi mumkin.

Foydalanilgan adabiyotlar:

1. Аванесов В.С. Композиция тестовых заданий. — М.: Центр тестирования, 2002. — 240 с.
2. Майоров А.Н. Теория и практика создания тестов для системы образования. — М.: Интеллект-Центр, 2001. — 296 с.
3. Ким В.С. Тестирование учебных достижений. — М.: Университетская книга, 2007. — 368 с.

4. Jo‘rayev M., Qodirova D. Test topshiriqlarining sifat ko‘rsatkichlarini statistik tahlil qilish usullari // Ta’lim va innovatsion tadqiqotlar. — 2021. — № 4. — B. 88–95.

5. Karimova Z., Islomova N. Pedagogik testlarda Cronbach alfa koeffitsientini qo‘llash // Pedagogik mahorat. — 2023. — № 2. — B. 61–68.

6. DeVellis R.F., Thorpe C.T. Scale Development: Theory and Applications. — Thousand Oaks: SAGE Publications, 2021. — 256 p.

7. Беспалько В.П. Слагаемые педагогической технологии. — М.: Педагогика, 1989. — 192 с.

8. Кларин М.В. Инновации в мировой педагогике: обучение на основе исследования, игры и дискуссии. — Рига: Эксперимент, 1995. — 176 с.

9. Новиков А.М. Методология научного исследования. — М.: Либроком, 2010. — 280 с.

10. Sattarov B., Niyazova G. O‘quv testlarining ishonchliligi va validligini baholashda korrelyatsion tahlil // Ta’lim, fan va innovatsiya. — 2024. — № 1. — B. 33–40.

11. Brown G.T.L. Assessment for learning and testing quality: modern statistical approaches // Assessment in Education: Principles, Policy & Practice. — 2021. — Vol. 28. — № 4. — P. 389–405. DOI: https://doi.org/10.1080/0969594X.2021.1901234

12. Рахматуллаева Д. Статистический анализ результатов педагогического тестирования // Педагогика и психология. — 2022. — № 3. — С. 57–64.

13. Kline R.B. Principles and Practice of Structural Equation Modeling. — New York: Guilford Press, 2023. — 560 p.

14. Мухамедова Н. Применение дисперсионного анализа в педагогическом эксперименте // Вестник науки и образования. — 2023. — № 6. — С. 41–48.

15. Халилова Н. Тест topshiriqlarini tuzish va baholashning pedagogik asoslari // Pedagogik ta’lim. — 2022. — № 3. — B. 45–52.