



Komil JALILOV

Baholash nazariyasi asoslari

(testologiya, pedagogik o'lchovlar)



Komil Jalilov

BAHOLASH NAZARIYASI ASOSLARI

(Testologiya, pedagogik o'lchovlar)

Toshkent
Akademnashr
2020

UO'K: 37.091.26

KBK: 74.202.5

J-21

J-21

Jalilov, Komil.

Baholash nazariyasi asoslari [Matn] / Komil Jalilov. – Toshkent: Akademnashr, 2020. – 256 b.

ISBN 978-9943-6498-4-2

UO'K: 37.091.26

KBK: 74.202.5

Ushbu metodik qo'llanmada baholash nazariyasi (testologiya, pedagogik o'lahovlar) fanining asosiy tushunchalari, tamoyillari, qonuniyatlari izchil bayon qilingan. Har bir tushunchaning mohiyatini ochib berishga hamda qo'llanilish holatlariga misollar keltirishga alohida e'tibor qaratilgan.

Kitob ta'lim sohasida faoliyat olib borayotgan pedagoglar va testolog mutaxassislar, pedagogika yo'nalishidagi oliy ta'lim muassasalarida tahsil olayotgan talabalar, magistrlar, malaka oshirish institutlari tinglovchilari, shuningdek, sohaga qiziquvchi keng kitobxonlar uchun mo'ljallangan.

Mas'ul muharrir:

Zulxumor Mirzayeva, filologiya fanlari doktori

Taqrizchilar:

Abdurahim Nosirov, O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Ta'lim sifatini nazorat qilish davlat inspeksiyasi bosh boshqarma boshlig'i, filologiya fanlari doktori, dotsent

Abduvali Ismailov, O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Ta'lim sifatini nazorat qilish davlat inspeksiyasi Ta'lim sifatini baholash bo'yicha xalqaro tadqiqotlarni amalga oshirish milliy markazi direktori

Nargiza Yusupova, O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Davlat test markazi bosh mutaxassisi

O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Ta'lim sifatini nazorat qilish davlat inspeksiyasi ilmiy-metodik kengashi tomonidan nashrga tavsiya qilingan (2020-yil 9-iyun, 3-sonli bayonnoma)

ISBN 978-9943-6498-4-2

© Komil Jalilov, 2020
© «Akademnashr», 2020

MUNDARIJA

KIRISH	5
1. BAHOLASH NAZARIYASIGA KIRISH	8
1.1. Pedagogik o'lovlar fani va uning asosiy tushunchalari.....	8
1.2. Baholashning maqsadlari.....	17
1.3. Test turlari	26
2. TESTNING SIFATI: VALIDLIK.....	37
2.1. Validlik tushunchasi va uning tadrji.....	37
2.2. Validlikning dalillari	46
2.2.1. Testning mazmuni bilan bog'liq dalillar	46
2.2.2. Test topshiruvchilarning test topshiriqlarini bajarishi jarayoni bilan bog'liq dalillar.....	51
2.2.3. Testning ichki strukturasi bilan bog'liq dalillar	52
2.2.4. Testning boshqa o'zgaruvchilar bilan bog'liqligi asosidagi dalillar	54
2.2.5. Testning oqibatlari bilan bog'liq dalillar.....	57
2.3. Validlikka xavf soluvchi omillar	62
3. ISHONCHLILIK	71
3.1. Ishonchlilik tushunchasi. Klassik test nazariyasi asoslari.....	71
3.2. Ishonchlilikni hisoblash usullari	76
3.3. Normal taqsimot. O'lchashning standart xatosi. Tasniflash barqarorligi	87
3.4. Testning asosiy tendensiyasi ko'rsatkichlari.....	95
3.5. O'lchashning standart xatosini hisoblashning muqobil usullari	98
4. TESTNI YARATISH.....	102
4.1. Testni yaratish bosqichlari	102
4.2. Yopiq test topshiriqlarini yaratish	110
4.2.1. Muqobil javobli test topshiriqlari.....	113
4.2.2. «To'g'ri-noto'g'ri» shaklidagi test topshiriqlari	135

4.2.3. Moslashtirishni talab qiluvchi test topshiriqlari.....	140
4.2.4. Ketma-ketlikni topishni talab qiluvchi test topshiriqlari	143
4.2.5. Yopiq test topshiriqlari yordamida yuqori darajadagi kognitiv ko'nikmalarni tekshirish.....	144
4.3. Ochiq test topshiriqlarini yaratish va baholash.....	150
4.4. Chet tilidan test topshiriqlarining o'ziga xos xususiyatlari (ingliz tili bo'yicha testlar misolida).....	167
5. TEST TOPSHIRIQLARINING STATISTIK TAHLILI	182
5.1. Statistik tahlilning ahamiyati	182
5.2. Test topshiriqlarining sifat ko'rsatkichlari	184
5.2.1. Test topshirig'ining qiyinlik darajasi	184
5.2.2. Test topshirig'ining diskriminatsiya indeksi	186
5.2.3. Test topshiruvchilar javoblarining distraktorlar bo'yicha taqsimoti.....	189
5.2.4. Test topshirig'ining DIF ko'rsatkichi.....	193
5.3. Mezonga mo'ljallangan testlarda test topshiriqlari tahlili	199
6. TESTNI O'TKAZISH VA UNING NATIJALARINI QAYTA ISHLASH. O'TISH BALLARINI BELGILASH.....	205
6.1. Testni o'tkazish	205
6.2. Test natijalarini yagona shkalaga keltirish	211
6.3. O'tish ballarini belgilash texnologiyalari.....	225
ILOVA.....	232
FOYDALANILGAN ADABIYOTLAR RO'YXATI.....	239
KITOBDA QO'LLANILGAN ASOSIY TUSHUNCHALAR VA ATAMALAR INDEKSI.....	244

KIRISH

Testologiya (pedagogik o‘lchovlar) **bilim, ko‘nikma, malaka va kompetensiyalarni baholash (o‘lchash)** uchun ilmiy asoslangan, sifat talablariga javob beradigan **o‘lchov vositalarini (testlarni) yaratish va qo‘llash haqidagi fan** sohasi sifatida pedagogika, psixologiya, sotsiologiya, statistika singari fanlar asosida XX asr boshlarida shakllandi. Hozirgi kunga kelib «klassik test nazariyasi», «zamonaviy test nazariyasi (IRT nazariyasi)» va «umumiylashtirish (generalizatsiya) nazariyasi» deb ataladigan bir qancha fundamental nazariyalarni va bu nazariyalarni amaliyotda qo‘llash bo‘yicha tavsiyalarni o‘z ichiga olgan bu soha xorijda shiddat bilan rivojlanmoqda, oliy ta‘lim muassasalarida alohida yo‘nalish sifatida o‘qitilmoqda, nazariy va amaliy tadqiqotlar olib borilmoqda, ilmiy darajalar berilmoqda.

Yurtimizda ham 1994-yilda oliy ta‘lim muassasalariga abituriyentlarni saralab olishda test tizimi joriy qilinganidan buyon test orqali baholash ta‘limning barcha bosqichlarida turli maqsadlarda keng qo‘llanilmoqda. Asosiy vazifasi test o‘tkazish va uning natijalarini tahlil qilib, ular asosida muhim

qarorlar qabul qilish bilan bog‘liq tashkilotlar (O‘zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Davlat test markazi, Ta‘lim sifatini nazorat qilish davlat inspeksiyasi, Xalq ta‘limi vazirligining Monitoring boshqarmasi) faoliyat yuritmoqda. O‘zbekistonning ta‘lim sifatini baholash bo‘yicha PIRLS, PISA, TIMMS singari xalqaro tadqiqotlarda ishtirok etishi kutilmoqda. Bularning barchasi **pedagogik o‘lchov vositalarini yaratish, ularni ilmiy asoslash va tahlil qilishga** qodir malakali mutaxassislar – **testologlarga** ehtiyojni oshiradi.

Shu bilan birga, testologiya sohasining nazariy asoslari va amaliy qo‘llanilishini izchil, tizimli ravishda bayon qiladigan manbalar o‘zbek tilida hali-hanuz yaratilmagan desak mubolag‘a bo‘lmaydi. Bu esa sohaga bevosita daxldor yoki soha bilan qiziqadigan, biroq ingliz, rus yoki boshqa tillardagi manbalardan foydalana olmaydigan mutaxassislarning imkoniyatlarini cheklaydi. Shuni hisobga olib testologiyaning asosiy tushuncha va qonuniyatlarini o‘zbek tilida ochib berish maqsadida mazkur kitob yaratildi.

Muallif kitob qo‘lyozmasini ko‘rib chiqib, uning sifatini yaxshilash borasida o‘z qimmatli maslahatlarini ayamagan hamkasblari – Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti dotsenti, filologiya fanlari doktori Zulxumor Mirzayeva, O‘zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Davlat test markazi bosh mutaxassisi Nargiza Yusupova, O‘zbekiston Respublikasi

Vazirlar Mahkamasi huzuridagi Ta'lim sifatini nazorat qilish davlat inspeksiyasi Ta'lim sifatini baholash bo'yicha xalqaro tadqiqotlarni amalga oshirish milliy markazi direktori Abduvali Ismailov, mazkur markaz bosh mutaxassisi Olimjon Eshqobilov, O'zbekiston Respublikasi Xalq ta'limi vazirligi boshqarma boshlig'i o'rinbosari Dilshoda Norboyevaga o'z minnatdorligini bildiradi.

Kitob Amerika Qo'shma Shtatlarining O'zbekistondagi elchixonasi «Alumni Outreach Grant» dasturi doirasidagi moliyaviy ko'magi bilan nashr etildi.

1. BAHOLASH NAZARIYASIGA KIRISH

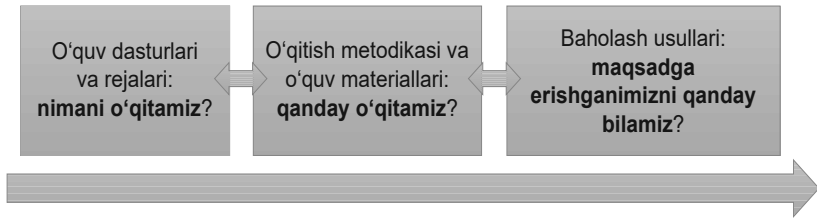
1.1. Pedagogik o'lvhovlar fani va uning asosiy tushunchalari

Ta'lim jarayoni uch asosiy **komponentdan** (tarkibiy qismdan) tashkil topadi:

- ta'lim standartlari, o'quv dasturlari va rejalari (curriculum);
- o'qitish metodikasi va o'quv materiallari (pedagogy);
- baholash usullari (assessment).

O'quv dasturlari va rejalari ta'lim maqsadlarini belgilaydi (nimani o'qitamiz? ya'ni «nimani o'qitamiz?») – degan mazmundagi savolga javob bo'ladi), shu maqsadlarga mos ravishda o'qitish metodikasi va o'quv materiallari tanlanadi (qanday o'qitamiz? ya'ni «qanday o'qitamiz?») – degan mazmundagi savolga javob bo'ladi). Baholash usullari esa ta'lim tizimi o'z oldiga qo'ygan maqsadlarga qay darajada erishganligi haqida ma'lumot beradi.

Agar bu uch komponent bir-biriga uyg'un holda ishlasa, ya'ni ta'lim maqsadlari aniq bo'lsa, shu maqsadlarga mos o'qitish usullari va materiallar tanlansa, baholash tizimi aynan ta'lim tizimi ishtirokchilari shu maqsadlarga qay darajada erishganliklarini aniqlab bera olsa, ta'lim tizimi **samarali** ishlayotgan bo'ladi va **sifatni** ta'minlab bera oladi.



Ta'lim jarayoni komponentlarining o'zaro bog'liqligi

O'quvchilarning bilim, ko'nikma, malaka va kompetensiyalarini baholashda ma'lum bir qonuniyatlar va talablar hisobga olinmog'i lozimki, bu qonuniyat va talablar pedagogik o'lchovlar fanining obyektini tashkil qiladi.

Pedagogik o'lchovlar – ta'lim oluvchilarda bilim, ko'nikma, malaka, kompetensiyalarning shakllanganlik darajasini o'lchash (aniqlash) usullari va vositalarini ishlab chiqadigan hamda amalda qo'llaydigan pedagogika sohasi.

«Pedagogik o'lchovlar», «baholash nazariyasi» va «testologiya», odatda, o'zaro muqobil, bir xil ma'no bildiradigan atamalar sifatida qo'llaniladi. Amerika Qo'shma Shtatlarida asosan «pedagogik o'lchovlar», Britaniyada «baholash» («baholash nazariyasi», «pedagogik baholash»), Rossiyada esa «testologiya» atamaları qo'llaniladi. Pedagogik o'lchovlar fanining asosiy objekti baholash jarayoni va unda qo'llaniladigan vositalardir.

Baholash – biron-bir domenning (bilim, ko‘nikma, malaka va kompetensiyalar majmuyining) ta’lim oluvchilarda (test topshiruvchilarda) shakllanganlik darajasini o‘lchash jarayoni.

Baholash jarayonining maqsadi **o‘lchash** ekan, o‘lchash nima ekanligini ko‘rib chiqishimiz lozim. Klassik test nazariyasining asoschilaridan Frederik Lord (*Frederick Lord*) va Melvin Novik (*Melvin Novick*) «o‘lchash» atamasiga quyidagicha ta’rif beradi: «o‘lchash – bu tajriba obyektining konkret xususiyatlariga raqamlar (ballar)ni birlashtirish jarayoni»¹. Misol uchun, biolog olim tomonidan ko‘l suvining bir tomchisida qancha bakteriyalar borligi aniqlanishi – bu o‘lchash jarayonidir. Ushbu misolda tajriba obyektini – ko‘l suvining bir tomchisi, ushbu obyektning konkret xususiyati – undagi bakteriyalar soni.

Mualliflar bu jarayonda uch bosqichni ajratib ko‘rsatadilar:

- birinchi: tajriba **obyektini** aniqlash;
- ikkinchi: o‘lchanadigan **xususiyatni** aniqlash;
- uchinchi: tajriba obyektining o‘lchanayotgan xususiyatini **raqam bilan ifodalash** qoidasini aniqlashtirishimiz lozim².

¹ Lord F. and Novick M. *Statistical Theories of Mental Test Scores*. – Reading, MA: Addison-Wesley Pub. 1968. – P.17.

² O‘sha manba. – P.16.

Masalan, biz kimningdir bo‘yini o‘lchamoqchi bo‘lsak, bu shaxs – tajriba obyekti, shaxsning bo‘yi (uzunligi) – o‘lchanadigan xususiyat, uzunlikning metr, dyuym yoki boshqa bir tizimda ifodalanishi – o‘lchanayotgan xususiyatni raqam bilan ifodalash qoidasi.

Pedagogikada o‘lchashning mohiyatini tushunish uchun domen va konstrukt tushunchalarini ham anglab olishimiz lozim bo‘ladi.

Domen – bilim, ko‘nikma, malaka, kompetensiyalar majmuyi.

Deylik, 9-sinf ona tili darslarida qo‘shma gaplar mavzusi o‘tildi: sodda va qo‘shma gaplarning farqlari, qo‘shma gaplarning turlari haqida tushuncha berildi, qo‘shma gaplarni tuzish hamda og‘zaki va yozma nutqda qo‘llash o‘rgatildi. Demak, 9-sinf uchun qo‘shma gaplar bilan bog‘liq mazkur bilim, ko‘nikma, malaka va kompetensiyalar majmuyi (yig‘indisi) domen hisoblanadi.

Odatda, ta‘limning har bir bosqichi uchun tegishli domen ta‘lim standartlari va o‘quv dasturlarida belgilab beriladi. Masalan, umumiy o‘rta ta‘limning Davlat ta‘lim standartlarida maktab o‘quvchilari egallashi lozim bo‘lgan bilim, ko‘nikma, malaka va kompetensiyalar majmuyi (domen) har bir fan bo‘yicha aniqlashtirib ko‘rsatilgan.

Konstrukt – o‘lchanishi lozim bo‘lgan xususiyat.

Biz biron-bir xususiyatning darajasini o‘lchash uchun o‘lchanadigan xususiyatga **mos** maxsus vositalarni – o‘lchov asboblari ishga solishimiz kerak. Masalan, uzunlikni o‘lchash uchun chizg‘ichdan yoki haroratni o‘lchash uchun termometrdan foydalanamiz (bunda uzunlik yoki harorat konstruktlar hisoblanadi). Tabiiy fanlarda konstruktarni o‘lchash uchun ishlatiladigan asboblari, odatda, **bevosita** va aniq o‘lchash imkonini beradi, masalan, stolning uzunligini o‘lchash uchun chizg‘ichni stolning ustiga qo‘yib uning uzunligini aniq santimetrlarda, dyuymlarda yoki boshqa o‘lchov birligida qanchaga teng ekanligini ko‘rishimiz mumkin. Ammo shunday holatlar bo‘ladiki, bevosita o‘lchashning imkoni mavjud emas: masalan, Yerning radiusi yoki koinotda sayyoralar orasidagi masofani o‘lchov asbobi yordamida bevosita hisoblab bo‘lmaydi.

Shunga o‘xshash, pedagogika va psixologiyada o‘quvchining biron-bir xususiyatini (masalan, uning ona tilida yozma nutqi shakllanganlik darajasini, matematik savodxonlik darajasini, biron-bir kasbga qiziqishini, chet tilini o‘rganishga layoqatini va hokazo) o‘lchamoqchi bo‘lsak, buni bevosita amalga oshirish imkoni mavjud emas. Bunday o‘lchash faqat **bilvosita** (kuzatish, suhbat, test va hokazo) orqali amalga oshirilishi mumkin. Masalan, psixologlar shaxsning temperamentini aniqlash uchun maxsus ishlab chiqilgan testlardan foydalanadilar.

Ta’lim jarayonida ham o‘quvchi yoki talabada biron-bir domenga tegishli bilim, ko‘nikma, malaka, kompetensiyalar qay darajada shakllanganligini bilmoqchi bo‘lsak, shu domen

qaysi bilim, ko‘nikma, malaka, kompetensiyalarni o‘z ichiga olishini hamda shu domenga xos qaysi konstrukt-larni (xususiyatlarni) o‘lchamoqchi ekanligimizni aniqlab olish lozim bo‘ladi. Shundan keyin o‘quvchining qanday xatti-harakati orqali (qanday topshiriqlarni bajarishi orqali) mazkur konstrukt-larni kuzatib baho berish mumkinligini aniqlash lozim.

Misol uchun, adabiyot darslarida ta‘lim oluvchi badiiy matnni tahlil qila olish-olmasligini bilmoqchi bo‘lsak, bu domen qanday konstrukt-larni o‘z ichiga olishini aniqlashtirishimiz lozim. Badiiy matnni tahlil qila olish domeni matnda ishlatilgan badiiy san’atlarni topish va sharhlash, muallif ochiq aytmagan fikrlarni tushunish, syujet elementlarini topish singari konstrukt-larni o‘z ichiga olishi mumkin. Demak, ta‘lim oluvchiga shu konstrukt-larni ko‘rsatish imkonini beruvchi topshiriqlarni bajartirish orqali uning adabiyot faniga tegishli bo‘lgan domenni qay darajada egallagani to‘g‘risida xulosa chiqarishimiz mumkin bo‘ladi.

Yoki chet tili darslarida o‘quvchi chet tilida eshitgan matnini tushungan yoki tushunmaganligini baholamoqchi bo‘lsak, tinglab tushunish kompetensiyasi (domen) qanday konstrukt-larni (xususiyatlarni) o‘z ichiga olishini aniqlashtirishimiz lozim (masalan, eshitilgan matnning umumiy ma’nosini anglab olish, matndagi ayrim detallarni ilg‘ab olish va ho-kazo). Shundan so‘ng ushbu konstrukt-larni kuzatib baholash imkonini beruvchi maxsus tuzilgan topshiriqlar orqali kompetensiya (domen) haqida xulosa chiqarish mumkin bo‘ladi.

Demak, pedagogikada baholash jarayoni **bilvosita xulosa chiqarishga** asoslangan. Bunda biz beradigan topshiriqlar **o‘lchov vositasi** vazifasini bajaradi va ta’lim oluvchi bu topshiriqlarni qanday bajarganligiga qarab u biron-bir **domenga** kiruvchi bilimlarni qay darajada egallagani, ko‘nikma, malaka va kompetensiyalar qay darajada shakllangani haqida xulosa chiqaramiz. Topshiriqlar baholanishi lozim bo‘lgan domenni to‘liq qamrab ololmaydi. Masalan, o‘quvchining adabiyotdan barmoq va aruz vaznlarini farqlay olishini tekshirmoqchi bo‘lsak, tabiiyki, barmoq va aruz vaznlariga tegishli barcha she’riy asarlarni bera olmaymiz. Faqat ayrim namunalar berib, qaysi namuna qaysi vaznga tegishli ekanligini so‘rashimiz mumkin. Demak, baholash jarayonining yana bir xususiyati **cheklangan namunalar orqali butun domen haqida xulosa chiqarishdir**.

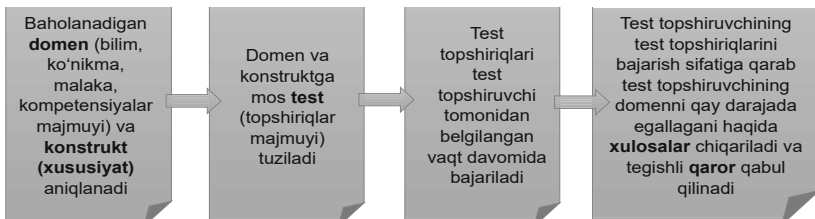
Shunday qilib, baholash baholanishi (o‘lchanishi) lozim bo‘lgan domen va konstrukt haqida xulosa chiqarish bo‘lib, bu xulosa bilvosita, cheklangan sondagi maxsus yaratilgan **topshiriqlar** yordamida chiqariladi. Baholash jarayonida qo‘llaniladigan topshiriqlar yig‘indisi **test** deyiladi. Test bitta yoki bir nechta **test topshiriqlaridan** iborat bo‘lishi mumkin.

Test – baholanishi lozim bo‘lgan domen va konstrukt haqida xulosa chiqarishga asos bo‘ladigan topshiriq yoki topshiriqlar yig‘indisi.

Shu o‘rinda aniqlik kiritish lozim: test – bu baholanishi lozim bo‘lgan bilim, malaka, ko‘nikma yoki kompetensiya haqida xulosa chiqarishga asos bo‘ladigan har qanday topshiriq yoki topshiriqlar majmuyi. Bu topshiriqlar turlicha ko‘rinishda bo‘lishi mumkin:

- yopiq test – berilgan javob variantlaridan to‘g‘risini tanlashni talab qiladigan topshiriq;
- ochiq test – javob yozishni talab qiladigan topshiriq;
- amaliy test – bajarib ko‘rsatishni talab qiladigan topshiriq.

Test shunday tuzilishi lozimki, uni bajarish aynan **baholanishi lozim bo‘lgan bilim, malaka, ko‘nikma yoki kompetensiyani ishga solishni talab qilsin**. Masalan, haydovchilik guvohnomasini berishdan oldin o‘quvchining transport vositasini boshqarish ko‘nikmasini qay darajada egallaganini baholash uchun transport vositasini haydatib ko‘rish kerak, «men mashinani qanday boshqaraman?» mavzusida insho yozdirish yoki bayon qildirish orqali bu ko‘nikmani baholab bo‘lmaydi. Yoki o‘quvchining adabiyot fanidan qofiya va radifni farqlay olishini ko‘rmoqchi bo‘lsak, she‘r berib, qofiya va radifni topishni so‘rash mumkin.



Test vositasida domen haqida xulosa chiqarish jarayoni

Obyektiv xulosalar chiqara olishimiz va test topshiruvchilarni bir-biri bilan solishtira olishimiz uchun test topshiriqlari va ularni bajarish talablari (sharoitlari) barcha o'quvchilar (test topshiruvchilar) uchun **teng** bo'lishi lozim. Masalan, transport vositasini boshqarish ko'nikmasini baholayotganda bir o'quvchi uchun hech qanday svetoforlari va chorrahalari yo'q tekis trassa, boshqa o'quvchi uchun harakat tirband bo'lgan markaziy ko'cha tanlansa, berilgan topshiriqni bajarish shartlari o'rtasidagi bunday farqlar obyektiv xulosalar chiqarish imkonini bermaydi. Yoki ona tilidan insho yozish so'ralgan testda bir o'quvchiga berilgan vaqt boshqa o'quvchiga berilgan vaqtdan farq qilsa, bu ham o'quvchilarni bir-biri bilan solishtirish va obyektiv xulosalar chiqarishga xalal beradi.

Testning shartlari: baholanishi lozim bo'lgan domen (bilim, malaka, ko'nikma yoki kompetensiya) haqida xulosa chiqarishga imkon berishi va uni bajarish shart-sharoitlari barcha test topshiruvchilar uchun teng bo'lishi lozim.

Asosiy xulosa:

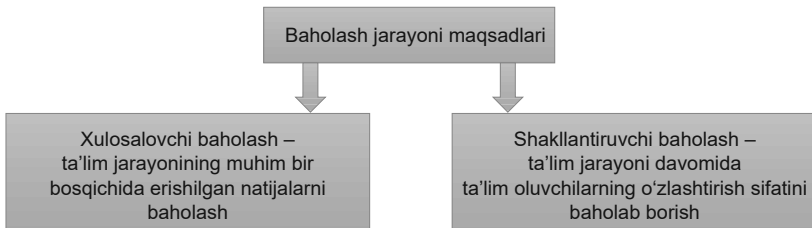
Baholash jarayonida cheklangan sondagi maxsus tuzilgan topshiriqlar (test topshiriqlari) yig'indisi (test) orqali ma'lum bir bilim, ko'nikma, malaka, kompetensiyalar majmuyining (domenning) test topshiruvchida qay darajada shakllangani haqida bilvosita xulosa chiqariladi.

1.2. Baholashning maqsadlari

Baholash jarayoni ma'lum bir **maqsadni** ko'zlab o'tkaziladi. Testdagi topshiriqlar soni, ularning tuzilishi, baholash mezonlari aynan baholashning maqsadidan kelib chiqqan holda tanlanadi. Bir maqsad uchun yaratilgan test boshqa bir maqsad uchun to'g'ri kelmasligi mumkin, masalan, o'quvchilarni chet tilini bilish darajasiga ko'ra guruhlarga ajratish maqsadida tuzilgan ko'p darajali (*multi-level*) testni chet tilini o'rganishning konkret bir darajasini baholash uchun ishlatib bo'lmaydi.

Yuqorida ko'rganimizdek, test – baholash jarayonida qo'llaniladigan vosita. Bu vosita yordamida ma'lum bir domen va bu domenga xos bo'lgan konstruktlar o'lchanadi, ular o'quvchida (test topshiruvchida) qay darajada shakllanganligi haqida xulosa chiqariladi. Bu xulosa esa, o'z navbatida, **biron-bir qaror qabul qilish uchun asos** bo'lib xizmat qiladi.

Baholash jarayonining maqsadidan kelib chiqqan holda ikki xil baholashni ajratib ko'rsatish mumkin: **xulosalovchi baholash** va **shakllantiruvchi baholash**.



Baholash jarayoni maqsadlari

Xulosalovchi baholash (o‘rganishni baholash) – ta’lim jarayonining ma’lum bir muhim bosqichida (masalan, modul yoki fan oxiri, semestr yoki o‘quv yili oxiri, ta’limning bir bosqichi nihoyasi yoki ta’limning keyingi bosqichiga o‘tishdan oldin) ta’lim oluvchining ta’limiy maqsadlarga qay darajada erishganligini o‘lchash jarayoni.

Xulosalovchi baholash natijasida chiqariladigan xulosalar asosida qabul qilinadigan qarorlar quyidagicha bo‘lishi mumkin:

1) Baho berish (monitoring): baholash natijasida chiqariladigan xulosalar o‘quvchi, o‘qituvchi, ta’lim yo‘nalishi, ta’lim muassasasi, ta’lim dasturi yoki butun boshli bir ta’lim tizimining faoliyatiga **baho berish** uchun ishlatilishi mumkin. Bunda ham olinadigan ma’lumot kim tomonidan ishlatilishiga qarab ikki xil baho berishni ajratib ko‘rsatish mumkin:

a) Ta’lim oluvchiga (o‘quvchiga, talabaga) baho berish: bunday baho berish, odatda, bevosita ta’lim berayotgan o‘qituvchi yoki ta’lim muassasasi tomonidan amalga oshiriladi va ta’lim oluvchining o‘quv dasturi yoki uning ma’lum bir qismini qay darajada o‘zlashtirgani haqida ma’lumot beradi. Bu ma’lumot ta’lim jarayonining turli ishtirokchilari tomonidan turli maqsadlarda ishlatilishi mumkin:

- ta’lim oluvchi tomonidan: fanni yoki modulni qayta o‘zlashtirish lozimligi haqida qaror qabul qilish maqsadida; bir ta’lim muassasasidan boshqasiga o‘tganda

o‘zlashtirilgan fanlar (modullar) to‘g‘risida ma’lumot taqdim qilish maqsadida va hokazo;

- ota-onalar tomonidan: farzandining o‘zlashtirishini kuza-
tib borish maqsadida; farzandiga qaysi fandan qo‘shim-
cha yordam kerakligini aniqlash maqsadida va hokazo;
- ta’lim muassasasi tomonidan: ta’lim oluvchini ta’lim-
ning keyingi bosqichiga o‘tkazish yoki o‘tkazmaslik
haqida qaror qabul qilish maqsadida; ta’lim oluvchiga
ta’lim muassasasini bitirganligi to‘g‘risida hujjat berish
yoki bermaslik haqida qaror qabul qilish maqsadida va
hokazo;

b) Ta’lim beruvchi (o‘qituvchi, ta’lim muassasasi), ta’lim yo‘nalishi (dasturi) yoki ta’lim tizimiga baho berish: bunday baho berish ichki (ta’lim muassasasi tomonidan o‘tkaziladigan) yoki tashqi (ta’lim sifatini nazorat qiluvchi davlat yoki nodavlat tashkilotlari tomonidan o‘tkaziladigan) bo‘lishi mumkin. Bunday baholash jarayoniga barcha ta’lim oluvchilar qamrab olinmasdan, ta’lim oluvchilarning ma’lum bir qismini baholash orqali o‘qituvchi, ta’lim muassasasi, ta’lim yo‘nalishi yoki butun boshli bir ta’lim tizimi qay dara-
jada samarali ishlayotgani haqida xulosalar chiqarilishi mum-
kin. Bu xulosalar esa, o‘z navbatida, qandaydir xatti-harakat-
lar uchun asos bo‘lib xizmat qiladi. Masalan:

- ta’lim beruvchilar (o‘qituvchilar, ta’lim muassasasi, ta’-
lim muassasasi rahbariyati) faoliyati haqida xulosalar
chiqarish;

- ta’lim dasturlariga o’zgartirishlar kiritish;
- ta’lim muassasasi faoliyatiga o’zgartirish kiritish yoki uni yopish;
- ta’lim tizimini isloh qilish va hokazo.

2) Tanlash: baholash natijasida chiqariladigan xulosalar test topshiruvchilarni **tanlash (saralash)** uchun ishlatiladi. Bu tanlash asosan ikki xil bo‘lishi mumkin:

a) Belgilangan talabga javob beradiganlarni tanlash: bunda minimal talablar o‘rnatiladi va baholashning maqsadi shu talablarga mos keladiganlarni ajratib olish bo‘ladi. Belgilangan minimal talablarga javob beradiganlarning barchasiga qandaydir hujjat (masalan, ta’limni muvaffaqiyatli tugatganligini tasdiqlovchi hujjat – diplom, guvohnoma, biron-bir faoliyat bilan shug‘ullanishga ruxsat beradigan hujjat – sertifikat, litsenziya) yoki huquq (masalan, ta’limning keyingi bosqichiga o‘tish huquqi) beriladi. Aksar xorij davlatlaridagi (AQSh, Yevropa, Janubi-Sharqiy Osiyo) oliy ta’lim muassasalari, odatda, bo‘lajak talabalarga qo‘yiladigan minimal talablarni e’lon qilishadi va shu talablarga javob beradigan barcha abituriyentlarni o‘qishga qabul qilishadi;

b) Cheklangan kvota sharoitida tanlash: bunday holat bir o‘ringa bir necha talabgor da’vo qilayotganda yuzaga keladi (masalan, oliy o‘quv yurtiga o‘qishga qabul qilish, xorijga malaka oshirishga yuborish uchun tanlash, bo‘sh ish o‘rniga nomzod tanlash va hokazo). Bunda test topshiruvchilarni bir-biriga solishtirish asosida qaror qabul qilinadi, ya’ni test

topshiruvchilarning ichidan eng yuqori natija ko'rsatganlari tanlab olinadi.

Shakllantiruvchi baholash (o'rganish uchun baholash) – ta'lim jarayoni davomida ta'lim oluvchilarning o'zlashtirish sifatini baholab borish.

Shakllantiruvchi baholash natijasida olinadigan xulosalar ta'lim jarayonini ta'minlash (ta'lim oluvchiga va o'qituvchiga ta'lim jarayoni qay darajada samarali kechayotgani haqida ma'lumot berish, keyingi qadamlarini belgilash) uchun ishlatiladi. Masalan:

- muntazam ravishda sinfdan o'quvchini baholab borish orqali u qaysi bilim, ko'nikma, malaka yoki kompetensiyalarni yaxshi o'zlashtirgani yoki o'zlashtirmagani, nimalar ustida ishlashi kerakligi haqida ma'lumot beriladi;
- biron-bir mavzu oxirida nazorat ishi olish orqali o'qituvchi o'quvchilari shu mavzuni qay darajada o'zlashtirganliklari, qanday xatolarga yo'l qo'yayotganlari haqida ma'lumot olishi va lozim bo'lsa, o'tilgan mavzuga yana qaytishi mumkin;
- o'qituvchi o'quvchilarning boshlang'ich bilim darajasini aniqlash (diagnostika) maqsadida test o'tkazishi va olingan ma'lumotlar asosida keyingi qadamlarini rejalashtirishi yoki o'quvchilarni guruhlariga ajratishi mumkin.

Demak, shakllantiruvchi baholash ta'lim beruvchilarga (o'qituvchilarga) va ta'lim oluvchilarga (o'quvchilarga, ta-

labalarga) asosan **jarayonni** baholash va lozim bo'lgan o'rinlarda unga o'zgartirish kiritish imkonini bersa, xulosalovchi baholash **natijani** baholash uchun ishlatiladi.

<i>Maqsad</i>	<i>Baholash natijasida olinadigan ma'lumot kim tomonidan ishlatilishi</i>	<i>Baholashning maqsadi</i>	<i>Misollar</i>
Baho berish (monitoring)			
Ta'lim oluvchiga baho berish	– Ta'lim oluvchi (o'quvchi, talaba) – Ta'lim muassasasi – Ota-ona	Ta'lim oluvchining o'quv dasturi yoki uning ma'lum bir qismini qay darajada o'zlashtirganini aniqlash	– Fan (modul)ni o'qitish yakunida topshiriladigan imtihon – Ta'lim muassasasini bitirish uchun topshiriladigan yakuniy davlat attestatsiyasi
Ta'lim beruvchiga, ta'lim dasturiga yoki ta'lim tizimiga baho berish	– Ta'lim muassasasi – Tashqi organlar	O'qituvchilar, ta'lim muassasasi, ta'lim dasturi, ta'lim tizimiga baho berish	Ta'lim sifatini nazorat qilish davlat inspeksiya tomonidan o'tkaziladigan attestatsiya
Tanlash			
Cheklangan kvota sharoitida tanlash	– Ta'lim muassasasi – Ish beruvchi – Grant beruvchi tashkilotlar	Nomzodlar ichidan eng yaxshilarini tanlab olish	Davlat test markazi tomonidan o'tkaziladigan oliy ta'lim muassasalariga saralash test sinovlari

Belgilangan talabga javob beradiganlarni tanlash	– Litsenziya va sertifikat beruvchi tashkilotlar – Ta’lim muassasasi	Nomzodlar ichidan belgilangan minimal talablarga javob beradiganlarni ajratib olish	– Davlat test markazi tomonidan o‘tkaziladigan chet tili bo‘yicha bilim dajasini aniqlash imtihonlari – Haydovchilik guvohnomasini olish uchun topshiriladigan imtihon
--	---	---	---

Xulosalovchi baholashning maqsadlari

<i>Maqsad</i>	<i>Baholash natijasida olinadigan ma’lumot kim tomonidan ishlatilishi</i>	<i>Baholashning maqsadi</i>	<i>Misollar</i>
Ta’lim jarayonini ta’minlash	– Ta’lim oluvchi – O‘qituvchi	Keyingi qadamlarni rejalashtirish	– Dars boshi yoki oxirida olinadigan testlar – Mavzu oxirida olinadigan nazorat ishlari – O‘quvchilarning boshlang‘ich bilim darajasini aniqlash uchun o‘tkaziladigan diagnostik testlar

Shakllantiruvchi baholashning maqsadi

Shakllantiruvchi baholash samarali bo'lishi uchun xulosalovchi baholashdan quyidagi jihatlari bilan farq qilishi lozim:

– xulosalovchi baholash asosan rasmiy shaklda (imtihon, sinov, kurs ishi topshirish va hokazo) o'tkazilsa, shakllantiruvchi baholash bevosita o'qituvchi tomonidan asosan **norasmiy shakllarda** (kuzatish, og'zaki savol-javob, kichik hajmdagi testlar, yozma ishlar va hokazo) o'tkaziladi;

– xulosalovchi baholashdan keyin, odatda, aynan shu ta'lim oluvchilar guruhi uchun ta'lim jarayoniga o'zgartirish kiritish imkoni mavjud bo'lmaydi (lekin xulosalovchi baholashdan olingan ma'lumotlar natijasida keyingi ta'lim oluvchilar guruhi uchun bunday o'zgarishlarni amalga oshirish mumkin), shakllantiruvchi baholash jarayonida olingan ma'lumotlar natijasida o'qituvchi ta'lim jarayoniga **kerakli o'zgartirishlar kirita oladi** (masalan, o'quvchilar bo'sh o'zlashtirgan mavzuni qayta o'tishi mumkin);

– xulosalovchi baholashdan keyin, odatda, qayta aloqa, fidbek (inglizcha «feedback»), ya'ni tahlil bo'lmaydi – test topshiruvchilarning natijalari e'lon qilinadi, lekin nega shunday natijalar chiqqaniga izoh berilmaydi yoki izoh berilsa ham, test topshiruvchi bu izohni keyingi qadamlarni rejalashtirish uchun ishlata olmaydi, shakllantiruvchi baholashda esa o'qituvchi ta'lim oluvchiga **izoh (tushuntirish) berishi lozim** – shunda ta'lim oluvchi o'z yutuq va kamchiliklarini anglab, keyingi qadamlarini rejalashtira oladi¹.

¹ Rowntree D. Assessing Students: How Shall We Know Them. – London: Kogan Page, 1987. – P.24.

Test tuzayotganda, uni ishlatayotganda va uning natijalarini talqin qilib, bu natijalar asosida biror-bir qarorga kelinayotganda, birinchi navbatda, baholashning maqsadini aniqlashtirib olish lozim. Umuman, har qanday testni ishlatishda quyidagi uch savolga¹ javob berish lozim:

1. Baholashning asosiy **maqsadi** nimadan iborat?
2. Baholashning **shakli** maqsadga mosmi?
3. Baholashdan ko‘zlangan maqsadga **erishildimi?**

Asosiy xulosa:

Xulosalovchi baholash asosan ta’lim oluvchiga yoki ta’lim oluvchi orqali ta’lim beruvchiga (o‘qituvchi, ta’lim muassasasi, ta’lim yo‘nalishi, ta’lim tizimi) baho berish (faoliyatini monitoring qilish) hamda ta’lim oluvchilarni saralash maqsadida rasmiy shaklda (imtihon, sinov, kurs ishi yoki diplom ishi topshirish va hokazo) o‘tkaziladi.

Shakllantiruvchi baholash asosan ta’lim jarayoni davomida bevosita o‘qituvchi tomonidan turli shakllarda o‘tkaziladi, bunday baholash natijasida olingan ma’lumot o‘qituvchi va ta’lim oluvchiga keyingi qadamlarini rejalashtirish uchun yordam beradi.

¹ Stobart G. Testing Times: The Uses and Abuses of Assessment. – London-New York: Routledge, 2008. – P.14.

1.3. Test turlari

Test bir yoki bir nechta test topshiriqlaridan iborat bo‘ladi. Test topshiruvchidan qanday harakatni amalga oshirish talab qilinayotganiga qarab test topshiriqlari ikki guruhga bo‘linadi: **yopiq test topshirig‘i** va **ochiq test topshirig‘i**.

Yopiq test topshirig‘i – berilgan javob variantlari ichidan to‘g‘risini tanlashni talab qiladigan test topshirig‘i.

Yopiq test topshiriqlari bir qancha shakllarda bo‘lishi mumkin:

- muqobil javobli test topshirig‘i;
- «to‘g‘ri-noto‘g‘ri» shaklidagi test topshirig‘i;
- moslashtirish;
- ketma-ketlikni topish.

Ochiq test topshirig‘i – test topshiruvchi javobni o‘zi yaratishi (yozishi, og‘zaki javob berishi, ijro etishi, bajarib ko‘rsatishi va hokazo) talab qiladigan test topshirig‘i.

Ochiq test topshiriqlari asosan uch shaklda bo‘lishi mumkin:

- qisqa javobni talab qiladigan test topshirig‘i;
- kengaytirilgan javobni talab qiladigan test topshirig‘i;
- amaliy test topshirig‘i (test topshiruvchi biron-bir mahsulot yaratishi yoki biron-bir amalni bajarib ko‘rsatishi talab qilinadigan test topshirig‘i).

Baholashning maqsadidan kelib chiqqan holda turli testlar qo'llanilishi mumkin. Testlarni tasniflashda (guruhlarga ajratishda) ularning har xil xususiyatlari asos qilib olinishi mumkin, shunga ko'ra, test turlari ham har xil bo'lishi mumkin.

Test natijalarini tahlil qilish usuliga ko'ra barcha testlar uch guruhga bo'linadi: **normaga mo'ljallangan test**, **mezon-ga mo'ljallangan test** va **ipsativ test**.

Normaga mo'ljallangan test – test topshiruvchining natijalari normalar (shu yoshdagi yoki o'rganishning shu bosqichidagi test topshiruvchilarning yirik guruhi natijalari) bilan solishtiriladi.

Normaga mo'ljallangan testda test topshiruvchi erishgan natija xuddi shu testni yoki aynan shu domenni baholaydigan boshqa testni topshirgan yirik guruhlar (**normativ guruhlar**) natijalari bilan solishtirilib, test topshiruvchining bilim darajasi boshqalarga nisbatan qanday ekanligi haqida xulosa chiqariladi. Masalan, Aliyev matematika bo'yicha testda 70 ta test topshirig'idan 33 tasini to'g'ri yechgan bo'lsa, bu natija xuddi shu (yoki shunday) testni topshirgan normativ guruh (yoshi, bilim darajasi jihatidan Aliyevga o'xshagan boshqa test topshiruvchilar) natijalari bilan solishtiriladi va Aliyevning matematika bo'yicha bilim darajasi boshqalarga nisbatan yuqoriroq, pastroq yoki o'xshash degan xulosaga kelinadi.

Normaga mo'ljallangan testda test topshiruvchi bo'yicha qabul qilinadigan qarorga boshqalarning natijalari ta'sir qiladi. Masalan, nufuzli oliy ta'lim muassasasi abituriyentlarning ichidan eng yuqori natija ko'rsatgan 20 foizini o'qishga qabul qilmoqchi bo'lsa, normaga mo'ljallangan test ishlatishi mumkin. Bunday holda test topshiruvchining o'qishga qabul qilinish yoki qilinmasligiga boshqa test topshiruvchilar qanday natijalarga erishganligi ham ta'sir qiladi.

Mezonga mo'ljallangan test – test topshiruvchining natijalari oldindan belgilangan mezonlar bilan solishtiriladi.

Mezonga mo'ljallangan testda maqsad test topshiruvchining baholanayotgan domenni qaysi darajada o'zlashtirganligini aniqlashdir. Bunda test topshiruvchining natijalari boshqalarning natijalari bilan solishtirilmaydi va test topshiruvchi bo'yicha chiqariladigan qarorga boshqalarning natijalari ta'sir qilmaydi. Masalan, ta'lim muassasasi ta'lim jarayoni so'nggida belgilangan talablarga javob beradigan o'quvchilarga diplom bermoqchi bo'lsa, mezonga mo'ljallangan test ishlatishi mumkin. Bunda bitiruvchilarning natijalari bir-biri bilan solishtirilmaydi, balki oldindan belgilab qo'yilgan mezonlar bilan solishtirilib, ularga diplom berish yoki bermaslik masalasi hal qilinadi.

Normaga mo'ljallangan testda test topshiruvchi **boshqa test topshiruvchilarga nisbatan** baholansa, mezonga mo'l-

jallangan testda test topshiruvchining ma'lum bir **domendan aynan nimani bilish-bilmasligi** (qila olish-olmasligi) aniqlanadi.

Normaga mo'ljallangan testlar asosan **solishtirish** uchun qo'llaniladi. Bu solishtirish turlicha bo'lishi mumkin: alohida olingan test topshiruvchilarni bir-birlari bilan yoki o'rtacha statistik normalar bilan solishtirish, shuningdek, test topshiruvchilar guruhlarini bir-birlari bilan yoki o'rtacha statistik normalar bilan solishtirish. Bunday solishtirish natijasida turli qarorlar qabul qilinishi, masalan, cheklangan kvotaga da'vo qilayotganlarning ichidan kuchliroq nomzodlarni tanlab olish mumkin.

Mezonga mo'ljallangan testlar esa test topshiruvchilar belgilangan **ta'limiy maqsadlarga qay darajada erishganlarini baholash** uchun qo'llaniladi. Bunday testlarni minimal talabga javob beradiganlarni tanlash, ta'lim muassasalarini attestatsiyadan o'tkazish, ta'lim oluvchilarni diagnostika qilish singari maqsadlarda ishlatish mumkin¹. Mezonga mo'ljallangan testlar ham, o'z navbatida, maqsadiga ko'ra ikki turga bo'linadi: **domenga mo'ljallangan testlar** va «**mastery**» testlar. **Domenga mo'ljallangan testlar** test topshiruvchining tekshirilayotgan **domening qancha qismini** o'zlashtirganligini tekshirishga qaratilgan. Misol uchun, test topshiruvchi bunday testni 100 foiz bajarsa, domenni 100 foiz o'zlashtirgan deyishimiz

¹ Кирейцева А.Н. Азбука тестирования. – Санкт-Петербург: Златоуст, 2013. – С.29.

mumkin. «**Mastery**» testlar («O‘zlashtirish saviyasi» testlari) esa test topshiruvchilarni **ikki yoki undan ortiq guruhlarga ajratish** (masalan, «biladiganlar – bilmaydiganlar», «yomon – o‘rtacha – yaxshi») uchun ishlatiladi.

Normaga mo‘ljallangan va mezonga mo‘ljallangan testlarni quyidagicha solishtirish mumkin¹:

	Normaga mo‘ljallangan testlar	Mezonga mo‘ljallangan testlar
1. Asosiy maqsadi	Test topshiruvchilarning o‘lchanayotgan domenni egalaganlik darajasini solishtirish	Test topshiruvchilarning o‘lchanayotgan domendan aynan nimani bilish-bilmasligini (qila olishi-qila olmasligini) aniqlash
2. Natijalarni qayta ishlash	Test topshiruvchining natijasi boshqa test topshiruvchilarning natijalari bilan solishtiriladi	Test topshiruvchining natijasi oldindan belgilab qo‘yilgan mezonlar bilan solishtiriladi
3. Qamrab olish diapazoni	Odatda, bilim, ko‘nikma, malaka, kompetensiyalarning keng diapazonini qamrab oladi	Odatda, bilim, ko‘nikma, malaka, kompetensiyalarning tor diapazonini qamrab oladi
4. Test topshiriqlarini tanlash	Haddan tashqari oson yoki haddan tashqari qiyin test topshiriqlari kiritilmaydi	Test topshiriqlarining oson-qiyinligi muhim rol o‘ynamaydi
5. Bajarish standartlari	Test topshiruvchining guruhga nisbatan o‘rni belgilanadi (masalan, kuchli 10 foizlikka kiradi)	Test topshiruvchining domenga nisbatan darajasi belgilanadi (masalan, 90 foiz materialni biladi)

¹ Майоров А.Н. Теория и практика создания тестов для системы образования: как выбирать, создавать и использовать тесты для целей образования. – Москва: Интеллект-центр, 2002. – С.39.

Ipsativ test – test topshiruvchining natijalari o‘zining shu domenni o‘lchaydigan oldingi test natijalari bilan solishtiriladi.

Ipsativ testda maqsad test topshiruvchining ma’lum bir vaqt oralig‘ida erishgan **yutuqlarini** aniqlash, ta’limdagi progressini o‘lchash hisoblanadi. Masalan, bir o‘quvchining ta’lim dasturini boshlashdan oldingi bilim darajasiga baho berish uchun kirish testi beriladi. Ta’lim dasturi yakunlangandan keyin xuddi shu o‘quvchi bajargan chiqish testi bilan kirish testi natijalari solishtirilib, o‘quvchi nimaga erishganligi haqida xulosaga kelinadi. Ayrim davlatlarning ta’lim sifati monitoringi tizimlarida maktabning o‘quvchiga bergan qo‘shimcha qiymatini (ya’ni o‘quvchining oldingi bilimlari va maktabda egallagan bilimlari o‘rtasidagi farqni) aniqlash va shu orqali maktab faoliyatiga baho berish maqsadida ipsativ testlardan foydalaniladi.

Test natijalarining qo‘llanish usuliga ko‘ra barcha testlar to‘rt guruhga bo‘linadi: **muvaffaqiyat testlari, egallash testlari, prognostik testlar (layoqatni aniqlash testlari)** va **diagnostik testlar**.

Muvaffaqiyat testlari test topshiruvchining konkret o‘quv materialini qay darajada o‘zlashtirganligini baholash uchun ishlatiladi va test topshiriqlari ta’lim jarayonida qo‘llanilgan o‘quv materiallari asosida tuziladi.

Masalan, fanni o‘qitish yakunida o‘qituvchi o‘quvchilari-ning fanni qay darajada o‘zlashtirganligini tekshirish maqsadida ta’lim jarayonida ishlatilgan darslik asosida test olishi mumkin.

Egallash testlari test topshiruvchining ma’lum bir ko‘nikma yoki kompetensiyalarni qay darajada egallaganligini baholash uchun ishlatiladi va test topshiriqlari ta’lim jarayonida qo‘llanilgan o‘quv materiallariga bog‘liq bo‘lmaydi.

Masalan, chet tilini bilish darajasini baholovchi xalqaro imtihonlar (IELTS, TOEFL singari) test topshiruvchining til ko‘nikmalarini (tinglab tushunish, o‘qish, yozish, gapirish) baholaydi va bunda test topshiruvchining imtihonga tayyorlanish uchun qaysi manbadan foydalanganligi hech qanday ahamiyatga ega emas.

Prognostik testlar (layoqatni aniqlash testlari) test topshiruvchining ma’lum bir sohaga (masalan, chet tilini o‘rganishga, matematik mantiqqa) layoqatini aniqlash uchun ishlatiladi.

Prognostik testlar, odatda, prognoz (bashorat) qilish maqsadida qo‘llaniladi: masalan, chet tili kursini boshlashdan oldin o‘qituvchi kimlarda til o‘rganishga layoqat borligini aniqlab olishi mumkin.

Diagnostik testlar test topshiruvchining bilimlaridagi bo'shliqlarni aniqlash uchun ishlatiladi.

Masalan, o'quvchini biron-bir imtihonga tayyorlayotgan o'qituvchi diagnostik test yordamida o'quvchining nimalarni bilmasliklarini aniqlab, shu asosda o'qitishni rejalashtirishi mumkin.

Testni bajarish uchun ajratilgan vaqtga ko'ra testlar uch guruhga bo'linadi: **tezkor test**, **kuchli test** va **aralash test**.

Tezkor test – bunda test topshirish vaqti qat'iy cheklangan bo'ladi, testga qiyinchilik darajasi bir xil (yoki bir-biriga juda yaqin) bo'lgan topshiriqlar kiritiladi, odatda, barcha topshiriqlarga bir xil ball beriladi, test topshiruvchining natijasi to'g'ri javoblar soniga bog'liq bo'ladi.

Kuchli test – bunda test topshirish vaqti qat'iy cheklangan bo'lmaydi, testga murakkab topshiriqlar kiritiladi va test topshiruvchining natijasi nafaqat to'g'ri javoblar soni, balki topshiriqlarning qanday yechilganligiga ham bog'liq bo'ladi.

Aralash test – bunday testlarda tezkor test va kuchli testlarning xususiyatlari birlashtiriladi: testga turli murakkablikdagi topshiriqlar kiritiladi va test uchun ajratilgan vaqt cheklangan bo'ladi, test topshiruvchining natijasi to'g'ri javoblar soniga bog'liq bo'ladi.

Qamrab olingan fan sohasiga ko'ra testlar ikki guruhga bo'linadi: **gomogen test** va **geterogen test**.

Gomogen test – faqat bir fan sohasi bo‘yicha test topshiriqlaridan iborat test.

Geterogen test – bir nechta turdosh fanlar sohalari bo‘yicha test topshiriqlaridan iborat test.

Kiritilgan test topshiriqlari turiga ko‘ra testlar ikki guruhga bo‘linadi: **monomorf test** va **polimorf test**.

Monomorf test – faqat bir turdagi test topshiriqlarini o‘z ichiga olgan test.

Polimorf test – har xil turdagi test topshiriqlarini o‘z ichiga olgan test.

Topshirish ssenariysiga ko‘ra testlar ikki guruhga bo‘linadi: **chiziqli test** va **adaptiv test**.

Chiziqli test – barcha test topshiruvchilarga bir xil test topshiriqlari oldindan belgilangan ketma-ketlikda beriladi.

Adaptiv test – har bir keyingi test topshirig‘i test topshiruvchining oldingi savolga bergan javobiga bog‘liq bo‘ladi: masalan, test topshiruvchi noto‘g‘ri javob bersa, keyingi savol osonroq, to‘g‘ri javob bersa, keyingi savol qiyinroq bo‘ladi. Adaptiv testlar kompyuterda olinadi va ularni o‘tkazish uchun maxsus dasturiy ta‘minot kerak bo‘ladi.

Bitta test qaysi jihati tasniflash uchun asos qilib olinganiga qarab yuqoridagi tasniflarning barchasiga kirishi mumkin. Masalan, Davlat test markazi tomonidan o'tkaziladigan oliy ta'lim muassasalari bakalavriatiga kirish test sinovlari:

- normaga mo'ljallangan test, chunki test topshiruvchilarning natijalari normalar bilan solishtirilib, shu asosda abituriyentlarning ichidan nisbatan kuchlilari aniqlanadi;
- muvaffaqiyat testi, chunki test materiallari abituriyentning o'rta yoki o'rta maxsus ta'lim o'quv dasturlarini qay darajada egallaganligini tekshiradi;
- aralash test, chunki test topshiriqlari turli murakkablikda bo'ladi va bu topshiriqlar belgilangan vaqt ichida bajarilishi lozim;
- gomogen test (to'g'rirog'i, gomogen testlar yig'indisi), chunki test bir necha fanlar blokini o'z ichiga oladi va har bir blok bir fan bo'yicha topshiriqlarni o'z ichiga oladi;
- monomorf test, chunki faqat yopiq test topshiriqlari ishlatiladi;
- chiziqli test, chunki test topshiriqlari ketma-ketligi oldindan belgilangan bo'ladi.

<i>Test natijalarini tahlil qilish usuliga ko'ra</i>	<i>Test natijalarining qo'llanish usuliga ko'ra</i>	<i>Testni bajarish uchun ajratilgan vaqtga ko'ra</i>	<i>Qamrab olingan fan sohasiga ko'ra</i>	<i>Kiritilgan test topshiriqlari turiga ko'ra</i>	<i>Topshirish ssenariysiga ko'ra</i>
<ul style="list-style-type: none"> – Normaga mo'ljallangan test – Mezonga mo'ljallangan test – Ipsativ test 	<ul style="list-style-type: none"> – Muvaffaqiyat testlari – Egallash testlari – Prognostik testlar (layoqatni aniqlash testlari) – Diagnostik testlar 	<ul style="list-style-type: none"> – Tezkor test – Kuchli test – Aralash test 	<ul style="list-style-type: none"> – Gomojen test – Geterojen test 	<ul style="list-style-type: none"> – Monomorf test – Polimorf test 	<ul style="list-style-type: none"> – Chiziqli test – Adaptiv test

Testlarning turlari

Asosiy xulosa:

Testlar turli xususiyatlariga ko'ra tasniflanishi hamda yopiq test topshiriqlari, ochiq test topshiriqlari va amaliy test topshiriqlaridan iborat bo'lishi mumkin.

2. TESTNING SIFATI: VALIDLIK

2.1. Validlik tushunchasi va uning tadriji

Oldingi bobda ko'rganimizdek, ta'lim jarayonida test biron-bir domenni, shu domen o'z ichiga olgan konstrukt-larni o'lchash uchun qo'llaniladi. Ta'lim oluvchining bilim darajasini, unda ko'nikma, malaka, kompetensiyalarning shakllanganlik darajasini bevosita ko'rish imkoni mavjud emas: o'qituvchi ularni faqat bilvosita (savol-javob, ku-zatish, topshiriqlar bajartirish orqali) ko'rishi, o'lchashi mumkin. Bunday o'lchash natijasida chiqarilgan xulosalar asosida esa biron-bir shaxs yoki tashkilot (masalan, o'qituv-chi, ta'lim muassasasi, ish beruvchi) tomonidan ma'lum bir qarorlar (masalan, diplom berish yoki bermaslik, o'qishga yoki ishga qabul qilish yoki qabul qilmaslik va hokazo) qa-bul qilinadi.

Qabul qilinadigan qarorlar asoslantirilgan bo'lishi uchun test asosida chiqariladigan xulosalar ham **valid (asosli)**¹ bo'li-shi lozim.

¹ «Valid» so'zi lotincha «validus» so'zidan olingan bo'lib, «kuchli; samarali; yaroqli; qonuniy» ma'nolarini bildiradi. Testologiyada bu atama tarjima qilinmagan holda ishlatiladi.

Validlik – testning sifatini belgilaydigan eng asosiy tushuncha. Validlik deganda testning maqsadidan kelib chiqqan holda test natijalari yordamida chiqariladigan xulosalarning qay darajada ilmiy va empirik jihatdan asosli ekanligi tushuniladi.

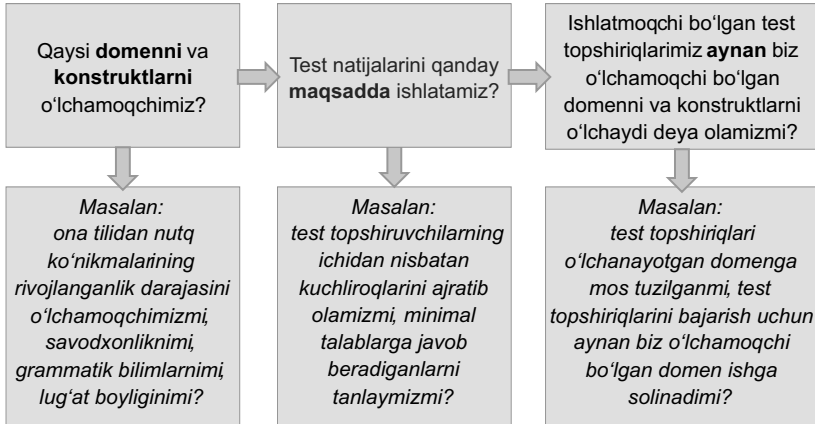
Demak, test natijalari asosida chiqarilgan xulosalar va ushbu xulosalar asosida qabul qilingan qarorlar asoslangan bo‘lishi lozim.

Masalan, ona tilidan morfologiya bo‘yicha yopiq shakldagi test topshiriqlarini yaxshi bajargan o‘quvchi haqida qanday xulosa chiqarishimiz mumkin? U tilshunoslikning morfologiya bo‘limiga tegishli qonun-qoidalarni yaxshi o‘zlashtirgan deya olamiz. U o‘zlashtirgan qoidalarini og‘zaki yoki yozma nutqda fikrini to‘g‘ri ifodalash uchun qo‘llay oladi deya olamizmi? Bu – bahsli masala. Boshqa bir misol. Nemis tilida B1 (quyi o‘rta) darajadagi matnlarni o‘qib tushunish bilan bog‘liq test topshiriqlarini a’lo darajada bajargan o‘quvchi B2 (yuqori o‘rta) darajadagi matnlarni ham tushunar ekan degan xulosaga kelishimiz mumkinmi?

Har qanday testdan oldin testni o‘tkazish uchun mas’ul shaxslar yoki tashkilotlar test bilan bog‘liq bir qancha jihatlarni aniqlab olishi lozim bo‘ladi, jumladan:

- qaysi **domenni** va **konstruktlarni** o‘lchamoqchimiz?
Bu domen va konstruktlar o‘z ichiga nimalarni oladi?
- test natijalari qanday **maqsadda** ishlatiladi?

- ishlatmoqchi bo‘lgan test topshiriqlarimiz aynan biz o‘lchamoqchi bo‘lgan domenni va konstruktarni o‘lchaydi deyishimiz uchun qanday **ilmiy dalillar** mavjud?



Testdan oldin aniqlanishi lozim bo‘lgan jihatlar

Misol uchun, ta’lim ingliz tilida olib boriladigan universitetga ingliz tili bo‘lmagan abituriyentlar hujjat topshirishdi. Universitet qabul komissiyasi abituriyentlarni o‘qishga qabul qilishdan oldin ularning ingliz tilida ma’ruzalar eshittish, ilmiy adabiyotlar bilan ishlash, kurs ishlari yozish, professor-o‘qituvchilar bilan muloqot olib borishga qay darajada tayyor ekanligini baholashi lozim. Buning uchun universitet abituriyentlarga ingliz tilida ilmiy matnlarni tinglab va o‘qib tushunish, ilmiy mavzularda fikrini og‘zaki va yozma bayon qilish qobiliyatini baholaydigan test berishi va abituriyentlarning tilni bilish darajasiga minimal talablar belgilab, bu

minimal talablarga javob beradiganlarni ajratib olishi kerak bo‘ladi. Bunday test natijalari universitetga qaysi abituriyent ta’limni ingliz tilida olib borishga tayyor ekanligi haqida asosli – valid xulosalar chiqarish imkonini beradi. Agar universitet abituriyentlarga ingliz tili grammatikasi bo‘yicha bilimlarni tekshirishga qaratilgan test bersa, bunday test abituriyentlarning ingliz tilidagi grammatik qoidalarni qay darajada o‘zlashtirgani haqida xulosa chiqarish imkonini beradi, lekin abituriyentlarning ingliz tilida ilmiy matnlarni tushunish yoki ilmiy mavzularda fikrlarini bayon qila olish qobiliyati haqida xulosa chiqarish imkonini bermaydi.

Boshqa bir misol: bo‘lajak tilshunoslar tayyorlaydigan ta’lim muassasasi talabalar ingliz tili grammatikasi kursini o‘qib bo‘lishganidan so‘ng ularning grammatikaga oid bilimlarni qay darajada o‘zlashtirganini baholamoqchi bo‘lsa, bunday vaziyatda ularga ingliz tili grammatikasi bo‘yicha bilimlarni tekshirishga qaratilgan test berish eng maqbul yo‘l sanaladi.

Demak, test natijasida chiqariladigan **xulosalar** qay darajada **valid (asosli)** bo‘lishi testning **maqsadiga** bog‘liq va bir vaziyatda valid xulosalar chiqarish imkonini beradigan test boshqa bir vaziyatda bunday imkonni bermasligi mumkin. Test natijasida chiqarilgan xulosalarning qay darajada asosli ekanligini aniqlashda testning maqsadi asosiy rol o‘ynaydi. Agar bitta test har xil maqsadlarda, har xil xulosalar chiqarish va har xil qarorlar qabul qilish maqsadida

qo‘llaniladigan bo‘lsa, testning har bir qo‘llanilishi natijasida chiqariladigan xulosalar asosli ekanligiga ishonch hosil qilinishi lozim.

Yuqorida ko‘rganimizdek, test, odatda, biror-bir qaror qabul qilish uchun ishlatiladi, ya’ni test natijalari asosida ta’lim oluvchilar (o‘quvchilar, talabalar), ta’lim beruvchilar (o‘qituvchilar, ta’lim muassasasi), hatto ta’lim tizimi uchun muhim bo‘lgan qarorlar qabul qilinadi. Bu qarorlar to‘g‘ri bo‘lishi uchun test natijasida chiqargan xulosalarimiz valid (asosli) bo‘lishi kerak. Xulosalar valid (asosli) deya olishimiz uchun **dalillar** kerak. Bu dalillar quyidagicha bo‘lishi mumkin:

- testning **mazmuni** bilan bog‘liq dalillar;
- test topshiruvchilarning test topshiriqlarini **bajarishi jarayoni** bilan bog‘liq dalillar;
- testning **ichki strukturasi** bilan bog‘liq dalillar;
- testning boshqa **o‘zgaruvchilar** bilan bog‘liqligi asosidagi dalillar;
- testning **oqibatlari** bilan bog‘liq dalillar¹.

Bu dalillar **validatsiya** jarayonida yig‘iladi.

Validatsiya – testdan chiqariladigan xulosalar asosli (valid) ekanligini tasdiqlovchi ilmiy dalillar keltirish jarayoni.

¹ Ovendahl N.V. Testwise: Understanding Educational Assessment. Vol.1. – Plymouth: Rowman & Littlefield Education, 2011. – P.87.

Demak, validlik:

– testning xususiyati emas: testning o‘zi valid (asosli) yoki novalid (asossiz) bo‘lmaydi, balki test natijasida chiqarilgan **xulosalar** valid (asosli) yoki novalid (asossiz) bo‘lishi mumkin;

– test natijalari **interpretatsiyasi (talqini)ga** bog‘liq – nima maqsadda test olgan edik va test natijalari nimani ko‘rsatyapti deb o‘ylaymiz?

– **nisbiy** tushunchadir: test natijasida chiqarilgan xulosalarimiz **ma‘lum bir darajada** valid yoki novalid bo‘lishi mumkin, mutaxassislarning vazifasi esa bu xulosalarning validligi imkon qadar yuqori bo‘lishini ta‘minlashdan iborat.

Validlik pedagogik o‘lchovlarning (testologiyaning) eng asosiy tushunchasi bo‘lishiga qaramay, bu tushunchaning ta‘ri-fi bo‘yicha mutaxassislar orasida yakdillik bo‘lmagan – fan-ning rivojlanishi davomida ta‘riflar o‘zgarib borgan. Validlik tushunchasi tadrijini bir qancha bosqichlarga bo‘lish mumkin:

- 1920 – 1950-yillar – «validlik» tushunchasining paydo bo‘lishi va rivojlanishi;
- 1950 – 1970-yillar – «validlik» tushunchasining frag-mentatsiyalanishi (parchalanishi);
- 1970 – 2000-yillar – «validlik» tushunchasining bir-lashtirilishi;
- 2000-yildan keyin – «validlik» tushunchasining zamo-naviy talqini¹.

¹Qarang: Newton P.E. and Shaw S.D. Validity in Educational and Psychological Assessment. – London: Sage, 2014.

«Validlik» tushunchasi dastlab paydo bo‘lganda uning ta’rifi quyidagicha bo‘lgan: «validlik testning o‘lchanishi lozim bo‘lgan xususiyatni qay darajada o‘lchashidir». Ushbu ta’rifdan ko‘rinib turibdiki, validlik dastlab testning xususiyati sifatida qaralgan, ya’ni testning o‘zi valid yoki novalid bo‘lishi mumkin deb hisoblangan va ushbu qarash 1970 – 1980-yillargacha davom etgan.

1950-yillarga kelib validlikning uch turi mavjud degan qarash shakllandi:

– mezoniy (kriterial) validlik – testning biror-bir mezon bilan mosligi, o‘z navbatida, mezoniy validlikning ikki turi mavjud deb qaralgan:

– prediktiv validlik – testning belgilangan mezon bo‘yicha natijalarni bashorat qila olish xususiyati. Masalan, oliy ta’lim muassasasiga kirish testlari bo‘lajak talabalarining tanlangan sohalarini qay darajada muvaffaqiyatli o‘zlashtira olishlarini bashorat qila olishi lozim. Agar kirish testlarida yuqori natijalarga erishib o‘qishga qabul qilingan talabalar o‘qishda ham yuqori natijalarga erishsa, prediktiv validlik talablari bajarilgan hisoblanadi;

– solishtirma validlik – test natijasining belgilangan mezon bo‘yicha boshqa o‘lchovlar natijalari bilan mosligi. Masalan, o‘quvchilarning ingliz tilidan bilim darajasini baholaydigan test ishlab chiqsak, uning xuddi shu domenni o‘lchaydigan shu paytgacha mavjud boshqa test (masalan, IELTS yoki TOEFL imtihonlari) bilan solishtirma validligini aniqlashimiz

mumkin. Buning uchun biz yaratgan testni IELTS imtihoni topshirgan guruhda sinab ko‘ramiz va test topshiruvchilarning bizning testimiz bo‘yicha natijalari IELTS testi bo‘yicha natijalariga o‘xshash bo‘lsa, solishtirma validlik talablari bajarilgan hisoblanadi;

- mazmuniy validlik – testning o‘lchanishi lozim bo‘lgan domen mazmuni bilan mosligi;
- konstruktiv validlik – testning o‘lchanishi lozim bo‘lgan konstrukt (xususiyatni) to‘g‘ri o‘lchashi.

Shuningdek, 1950-yillarda qanday maqsadda test o‘tkazilishi va testning qaysi turi ishlatilishidan kelib chiqqan holda validlikning ma’lum bir turi sharti bajarilishi kerak degan tushuncha shakllandi. Misol uchun, prognostik testlar uchun prediktiv validlik, muvaffaqiyat testlari uchun mazmuniy validlik, psixologik testlar konstruktiv validlik shartlari bajarilsa kifoya degan qarash hukmron edi.

1970 – 1980-yillarga kelib hozirda pedagogik o‘lchovlar fanining mumtoz namoyandalari hisoblanadigan Li Kronbax (*Lee Cronbach*) va Samuel Messik (*Samuel Messick*) validlik testning emas, balki test asosida chiqariladigan xulosalarning xususiyati degan fikrni ilgari surishdi. Ularning fikricha, testning o‘zi valid yoki novalid bo‘la olmaydi, chunki test bir vosita, xolos. Bu vosita turli xil vaziyatlarda va turli maqsadlar uchun ishlatilishi mumkin. Misol uchun, skalpel mohir jarroh qo‘lida kimningdir hayotini saqlab qolishga xizmat qilsa, johilning qo‘lida kimningdir hayotiga nuqta qo‘yishi mum-

kin, lekin bunda ayb skalpelda emas, balki undan noto'g'ri maqsadda foydalanganlarda bo'ladi.

2000-yillarga kelib «validlikning turlari» tushunchasidan voz kechildi va uning o'rniga validlikning yuqorida qayd etilgan besh turdagi dalili tushunchasi rasman qabul qilindi. Bu pedagogik o'lchovlar bilan shug'ullanuvchi tashkilotlar uchun rasmiy hujjat bo'lib xizmat qiladigan «Ta'limiy va psixologik testlar standartlari»da¹ ham o'z aksini topdi. Endilikda test tuzish va o'tkazish bilan shug'ullanadigan tashkilotlar o'z testlari asosida chiqariladigan xulosalar asosli (valid) ekanligini isbotlashi uchun validatsiya jarayonida besh xil dalil keltirishlari talab qilinadi.

Asosiy xulosa:

Test natijasida chiqariladigan xulosalarning qay darajada valid (asosli) ekanligi nima maqsadda test olganligimizga va test natijalarini nima maqsadda ishlatishimizga bog'liq.

¹ Qarang: American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. – Washington, D.C.: American Educational Research Association, 2014.

2.2. Validlikning dalillari

Test asosida asosli (valid) xulosalar chiqara olishimiz va bu xulosalar asosida to‘g‘ri qarorlar qabul qila olishimiz uchun testni tuzish, uni o‘tkazish, uning natijalarini qayta ishlash va e‘lon qilish bilan shug‘ullanuvchi tashkilotlar **validlikning bir qancha dalillarini** keltirishlari kerak.

2.2.1. Testning mazmuni bilan bog‘liq dalillar

Testning mazmuni bilan bog‘liq dalillar keltirish uchun **testning mazmuni va test yordamida baholanishi (o‘lchanishi) lozim bo‘lgan konstrukt** tahlil qilinadi. Bunda, birinchi navbatda, quyidagi savolga e‘tibor qaratiladi:

– **Test orqali o‘lchanayotgan konstrukt qanchalik muhim va testning maqsadi bilan qay darajada bog‘liq?**

Masalan, maktab ta‘limining oxirgi yilida bitiruvchilarga adabiyot fanidan test berilishi kerak deylik. Bu test bitiruvchilarning maktab adabiyot dasturini qay darajada o‘zlashtirganligini baholab berishi lozim, ma‘lumki, bunday testning natijalari bitiruvchilarning shahodatnomasida yozib qo‘yiladi. Shunday ekan, bu testga nimalarni kiritamiz? Adiblar biografiyasiga oid ma‘lumotlarni o‘z ichiga olgan test topshiriqlarinimi? Asarlardan parchalarni o‘z ichiga olgan test topshiriqlarinimi? Adabiyot nazariyasiga oid test topshiriqlarinimi? Bu savolga javob berish uchun maktab

adabiyot fani (domen) qaysi konstruktarni o'z ichiga olgan va bu konstrukt qay darajada muhim degan savolga javob berishimiz kerak bo'ladi. Testimiz validlikning test mazmuni bilan bog'liq talablariga javob berishi uchun u maktab adabiyot dasturiga kiritilgan muhim konstruktarni adekvat darajada qamrab olishi lozim.

Demak, validlikning shartlaridan biri testning baholanishi (o'lchanishi) lozim bo'lgan domen va konstruktarni qamrab olishidir. Odatda, domen ta'lim standartlarida belgilab beriladi, chunki ta'lim standartlarida ta'limning maqsadi – o'quvchilar egallashi lozim bo'lgan bilim, ko'nikma, malaka, kompetensiyalar yoritilgan bo'ladi. Test aynan shu bilim, ko'nikma, malaka, kompetensiyalarni baholashi lozim. Ta'limda ishlatiladigan testlar, odatda, cheklangan vaqt ichida bajarilishi talab qilinadi, shuning uchun testga baholanishi lozim bo'lgan domenning **reprezentativ namunalari** (domen haqida asosli xulosalar chiqarishga imkon beradigan muhim namunalari) kiritiladi.

Test mazmuni bilan bog'liq validlik talablariga javob berishi uchun, birinchi navbatda, o'lchanishi lozim bo'lgan domen va konstrukt aniq belgilab olinishi lozim. Masalan, 9-sinf o'quvchilari uchun tarix fanidan test tuzmoqchi bo'lsak, «o'quvchilar tarix fanidan nimalarni bilishi va qanday ko'nikmalarga ega bo'lishi lozim?» – degan savolga javob berilishi kerak. O'quvchilar egallashi zarur bo'lgan bilim, ko'nikma, malaka, kompetensiyalar ro'yxati tuzilib, ush-

bu ro'yxatdagi har bir element qanday shakldagi va qancha miqdordagi test topshiriqlari yordamida tekshirilishi (baholanishi) hal qilinadi. Bu maxsus hujjat – **test spetsifikatsiyasida (tafsilotida)** aks etadi.

Test spetsifikatsiyasi (tafsiloti) – testning maqsadi, qo'llanilishi, strukturasi, shuningdek, test topshiriqlariga, test o'tkazish jarayoniga va test natijalarini hisoblashga qo'yiladigan talablarni o'z ichiga olgan hujjat.

Testning mazmunini belgilayotganda bir qancha **tamoyillar**¹ hisobga olinishi lozim:

- test mazmunining test maqsadiga muvofiqligi;
- tekshirilayotgan bilimning muhimligi – testga fanning muhim elementlari kiritiladi. Masalan, maktab bitiruvchisining XX asr o'zbek adabiyotini qay darajada bilishini baholamoqchi bo'lsak, Abdulla Oripovning ilk she'rlar to'plami qanday nomlanganligini bilish muhimmi? Test topshiruvchining bu ma'lumotni bilishi yoki bilmasligi asosida uning XX asr o'zbek adabiyoti, xususan, Abdulla Oripov she'riyatini tushunishi haqida xulosa chiqara olamizmi?
- test topshiriqlari mazmunining to'g'riligi – testga faqat to'g'ri ma'lumotlar kiritilishi kerak, noto'g'ri va bahstalab ma'lumotlarni testga kiritib bo'lmaydi;

¹ Воробьева С.В. Современные средства оценивания результатов обучения в общеобразовательной школе. Изд. 2-е, пер. и доп. – Москва: Юрайт, 2019. – С.396.

– test mazmunida fanning hozirgi rivojlanish darajasi aks etganligi – testga eskirgan, o‘z ahamiyatini yo‘qotgan ma’lumotlar kiritilmaydi;

– test mazmunining kompleksliligi va to‘g‘ri balanslashtirilganligi – test tekshirilayotgan domenga kiruvchi muhim elementlarning barchasini qamrab olishi kerak;

– test mazmunida fan mazmunining reprezentativ aks etganligi – tekshirilayotgan domenga kiradigan har bir elementni baholash va u haqda asosli xulosa chiqarish uchun yetarli darajada test topshiriqlari bo‘lishi lozim. Masalan, maktab kimyo dasturining qay darajada o‘zlashtirilganligini tekshirmoqchi bo‘lsak, shu maqsadda tayyorlangan test quyidagi kabi talablarga javob berishi muhim sanaladi: «Testda ushbu fanning muhim bo‘limlari yetarlicha aks etganmi?», «Testga organik kimyodan bitta topshiriq kiritilgan bo‘lsa, bu topshiriq test topshiruvchining organik kimyoni qay darajada o‘zlashtirgani haqida asosli xulosa chiqarish imkonini beradimi?».

– test mazmunining tizimliligi – testga kiritilgan topshiriqlar fan mazmunini tizimli ravishda aks ettirishi zarur;

– test shaklining test mazmuniga bog‘liqligi. Masalan, test topshiruvchining ona tilida fikrni yozma ifodalay olish ko‘nikmasini baholamoqchi bo‘lsak, buni yopiq test topshirig‘i yordamida amalga oshirish mumkinmi degan savolga javob bera olishimiz lozim;

– test mazmunining variativligi – agar test har xil kunlarda o‘tkaziladigan bo‘lsa, test topshiriqlari test topshiruvchilarga

oldindan ma'lum bo'lib qolmasligi uchun **parallel test formalari** ishlatilishi mumkin.¹

Parallel test formalari – bir spetsifikatsiya asosida tuzilgan, aynan bir xil domen va konstruktarni o'Ichaydigan, qiyinchilik va boshqa statistik parametrlari o'xshash bo'lgan, lekin har xil test topshiriqlaridan iborat test variantlari¹².

Test mazmuni bilan bog'liq validlik talablariga javob berishni ta'minlashda bir qancha strategiyalarni qo'llash mumkin:

– testni tuzish jarayonida mazmun bilan bog'liq talablarni hisobga olish – bu, odatda, test tuzuvchilar tomonidan, lozim bo'lganda tashqi mutaxassislarni jalb qilgan holda amalga oshiriladi;

– testni tuzib bo'lgandan keyin mazmun bilan bog'liq talablar asosida testga baho berish – bu ko'pincha tashqi mutaxassislar tomonidan amalga oshiriladi.

Odatda, testning mazmun bilan bog'liq validlik talablariga javob berishini ta'minlash uchun har ikkala strategiya qo'llaniladi.

¹ Klassik test nazariyasiga ko'ra, test formalari parallel hisoblanishi uchun: a) har bir test topshiruvchining bu test formalarni qo'llaganda olgan haqiqiy bahosi teng bo'lishi; b) xatolar dispersiyalari teng bo'lishi lozim. Demak, parallel test formalarida natijalarning o'rtacha arifmetik ko'rsatkichlari va dispersiyalar teng bo'ladi (Qarang: Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов.– Москва: Логос, 2010. – С.166).

Haqiqiy baho tushunchasi va testning statistik ko'rsatkichlari keyingi bobda yoritilgan.

2.2.2. Test topshiruvchilarning test topshiriqlarini bajarishi jarayoni bilan bog‘liq dalillar

Test topshiruvchilarning test topshiriqlarini bajarishi jarayoni bilan bog‘liq dalillar keltirish uchun **test topshirig‘ini bajarish jarayonida test topshiruvchilar nima qilishlari** tahlil etiladi. Bunda quyidagi savolga e‘tibor qaratiladi:

– **Test topshiruvchilarning test topshirig‘ini bajarish jarayonidagi amallari test orqali o‘lchanayotgan konstrukt bilan qay darajada bog‘liq?**

Masalan, testning maqsadi «O‘tkan kunlar» romanini tahlil qilish ko‘nikmasini baholash bo‘lsa va testda asar dastlab qaysi jurnalda, qachon chop etilganligi so‘ralsa, bu topshiriqni bajarish (savolga javob berish) uchun test topshiruvchi nima qiladi? Bu topshiriqni bajarish uchun haqiqatan ham asarni tahlil qilish ko‘nikmasi ishga solinadimi yoki xotirani ishga solish kerak bo‘ladimi? Test tuzuvchilar testga test topshiriqlarini kiritayotganda bu topshiriqlarni bajarish uchun **qanday ko‘nikmalar (malakalar, kompetensiyalar) ishga solinishi lozimligini** tahlil qilishlari, bu ko‘nikmalar (malakalar, kompetensiyalar) testning maqsadi bilan qay darajada mos ekanligiga baho berishlari lozim.

Test topshiriqlarini bajarish jarayoni bilan bog‘liq validlik talablariga javob berishni ta‘minlashda bir qancha strategiyalarni qo‘llash mumkin: testdan keyin test topshiruvchilar

bilan suhbatlashish (suhbat davomida «Test topshiriqlarini bajarish uchun nima qildingiz? Qanday ko‘nikmalarni ishga soldingiz?» singari savollar berish mumkin), test topshirish jarayonini hamda bu jarayonda test topshiruvchilar nima qilayotganlarini kuzatib tahlil qilish va hokazo.

2.2.3. Testning ichki strukturasi bilan bog‘liq dalillar

Testning ichki strukturasi bilan bog‘liq dalillar keltirish uchun test komponentlari (qismlari) o‘rtasidagi bog‘liqlik yoki alohida olingan test topshiriqlari va butun test o‘rtasidagi bog‘liqlik tahlil qilinadi.

Ma‘lumki, test biron-bir domen va konstruktini (yoki konstruktarni) o‘lchash uchun ishlatiladi. Test yordamida o‘lchanayotgan konstruktlar bitta yoki bir nechtaligiga qarab test **bir o‘lchamli** yoki **ko‘p o‘lchamli** bo‘lishi mumkin.

Bir o‘lchamli test – faqat bitta konstruktini o‘lchash uchun mo‘ljallangan test.

Ko‘p o‘lchamli test – bir-biriga bog‘liq bir qancha konstruktarni o‘lchash uchun mo‘ljallangan test.

Masalan, agar test faqat chet tili grammatikasiga oid bilimlarni tekshiishga qaratilgan bo‘lsa, bunday test bir o‘lchamli test hisoblanadi. Agar chet tili bo‘yicha test tinglab tushunish, o‘qib tushunish, yozish, gapirish singari bo‘limlarni o‘z ichi-

ga olsa, bunday test ko'p o'lchamli testlarga misol bo'la oladi, chunki bunda testning har bo'limi bir konstruktни tekshiradi.

Ko'p o'lchamli testlarda test komponentlari (qismlari) o'rtasidagi bog'liqlikni ko'rish uchun faktor tahlili usuli qo'llaniladi. Bunda test topshiriqlari qaysi konstruktни tekshirishiga qarab guruhlariga ajratiladi va guruhlar o'rtasidagi bog'liqliklar matematik yo'l bilan tahlil etiladi. Test natijalari e'lon qilinganda ham testning umumiy natijasi bilan birga har bir komponent natijasi e'lon qilinadi. Masalan, ingliz tilini bilish darajasini baholash uchun mo'ljallangan IELTS va TOEFL imtihonlarida test topshiruvchining umumiy natijasi bilan birga testning har bir qismida u to'plagan ballar ham e'lon qilinadi. Bu test topshiruvchining test qamrab olgan har bir konstrukt bo'yicha erishgan natijasini ko'rish imkonini beradi.

Bir o'lchamli testlarda alohida olingan test topshiriqlari va butun test o'rtasidagi bog'liqlik (korrelyatsiya) tahlil etiladi. Bunda bir qancha parametrlarga, xususan, testning ishonchlik darajasiga, test topshiruvchining umumiy natijasi va uning alohida olingan test topshiriqlariga bergan javoblari korrelyatsiyasiga hamda test topshiriqlarining **DIF ko'rsatkichiga** ham e'tibor qaratiladi. Masalan, nisbatan qiyinroq test topshiriqlariga to'g'ri javob bergan test topshiruvchining umumiy test natijasi ham nisbatan yuqoriroq bo'lishi darkor. Shu bilan birga, test topshiruvchilarning biron-bir jihati (masalan, jinsi, tili, millati, dini) bo'yicha boshqalardan ajralib turadi-

gan guruhi ayrim test topshiriqlariga boshqalarga qaraganda yaxshiroq yoki yomonroq javob bersa, bunday test topshiriqlariga alohida e'tibor qaratish kerak bo'ladi, chunki ular validlikka salbiy ta'sir ko'rsatishi mumkin.

2.2.4. Testning boshqa o'zgaruvchilar bilan bog'liqligi asosidagi dalillar

Testning boshqa o'zgaruvchilar bilan bog'liqligi asosidagi dalillar keltirish uchun testning boshqa ko'rsatkichlar, masalan, test topshiruvchilarning xuddi shu konstruktini o'lchaydigan boshqa testlardan olgan natijalari bilan korrelyatsiyasi tahlil etiladi.

Korrelyatsiya – ikki o'zgaruvchi o'rtasidagi bog'liqlik. Agar ikki o'zgaruvchi bir xil harakatlansa (ya'ni bir o'zgaruvchi oshganda ikkinchisi ham oshsa yoki bir o'zgaruvchi kamayganda ikkinchisi ham kamaysa), bu holat ijobiy korrelyatsiya deyiladi. Agar ikki o'zgaruvchi bir-biriga teskari harakatlansa (ya'ni bir o'zgaruvchi oshganda ikkinchisi kamaysa yoki bir o'zgaruvchi kamayganda ikkinchisi oshsa), bu holat salbiy korrelyatsiya deyiladi. Korrelyatsiya -1 dan $+1$ gacha ko'rsatkichda o'lchanadi, bunda -1 – o'ta salbiy korrelyatsiya, 0 – korrelyatsiyaning (bog'liqlikning) yo'qligi, $+1$ – ideal ijobiy korrelyatsiya.

Masalan, Aliyev kimyo bo'yicha testdan past natija oldi. Biroq u shu paytgacha kimyo darslarida yaxshi baholar olib kelgan, bu fan bo'yicha olimpiadalarda muvaffaqiyatli qatnashgan. Agar biz Aliyevning kimyo bo'yicha test natijasini A o'zgaruvchi, uning xuddi shu konstrukt (ya'ni kimyo fani bo'yicha bilimlarini) tekshiradigan boshqa o'lchovlar natijalarini (kimyo darslarida olgan baholari, kimyo olimpiadasidagi natijalarini) B o'zgaruvchi deb olsak, A va B o'zgaruvchilar o'rtasidagi korrelyatsiya manfiy bo'lyapti. Bu holat kimyo fani testi asosida chiqaradigan xulosalarimiz validligini shubha ostiga qo'yadi.

Testning boshqa o'zgaruvchilar bilan bog'liqligi asosidagi dalillar keltirishda **konvergent** va **diskriminant** dalillarga e'tibor qaratish lozim.

Konvergent dalillar – aynan bir xil yoki bir-biriga yaqin konstruktarni o'lchaydigan testlar natijalari orasidagi korrelyatsiyaga asoslangan dalillar.

Diskriminant dalillar – har xil konstruktarni o'lchaydigan testlar natijalari orasidagi korrelyatsiyaga asoslangan dalillar.

Masalan, o'qish savodxonligini tekshirishga qaratilgan test yaratildi. Bu test va xuddi shu konstrukt, ya'ni o'qish savodxonligini tekshiradigan boshqa test natijalari o'rtasidagi korrelyatsiyaga asoslangan dalillar konvergent dalillar hisoblanadi. Bu test va boshqa konstrukt, masalan, mantiqiy fikrlash

va xulosalar chiqara olish qobiliyatini tekshiradigan boshqa test natijalari o'rtasidagi korrelyatsiyaga asoslangan dalillar diskriminant dalillar hisoblanadi.

Test natijalarining boshqa o'zgaruvchilar bilan korrelyatsiyasini tahlil qilayotganimizda bu o'zgaruvchilar testimiz natijalari bilan bir (yoki deyarli bir) paytda olinishi yoki keyinroq olinishi mumkin. Shunga qarab dalillarimiz **solishtirma** yoki **prediktiv** bo'lishi mumkin.

Solishtirma dalillar – test va bu test bilan bir (yoki deyarli bir) paytda olingan boshqa o'lchovlar natijalarining korrelyatsiyasiga asoslangan dalillar.

Prediktiv dalillar – test va bu testdan keyin olingan boshqa o'lchovlar natijalarining korrelyatsiyasiga asoslangan dalillar.

Masalan, matematika bo'yicha test yaratdik. Bu testimizni topshirganlarning natijalari va test topshiruvchilarning matematika darslarida olgan baholari yoki matematika bo'yicha bilimlarni sinaydigan boshqa testdan erishgan natijalari o'rtasida ijobiy korrelyatsiya mavjud bo'lsa, demak, matematika testimiz uchun solishtirma dalillar keltirgan bo'lamiz. Agar testimizni topshirganlarning natijalari va ularning keyinchalik matematika bo'yicha bilimlarni ishga solishni talab qiladigan vaziyatlarda (masalan, oliy matematika yoki statistika kursida) erishgan natijalari o'rtasida ijobiy korrelyatsiya mavjud bo'lsa, demak, matematika testimiz uchun prediktiv

dalillar keltirgan bo‘lamiz. Ko‘rib turganimizdek, prediktiv dalil keltirilishi uchun ma‘lum bir vaqt o‘tishi kerak.

2.2.5. Testning oqibatlari bilan bog‘liq dalillar

Test ma‘lum bir maqsadlarni ko‘zlab tuziladi va o‘tkaziladi. Testni qo‘llash (joriy qilish) esa ma‘lum bir **oqibatlarga** olib keladi, ya‘ni test ta‘lim jarayonida va jamiyatda ma‘lum bir o‘zgarishlarga sabab bo‘ladi. Bu oqibatlar **ko‘zda tutilgan** yoki **ko‘zda tutilmagan, ijobiy** yoki **salbiy** bo‘lishi mumkin.

1994-yildan O‘zbekistonda abituriyentlarni oliy ta‘lim muassasalariga saralab olishda test tizimi ishlatila boshlandi, bu testlarni tuzish, o‘tkazish, natijalarini qayta ishlash va e‘lon qilish vakolatlari Vazirlar Mahkamasi huzuridagi Davlat test markaziga berildi. Test tizimini joriy qilishdagi **ko‘zlangan maqsadlardan** biri oliy ta‘lim muassasalariga qabul qilishdagi subyektivlik va korrupsiyani yo‘qotish edi. Test tizimining joriy qilinishi orqali ko‘zda tutilgan bu maqsadga qay darajada erishildi? Test tizimining joriy qilinishi boshqa, ko‘zda tutilmagan oqibatlarga olib kelmadimi? Test tizimining joriy qilinishi oqibatlari ijobiy yoki salbiy bo‘ldimi? Testni qo‘llash (joriy qilish) oqibatlari tahlil etilayotganda shu kabi savollarga javob izlanishi lozim bo‘ladi.

Masalan, ona tili va adabiyotdan test topshirayotgan abituriyentlarning bilimlari imtihonning insho yoki og‘zaki savol-javob turlariga qaraganda obyektiv baholanadi. Bu – test-

ning ko‘zda tutilgan oqibati va u ijobiy oqibatdir. Shu bilan birga, ona tili va adabiyotdan test topshirayotgan abituriyentlar asosiy e‘tiborlarini ona tilidan qoidalarni, adabiyotdan esa fakt va detallarni yodlashga qaratishdi. Bu ularning og‘zaki va yozma nutq kompetensiyalari, badiiy asarni tushunish va tahlil qilish kompetensiyalariga salbiy ta‘sir qildi. Bu – testning ko‘zda tutilmagan oqibati va u salbiy oqibatdir.

Testning oqibatlari bilan bog‘liq dalillar keltirish uchun test joriy qilingandan keyin yuzaga kelgan o‘zgarishlar tahlil etilishi va «Bu oqibatlar ko‘zda tutilgan edimi yoki ko‘zda tutilmagan edimi? Ijobiy yoki salbiy bo‘ldimi?» singari savollarga javob berilishi lozim. Test validlik talablariga javob berishi uchun uning ko‘zda tutilmagan, **salbiy oqibatlari imkon qadar minimallashtirilishi** kerak.

Shuni esda tutish joizki, **yuqori ta‘sir kuchiga ega testlarning** oqibatlari boshqa testlarga nisbatan kuchliroq bo‘ladi.

Yuqori ta‘sir kuchiga ega test – natijalari muhim qarorlar qabul qilish uchun ishlatiladigan test.

Yuqori ta‘sir kuchiga ega testlarga misol sifatida oliy ta‘lim muassasalariga kirish testlarini keltirish mumkin, chunki bu test sinovlarini har yili bir necha yuz ming abituriyentlar topshiradi va bu testlar natijasida abituriyentlar uchun muhim hisoblanadigan qarorlarga (o‘qishga qabul qilinishi yoki qabul qilinmasligi, agar qabul qilinsa, davlat granti yoki to‘lov-kontrakt asosida o‘qishi) kelinadi.

Yuqori ta'sir kuchiga ega testlar ta'lim jarayoni qatnashchilariga va ular orqali butun jamiyatga ta'sir o'tkazishi mumkin (bunday ta'sir inglizcha «**washback**» termini bilan ataladi). Testlar quyidagilarni **o'zgartirishi** mumkin:

- ta'lim jarayoni qatnashchilarining (ta'lim oluvchilar, ta'lim beruvchilar) ta'lim jarayonining turli jihatlariga (ta'lim mazmuni, ta'lim jarayonida qo'llaniladigan uslublar va hokazo) bo'lgan **munosabati**;
- ta'lim **jarayonining** o'zi (ta'lim jarayonining qay tarzda tashkil qilinishi);
- ta'lim jarayonining **natijasi** (nima va qay darajada o'rganildi?)¹.

Test topshiruvchilar o'zlari uchun muhim hisoblangan testlarda yaxshi natija qayd etishni xohlaydilar, demak, ularning asosiy e'tiborlari aynan shu **test orqali tekshiriladigan domenni** o'zlashtirishga qaratiladi va test qamrab olmagan domen o'zlashtirilmaydi².

Misol uchun, O'zbekistondagi oliy ta'lim muassasalariga kirish testlari:

- ta'lim oluvchilarning o'qitilayotgan fanga bo'lgan munosabatini o'zgartiradi – abituriyentlar testga kiritilgan fanlarni testdan muvaffaqiyatli o'tish maqsadida o'rgana boshlashadi;
- ta'lim jarayoniga ta'sir qiladi – test topshiriqlarini yechish ta'lim jarayonining muhim qismiga aylanadi va rasmiy

¹ Alderson J. and Wall D. Does Washback Exist? // Applied Linguistics. 1993. 14:2. – P.115 – 129.

² Frederiksen N. The Real Test Bias, ETS Research Report Series RR 81-40. – Princeton, NJ: Educational Testing Service, 1981.

ta'limdan ko'ra norasmiy ta'limga (o'quv markazlari, repertorlar) ko'proq e'tibor qaratiladi;

– ta'lim jarayonining natijasini belgilaydi – faqat testda qamrab olingan bilim va ko'nikmalar o'zlashtiriladi (masalan, ona tilida yozma nutq ko'nikmasini shakllantirishga e'tibor berilmaydi, chunki bu ko'nikma oliy ta'lim muassasalariga kirish testlarida tekshirilmaydi).

Shu bilan birga, natijalari **test topshiruvchiga bevosita ta'sir qiladigan** testning ta'siri kuchliroq bo'ladi¹. Masalan, oliy ta'lim muassasalariga kirish test sinovlari natijalari test topshiruvchiga bevosita ta'sir qiladi, demak, bunday testning ta'siri kuchliroq bo'ladi. Ta'lim muassasasini attestatsiyadan o'tkazish maqsadida o'tkazilayotgan testning natijalari esa test topshiruvchiga bevosita ta'sir qilmaydi, demak, bunday testning ta'siri kuchsizroq bo'ladi.

Test oqibatlarining ijobiy yoki salbiy bo'lishi asosan **test orqali tekshirilayotgan domen (konstruktlar) ta'limning maqsadlari bilan qay darajada uyg'un ekanligiga bog'liq**². Agar test orqali tekshirilayotgan domen (konstruktlar) ta'limning maqsadlari bilan uyg'un bo'lsa, ta'lim oluvchilar testga tayyorlanish asnosida ta'limiy maqsadlarga erishadilar. Aks holda ta'lim oluvchilarning asosiy e'tibori testga qaratiladi

¹ Shohamy E., Donitsa-Schmidt S. and Ferman I. Test Impact Revisited: Washback Effect Over Time // Language Testing. 1996. 13:3. – P.298 – 317.

² Green A. Watching for Washback: Observing the Influence of the International English Language Testing System Academic Writing Test in the Classroom // Language Assessment Quarterly. 2006. 3:4. – P.333 – 368.

va ta'limiy maqsadlar e'tibordan chetda qolib ketadi. Misol uchun, chet tilidan ta'limning maqsadi ta'lim oluvchilarda muloqot qila olish (o'qib tushunish, tinglab tushunish, yozish, gapirish) ko'nikmalarini rivojlantirish bo'lsa va yuqori ta'sir kuchiga ega test ham aynan shu ko'nikmalarni tekshirsa, ta'lim oluvchilar e'tiborlarini shu ko'nikmalarni rivojlantirishga qaratadilar. Agar test chet tilidan faqat grammatikaga oid bilimlarni tekshirsa va muloqot qila olish ko'nikmalari testda qamrab olinmasa, ta'lim oluvchilar ham e'tiborlarini grammatikaga qaratadilar va muloqot qila olish ko'nikmalari e'tibordan chetda qoladi.

Asosiy xulosa:

Test asosida chiqariladigan xulosalarimiz asosli (valid) deya olishimiz uchun:

- test orqali o'lchanoyotgan domen va konstruktlar testning maqsadi bilan bog'liq bo'lishi lozim;
- test topshiruvchilarning test topshiriqlarini bajarish jarayonidagi amallari o'lchanayotgan konstruktarga mos bo'lishi kerak;
- testning qismlari o'rtasida, test topshiriqlari va test o'rtasida ichki bog'liqlik bo'lishi kerak;
- testning natijalari va boshqa o'zgaruvchilar o'rtasida korrelyatsiya bo'lishi kerak;
- testning ko'zda tutilmagan, salbiy oqibatlari imkon qadar minimallashtirilishi kerak.

2.3. Validlikka xavf soluvchi omillar

Test asosida chiqariladigan xulosalar asosli (valid) bo‘lishi va bu xulosalar asosida to‘g‘ri qarorlar qabul qilinishi uchun bir qancha turdagi dalillar keltirilishi kerak. Shu bilan birga, **validlikka xavf soluvchi bir qator omillar** (testologiya fanida «validlikka xavflar» deb ataladi) aniqlanishi va bartaraf qilinishi lozim.

Test biror-bir konstruktни o‘lchash (baholash) uchun tuziladi. Validlikka xavf soluvchi omillar ham test o‘lchashi lozim bo‘lgan **konstruktga bog‘liq bo‘lmagan o‘zgaruvchilarning** test natijalariga ta’sir etishi yoki testda **konstruktning yetarli darajada aks etmasligi** bilan bog‘liq.

Konstruktga bog‘liq bo‘lmagan o‘zgaruvchilar – testning natijalariga test o‘lchashi lozim bo‘lgan konstrukt-lardan boshqa omillarning ta’sir qilishi.

Bu omillar turli xil bo‘lishi mumkin: testning mazmuni, test topshirish sharoitlari, testni baholash va natijalarni hisoblash jarayonlari va boshqa turli xil holatlar bilan bog‘liq omillar. Bu omillar ta’sirida alohida olingan test topshiruvchining, test topshiruvchilar guruhining yoki barcha test topshiruvchilarning natijalari yuqoriroq yoki pastroq chiqadi va biz test o‘lchashi lozim bo‘lgan konstruktlar haqida noto‘g‘ri xulosa chiqaramiz.

Testning mazmuni bilan bog‘liq omillar – testga o‘lchanishi lozim bo‘lgan konstrukt va domenga aloqasi bo‘lmagan topshiriqlar yoki topshiriqlar elementlarining kiritilishi. Masalan, biz test topshiruvchilarning mumtoz g‘azallarni tahlil qilish ko‘nikmasini o‘lchamoqchimiz (baholamoqchimiz). Testga «Boborahim Mashrabning «O‘rtar» radifli g‘azalida ishlatilgan badiiy san’atlarni toping va ularning ishlatilishini tahlil qiling» topshirig‘i kiritildi, lekin g‘azal matni berilmadi. Bunday holatda she‘riy san’atlarni yaxshi bilgan test topshiruvchi g‘azal matnini eslay olmagan (xotirasi pand bergani) uchun topshiriqni bajara olmasligi mumkin. Ya’ni biz o‘lchamoqchi bo‘lgan konstruktga (badiiy matnni tahlil qilish ko‘nikmasiga) boshqa omil (xotira) ta’sir qilyapti. Bunday holatda biz she‘riy san’atlarni yaxshi bilgan, lekin xotirasi pand bergan test topshiruvchi haqida noto‘g‘ri xulosaga kelishimiz mumkin.

Boshqa bir misol. Deylik, biz test topshiruvchilarning matematika bo‘yicha bilimlarini o‘lchamoqchimiz. Testga kiritilgan matnli masalalar shunday murakkab tilda yozilgan-ki, ularni tushunish uchun o‘qish ko‘nikmasi yaxshi rivojlangan bo‘lishi kerak. Bunday holatda ham matematika bo‘yicha yaxshi bilimga ega, lekin o‘qish ko‘nikmasi rivojlanmagan test topshiruvchi past natija ko‘rsatishi mumkin va biz u haqda «matematikani yaxshi bilmas ekan» degan noto‘g‘ri xulosaga kelishimiz mumkin.

Testni bajarish uchun test topshiruvchidan test yordamida o‘lchanishi lozim bo‘lgan konstruktga aloqasi bo‘lmagan

maxsus bilimlarning talab qilinishi ham testning mazmuni bilan bog‘liq omillarga kiradi. Masalan, biz test topshiruvchining ingliz tilida gapirish ko‘nikmasini baholamoqchimiz. Deylik, testga «Amerika futboli haqida gapirib bering», «London shahrining rayonlarini tasvirlab bering» singari topshiriqlar kiritildi. Ingliz tilini yaxshi biladigan test topshiruvchi Amerika futboli yoki London shahrining rayonlari haqida maxsus bilimi yo‘qligi tufayli testda yaxshi natija ko‘rsata olmadi. Demak, bunda ham testning natijalariga test o‘lchashi lozim bo‘lgan konstruktlerden boshqa omillar ta‘sir qiladi.

Shuningdek, test topshiriqlari va test yo‘riqnomalarining (instruksiyalarining) tili ham ravon va barcha test topshiruvchilar uchun tushunarli bo‘lishi lozim. Misol uchun, test yo‘riqnomalari haddan ziyod murakkab tilda yozilgan bo‘lsa va test topshiruvchilar nima qilishlari kerakligini tushunmasalar, bu ham ularning natijalariga ta‘sir qilishi va noto‘g‘ri xulosalarga olib kelishi mumkin.

Test topshirish sharoitlari bilan bog‘liq omillar – turli test topshiruvchilar uchun turli sharoitlar yaratilganligining test natijalariga ta‘sir qilishi. Masalan, ikkita guruh aynan bir testni ikkita auditoriyada topshirdi. Bir guruhning auditoriyasi shinam, yorug‘, konditsioner bilan jihozlangan, boshqa bir guruhning auditoriyasi esa tor va dim edi. Ikkinchi guruh uchun mavjud noqulay sharoit salbiy ta‘sir qilishi va ularning natijalari birinchi guruh natijalariga nisbatan pastroq chiqishi mumkin. Yoki bir guruhga mas‘ul kuzatuvchilar (nazoratchi-

lar) ko‘ngli bo‘shlik qilib test topshiruvchilarga ko‘proq vaqt bersa, boshqa guruhdagi kuzatuvchilar esa qattiqqo‘l bo‘lsa, bunda ham birinchi guruhning natijalari yaxshiroq chiqishi mumkin. Har ikkala holda ham birinchi guruhning bilimi ikkinchi guruhning bilimiga qaraganda yaxshiroq bo‘lmasligi mumkin, lekin birinchi guruh uchun yaratilgan sharoitlar (ya’ni test orqali o‘lchanayotgan konstruktarga bog‘liq bo‘lmagan omillar) tufayli birinchi guruhning natijalari yaxshiroq chiqishi mumkin. Bunday holatlarning oldini olish uchun testning muhim shartlaridan biri – **test topshirish sharoitlarining barcha test topshiruvchilar uchun birxilligiga** qattiq e’tibor berish lozim. Buning uchun test topshiriladigan auditoriyalar bir xil jihozlanadi, kuzatuvchilar (nazoratchilar) belgilangan talablar asosida tanlab olinadi va o‘qitiladi, test topshirish qoidalariga (reglamentiga) barchaning bir xilda rioya qilishi ta’minlanadi.

Testni baholash va natijalarni hisoblash jarayonlari bog‘liq omillar – testni baholash va testning natijalarini hisoblash jarayonlarida turlicha yondashuvlarning qo‘llanilishi natijasida ayrim test topshiruvchilarning yoki test topshiruvchilar guruhining yuqori (yoki past) natija ko‘rsatishi. Masalan, testga ochiq test topshiriqlari (insho, bayon, og‘zaki javob va hokazo) kiritilgan bo‘lsa, bunday topshiriqlar mutaxassislar (ekspertlar) tomonidan baholanadi (testologiyada ochiq test topshiriqlarini baholash uchun jalb qilingan ekspertlar «baholovchilar» deb nomlanadi). Qaysidir baholovchi boshqa hamkasblariga nisbatan qattiqqo‘l bo‘lishi mumkin

va bu holat aynan shu baholovchi tomonidan baholangan test topshiruvchilarning natijalariga ta'sir qilishi mumkin. Deylik, Aliyev va Valiyev testni bir xilda bajarishsa, lekin Aliyevning ishi qattiqqo'1 baholovchiga tushib qolsa, Aliyev Valiyevga qaraganda pastroq ball olishi mumkin. Yoki aynan bir baholovchi ish kunining boshida o'ziga ajratilgan ishlarni hafsala bilan tekshirishi, ish kuni oxirida esa charchaganidan e'tiborsizlik bilan tekshirishi va sifatsiz ishlarga ham yuqori ball qo'yib ketishi mumkin. Shuning uchun ham test bilan shug'ullanuvchi tashkilotlar **baholovchilar orasidagi ishonchlilik** va **baholovchining ichki ishonchliligi** ko'rsatkichlariga alohida e'tibor berishadi. Bu ko'rsatkichlar maxsus formulalar yordamida hisoblanadi va tahlil etiladi.

Baholovchilar orasidagi ishonchlilik – turli baholovchilarning ayni bir ishga (yoki bir xil darajadagi, bir xil sifatga ega bo'lgan ishlarga) bergan baholarining yaqinligi.

Baholovchining ichki ishonchliligi – alohida olingan baholovchining ayni bir ishga (yoki bir xil darajadagi, bir xil sifatga ega bo'lgan ishlarga) turli vaziyatlarda bergan baholari yaqinligi.

Bir xil darajadagi (bir xil sifatga ega bo'lgan) ishlarning turli baholovchilar tomonidan turli vaziyatlarda har xil baholanishining oldini olish uchun **baholash mezonlari** ishlab chiqiladi va barcha baholovchilarning bu mezonlarni bir xilda tushunib, bir xilda qo'llashlari doimiy nazorat qilib borilishi lozim.

Shu bilan birga, testning natijalarini qayta islash jarayonida ham validlikka xavf soluvchi holatlarga yo‘l qo‘yilishi mumkin. Masalan, ayrim test topshiruvchilarning natijalari test orqali tekshirilayotgan konstruktga bog‘liq bo‘lmagan mezonlar asosida o‘zgartirilishi mumkin. O‘zbekiston oliy ta‘lim muassasalariga kirish test sinovlarida yaqingacha qo‘llanilgan muddatli harbiy xizmatni o‘tab qaytgan abituriyentlarga ball qo‘shib berish amaliyoti bunga misol bo‘la oladi. Bunday holda bilim darajasi bir xil bo‘lgan ikki abituriyentdan muddatli harbiy xizmatni o‘tab qaytgani yuqoriroq ball oladi, vaholanki, muqobil harbiy xizmat test orqali tekshirilayotgan konstruktga aloqasi yo‘q omil hisoblanadi. Bu esa test natijalari asosida chiqariladigan xulosalarning validligiga va qarorlarning to‘g‘riligiga salbiy ta‘sir qiladi.

Testni yaratish, uning natijalarini qayta ishlash va e‘lon qilish bilan shug‘ullanuvchi tashkilotlar konstruktga bog‘liq bo‘lmagan o‘zgaruvchilarning test natijalariga ta‘sirini minimallashtirish uchun barcha choralarni ko‘rishlari lozim.

Test asosida chiqariladigan xulosalarning validligiga xavf soluvchi omillardan yana biri **testda konstruktning yetarli darajada aks etmasligidir.**

Konstruktlarning yetarli darajada aks etmaganligi – testda tekshirilishi lozim bo‘lgan domen va konstruktning muhim elementlari tushib qolishi.

Masalan, tilning vazifasi og‘zaki va yozma muloqotni ta’minlash ekan, ona tili yoki chet tili testlari test topshiruvchining faqat grammatika bo‘yicha bilimlarini tekshirsa va fikrini og‘zaki yoki yozma bayon qila olish ko‘nikmasini tekshirmasa, bunday testlar til domenini to‘liq qamrab olgan deyishimiz mumkinmi? Boshqa bir misol. Deylik, maktabda fizikaning mexanika, optika, termodinamika bo‘limlari o‘tildi. Shu bilan birga, fizikaning nazariy qonunlari hayotiy vaziyatlarda qanday tatbiq qilinishi o‘rgatildi. Maktab bitiruvchilarining fizika kursini qay darajada o‘zlashtirganligini tekshirishga qaratilgan testga esa fizikaning maktabda qamrab olingan bo‘limlaridan biri kiritilmadi. Yoki faqat nazariy savollar kiritilib, amaliy savollar kiritilmadi. Bunday holda testimizda tekshirishi lozim bo‘lgan konstruktlar yetarli darajada aks etmagan bo‘ladi.

Bunday holatning oldini olish uchun test qamrab olishi lozim bo‘lgan domen va konstruktlar tahlil qilinib, test tafsilotlarida aks ettiriladi va shu tafsilotlar asosida test tuziladi.

Asosiy xulosa:

Test asosida chiqariladigan xulosalarimiz asosli (valid) deya olishimiz uchun:

- testning natijalariga test orqali o‘lchanayotgan konstruktдан boshqa omillar ta’sir qilmasligi kerak;
- test o‘lchanishi lozim bo‘lgan domen va konstruktning barcha muhim elementlarini qamrab olishi lozim.

<i>Validlik bilan bog'liq dalillar</i>	<i>Dalil keltirish jarayonida qanday savollarga javob izlanadi?</i>	<i>Dalillar qanday yo'llar bilan olinadi?</i>
<p>Testning mazmuni bilan bog'liq dalillar</p>	<p>Test orqali o'lchanayotgan konstrukt qay darajada muhim? Bu konstrukt testning maqsadi bilan qay darajada bog'liq? Testda konstrukt yetarli darajada aks etganmi? Testga baholanishi lozim bo'lgan domenning reprezentativ (muhim) namunalari kiritilganmi? Test spetsifikatsiyasi mavjudmi? Test natijalariga konstruktga bog'liq bo'lmagan o'zgaruvchilar ta'sir ko'rsatmaydimi?</p>	<p>Mutaxassislar ekspertizasi</p>
<p>Test topshiruvchilarning test topshiriqlarini bajarishi jarayoni bilan bog'liq dalillar</p>	<p>Test topshiruvchilarning test topshirig'ini bajarish jarayonidagi amallari test orqali o'lchanayotgan konstrukt bilan qay darajada bog'liq? Test topshiruvchilar test topshiriqlarini bajarish jarayonida test orqali tekshirilishi lozim bo'lgan ko'nikmalarni ishga soladilarmi? Test natijalariga konstruktga bog'liq bo'lmagan o'zgaruvchilar ta'sir ko'rsatmaydimi?</p>	<p>Test topshiruvchilar bilan testdan keyin suhbat Test topshirish jarayonini kuzatish</p>
<p>Testning ichki strukturasi bilan bog'liq dalillar</p>	<p>Test topshiriqlari va test o'rtasidagi bog'liqlik (korrelyatsiya) ko'rsatkichlari qanday? Test topshiriqlarining DIF ko'rsatkichlari qanday?</p>	<p>Statistik tahlil</p>

Testning boshqa o'zgaruvchilar bilan bog'liqligi asosidagi dalillar	Testning xuddi shu konstruktini o'lchaydigan boshqa o'zgaruvchilar bilan korrelyatsiya ko'rsatkichlari qanday? Testning boshqa konstruktini o'lchaydigan o'zgaruvchilar bilan korrelyatsiya ko'rsatkichlari qanday?	Statistik tahlil
Testning oqibat-lari bilan bog'liq dalillar	Testning qo'llanilishi qanday oqibatlar-ga olib keldi? Bu oqibatlar ko'zda tutil-ganmi yoki ko'zda tutilmaganmi? Ijo-biyimi yoki salbiyimi?	Ilmiy tad-qiqotlar

Validlikning dalillari va bu dalillarni olish yo'llari

3. ISHONCHLILIK

3.1. Ishonchlilik tushunchasi. Klassik test nazariyasi asoslari

Oldingi boblarda ko'rganimizday, test – o'lchov vositasi, bu vosita yordamida bevosita o'lchashning imkoni bo'lmagan xususiyatlar (test topshiruvchilarda biron-bir domenning egallanganlik darajasi, biron-bir konstruktning shakllanganlik darajasi) o'lchanadi (baholanadi). Bu o'lchov natijalari asosida ma'lum bir xulosalarga kelinadi va qarorlar qabul qilinadi. Bu xulosalar asosli bo'lishi uchun o'lchov vositasi (test) ma'lum bir talablarga javob bermog'i lozim. Shunday talablardan biri **ishonchlilik** hisoblanadi.

Ishonchlilik – o'lchov vositasida olingan natijalarning barqarorligi.

Tasavvur qilaylik, bizda tarozi bor. Tarozimiz bitta jismni (masalan, toshni) o'lchaganda har safar har xil natija ko'rsatasa, masalan, bir safar 1 kg, ikkinchi safar 1,5 kg, boshqa safar 0,75 kg ko'rsatasa, bu tarozining ko'rsatkichlari shubha uyg'otadi. Shunga o'xshash, bir guruhda test olganimizda Aliyev 100 dan 80 ball, Valiyev 60 ball, Salimov 40 ball olsa, ertasi kuni xuddi shu testdan Aliyev 40 ball, Valiyev 80 ball,

Salimov 50 ball olsa, bu testimiz o‘lchov vositasi sifatida ishonchli emas.

Ishonchlilik validlikning muhim bo‘lgan sharti, chunki validlikka dalillar keltirilayotganda testning ichki strukturasi bilan bog‘liq dalillar ishonchlilik ko‘rsatkichlarini ham o‘z ichiga olishi lozim. Shu bilan birga, **ishonchlilikning o‘zi validlik uchun yetarli emas**. Boshqacha qilib aytganda, **ishonchlilik talablariga javob bermaydigan testlar validlik talablariga ham javob bermaydi, lekin testlarning ishonchlilik talablariga javob berishi ular validlik talablariga ham javob beradi degani emas**¹. Test qanchalik ko‘p ta’sir kuchiga ega bo‘lsa, uning ishonchliligi ham shunchalik yuqori bo‘lishi talab etiladi, chunki bunday test asosida ko‘pchilik uchun muhim bo‘lgan qarorlar qabul qilinadi.

Klassik test nazariyasida ishonchlilik **o‘lchashning xatosi** tushunchasiga asoslanadi. Bu tushuncha o‘rta asrlarda mashhur astronom Galileo Galiley tomonidan kiritilgan. Uning fikricha, yulduzlarni kuzatayotgan turli astronomlar bir yulduzdan Yerning markazigacha bo‘lgan masofani turlicha hisoblashlari mumkin: kimdir bu masofani asliga qaraganda ko‘proq deb hisoblasa, kimdir kamroq deb hisoblashi mumkin. Agar barcha astronomlarning o‘lchovlari qo‘shilib, o‘rtacha arifmetik qiymati olinsa, astronomlar-

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.121.

ning xatolari 0 (nol)ga tenglashadi. Keyinchalik XIX asrda Karl Fridrix Gauss tomonidan bu holat matematik yo‘l bilan isbotlangan.

Xuddi shunday, agar bir test topshiruvchi ayni bir testni bir necha marta topshirsa, uning bu testda olgan natijasi har safar har xil chiqishi mumkin. Buning sabablari ko‘p: test topshiruvchining test kuni o‘zini qanday his qilganligi, test topshirilgan auditoriyadagi sharoitlar, test topshirish jarayonidagi nazorat va hokazo. Bunday omillar natijasida test topshiruvchining testda olgan bahosi uning haqiqiy bilim darajasidan yuqoriroq yoki pastroq bo‘lishi mumkin.

Demak, test topshiruvchining test topshirganda **olgan bahosi** uning **haqiqiy bahosi** (bilimiga berilgan haqiqiy baho) va **xatolikdan** (turli tashqi omillarning test topshiruvchining bahosiga ta’siridan) iborat bo‘ladi.

Buni quyidagicha formula bilan ifodalash mumkin¹:

$$X = T + E,$$

bunda: X = test topshiruvchining testda olgan bahosi;

T = test topshiruvchining haqiqiy bahosi (bilimiga berilgan haqiqiy baho);

E = xatolik (turli tashqi omillarning test topshiruvchining bahosiga ta’siri).

Demak, har qanday testda test topshiruvchining olgan bahosi uning haqiqiy bilimlar darajasini aks ettirmaydi, chunki

¹ McDonald R.P. Test Theory: A Unified Treatment. – Mahwah, NJ: Lawrence Erlbaum Associates, 1999. – P.96.

bu bahoda xatoning ham ulushi mavjud. Masalan, bir guruh test topshiruvchilar adabiyotdan test topshirishdi. Aliyev 10 ta test topshirig‘idan 7 tasining javobini bilardi, lekin o‘sha kuni kasal bo‘lganligi sababli faqat 5 ta topshiriqni to‘g‘ri yechdi. Demak, Aliyevning testda olgan bahosi 5 bo‘lsa, uning haqiqiy bahosi 7 ga, xatolik esa -2 ga teng: $5 = 7 - 2$. Valiyev esa 4 ta test topshirig‘ining javobini bilardi, lekin tasodifan yana 3 ta test topshirig‘ining javobini to‘g‘ri topdi, shuning uchun testda 10 balldan 7 ball oldi. Demak, Valiyevning testda olgan bahosi 7 bo‘lsa, uning haqiqiy bahosi 4 ga, xatolik esa 3 ga teng: $7 = 4 + 3$.

Tasavvur qilaylik, biz ayni testni ayni test topshiruvchilar guruhida cheksiz marta o‘tkazdik. Test topshiruvchilarimiz har safar har xil natija ko‘rsatishdi. Agar biz test topshiruvchining barcha testlarda ko‘rsatgan natijalarining o‘rtacha arifmetik qiymatini olsak, bu uning haqiqiy bahosi bo‘ladi, chunki xatolik 0 ga tenglashadi. Bu haqiqiy bahoning klassik modelida qabul qilingan asosiy tushunchalardan biridir¹.

Agar testning ishonchlilik darajasi yuqori bo‘lsa, test topshiruvchining har safar test topshirganda olgan baholari (X) bir-biriga yaqin bo‘ladi va bu baholar uning haqiqiy bahosiga (T) yaqin bo‘ladi. Amaliyotda test topshiruvchiga aynan bitta testni bir necha marta topshirtirib, olgan baholarining o‘rtacha arifmetik qiymatini chiqarish imkoni mavjud emas. Shuning

¹ Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов. – Москва: Логос, 2010. – С.161.

bo‘lishi mumkin. Tizimli xatoliklarni aniqlash va kamaytirish uchun testni tuzish, ishlatish va natijalarni hisoblash bilan bog‘liq jarayonlarni tahlil etish hamda validlik talablariga rioya qilish lozim bo‘ladi.

Asosiy xulosa:

Test asosida chiqariladigan xulosalarimiz asosli (valid) bo‘lishi uchun test ishonchlilik talablariga javob berishi, ya’ni testning natijalari barqaror bo‘lishi lozim.

Test natijasida olingan baho test topshiruvchining haqiqiy bilimini aks ettirmaydi, chunki bu bahoda xatolik mavjud.

Tizimli xatolik o‘lchov natijalariga doimo ta’sir qiladi, tasodifiy xatolik esa o‘lchov natijalariga konkret bir vaziyatda ta’sir qiladi.

3.2. Ishonchlilikni hisoblash usullari

Demak, testni ishonchli deya olishimiz uchun testning natijalari barqaror bo‘lishi, ya’ni testni har safar ayni bir test topshiruvchilar guruhida ishlatganimizda olinadigan natijalar bir-biriga yaqin bo‘lishi kerak. Ishonchlilik 0 dan 1 gacha ko‘rsatkichda o‘lchanadi: ko‘rsatkich 1 ga qanchalik yaqin bo‘lsa, test shunchalik ishonchli hisoblanadi. Testning ishonchliligi 0,9 dan yuqori bo‘lsa, bu juda yaxshi ko‘rsatkich, 0,8 – 0,9 oralig‘ida bo‘lsa, yaxshi ko‘rsatkich, 0,7 – 0,8 oralig‘ida bo‘lsa,

qoniqarli ko'rsatkich hisoblanadi. Yuqori ta'sir kuchiga ega (ya'ni muhim qarorlar qabul qilish uchun ishlatiladigan) testlarning ishonchlilik ko'rsatkichi yuqori bo'lishi talab qilinadi.

Testning ishonchliligini hisoblashda uch xil usuldan foydalanish mumkin.

Birinchi usul «**test – qayta test**» usuli deb nomlanadi. Test topshiruvchilarga test beriladi. Ma'lum bir vaqt o'tgandan keyin xuddi shu test topshiruvchilarga xuddi shu test qayta beriladi. Shundan so'ng har ikkala holatdagi natijalar korrelyatsiyasi hisoblanadi. Agar bu korrelyatsiya ko'rsatkichi yuqori bo'lsa (+1 ga yaqin bo'lsa), test ishonchli sanaladi.

Bu usul bo'yicha ishonchlilikni (testologiyada ishonchlikning bu turi **barqarorlik koeffitsiyenti** deyiladi) hisoblash uchun¹:

1. Test topshiruvchilarning birinchi marta test topshirganda olgan ballari yig'indisini hisoblaymiz ΣX_i .

2. Test topshiruvchilarning ikkinchi marta test topshirganda olgan ballari yig'indisini hisoblaymiz ΣY_i .

3. Har bir test topshiruvchining birinchi va ikkinchi marta test topshirganda olgan ballarini o'zaro ko'paytirib chiqamiz, hosil bo'lgan sonlar yig'indisini hisoblaymiz $\Sigma X_i Y_i$.

4. Har bir test topshiruvchining birinchi marta test topshirganda olgan ballarini kvadratga oshirib, yig'indisini hisoblaymiz $\Sigma (X_i)^2$.

¹ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.219.

5. Har bir test topshiruvchining ikkinchi marta test topshirganda olgan ballarini kvadratga oshirib, yig'indisini hisoblaymiz $\Sigma(Y_i)^2$.

Hosil bo'lgan qiymatlarni quyidagi formulaga qo'yamiz:

$$r = \frac{(N \times 3) - (1 \times 2)}{\sqrt{(N \times 4 - 1^2)} \times \sqrt{(N \times 5 - 2^2)}}$$

bunda: $r =$ «test – qayta test» usuli bo'yicha ishonchlilik ko'rsatkichi;

$N =$ test topshiruvchilarning umumiy soni;

$1 - 5 =$ yuqoridagi 5 ta yig'indining qiymatlari mos ravishda 1 dan 5 gacha raqamlar o'rniga qo'yib chiqiladi.

Misol uchun, 10 ta test topshiruvchidan aynan bir testni ma'lum bir vaqt oralig'ida (deylik, ikki hafta o'tib) ishlatgan holda ikki marta test oldik. Ularning natijalarini jadval ko'rinishida quyidagicha aks ettirishimiz mumkin:

Test topshiruvchining raqami i	Birinchi testda qayd etilgan natijalar X_i	Ikkinchi testda qayd etilgan natijalar Y_i
1	6	5
2	2	4
3	1	2
4	9	7
5	4	6
6	4	3
7	5	7
8	4	6
9	9	7
10	6	8

Endi yuqoridagi 5 amalni bajaramiz:

1. Test topshiruvchilarning birinchi marta test topshirganda olgan ballari yig'indisi $\sum X_i = 50$.

2. Test topshiruvchilarning ikkinchi marta test topshirganda olgan ballari yig'indisini hisoblaymiz $\sum Y_i = 55$.

3. Har bir test topshiruvchining birinchi va ikkinchi marta test topshirganda olgan ballarini ko'paytiramiz ($6 \times 5 = 30$, $2 \times 4 = 8$ va hokazo), hosil bo'lgan sonlar yig'indisi $\sum X_i Y_i = 309$.

4. Har bir test topshiruvchining birinchi marta test topshirganda olgan ballari kvadratini olamiz ($6^2 = 36$, $2^2 = 4$ va hokazo), hosil bo'lgan sonlar yig'indisi $\sum (X_i)^2 = 312$.

5. Har bir test topshiruvchining ikkinchi marta test topshirganda olgan ballari kvadratini olamiz ($5^2 = 25$, $4^2 = 16$ va hokazo), hosil bo'lgan sonlarning yig'indisi $\sum (Y_i)^2 = 337$.

Olgan qiymatlarimizni formuladagi raqamlar o'rniga qo'yib chiqamiz, bunda hosil bo'lgan yakuniy natija 0 – 9 oralig'idagi son bo'ladi:

$$r = \frac{(10 \times 309) - (50 \times 55)}{\sqrt{10 \times 312 - 50^2} \times \sqrt{10 \times 337 - 55^2}} = \frac{340}{\sqrt{620} \times \sqrt{345}} \approx 0,78.$$

Demak, ko'rib chiqqan misolimizda testimizning ishonchiligi taxminan 0,78 ga teng.

Bu usulda test va qayta test orasidagi **vaqt omilini** hisobga olish lozim. Bir tarafdand, bu vaqt haddan tashqari qisqa bo'lsa, test topshiruvchining birinchi marta test topshirgani ikkinchi marta test topshirgandagi natijalariga ta'sir qilishi mumkin (masalan, test topshiriqlari va unga bergan javob-

lari esida turishi mumkin). Ikkinchi tarafdan, agar bu vaqt haddan tashqari ko‘p bo‘lsa, test topshiruvchi birinchi safar testni topshirganda bilmagan narsalarini o‘rganib olishi yoki, aksincha, bilgan narsalarini esidan chiqarishi mumkin. Shuningdek, test topshiruvchilar ayni bir testni ikkinchi marta topshirganda birinchi martasiga nisbatan e‘tiborsizroq, hafsalasizroq bo‘lishlari mumkin. Bu esa birinchi va ikkinchi marta topshirilgan test natijalari orasidagi farqqa ta’sir qilishi ehtimoldan xoli emas.

Testning ishonchliligini hisoblashda qo‘llanilishi mumkin bo‘lgan keyingi usul **«parallel test formalari» usuli deb nomlanadi**. Bunda test topshiruvchilarga test beriladi. Shu vaqtning o‘zida yoki ma’lum bir vaqt (odatda, bir necha kun) o‘tgandan keyin xuddi shu test topshiruvchilarga oldingi testning parallel formasi beriladi. Shundan so‘ng «test – qayta test» usulida qo‘llanilgan formula yordamida ishonchlilik ko‘rsatkichi hisoblanadi. Testologiyada ishonchlilikning bu turi **ekvivalentlik koeffitsiyenti** deyiladi.

Yuqoridagi ikki usul amalda bir qator qiyinchiliklarga ega (aynan bir test topshiruvchilar guruhidan ikki marta test olish zarurati, testning parallel formalari yaratish va hokazo).

Testning ishonchliligini hisoblashda qo‘llaniladigan uchinchi usul **testning ichki barqarorligini hisoblashga** asoslanadi. Bu usulning yuqoridagi ikki usuldan farqi – testning ishonchliligini aniqlash uchun test topshiruvchilardan bir necha marta test olish shart emas, maxsus formulalar bir marta olingan test

natijalari yordamida ishonchlilikni hisoblash imkonini beradi. Bunda ham, o'z navbatida, bir qancha usullar mavjud:

a) «ikkita yarim» usuli. Bunda test ikkita «yarim test»ga bo'linadi. Masalan, test 30 ta test topshirig'idan iborat bo'lsa, toq raqamlardagi test topshiriqlari (1, 3, 5, ... 27, 29) bitta test deb olinadi, juft raqamlardagi test topshiriqlari (2, 4, 6, ... 28, 30) esa ikkinchi test deb olinadi. Har «ikki» testing (ya'ni butun testimizning har ikkala – toq va juft yarmining) ekvivalentlik koeffitsiyenti quyidagi Spirman-Braun (*Spearman-Brown*) formulasi yordamida hisoblanadi:

$$r_a(E) = \frac{2r_{AB}}{1 + r_{AB}}$$

bunda: $r_a(E)$ = Spirman-Braun bo'yicha ishonchlilik ko'rsatkichi;

r_{AB} = toq va juft yarimlarning natijalari orasidagi korrelyatsiya (r Pirson korrelyatsiyasi) koeffitsiyenti.

Toq va juft yarimlarning natijalari orasidagi korrelyatsiyani (r_{AB}) hisoblash uchun «test – qayta test» usulida qo'llanilgan formula ishlatiladi, bunda testning toq yarmi natijalari birinchi test natijalari o'rniga, juft yarmi natijalari ikkinchi test natijalari o'rniga olinadi.

Misol uchun, test topshiruvchilar guruhida test o'tkazdik va testing toq va juft yarimlari o'rtasidagi korrelyatsiya 0,60 ga teng bo'ldi. Shunda Spirman-Braun formulasi yordamida testning ishonchlilik ko'rsatkichi 0,75 ga teng bo'ladi:

$$r_{\alpha}(E) = \frac{2 \times 0,60}{1 + 0,60} = \frac{1,20}{1,60} = 0,75$$

Bu formulani qoʻllayotganda testning ikkita yarmi dispersiyasi oʻrtasidagi farq 10 – 20% dan oraligʻida boʻlishi kerak. Agar bu farq 20% dan oshsa, Rulon (*Rulon*) formulasi qoʻllaniladi¹:

$$r_{\alpha}(E) = 1 - \frac{S_{A-B}^2}{S_X^2}$$

bunda: $r_{\alpha}(E)$ = Rulon boʻyicha ishonchlilik koʻrsatkichi;

S_{A-B}^2 = toq va juft yarimlarning natijalari orasidagi farq dispersiyasi;

S_X^2 = butun test dispersiyasi.

Izoh: dispersiya – oʻlchov natijalarining oʻrtacha arifmetik qiymatdan farqlanish darajasi². Ixtiyoriy kattalikning (sonning) dispersiyasi quyidagi formula boʻyicha hisoblanadi:

$$D(X) = M((X - M(X))^2)$$

bunda: M = dispersiyasi hisoblanayotgan kattalikning oʻrtacha qiymati.

b) Kuder-Richardson (*Kuder-Richardson*) formulasi.

Bu formula **dixotom** (yaʼni javobi «toʻgʻri – notoʻgʻri» sifatida baholanadigan) test topshiriqlaridan iborat testlarning ishonchliligini hisoblash uchun ishlatiladi. Bu formularning

¹ Пермяков О.Е., Максимова О.А. Основы технологии комплексной экспертизы качества педагогических тестов. – Томск: ТОИПКРО, 2008. – С.77.

² Балыхина Т.М. Словарь понятий и терминов тестологии. – Москва: Русский язык, 2006. – С.69.

ikki xil varianti mavjud: K-R20 va K-R21. K-R20 formulasi aniqroq, K-R21 formulasi esa hisoblash uchun osonroqdir.

K-R21 formulasi:

$$r_{\alpha} = \frac{K}{K-1} \left(1 - \frac{M(K-M)}{Ks^2} \right)$$

bunda: r_{α} = Kuder-Richardson bo'yicha ishonchlilik ko'rsatkichi;

K = testdagi test topshiriqlari soni;

M = test natijalari yig'indisining o'rtacha arifmetik ko'rsatkichi;

s = test natijalari to'plamining standart og'ishi.

K-R20 formulasi:

$$r_{\alpha} = \frac{K}{K-1} \left(1 - \frac{\sum pq}{s^2} \right)$$

bunda: r_{α} = Kuder-Richardson bo'yicha ishonchlilik ko'rsatkichi;

K = testdagi test topshiriqlari soni;

$\sum pq$ = javobi dixotom («to'g'ri – noto'g'ri» sifatida baholangan) test topshiriqlari dispersiyasi yig'indisi;

s^2 = butun test dispersiyasi.

c) Kronbax (Cronbach) alfasi. Bu usul **politom** (ya'ni javobi ko'p ballik tizimda baholanadigan) test topshiriqlaridan iborat testlarning ishonchligini hisoblash uchun ishlatiladi.

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum s_i^2}{s_x^2} \right)$$

bunda: α = Kronbax bo'yicha ishonchlilik ko'rsatkichi (Kronbax alfasi);

K = testdagi test topshiriqlari soni;

s_i^2 = alohida olingan test topshirig'i dispersiyasi;

s_x^2 = butun test dispersiyasi.

Masalan, test 4 ta insho shaklidagi (kengaytirilgan javob talab qiladigan) test topshirig'idan iborat. Har bir test topshirig'i 0 dan 10 gacha ball bilan baholanadi. Agar 4 ta test topshirig'i dispersiyalari mos ravishda 9, 4,8, 10,2 va 16 ga, butun test dispersiyasi 100 ga teng bo'lsa, bunday testning ishonchliligi (Kronbax alfasi) 0,8 ga teng bo'ladi:

$$\alpha = \frac{4}{4-1} \left(1 - \frac{9 + 4,8 + 10,2 + 16}{100} \right) = 0,8$$

Shuni hisobga olish lozimki, testning ichki barqarorligi asosida ishonchlilikni hisoblashda test bir o'lchamli (bitta konstruktini tekshirishga qaratilgan) bo'lishi darkor. Shuningdek, ichki barqarorlik formulalari **tezkor testlarning** ishonchliligini asliga qaraganda ancha yuqori, aralash testlarda esa nisbatan yuqori qilib ko'rsatadi. Tezkor testlarning ishonchliligi «test – qayta test» yoki «parallel test formalari» usullarida hisoblangani maqsadga muvofiq. Ta'lim jarayonida ishlatiladigan testlarning aksariyati aralash testlar bo'lgani,

ya'ni tezkor testlarning ham xususiyatlariga ega bo'lgani uchun ishonchlilikni hisoblashning bir qancha usullaridan foydalanib xulosa chiqarish tavsiya qilinadi¹.

Testning ishonchliligi bilan bog'liq quyidagi **tamoyillarni** yodda tutish lozim²:

1. Test topshirgan guruh qanchalik **geterogen** (turli xil test topshiruvchilardan iborat) bo'lsa, testning ishonchliligi shunchalik ortadi. Aslida, nazariy jihatdan olganda, testning o'zi ishonchli yoki ishonchsiz bo'lmaydi. Ishonchlilik – bu **test topshiruvchilar o'ziga xos guruhining testda ko'rsatgan natijalari barqarorligi**³. Shuning uchun ham testni aprobatsiyadan o'tkazayotganda aprobatsiyada qatnashayotganlar guruhi tarkibi va variativligi jihatdan haqiqiy test topshiruvchilar guruhiga yaqin bo'lishi kerak.

2. Testda ochiq test topshiriqlari mavjud bo'lgan hollarda **baholovchilar orasidagi ishonchlilik** va **baholovchining ichki ishonchliligi** testning ishonchliligiga ta'sir qiladi. Agar baholovchilar bir xil darajada bajarilgan ishlarga har xil baho qo'yishga moyil bo'lsalar, bu testning ishonchliligini kamaytiradi.

3. Umuman olganda, **uzunroq** testlarning, ya'ni test topshiriqlari soni ko'proq bo'lgan testlarning ishonchliligi yu-

¹ Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.343.

² O'sha manba. – P.343.

³ Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов. – Москва: Логос, 2010. – С.204.

qoriroq bo'ladi, chunki ko'proq sondagi test topshiriqlari o'lchanayotgan domen va konstruktни aniqroq o'lchash imkonini beradi.

Test uzunligining test ishonchliligiga ta'sirini quyidagi formula bo'yicha hisoblash mumkin:

$$\alpha_r = \frac{k\alpha}{1 + (k - 1)\alpha}$$

bunda: α_r = testning uzaytirilgandan keyingi ishonchliligi;

k = testning necha marta uzaytirilishi;

α = testning uzaytirilgandan oldingi ishonchliligi.

Misol uchun, testning ishonchliligi 0,7 ga teng bo'lgan bo'lsa va testning uzunligini (ya'ni testdagi topshiriqlar sonini) 3 barobarga oshirsak, testning ishonchliligi 0,88 ga teng bo'ladi:

$$\alpha_r = \frac{3 \times 0,7}{1 + (3 - 1) \times 0,7} = \frac{2,1}{2,4} = 0,88.$$

Albatta, bunda testning uzunligi testga har qanday topshiriqlarni qo'shish hisobiga emas, balki test spetsifikatsiyasiga mos keluvchi sifatli test topshiriqlarni qo'shish hisobiga uzaytirilishi lozim. Shuningdek, testning uzunligini oshirish haqida qaror qabul qilayotganda boshqa omillar (masalan, testga ajratiladigan vaqt) ham hisobga olinishi kerak bo'ladi.

4. Haddan tashqari **oson** yoki haddan tashqari **qiyin** testlarning ishonchliligi past bo'ladi, chunki bunday testlarda ballar taqsimoti gomogen bo'ladi.

Testni, ayniqsa, yuqori ta'sir kuchiga ega testni yaratishga mas'ul tashkilot testning ishonchliligi bilan bog'liq natijalarni e'lon qilib borishi darkor. Masalan, AQSh oliy ta'lim muassalariga qabul qilishda ishlatiladigan SAT testining yaratuvchisi bo'lgan College Board tashkilotining internetdagi rasmiy saytida testning umumiy ishonchliligi, subtestlar (testning alohida olingan bo'limlari) ishonchliligi, o'lchashning standart xatosi va boshqa ma'lumotlar keltirilgan va izohlangan.

Asosiy xulosa:

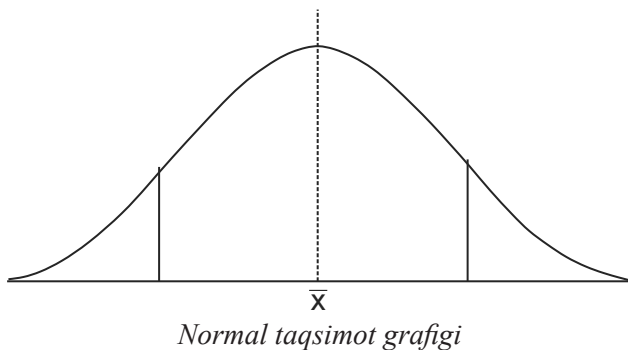
Testning ishonchliligini hisoblashning bir qancha usullari mavjud. «Test – qayta test» va «parallel test formalari» usullarini qo'llash uchun aynan bir test topshiruvchilar guruhidan ikki marta test olish va test natijalari o'rtasidagi korrelyatsiyani ko'rish kerak. Spirman-Braun, Kuder Richardson yoki Kronbax alfa formulalari uchun bir marta test o'tkazish yetarli. Ishonchlilik ko'rsatkichi 0,8 va undan yuqori bo'lgan hollarda test ishonchli hisoblanadi.

3.3. Normal taqsimot. O'lchashning standart xatosi. Tasniflash barqarorligi

Test topshiruvchining testda olgan bahosi uning haqiqiy bahosi emas, chunki har qanday testda olingan baho haqiqiy baho va xatolikdan iborat. Agar test ishonchli bo'lsa,

test topshiruvchi testni qayta-qayta topshirganida ham uning natijalari bir-biriga yaqin bo‘laveradi, lekin bu natijalar ham uning haqiqiy bahosi emas. Test topshiruvchining haqiqiy bahosini bilish uchun o‘lchashning standart xatosini bilishimiz kerak. O‘lchashning standart xatosi va uning qo‘llanilishini tushunish uchun, birinchi navbatda, normal taqsimot haqida tushunchaga ega bo‘lish lozim.

XVIII asr oxiri – XIX asrning birinchi yarmida yashagan buyuk nemis olimi Karl Fridrix Gauss birinchi bo‘lib normal taqsimot qonunini tasvirlab bergan. Uning fikricha, tabiatdagi kattaliklarning qiymatlarini o‘rtacha qiymat atrofida grafik shaklida joylashtirsak, qo‘ng‘iroqsimon grafik hosil bo‘ladi¹:



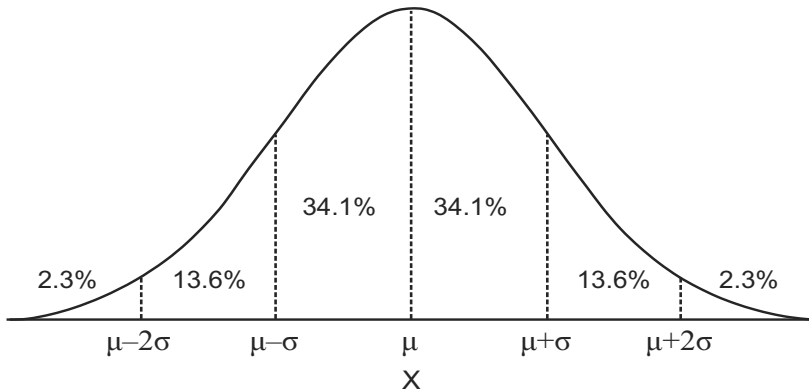
Masalan, biron-bir guruhdagi kishilarning bo‘yini grafik shaklida tasvirlamoqchi bo‘lsak, o‘rta bo‘yli kishilar soni ko‘proq bo‘ladi va grafikning o‘rtasida joylashadi, bo‘yi past

¹ Кирейцева А.Н. Албука тестирования. – Санкт-Петербург: Златоуст, 2013. – С.27.

va bo'yi yuqori kishilarning soni kamroq bo'ladi, grafikning chap va o'ng tomonlariga joylashadi. Bu holat **normal taqsimot** deyiladi.

Normal taqsimot – ko'p sonli tasodifiy o'zgaruvchilarning simmetrik shaklda joylashuvi.

Normal taqsimot grafigini quyidagicha teng bo'laklarga bo'lish mumkin:



Normal taqsimot grafigi

Grafikdan ko'rinib turibdiki, guruhimizda bo'yi o'rtacha va o'rtachadan bir pog'ona yuqori bo'lganlari 34,1 foizni, bo'yi o'rtacha va o'rtachadan bir pog'ona past yoki yuqori bo'lganlari 68,2 foizni tashkil qiladi. O'rtacha ko'rsatkichdan farq **standart og'ish** deyiladi.

Standart og‘ish – guruhda guruh a‘zolari ko‘rsatkichlarining guruhning o‘rtacha ko‘rsatkichidan farqini ko‘rsatuvchi qiymat.

Standart og‘ish, odatda, dispersiyaning musbat kvadrat il-diziga teng va matematikada σ (grekcha: *sigma*) harfi bilan belgilanadi.

Yuqoridagi misolda ko‘rganimizdek, guruhimizda bo‘yi o‘rtacha (μ) va o‘rtachadan bir standart og‘ishga past ($\mu-\sigma$) yoki yuqori ($\mu+\sigma$) bo‘lganlar 68,2 foizni tashkil etadi. Bo‘yi o‘rtachadan ikki standart og‘ishga past ($\mu-2\sigma$) yoki yuqori ($\mu+2\sigma$) bo‘lganlar 27,2 foizni tashkil etadi. Shunday qilib, ak-sariyatining (95,4 foiz) bo‘yi o‘rtachadan maksimum 2 stan-dart og‘ishga farq qiladi.

Xuddi shunday, klassik test nazariyasiga ko‘ra, test olgani-mizda testning natijalarini grafik shaklida ifodalasak, normal taqsimot grafigi hosil bo‘lishi kerak: o‘rtacha va o‘rtachaga yaqin ball olganlar ko‘pchilikni, haddan tashqari past yoki yuqori ball olganlar esa kamchilikni tashkil qilishi lozim.

Yuqorida ko‘rganimizdek, klassik test nazariyasiga ko‘ra, alohida olingan test topshiruvchining konkret testda olgan bahosi uning haqiqiy bahosi va o‘lchash xatoligidan iborat. Agar biz test topshiruvchidan juda ko‘p (yoki cheksiz marta) test olsak, uning olgan baholari ham normal taqsimot grafigi singari qo‘ng‘iroqsimon shaklda joylashar edi. Test topshiruv-chining konkret bir testda olgan bahosida uning haqiqiy ba-

hosi va xatolikning ulushini bilish uchun esa **o‘lchashning standart xatosini** bilishimiz kerak.

O‘lchashning standart xatosi – test topshiruvchini ayni bir test yordamida qayta-qayta baholaganda test topshiruvchi olgan baholari uning haqiqiy bahosiga qay darajada yaqinligini ko‘rsatuvchi qiymat.

O‘lchashning standart xatosi quyidagi formula yordamida hisoblanadi:

$$s_e = s_x \sqrt{1 - r_{xx}}$$

bunda: S_e = o‘lchashning standart xatosi;

S_x = test natijalarining standart og‘ishi;

r_{xx} = testning ishonchliligi.

Misol uchun, testning ishonchliligi 0,91 ga, standart og‘ishi 7,5 ga teng bo‘lsa, o‘lchashning standart xatosi 2,25 ga teng bo‘ladi:

$$s_e = 7,5 \sqrt{1 - 0,91} = 7,5 \sqrt{0,09} = 7,5 \times 0,3 = 2,25$$

Agar test topshiruvchimiz testda 37 ball olgan bo‘lsa, uning haqiqiy bahosi 68,2% holda $37 \pm 2,25$ ga teng, 95,4% holda esa $37 \pm 2 \times 2,25$ ga teng. Ya’ni xuddi shu testni qayta va qayta ishlatganimizda 68,2% holda test topshiruvchimizning bahosi 34,75 va 39,25 oralig‘ida bo‘lardi, 95,4% holda esa 32,5 va 41,5 oralig‘ida bo‘lardi.

Demak, test topshiruvchining konkret bir testdan olgan bahosini 68,2% ishonch bilan uning haqiqiy bahosidan bir pog‘ona (o‘lchashning standart xatosi ko‘rsatkichida) yuqori yoki past, 95,4% ishonch bilan uning haqiqiy bahosidan ikki pog‘ona yuqori yoki past deyishimiz mumkin. Test topshiruvchining haqiqiy bahosini hech qachon bila olmaymiz, ya’ni hech qachon 100% ishonch bilan test topshiruvchining testda olgan bahosi uning haqiqiy bahosiga teng deya olmaymiz. Shu bilan birga, **o‘lchashning standart xatosi qanchalik past bo‘lsa, test topshiruvchining olgan bahosi uning haqiqiy bahosiga aksar hollarda shunchalik yaqin bo‘ladi.** Testni yaratishga mas’ul tashkilot testning ishonchlilik ko‘rsatkichlari bilan birga o‘lchashning standart xatosini ham e’lon qilishi lozim.

Bilamizki, test natijalari asosida ma’lum bir qarorlar qabul qilinadi. Masalan, Davlat test markazi (DTM) o‘tkazadigan testlar natijalari asosida abituriyentlar oliy ta’lim muassasalariga qabul qilinadilar. Kirish balidan past ball to‘plaganlar OTMga qabul qilinmaydi. Deylik, Aliyev OTMga topshirib, kirish imtihonida 148 ball oldi. Kirish bali esa 150 edi. Lekin, yuqorida ko‘rganimizdek, 148 ball Aliyevning haqiqiy bahosi emas. Agar DTM ishlatgan testda o‘lchashning standart xatosi 3 ga teng bo‘lsa, 68,2% ishonch bilan Aliyevning haqiqiy bahosi 145 – 151 oralig‘ida, 95,4% ishonch bilan uning haqiqiy bahosi 142 – 154 oralig‘ida deyishimiz mumkin. Aliyev xuddi shu testni boshqa kuni topshirganda 148 balldan

ko‘proq ham, kamroq ham olishi mumkin edi. Demak, kirish imtihonidan o‘ta olmagan Aliyev xuddi shu testni boshqa kuni topshirganida o‘tishi ham mumkin edi. Bunday holat **tasniflash xatosi** deyiladi.

Tasniflash xatosi – testning test topshiruvchini o‘tish baliga nisbatan noto‘g‘ri tasniflashi.

Tasniflash xatosi ikki xil bo‘lishi mumkin: **yolg‘on ijobiy** – haqiqiy bahosi o‘tish balidan past bo‘lgan test topshiruvchining testdan o‘tgan deb tasniflanishi va **yolg‘on salbiy** – haqiqiy bahosi o‘tish balidan yuqori bo‘lgan test topshiruvchining testdan o‘tmagan deb tasniflanishi.

Testda o‘tish bali test topshiruvchilarning o‘rtacha baliga qanchalik yaqin bo‘lsa, tasniflash xatosi ehtimoli shunchalik ko‘proq bo‘ladi¹. Masalan, agar test topshiruvchilarning o‘rtacha bali 25 ga, o‘tish bali esa 27 ga teng bo‘lsa, yolg‘on ijobiy va yolg‘on salbiy holatlar ko‘proq bo‘ladi, chunki o‘tish baliga yaqin ball to‘plaganlar soni ko‘p bo‘ladi. Agar test topshiruvchilarning o‘rtacha bali 25 ga, o‘tish bali esa 45 ga teng bo‘lsa, yolg‘on ijobiy va yolg‘on salbiy holatlar kamroq bo‘ladi, chunki o‘tish baliga yaqin ball to‘plaganlar soni kam bo‘ladi. Birinchi holatda testing **tasniflash barqarorligi** pastroq, ikkinchi holatda esa yuqoriroq bo‘ladi.

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.124.

Tasniflash barqarorligi – test natijalariga ko‘ra to‘g‘ri tasniflangan test topshiruvchilar foizi.

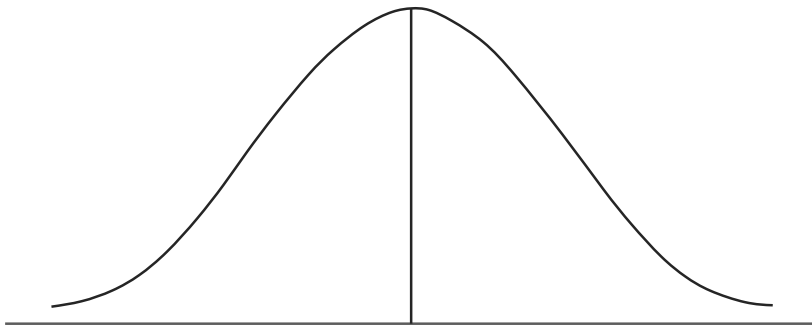
Masalan, bir testni test topshiruvchilar guruhi ikki marta topshirsa va har ikkala holatda ham «o‘tmadi» deb baholanganlar 27 foizni, har ikkala holatda ham «o‘tdi» deb baholanganlar 62 foizni, bir holatda «o‘tdi», ikkinchi holatda «o‘tmadi» (yoki teskari) deb baholanganlar 11 foizni tashkil qilsa, bu testning tasniflash barqarorligi 89% ga teng. Ya’ni bu testni boshqa test topshiruvchilar guruhida ishlatganimizda ham 89% test topshiruvchilar «o‘tdi-o‘tmadi»ga to‘g‘ri ajratiladi. Tasniflash barqarorligini hisoblash uchun testni bir test topshiruvchilar guruhida bir necha marta o‘tkazib, natijalarni solishtirish kerak.

Asosiy xulosa:

Test topshiruvchining haqiqiy bahosi aksar hollarda uning testda olgan bahosidan eng ko‘pi bilan o‘lchashning standart xatosining ikki baravarigacha past yoki yuqori bo‘lishi mumkin. O‘lchashning standart xatosi qanchalik kichik bo‘lsa, test topshiruvchining testda olgan bahosi uning haqiqiy bahosiga shunchalik yaqin bo‘ladi. Tasniflash barqarorligi yuqori bo‘lgan testlar ko‘proq test topshiruvchilarni «o‘tdi-o‘tmadi»ga to‘g‘ri ajratib beradi.

3.4. Testning asosiy tendensiyasi ko'rsatkichlari

Testning sifati bilan bog'liq muhim ko'rsatkichlarga testning **asosiy tendensiyasi** bilan bog'liq ko'rsatkichlar ham kiradi. Bu ko'rsatkichlar normal taqsimot qonuniga asoslanadi. Yuqorida ko'rganimizdek, test natijalari grafik shaklida ifodlanganda qo'ng'iroqsimon shakl hosil bo'lishi kerak: o'rtacha va o'rtachaga yaqin ball olganlar ko'pchilikni tashkil qilib, juda past yoki juda yuqori ball olganlar kamchilikni tashkil qilishi lozim. Bunday holatda testning asosiy tendensiya ko'rsatkichlari – **o'rta arifmetik, mediana va moda** o'zaro teng bo'ladi.



O'rta arifmetik
Mediana
Moda

*Normal taqsimot grafigida moda, mediana
va o'rtacha arifmetikning joylashuvi*

O'rta arifmetik – bu test topshiruvchilar natijalarining o'rtacha qiymati. O'rta arifmetikni hisoblash uchun test top-

shiruvchilar ballarining yig'indisi olinadi va test topshiruvchilar soniga bo'linadi.

Misol uchun, 7 kishi test topshirdi va ularning ballari quyidagicha bo'ldi:

10, 9, 9, 8, 7, 6, 5

O'rta arifmetikni hisoblash uchun:

– jami ballar yig'indisini olamiz: $10+9+9+8+7+6+5=54$;

– natijani test topshiruvchilar soniga bo'lamiz: $54/7=7,7$.

Demak, o'rta arifmetik 7,7 ga teng.

Boshqa bir misol: 8 kishi test topshirdi va ularning ballari quyidagicha bo'ldi:

8, 7, 7, 6, 5, 4, 4, 3

O'rta arifmetikni hisoblash uchun:

– jami ballar yig'indisini olamiz: $8+7+7+6+5+4+4+3=44$;

– natijani test topshiruvchilar soniga bo'lamiz: $44/8=5,5$.

Demak, o'rta arifmetik 5,5 ga teng.

Mediana – test natijalarini kamayish ketma-ketligida joylashtirganimizda ketma-ketlikni teng ikkiga ajratuvchi son.

Yuqoridagi birinchi misolimizda: 10, 9, 9, 8, 7, 6, 5

mediana 8 ga teng, chunki 8 sonigacha 3 ta raqam, 8 sonidan keyin ham 3 ta raqam joylashgan.

Yuqoridagi ikkinchi misolimizda: 8, 7, 7, 6, 5, 4, 4, 3

mediana 5,5 ga teng, chunki o'rtadagi 6 va 5 sonlarigacha 3 ta raqam, bu sonlardan keyin ham 3 ta raqam joylashgan, o'rtada ikkita son bo'lib qolganda esa ularning o'rta arifmetik qiymati olinadi: $(6+5)/2=5,5$.

Moda – raqamlar ketma-ketligida eng ko‘p uchraydigan raqam. Ayrim hollarda raqamlar ketma-ketligida bittadan ortiq moda bo‘lishi mumkin.

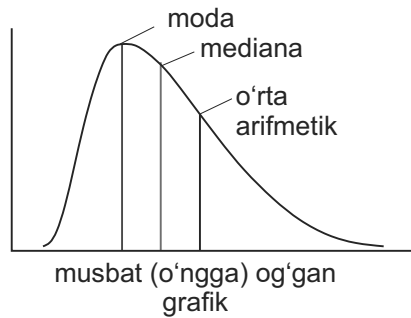
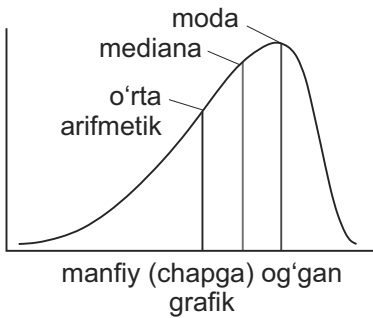
Yuqoridagi birinchi misolimizda: 10, 9, 9, 8, 7, 6, 5

moda 9 ga teng, chunki 9 raqami ikki marta uchragan, qolgan raqamlar bir martadan uchragan.

Yuqoridagi ikkinchi misolimizda: 8, 7, 7, 6, 5, 4, 4, 3

ikkita moda mavjud: 7 va 4. Bu har ikki raqam ikki martadan uchragan, qolgan raqamlar esa bir martadan uchragan.

Agar test test topshiruvchilar guruhi uchun juda **oson** bo‘lsa, yuqori ballar ko‘p, **mediana va moda o‘rta arifmetikdan katta** bo‘ladi. Bunday test natijalari grafigi **manfiy og‘gan (chapga og‘gan)** deyiladi. Agar test test topshiruvchilar guruhi uchun juda **qiyin** bo‘lsa, past ballar ko‘p, **moda va mediana o‘rta arifmetikdan kichik** bo‘ladi. Bunday test natijalari grafigi **musbat og‘gan (o‘ngga og‘gan)** deyiladi.



Asosiy xulosa:

Normal taqsimot grafigida moda, mediana va orta arifmetiklar o'zaro teng bo'ladi. Juda oson testda mediana va moda o'rta arifmetikdan katta, juda qiyin testda moda va mediana o'rta arifmetikdan kichik bo'ladi.

3.5. O'lchashning standart xatosini hisoblashning muqobil usullari

Demak, klassik test nazariyasida test topshiruvchining haqiqiy bahosi – bu test topshiruvchidan ayni bir testni (yoki testning parallel formalarini) ishlatgan holda qayta-qayta test olganimizda qayd etilgan natijalarning o'rtacha qiymati. Biroq shunday holatlar bo'ladiki, test ma'lum bir domenni (konstruktni) o'lchaydigan topshiriqlarning katta bazasidan tasodifiy holda yig'ilgan cheklangan sondagi topshiriqlar yig'indisi bo'lishi mumkin. Masalan, tarix fanidan 1000 ta test topshirig'idan iborat test topshiriqlari bazasi yaratildi. Tarix kursi yakunida talabalarni baholash maqsadida ulardan kompyuterda test olindi. Kompyuter har bir talabaga bazadagi 1000 ta test topshirig'idan tasodifiy tarzda 100 tasini berdi. Demak, har bir test topshiruvchi o'zining individual test formasini (variantini) yechadi va kompyuter topshiriqlarni tasodifiy yo'l bilan tanlagani uchun bu variantlar **parallel emas**. Bunday holda test topshiruvchining haqiqiy

bahosi va o'lchashning standart xatosini hisoblash uchun boshqacha yo'l tutiladi.

Bunday holda test topshiruvchining haqiqiy bahosi quyidagicha hisoblanadi:

$$T_a = nP_a$$

bunda: T_a = test topshiruvchining haqiqiy bahosi;

n = test topshiruvchiga berilgan test topshiriqlari soni;

P_a = test topshiruvchi bazadan yecha oladigan test topshiriqlari foizi.

Masalan, kompyuter Aliyevga bazadagi 1000 ta test topshirig'idan 100 tasini berdi. Aliyev 80 ball oldi. Agar Aliyev testni ertaga yoki boshqa kun topshirganida, kompyuter unga bazadan boshqa 100 ta test topshirig'ini berardi va Aliyev 78 ball olishi mumkin edi. Boshqa bir safar esa 72 ball olishi mumkin edi. Agar unga bazadagi 1000 ta test topshirig'ining hammasini yechtirganimizda u 750 tasini (75% ni) yecha olardi. Aliyevning haqiqiy bahosi $100 \times 75\% = 75$ ga teng. Biroq biz Aliyevga (yoki boshqa test topshiruvchilarga) bazadagi barcha test topshiriqlarini yechtira olmaymiz. Bunday holda baholashning standart xatosi **binominal taqsimot** formulasi asosida aniqlanadi¹.

Binominal taqsimot – tasodifiy tajribalar ketma-ketligida «omadli» natijalarning taqsimoti, bunda har bir tajribada «omadli» natija ehtimoli o'zgarmas bo'ladi.

¹ Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов. – Москва: Логос, 2010. – С.178.

Tasavvur qiling, biz tangani 50 marta tashlasak, u har safar yo yuza tarafi bilan, yo orqa tarafi bilan tushadi. Tanga 50 martadan 49 marta yuza tarafi bilan, 1 marta orqa tarafi bilan tushishi mumkin yoki 48 marta yuza tarafi bilan, 2 marta orqa tarafi bilan tushishi mumkin va hokazo. Biz bilamizki, har safar tangani tashlaganimizda faqat ikki xil natija kutilishi mumkin va bu ikki natijadan birini olishimiz ehtimoli mavjud. Testda ham kompyuter tanlagan test topshirig‘i Aliyev biladigan yoki bilmaydigan bo‘lishi mumkin. Aliyevga 100 martadan 1 marta u bilmaydigan test topshirig‘i, 99 marta u biladigan test topshirig‘i tushishi mumkin.

Real vaziyatda Aliyev faqat bir marta test topshiradi. Demak, biz faqat uning aynan shu test formasidan (variantidan) olgan bahani bilamiz. Bu natijani quyidagi formulaga qo‘yib o‘lchashning standart xatosini aniqlashimiz mumkin:

$$\sigma_a = \sqrt{\frac{X_{ga}(n - X_{ga})}{n - 1}}$$

bunda: X_{ga} = test topshiruvchining unga tushgan test variantida olgan bahosi;

n = test topshiruvchiga berilgan test topshiriqlari soni.

Masalan, Aliyev unga berilgan 100 ta test topshirig‘idan 80 tasini to‘g‘ri yechib, 80 ball olgan bo‘lsa, o‘lchashning standart xatosi 4,02 ga teng:

$$\sigma_a = \sqrt{\frac{80(100 - 80)}{100 - 1}} = \sqrt{\frac{1600}{99}} = \sqrt{16,16} = 4,02$$

Binominal formula asosida o'lchashning standart xatosi hisoblanganda quyidagilarni yodda tutish lozim:

– klassik test nazariyasida o'lchashning standart xatosi bir marta hisoblansa, binominal o'lchashning standart xatosi har xil test formalari uchun har xil qiymatga ega bo'lishi mumkin;

– binominal o'lchashning standart xatosi X_{ga} ning qiymati n ning oraliq qiymatiga yaqin bo'lganda kattaroq, n ning minimal va maksimal qiymatlariga yaqin bo'lganda esa kichikroq bo'ladi.

Asosiy xulosa:

Test topshiruvchi test topshiriqlarining katta bazasidan tasodifiy yo'l bilan tanlangan cheklangan sondagi test topshiriqlarini yechganida o'lchashning standart xatosi binominal taqsimot formulasi asosida aniqlanadi.

4. TESTNI YARATISH

4.1. Testni yaratish bosqichlari

Har qanday testni yaratish tizimli jarayon bo‘lib, bu jarayon bir qancha **bosqichlarni** o‘z ichiga oladi¹:

1. Testning maqsadini aniqlash.
2. Domen va konstruktarni aniqlash.
3. Test spetsifikatsiyasini yaratish.
4. Test spetsifikatsiyasi asosida test topshiriqlarini shakllantirish.
5. Test topshiriqlarini ekspertlar (mutaxassislar) yordamida tahlil etish va lozim bo‘lgan o‘rinlarda tuzatishlar kiritish.
6. Test topshiriqlarining dastlabki aprobatsiyasini o‘tkazish va lozim bo‘lgan o‘rinlarda test topshiriqlariga tuzatishlar kiritish.
7. Test topshiriqlarini katta miqyosda aprobatsiyadan o‘tkazish.
8. Aprobatsiya natijalari asosida test topshiriqlarini statistik tahlil etish, talabga javob bermaydigan topshiriqlarni chiqarib tashlash.

¹ Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов. – Москва: Логос, 2010. – С.105.

9. Testning yakuniy varianti (formasi) uchun validlik va ishonchlilik tadqiqotlarini o'tkazish.

10. Test natijalari talqini uchun tavsiyalar tuzish.

Ushbu bosqichlar har qanday o'lchov vositasi (test) yaratilishida majburiy bosqichlar bo'lib, ayrim hollarda bosqichlarni bir necha marta qayta amalga oshirish kerak bo'lishi mumkin. Masalan, test topshiriqlarini katta miqyosda aprobatsiyadan o'tkazgandan keyin talabga javob bermaydigan topshiriqlarni qayta ishlab, takroriy aprobatsiyadan o'tkazish mumkin. Validlik va ishonchlilik talablariga javob beradigan testlarni yaratish uchun bir necha yilgacha vaqt talab qilinishi mumkin.

Birinchi bosqichda testning **maqsadi** aniqlanadi. Testning maqsadi uning qanday shaklda bo'lishi, qanday xarakteristikalariga ega bo'lishini aniqlab beradi. Masalan, o'quvchilarning bilimlaridagi bo'shliqlarni aniqlash maqsadida ishlatiladigan testga, odatda, o'quvchilar ko'proq xato qiladigan nuqtalar kiritiladi. Fanni o'qitishning oxirida o'zlashtirganlikni baholashga qaratilgan testda fanning muhim bo'limlari va qonuniyatlari aks etishi lozim.

Ikkinchi bosqichda baholanishi (o'lchanishi) kerak bo'lgan **domen va konstruktlar** aniqlanadi. Buning uchun bir qancha usullar qo'llanilishi mumkin:

– ta'lim standartlarida belgilangan maqsadlarni tahlil etish. Masalan, maktab bitiruvchilari uchun ona tilidan test yaratmoqchi bo'lsak, birinchi navbatda, ta'lim standartlarida mak-

tab bitiruvchilari bu fandan qanday bilim, ko‘nikma, malaka, kompetensiyalarga ega bo‘lishlari kerakligi belgilangan degan savolga javob izlashimiz darkor. Ta‘lim standartlarida belgilangan barcha maqsadlarni ham o‘lchash (baholash) imkoni bo‘lmasligi mumkin, bunday holda aynan qaysi ta‘limiy maqsadlar o‘lchanishini aniqlab olish lozim¹;

– soha mutaxassislarining fikrini o‘rganish. Masalan, pedagogik mahoratni baholaydigan test yaratmoqchi bo‘lsak, pedagogik tajribaga ega, malakali o‘qituvchilar o‘rtasida so‘rov o‘tkazib, pedagog uchun muhim xususiyatlarni aniqlab olishimiz mumkin;

– o‘lchanishi lozim bo‘lgan konstruktlar qo‘llaniladigan vaziyatlarni kuzatish. Masalan, mehmonxona xodimlarini baholash uchun test yaratishda bevosita mehmonxona xodimlari ishini kuzatish va bu kuzatish jarayonida ular uchun kerakli ko‘nikmalarni aniqlashtirib olish mumkin;

– ilmiy tadqiqotlarni tahlil qilish. Masalan, o‘qish savodxonligini baholaydigan test tuzish uchun «o‘qish savodxonligi nima?» degan savolga javob izlashimiz, buning uchun esa shu sohadagi ilmiy tadqiqotlarni o‘rganib chiqishimiz talab etiladi.

Aksar hollarda domen va konstruktarni aniqlashda bir emas, bir necha usullar birdaniga qo‘llaniladi. Test vaqt va hajm jihatidan cheklanganligi sababli u orqali domen va

¹ Чельщикова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002. – С.86.

konstruktlarning barchasini qamrab olish yoki barchasini teng darajada aks ettirish imkoni bo'lmasa, bunday holda tekshirilishi lozim bo'lgan domen va konstruktlarning (hamda ularning tarkibiy qismlarining) muhimlik darajasi bo'yicha iyerarxiasini tuzib, qaysi elementlar testda qay darajada aks etishini aniqlab olish lozim. Shuningdek, bir-biriga yaqin elementlarni birlashtirish ham samarali bo'lishi mumkin¹.

Uchinchi bosqichda test spetsifikatsiyasi yaratiladi. Spetsifikatsiyada baholanishi lozim bo'lgan domen va konstrukting har bir elementi qanday va nechta sondagi test topshiriqlari yordamida tekshirilishi aks etadi. Test spetsifikatsiyasini yaratishda nafaqat domenning mazmuniga kiruvchi tarkibiy qismlarni, balki test topshiruvchidan talab qilinadigan kognitiv operatsiya turini ham hisobga olish kerak. Buning uchun kognitiv jarayonlarni tasniflaydigan nazariy model ishlatiladi. Ta'lim va baholash jarayonida keng qo'llaniladigan shunday nazariy modellardan biri AQSh olimi Benjamin Blum (*Benjamin Bloom*) tomonidan yaratilgan bo'lib, **Blum taksonomiyasi** deb ataladi.

¹ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.102.

	Kognitiv ko'nikma	Izoh	Misollar
6.	Baho berish	Ichki yoki tashqi mezonlar asosida holat, voqea-hodisaga baho berish	– Asarga taqriz yozing. – Taklif qilingan yechimning samaradorligini baholang.
5.	Sintez	Tarkibiy qismlarni yangicha shaklda birlashtirish asosida yangi narsa yaratish, muammoning yangi yechimini taklif qilish	– Matnning qisqacha mazmunini yozing. – Asarga sikvel yozing. – Berilgan masalani yechadigan kompyuter dasturi tuzing.
4.	Analiz (tahlil)	Murakkab tushunchani tarkibiy qismlarga ajratish va qismlar orasidagi bog'liqlikni ko'ra olish	– Matndan asosiy va ikkilamchi g'oyalarni ajrating. – Ikkinchi jahon urushining SSSRdagi ijtimoiy hayotga ta'sirlarini sanang. – Berilgan o'simlikning gul va barglarini o'rganib, uning qaysi turkumga mansubligini aniqlang.
3.	Qo'llash	Olingan bilim asosida yangi muammolarni yechish	– Berilgan asarga muqobil (boshqacha) yakun yozing. – Elektr o'tkazuvchining qarshiligini hisoblang.
2.	Tushunish	Olingan bilimni talqin qilish, xulosalar chiqarish, boshqa shaklga solish	– Berilgan gapda undalmani toping. – Aruz va barmoq vaznlarining farqlarini sanang. – Berilgan xususiyatlar asosida davlat boshqaruvi shaklini aniqlang.
1.	Bilish	O'rganilgan ma'lumotni eslash va qaytarish	– «Kesim» tushunchasiga ta'rif bering. – «O'tkan kunlar» romani qachon yozilganligini ayting. – Ikkinchi jahon urushining asosiy janglarini sanang.

Blum taksonomiyasi bo'yicha kognitiv ko'nikmalar

Blum taksonomiyasi bo'yicha kognitiv ko'nikmalar (bunda bilish – eng quyi darajadagi ko'nikma, baho berish – eng yuqori darajadagi ko'nikma)

Blum taksonomiyasi – kognitiv (bilish jarayoni bilan bog‘liq) ko‘nikmalarning iyerarxik (ketma-ketlikka asoslangan) modeli.

Blum taksonomiyasida kognitiv ko‘nikmalar ikki darajaga bo‘linadi: quyi darajadagi kognitiv ko‘nikmalar va yuqori darajadagi kognitiv ko‘nikmalar¹. Bilish, tushunish va qo‘llash jarayonlari quyi darajadagi kognitiv ko‘nikmalar hisoblansa, tahlil (analiz), sintez va baho berish yuqori darajadagi kognitiv ko‘nikmalardir. Test spetsifikatsiyasini tuzishda taksonomiyadagi ko‘nikmalarning fan mavzulari (bo‘limlari) kesimida aks etishini hisobga olish lozim, buning uchun har bir **mazmun-mundarija doirasida qaysi darajadagi kognitiv ko‘nikmalar baholanishi kerakligini** aniqlashtirib olish dar-
kor². Masalan, adabiyot fanidan test topshiruvchi «O‘tkan kunlar» romani qachon yozilganligini bilishi yoki undagi qahramonlarni ajrata olishi – bir daraja, romanda qanday muammolar ko‘tarilganligini tahlil qilib bera olishi yoki roman qahramonlarining xatti-harakatlariga baho berishi – boshqa daraja. Test spetsifikatsiyasida domenning har bir elementi qaysi darajada yoki qaysi darajagacha baholanishi aniq ko‘rsatilishi lozim.

Test spetsifikatsiyasi yaratilayotganida mavzular va ko‘nikmalar kesimini jadval shaklida ifodalash mumkin:

¹ Petty G. Teaching Today: A Practical Guide. 4th ed. – Cheltenham: Nelson Thornes, 2009. – P.9.

² Brookhart S.M. How To Assess Higher-Order Thinking Skills in Your Classroom. – Alexandria, VA: ASCD, 2010. – P.18.

Mavzular	Test topshiriqlari soni				Mavzu bo'yicha jami
	Tushunish	Qo'llash	Tahlil	Sintez	
1-mavzu	1	2	2	1	6
2-mavzu	1	2	2	1	6
3-mavzu	1	2	2	1	6
Jami	3	6	6	3	18

Test topshiriqlari sonining mavzular va ko'nikmalar kesimida taqsimlanishi namunaviy jadvali

Domen va konstruktarni tekshirishga mos vositalar (test topshiriqlari turlari) tanlanayotganda «tekshirilayotgan konstrukt (xususiyatni) ko'rishimiz uchun qanday **dalil** maqbul hisoblanadi?» va «berilgan test topshirig'ini bajarish jarayonida test topshiruvchi biz tekshirmoqchi bo'lgan kognitiv ko'nikmani ishga soladimi?» – degan savollarga javob berish maqsadga muvofiq¹. Masalan, test topshiruvchida badiiy matnni tahlil etish ko'nikmasi shakllanganligiga ishonch hosil qilmoqchi bo'lsak, test topshiruvchi qanday amalni bajarishi kerak? Nimani biz unda badiiy matnni tahlil etish ko'nikmasi shakllanganligiga maqbul dalil sifatida qabul qilamiz? Agar biz badiiy matndan parcha berib, uning qaysi asardan olinganligini so'rasak, test topshiruvchi bu topshiriqni bajarish jarayonida biz tekshirmoqchi bo'lgan tahlil ko'nikmasini ishga soladimi? Shu savollarga javob bera olsak, testga qanday topshiriqlar kiritish lozimligi oydinlashadi.

¹ Brookhart S.M. How To Assess Higher-Order Thinking Skills in Your Classroom. – Alexandria, VA: ASCD, 2010. – P.19.

Test spetsifikatsiyasi qanchalik batafsil va tushunarli shaklda tuzilsa, ushbu spetsifikatsiya asosida test topshiriqlarini tuzish va yakuniy testni (testning parallel formalarini) shakllantirish shunchalik yengillashadi.

Keyingi bosqichlarda test topshiriqlari yaratiladi va ekspertlar (mutaxassislar) tomonidan **tahlil etiladi**. Test topshiriqlarini tahlil etishda ularning test spetsifikatsiyasiga mosligi, test topshiriqlarini yaratish qoidalariga mosligi, tushunarligi, aniqligi singari jihatlariga e'tibor beriladi. Bunda bir qancha ekspertlar qatnashib, har bir ekspert test topshirig'ining bir jihatini tahlil qilib berishi yoki barcha ekspertlar har tomonlama tahlil qilishlari mumkin. Bunda ekspertlar oldiga aniq masalalar qo'yish ekspertiza jarayonining samaradorligini oshiradi.

Oltinchi bosqichda test topshiriqlarining **dastlabki aprobatsiyasi** o'tkaziladi.

Aprobatsiya – testning psixometrik xarakteristikalarini (ya'ni testning shaxsdagi xususiyatlarni qay darajada o'lchay olishini) aniqlash hamda testdagi noaniqliklar va xatolarni aniqlash maqsadida o'tkaziladigan jarayon.

Dastlabki aprobatsiya kichik guruhda o'tkaziladi va bunda aprobatsiya jarayonida qatnashayotganlar test topshiriqlarini qanday yechayotganliklari kuzatiladi. Aprobatsiyadan keyin qatnashuvchilar bilan suhbatlashib, ularning test topshiriqlari haqidagi fikrlarini o'rganish, shuningdek, test topshiriqlarini

statistik tahlil etib, lozim bo'lgan o'rinlarda ularga o'zgartirishlar kiritish maqsadga muvofiq.

Keyingi bosqichlarda test topshiriqlari katta miqyosda **aprobatsiyadan** o'tkaziladi. Aprobatsiyada qatnashuvchilar o'z xususiyatlari jihatdan haqiqiy test topshiruvchilarga yaqin bo'lishi, aprobatsiya sharoitlari ham imkon qadar haqiqiy testdan farq qilmasligi lozim. Aprobatsiyadan keyin test topshiriqlari **statistik tahlil** etiladi, talabga javob bermaydigan topshiriqlar chiqarib tashlanadi yoki qayta ishlanadi. Qayta ishlangan topshiriqlar qayta aprobatsiyaga olib chiqilishi lozim bo'ladi. Shundan so'ng testning yakuniy formasi (formalari) **validlik va ishonchlilik** talablari nuqtayi nazaridan **tahlil** etiladi. Test bu talablarga javob berishiga ishonch hosil qilingandan keyin test natijalarini qanday talqin etish va ulardan qanday xulosalar chiqarish bo'yicha tavsiyalar tuziladi.

Asosiy xulosa:

Testni yaratish jarayoni bir qancha muhim bosqichlardan iborat.

4.2. Yopiq test topshiriqlarini yaratish

Demak, testning maqsadi, test orqali o'lchanishi lozim bo'lgan domen va konstruktlar aniqlab olingandan keyin test spetsifikatsiyasi yaratiladi va bu spetsifikatsiya asosida test

topshiriqlari yoziladi. Testda ishlatish mumkin bo'lgan test topshiriqlari turlaridan biri – **yopiq test topshiriqlari**, ya'ni test topshiruvchidan to'g'ri javobni berilgan muqobil javob variantlari ichidan **tanlashini** talab qiladigan test topshiriqlari. Bunday test topshiriqlarining eng katta afzalligi test natijalarini maxsus qurilmalar yordamida **tez qayta ishlash** hamda tekshirishda **subyektivlikni yo'qotish** imkonini berishidir.

Yopiq test topshiriqlari **samarali** bo'lishi, o'z vazifasini bajarishi va **validlik talablariga** javob berishi uchun ularni tuzishda bir qancha qoidalarga rioya qilinishi lozim. Testologiya talablariga muvofiq, to'g'ri tuzilgan yopiq test topshiriqlari tekshirilishi kerak bo'lgan domenni keng va har tomonlama qamrab olish imkonini beradi, nafaqat ma'lumotni bilish (eslash) va qo'llashni, balki yuqori darajadagi kognitiv ko'nikmalarni ham tekshirishi mumkin. Shu bilan birga, bunday test topshiriqlari yordamida yuqori darajadagi kognitiv ko'nikmalar integratsiyasini hamda kognitiv bo'lmagan ko'nikmalarni (masalan, yozish yoki gapirish ko'nikmalari, amaliy ko'nikmalar va hokazo) tekshirish deyarli imkonsiz – bunday maqsadlar uchun ochiq test topshiriqlarini qo'llash maqsadga muvofiq¹.

Yopiq test topshiriqlari bir qancha shakllarda bo'lishi mumkin: muqobil javobli test topshirig'i, «to'g'ri-noto'g'ri»

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.290 – 291.

shaklidagi test topshirig‘i, moslashtirishni talab qiluvchi test topshirig‘i va ketma-ketlikni topishni talab qiluvchi test topshirig‘i. Test topshirig‘i turini tanlashda **moslik** tamoyiliga amal qilish talab etiladi: tanlangan test topshirig‘i turi baholamayotgan ko‘nikmani ko‘rsata olishi zarur¹. Masalan, maqsad test topshiruvchining berilgan matnning qisqacha mazmunini yoza olish ko‘nikmasini baholash bo‘lsa, buni yopiq test topshirig‘i yordamida amalga oshirib bo‘lmaydi.

<i>Yopiq test topshiriqlarining turlari</i>	<i>Izoh</i>	<i>Misollar</i>
Muqobil javobli test topshirig‘i	Berilgan muqobil javob variantlarining ichidan to‘g‘risini tanlash talab qilinadi	O‘tkir Hoshimov qalamiga mansub romanni toping. A) «Diyonat» B) «Tushda kechgan umrlar» C) «Odam bo‘lish qiyin» D) «Shaytanat»
«To‘g‘ri-noto‘g‘ri» shaklidagi test topshirig‘i	Berilgan jumlaning (hukmning) to‘g‘ri yoki noto‘g‘riligini topish talab qilinadi	Quyida berilgan hukmlarning to‘g‘ri yoki noto‘g‘riligini belgilang. 1. Ikkinchi jahon urushi 1941-yil 22-iyunda Germaniyaning SSSRga hujum qilishi bilan boshlandi. A) To‘g‘ri B) Noto‘g‘ri 2. Ikkinchi jahon urushi 1945-yil 2-sentyabrda Yaponiyaning taslim bo‘lishi bilan yakunlandi. A) To‘g‘ri B) Noto‘g‘ri

¹ Fein M. Test Development: Fundamentals for Certification and Evaluation. – Alexandria, VA: American Society for Training and Development, 2012. – P.94.

Moslashtirishni talab qiluvchi test topshirig'i	Berilgan ikki ro'yxatdan bir-biriga qaysidir jihati bilan mos tushuvchi javoblarni topish talab qilinadi	Quyida davlatlar va poytaxtlar nomlari berilgan. Har bir davlatning poytaxtini toping. Har bir javob faqat bir marta ishlatilishi mumkin. Diqqat: A – E javoblardan biri ortiqcha! <i>Davlatlar</i> <i>Poytaxtlar</i> 1. AQSh A) London 2. Avstraliya B) Kanberra 3. Kanada C) Ottava 4. Yangi Zelandiya D) Vashington E) Vellington
Ketma-ketlikni topishni talab qiluvchi test topshirig'i	Berilgan javob variantlarini to'g'ri ketma-ketlikda qo'yib chiqish talab qilinadi	Alisher Navoiy «Xamsa»siga kiruvchi dostonlarni to'g'ri ketma-ketlikda qo'yib chiqing. A) «Farhod va Shirin» B) «Saddi Iskandariy» C) «Hayrat ul-abror» D) «Layli va Majnun» D) «Sab'ayi sayyor» 1) 2) 3) 4) 5)

Yopiq test topshiriqlari turlari

4.2.1. Muqobil javobli test topshiriqlari

Muqobil javobli test topshiriqlari **o'zak** va uchta yoki undan ortiq **muqobil javoblardan** iborat bo'ladi. Muqobil javoblar ichida to'g'ri javob **kalit**, noto'g'ri javoblar **distraktorlar** (chalgituvchi javoblar, inglizcha «*distract*» – «*chalgitmoq*» so'zidan) deyiladi.

Test topshirig'ining **o'zagi** ikki xil ifodalanishi mumkin:

– **to'liqsiz gap** shaklida – bunda muqobil javoblarda berilgan so'zlar (jumlar) gapni to'ldiradi;

– **so‘roq gap** shaklida – bunda muqobil javoblarda berilgan so‘zlar, jumlar yoki gaplar berilgan so‘roqqa javob bo‘ladi.

Test topshirig‘ining **to‘g‘ri javobi (kaliti)** ikki xil bo‘lishi mumkin¹:

– **yagona** to‘g‘ri javob – bunda berilgan muqobil javoblardan faqat bittasi to‘g‘ri javob bo‘ladi;

– **eng maqbul** to‘g‘ri javob – bunda berilgan muqobil javoblardan bittasi eng maqbul to‘g‘ri javob bo‘lib, qolgan muqobil javoblar qaysidir jihati bilan qo‘yilgan savolga to‘liq javob bo‘la olmaydi.

O‘zbekiston Respublikasining poytaxti ... shahridir.

- A) Toshkent ← **kalit** } **muqobil javoblar** ← **o‘zak**
B) Nurafshon }
C) Samarqand } **distraktorlar**
D) Buxoro }

O‘zbekiston Respublikasida konstitutsiyaviy qonunlar qaysi organ tomonidan qabul qilinadi?

- A) Vazirlar Mahkamasi B) Oliy Majlis
C) Konstitutsiyaviy sud

Muqobil javobli test topshirig‘i namunasi

Muqobil javobli test topshiriqlarini tuzishda quyidagi **qoidalar va tavsiyalarga** rioya qilmoq lozim:

¹ Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013. – P.63.

1. Har bir test topshirig‘i test spetsifikatsiyasiga muvofiq o‘lchanayotgan domenning (konstruktning) **konkret bir elementini** baholash maqsadida tuzilishi lozim¹. Ya’ni har bir test topshirig‘ining maqsadi – domenning (konstruktning) qaysi elementini tekshirishga qaratilganligi aniq bo‘lishi, shu bilan birga, bitta test topshirig‘i orqali konstruktning bir nechta elementini qamrab olishga harakat qilinmasligi kerak.

Misol:

Yurtimizdagi qaysi shaharda Kamoliddin Behzod nomidagi memorial bog‘ barpo etilgan?²

A) Buxoro B) Toshkent C) Andijon D) Samarqand

Ushbu test topshirig‘i tarix fani bo‘yicha test topshiriqlari to‘plamiga kiritilgan bo‘lsa-da, o‘lchanayotgan domenning (tarix fanining) qaysi elementini baholashga qaratilganligi noaniq.

Yana bir misol:

Oddiy kitobxon aqlli ikkita qahramonning qismatidan hayajonlanib, uzoq yillar eslab yuradigan yozuvchi – haqiqiy yozuvchi.

Ushbu gapdagi ko‘makchi morfemalar tarkibida nechta o‘rinda quyi keng unlilar ishtirok etgan?

A) 8 B) 10 C) 7 D) 9

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

² Davlat test markazi. Tarix. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.53.

Ushbu test topshirig'ida ham ko'makchi morfemalarni, ham ular tarkibidagi quyi keng unlilarni topish so'ralmoqda, ya'ni test topshiruvchi bitta topshiriqqa to'g'ri javob berishi uchun ham morfemikadan, ham fonetikadan bilimini ishga solishi kerak bo'ladi. Bu holda muammoning yechimi – ikkita alohida test topshirig'i tuzib, bir test topshirig'ida ko'makchi morfemalarni, ikkinchisida quyi keng unlilarni so'rash.

Yana bir misol:

Alovuddin Muhammad Talos vodiysida qoraxitoylarni mag'lubiyatga uchratgan yilda qaysi shahar aholisi xorazmshohga qarshi qo'zg'olon ko'targan?¹

A) Buxoro B) Samarqand C) Toshkent D) O'tror

Ushbu test topshirig'ida ham yuqoridagi qoida buzilgan – test topshiruvchidan ham Alovuddin Muhammad Talos vodiysida qoraxitoylarni mag'lubiyatga uchratgan yilni topish, ham shu yili aholisi xorazmshohga qarshi qozg'olon ko'targan shaharni topish talab qilinmoqda. Test topshiruvchi yuqoridagi topshiriqqa to'g'ri javob bera olmasa, uning aynan nimani – yilni yoki shaharni bilmasligi haqida asosli xulosa chiqarib bo'lmaydi.

¹ Davlat test markazi: Tarix fanidan namunaviy savollar (2018). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

Yana bir misol:

Quyidagilardan rasmda tasvirlangan shaxs ko'rsatmasi va homiyligida yozilgan «Ravzat us-safo» asari muallifini toping¹.

- A) Mirxond B) Munis
C) Xondamir D) Abdurahmon Jomiy



Ushbu test topshirig'i tarix fani bo'yicha test topshiriqlari to'plamiga kiritilgan bo'lsa-da, o'lchanayotgan domenning (tarix fanining) qaysi elementini baholashga qaratilganligi noaniq. Shuningdek, test topshirig'iga javob berish uchun test topshiruvchi oldin rasmdagi shaxsni tanishi, so'ngra uning homiyligida yozilgan asar muallifini topishi kerak.

2. Har bir test topshirig'i baholanayotgan domen uchun **muhim bo'lgan bilimga** asoslanishi lozim, ya'ni test topshirig'i muhim bo'lmagan, mayda-chuyda detallarga asoslanmasligi kerak². Shu bilan birga, test topshirig'i **haddan tashqari umumiy** yoki **haddan tashqari xususiy** mazmun asosida tuzilmasligi zarur³. Haddan tashqari umumiy mazmun

¹ Davlat test markazi. Tarix. 2019-yil. Test topshiriqlari to'plami: amaliy qo'llanma. – Toshkent: Davr-press, 2019. – B.115.

² Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

³ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

asosida tuzilgan test topshirig‘iga tekshirilayotgan domenni yaxshi o‘zlashtirmagan, yuzaki bilimga ega test topshiruvchi ham to‘g‘ri javob berishi mumkin. Haddan tashqari xususiy mazmun asosida tuzilgan test topshirig‘iga tekshirilayotgan domenni yaxshi o‘zlashtirgan test topshiruvchi ham javob bera olmasligi mumkin.

Misol:

Anbar Otin qaysi buyuk mumtoz shoiramizning jiyani bo‘lgan?

- A) Nodirabegim B) Jahon otin Uvaysiy
C) Zebiniso D) Dilshodi Barno

Ushbu test topshirig‘i baholanayotgan domen – adabiyot uchun muhim bo‘lmagan ma’lumot asosida tuzilgan: Anbar Otin kimning jiyani bo‘lganligini bilish uning ijodini tushunishda hech qanday rol o‘ynamaydi.

Yana bir misol:

O‘zbekistonda «Fayzifilm» xususiy kinostudiyasi qachon ro‘yxatga olingan?¹

- A) 1992-yil B) 1991-yil C) 1993-yil D) 1999-yil

Ushbu test topshirig‘ida ham yuqoridagi qoida buzilgan: «Fayzifilm» kinostudiyasi qachon ro‘yxatga olinganligini

¹ Davlat test markazi: Tarix fanidan namunaviy savollar (2018). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

mustaqil O‘zbekiston tarixidagi muhim, ahamiyatga molik voqea deb bo‘lmaydi, demak, o‘quvchining bu ma‘lumotni bilish-bilmasligi tarixning shu davrini yaxshi bilish-bilmasligi haqida valid (asosli) xulosa chiqarish imkonini bermaydi.

Yana bir misol:

– *Ishga kirganiga to‘rt oy bo‘ldi-yu, to‘rt yillik charchadi bolam! – dedi ona. – To‘rt yillik qaridi bolam! ... Yo‘q, aravakashligi qursin! Boshqa ishga beramiz ... bo‘lmadi bu.*

«Kecha va kunduz» romanidan olingan ushbu parchada onasi tomonidan qayg‘urilayotgan aravakash obrazi qaysi javobda berilgan?¹

A) Hakimjon B) O‘lmasjon C) Zunnun D) Xolmat

Ushbu test topshirig‘i haddan tashqari xususiy mazmun asosida tuzilgan bo‘lib, testning maqsadi test topshiruvchining «Kecha va kunduz» romanini tushunishini tekshirishga qaratilgan bo‘lsa-da, mazkur asarni o‘qigan, u haqda tushunchaga ega test topshiruvchi ham bu topshiriqqa javob bera olmaydi.

3. Har bir test topshirig‘i **mustaqil** bo‘lishi, ya‘ni bir test topshirig‘iga javob berish uchun oldin boshqa bir test topshirig‘iga javob berish talab qilinmasligi kerak². Shu bilan bir-

¹ Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.112.

² Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

ga, testda ishlatilgan bir test topshirig'ining javobi xuddi shu testdagi boshqa test topshirig'ining javobini topishga yordam bermasligi lozim.

Misol:

1. Shekspirning Shotlandiya tarixi voqealariga bag'ishlangan fojiasi qanday nomlanadi?

- A) «Hamlet» B) «Qirol Lir»
C) «Makbet» D) «Romeo va Julyetta»

2. Ushbu asarda bosh qahramonning qanday g'ayritabiiy kuch bilan uchrashuvi uning keyingi xatti-harakatlari uchun zamin hozirlaydi?

- A) otasining arvohi B) uch alvasti
C) dahshatli bo'ron D) chaqmoq

Ushbu misolda keyingi test topshirig'iga to'g'ri javob berish uchun oldingi test topshirig'iga to'g'ri javob berish kerak, chunki har ikkala test topshirig'i mazmunan bir-biri bilan bog'liqdir.

Yana bir misol:

1. Abayning inson hayotida muhim o'rin tutadigan aql, ilm-ziyo, insof va adolat, yaxshi xulq, muomala madaniyati, din, e'tiqod, davlatni idora qilish, xalqlar, elatlarning ahil, do'stona munosabatda bo'lishlari haqida mulohaza yuritgan nasriy asari qaysi javobda berilgan?¹

¹ Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to'plami: amaliy qo'llanma. – Toshkent: Davr-press, 2019. – B.117.

- A) «Naqliya soʻzlar»
B) «Boy bolasi bilan kambagʻal bolasi»
C) «Iskandar» D) «Ilm oʻrgan»

2. Abayning «Naqliya soʻzlar» («Nasihatlar») asari kim tomonidan oʻzbekchaga tarjima qilingan?

- A) Asil Rashidov B) Nosir Fozilov
C) Usmon Qoʻchqor D) Nizom Komil

Yuqoridagi misolda test topshiruvchi keyingi test topshirigʻida berilgan savoldan taxmin qilib oldingi test topshirigʻining javobini bilmagan holda ham toʻgʻri topishi mumkin.

4. Test topshirigʻida ishlatilgan **tushuntirishlar va misollar** test topshiruvchi uchun **yangi (notanish)** boʻlishi kerak¹. Aks holda, test topshiruvchi tekshirilayotgan domenni bilgani uchun emas, test topshirigʻida berilgan tushuntirish yoki misolni eslab qolgani uchun toʻgʻri javob berishi mumkin. Misol uchun, matematika darslarida kvadrat tenglamani yechish oʻrgatilgan boʻlsa, testga dars jarayonida ishlatilgan, oʻquvchilarning qoʻlidagi materiallarda berilgan kvadrat tenglama emas, ular uchun notanish boʻlgan tenglama kiritilishi lozim. Shunda oʻquvchi testni yechish uchun dars jarayonida oʻrgangan bilimlarini ishga solishi kerak boʻladi. Boshqa bir misol. Ingliz tili darslarida oʻquvchilarga matn bilan ishlash,

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

matnni tushunish o'rgatilgan bo'lsa, testda o'quvchilarning matnni tushunish ko'nikmalarini baholash uchun yangi (dars jarayonida qo'llanilmagan), lekin shakli, qiyinchilik darajasi va boshqa xususiyatlari bilan darsda ishlatilgan matnlarga o'xshash matnlardan foydalanish darkor.

Misol:

Abayning inson hayotida muhim o'rin tutadigan aql, ilm-ziyo, insof va adolat, yaxshi xulq, muomala madaniyati, din, e'tiqod, davlatni idora qilish, xalqlar, elatlarning ahil, do'stona munosabatda bo'lishi haqida mulohaza yuritgan nasriy asari qaysi javobda berilgan?¹

A) «Naqliya so'zlar»

B) «Boy bolasi bilan kambag'al bolasi»

C) «Iskandar»

D) «Ilm o'rgan»

Darslikda berilgan: «Abay ijodiy merosining salmoqli qismini xalqni ilm-ma'rifatga, madaniyatga da'vat qiluvchi «Naqliya so'zlar» majmuasi tashkil etadi. Hassos adib va shoir ushbu asarida inson hayotida muhim o'rin tutadigan aql, ilm-ziyo, insof va adolat, yaxshi xulq, muomala madaniyati, din, e'tiqod, davlatni idora qilish, xalqlar, elatlarning ahil, do'stona munosabatda bo'lishlari haqida mulohaza yuritadi»².

¹ Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to'plami: amaliy qo'llanma. – Toshkent: Davr-press, 2019. – B.117.

² Ahmedov S. va boshqalar. Adabiyot. Umumiy o'rta ta'lim maktablarining 6-sinfi uchun darslik-majmua. 2-qism. Qayta ishlangan 4-nashri. – Toshkent: Ma'naviyat, 2017. – B.70.

Ko‘rib turganimizdek, test topshirig‘i Abayning «Naqliya so‘zlar» asarining asosiy g‘oyasini tushunishni tekshirishga qaratilgan bo‘lsa-da, mazkur asarni o‘qimagan test topshiruvchi ham darslikda berilgan ma‘lumotni eslab qolgan bo‘lsa, to‘g‘ri javobni topa oladi.

5. Test topshirig‘ida subyektiv munosabat, subyektiv baho ifodalanmasligi kerak¹. Subyektiv bahoni o‘z ichiga olgan test topshirig‘i bahsli bo‘lib qoladi, chunki bunday test topshirig‘ining yagona to‘g‘ri javobi mavjud bo‘lmay qoladi.

Misol:

Qaysi Xiva xoni davrida qozoq cho‘llarining Rossiya imperiyasi va Xiva xonligi o‘rtasida bir-birini ajratib turuvchi hudud bo‘lib qolishi uchun astoydil harakat qilingan?²

- A) Sherg‘ozixon B) Olloqulixon
C) Rahimqulixon D) Muhammad Rahimxon I

Test topshirig‘i o‘zagida ishlatilgan «astoydil harakat qilish» subyektiv baho berishdir va hozirgi zamondan turib o‘sha paytdagi harakatlarning «astoydil» yoki «astoydil emasligi»ga baho berish mumkin emas. Bundan tasqari, test topshirig‘ini tuzgan muallif «astoydil» deb baho bergan ha-

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

² Davlat test markazi: Tarix fanidan namunaviy savollar (2018). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

rakatlarga test topshiruvchi «astoydil» deb baho bermasligi va bu bahslarga sabab bo‘lishi mumkin.

6. Muqobul javobli test topshirig‘ining **murakkab** shakli mavjud bo‘lib, bunday test topshirig‘ining o‘zagida bir nechta variant beriladi va muqobil javoblarda bu variantlar har xil guruhlanadi. Bunday shakldagi test topshiriqlarini ishlatish **tavsiya qilinmaydi**, chunki¹:

– bunday test topshiriqlari muqobil javobli test topshirig‘ining odatiy shakllariga nisbatan qiyinroq hisoblanadi;

– test topshiruvchi berilgan variantlardan birining to‘g‘ri yoki noto‘g‘riligini aniq bilsa, mos kelmaydigan muqobil javoblarni chiqarib tashlash yo‘li bilan to‘g‘ri javobni topishi mumkin, bu esa tekshirilayotgan konstruktga bog‘liq bo‘lmagan faktorning test natijalariga ta‘sirini oshiradi;

– bunday test topshiriqlarining statistik ko‘rsatkichlari, odatda, past bo‘ladi, bu esa, o‘z navbatida, testning ishonchliligiga salbiy ta‘sir qiladi;

– bunday test topshiriqlarini tuzish va tahrir qilish ancha qiyin;

– bunday test topshiriqlariga javob berish ancha vaqtni talab qiladi, bu esa test orqali berilgan vaqt ichida qamrab olinishi mumkin bo‘lgan domen (konstrukt) elementlari sonini kamaytiradi va test natijasida chiqariladigan xulosalar validligiga salbiy ta‘sir ko‘rsatadi.

¹ Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013. – P.72.

Muqobil javobli test topshirig‘ining murakkab shakliga misollar:

XIX asrning 80-yillarida bo‘lib o‘tgan qaysi voqealar Xiva xonligining Chor Rossiyasiga siyosiy qaramligini yanada kuchaytirdi?¹

1. Kaspiyorti temiryo‘lining qurilishi
2. Amudaryo flotiliyasining qurilishi
3. Amudaryo bo‘limining ochilishi
4. Xiva xonini nazorat qilish maqsadida 7 kishidan iborat kengash tuzilishi

A) 1, 2 B) 1, 3 C) 3, 4 D) 2, 4

Quyidagi qaysi asarlar roman janrida yozilmagan?²

1) «Oqqushlar, oppoq qushlar» (O.Yoqubov) 2) «Jar yoqasidagi chaqmoq» (A.Muxtor)

3) «Adolat manzili» (O.Yoqubov) 4) «Sevgim, sevgilim» (O‘.Umarbekov)

5) «Davr mening taqdirimda» (A.Muxtor)

A) 2, 4 B) 1, 2, 3 C) 1, 3, 5 D) 4, 5

Qadimda Buyuk ipak yo‘li orqali Xitoyga So‘g‘diyona (a) va Farg‘ona (b)dan nimalar olib borilgan?³

¹ Davlat test markazi: Tarix fanidan namunaviy savollar (2018). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

² Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.125.

³ Davlat test markazi: Tarix fanidan namunaviy savollar (2018). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

1) jun gazlama; 2) tuyalar; 3) gilam; 4) zotli otlar; 5) zeb-zinyatlar.

A) $a - 1, 2, 4, b - 3, 5$ B) $a - 1, 3, 5, b - 4$

C) $a - 2, 4, 5, b - 3$ D) $a - 1, 3, 4, b - 5$

Muqobil javobli test topshirig‘ining murakkab shakli o‘rni-ga yopiq test topshiriqlarining boshqa shakllarini (masalan, «to‘g‘ri-noto‘g‘ri» shaklidagi test topshiriqlari yoki moslash-tirishni talab qiluvchi test topshiriqlari) ishlatish maqsadga muvofiq. Misol uchun, yuqorida keltirilgan test topshiriqla-rini quyidagi shakllarda berish mumkin:

1 – 4-savollar: XIX asrning 80-yillarida bo‘lib o‘tgan qu-yidagi voqealardan qaysilari Xiva xonligining Chor Rossiya-siga siyosiy qaramligini yanada kuchaytirdi? Agar berilgan voqea siyosiy qaramlikni kuchaytirgan bo‘lsa, A variantni, aks holda, B variantni belgilang.

1. Kaspiyorti temiryo‘lining qurilishi

A) kuchaytirdi B) kuchaytirmadi

2. Amudaryo flotiliyasining qurilishi

A) kuchaytirdi B) kuchaytirmadi

3. Amudaryo bo‘limining ochilishi

A) kuchaytirdi B) kuchaytirmadi

4. Xiva xonini nazorat qilish maqsadida 7 kishidan iborat kengash tuzilishi

A) kuchaytirdi B) kuchaytirmadi

1 – 5-savollar: Qadimda Buyuk ipak yo‘li orqali Xitoyga olib borilgan quyidagi mollarning kelib chiqish hududini toping. Diqqat: A va B variantlardan har biri bir necha martadan ishlatilishi mumkin.

Mollar

1. jun gazlama
2. tuyalar
3. gilam
4. zotli otlar
5. zeb-ziyatlar

Hududlar

- A) So‘g‘diyona
- B) Farg‘ona

7. Test topshirig‘ining o‘zagida savol yoki muammo **aniq** ifodalanishi, test topshiruvchi muqobil javoblarni o‘qimasdan ham nima so‘ralayotganini aniq tasavvur qila olishi lozim¹.

Misol:

Samarqand ...

- A) O‘zbekiston SSRning poytaxti bo‘lgan.
- B) Amir Temur tomonidan qayta qurilgan.
- C) «Islom madaniyati markazi» deb e’tirof etilgan.
- D) Qo‘qon xonligi tarkibiga kirgan.

Ushbu test topshirig‘ining o‘zagi bitta so‘zdan iborat bo‘lib, test topshiruvchi undan nima so‘ralayotganini tasavvur qilishi uchun barcha muqobil javoblarni o‘qib chiqishi, test

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.244.

topshirig‘iga javob berish uchun esa muqobil javoblarni qayta o‘qib chiqishi talab etiladi. Bu esa test topshirig‘ini yechishda ortiqcha vaqt sarflanishiga olib keladi.

8. Test topshirig‘ining o‘zagida **inkor** so‘zlarni ishlatmaslikka harakat qilish darkor¹. Inson miyasi ma’lumotni qabul qilishidagi o‘ziga xoslik (inkor, salbiy ma’lumotlarga e’tibor bermaslikka moyillik) natijasida test topshiruvchilar test topshirig‘ining o‘zagidagi inkor so‘zlarga e’tibor bermasliklari mumkin. Agar test topshirig‘ining o‘zagida inkor so‘zni ishlatish juda ham zarur bo‘lsa, inkor so‘zlar test topshiruvchining ko‘ziga tashlanadigan shaklda **ajratib ko‘rsatilishi** zarur.

Misol:

Qaysi javobdagi baytda tazod she’riy san’ati qo‘llanilmagan?²

A) Nastarin yuvib yuzni, yosuman tuzib o‘zni, Nargis ochibon ko‘zni intizori yoronlar.

B) Vasl uyin obod qildim, buzdi hijron oqibat, Seli g‘amdin bu imorat bo‘ldi vayron oqibat.

C) Doda keldim, ey salotin sarvari, dodim eshit, Sen shah-u, men benavo, lutf ayla, faryodim eshit.

D) Bevafolar zulm-u bedodig‘a ermasmen harif, Ey sinuq ko‘nglum, agar xud shisha, gar po‘lod bo‘l.

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

² Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.121.

Yuqorida aytilganidek, test topshirig'ini bunday shaklda berish tavsiya qilinmaydi, chunki aksar test topshiruvchilar topshiriq o'zagidagi «qo'llanilmagan» so'zini avtomatik tarzda «qo'llanilgan» shaklida o'qib ketishadi.

9. Topshiriq o'zagi imkon qadar qisqa, lo'nda, ortiqcha ma'lumotardan xoli bo'lishi shart. Shu bilan birga, muqobil javoblarda **takrorlanadigan** so'zlar ham topshiriq **o'zagiga** chiqarilishi lozim¹. Topshiriq o'zagi yetti-sakkiz so'zdan iborat bo'lgan bitta gapda ifodalanishi, bu gapda ergash gaplar soni bittadan oshmasligi tavsiya qilinadi².

10. Asosiy, tekshirilayotgan ma'lumot topshiriq o'zagida aks etishi kerak³.

11. Test topshirig'i grammatik, uslubiy, imlo xatolarisiz tuzilgan bo'lishi zarur⁴.

12. Muqobil javoblarni tuzayotganda quyidagilarga e'tibor berish lozim⁵:

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

² Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002. – С.119.

³ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

⁴ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

⁵ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

– tadqiqotlar **uchta** muqobil javob optimal ekanligini ko‘rsatadi;

– muqobil javoblardan faqat **bittasi** to‘g‘ri javob bo‘lishi kerak;

– testda to‘g‘ri javoblar **turli o‘rinlarda** joylashishi va to‘g‘ri javoblarning joylashuvida qandaydir **qolip** bo‘lmasligi (masalan, hamma A muqobil javoblar to‘g‘ri javob bo‘lib qolmasligi yoki to‘g‘ri javoblar A, B, C, D shaklida joylashib qolmasligi) darkor;

– muqobil javoblar mantiqiy **ketma-ketlikda** (masalan, alifbo tartibida, xronologik yoki o‘sib borish tartibida) joylashishi zarur;

– muqobil javoblar ma‘no jihatidan bir-birini **qamrab olmasligi** kerak;

– muqobil javoblar **grammatik bir xil shaklda** bo‘lishi lozim;

– muqobil javoblar ma‘no jihatidan **gomogen** (bir turkumga mansub) bo‘lishi zarur;

– muqobil javoblarning **uzunligi** bir-biriga yaqin bo‘lishi darkor;

– «javoblarning barchasi to‘g‘ri», «javoblarning barchasi noto‘g‘ri», «A va B javoblar to‘g‘ri» singari muqobil javoblardan **foydalanmaslik** lozim;

– to‘g‘ri javobda (kalitda) «doim», «hech qachon», «eng» singari so‘zlarni, shuningdek, topshiriq o‘zagida ishlatilgan so‘zlarni (yoki ma‘no jihatidan ularga yaqin so‘zlarni) **ishlatmaslik** kerak;

– distraktorlar (chalg'ituvchi javoblar) asosli tanlanishi, tekshirilayotgan domenni yaxshi o'zlashtirmagan test topshiruvchiga **to'g'ri javobday** ko'rinishi darkor. Buning uchun distraktorlar shu savolga fanni (mavzuni) yaxshi o'zlashtirmagan o'quvchilar javob berishda yo'l qo'yadigan **tipik (odatiy)** xatolar asosida tuzilishi maqsadga muvofiq.

Misol:

Italiyada va boshqa mamlakatlarda mohir ertakchi adib sifatida tanilgan, «Chippolinoning sarguzashtlari», «Telefonda aytilgan ertaklar», «Jelsomino yolg'onchilar mamlakatida» kabi hikoyalar turkumini yaratgan yozuvchini aniqlang¹.

- A) Janni Rodari B) Jonatan Svift
C) Hans Kristian Andersen D) Antuan de Sent Ekzyuperi

Ushbu test topshirig'i o'zagidagi «Italiyada va boshqa mamlakatlarda mohir ertakchi adib sifatida tanilgan» ma'lumoti ortiqcha, chunki so'ralayotgan adibni bu ma'lumotsiz, berilgan asarlaridan ham topish mumkin. Bundan tashqari, «Italiyada», «Jelsomino», «Chippolino» so'zlaridan italyan adibi haqida gap ketayotganligi ravshan, muqobil javoblarda esa faqat bita italyan adibini (Janni Rodari) ko'rishimiz mumkin. O'zakda berilgan asarlardan birontasini ham o'qimagan, umuman, Janni Rodari ijodi bilan tanish bo'lmagan test topshiruvchi ham qiyalmasdan A javob to'g'ri ekanligini topa oladi.

¹ Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to'plami: amaliy qo'llanma. – Toshkent: Davr-press, 2019. – B.119.

Yana bir misol:

«Stounhenj» obidasi qayerda joylashgan?

- A) Buyuk Britaniya B) Angliya
C) Shimoliy Amerika D) Meksika

Ushbu test topshirig‘ida muqobil javoblar bir-birini qamrab olgan, ya’ni Angliya Buyuk Britaniyaga kiradi, Meksika esa Shimoliy Amerikaga kiradi. Shuningdek, Angliya Buyuk Britaniyaga kirganligi sababli A va B javoblar to‘g‘ri.

Yana bir misol:

«Test» so‘zini kim birinchi marta qo‘llagan?

- A) Testlardan foydalanish haqidagi dastlabki ma’lumotlar ingliz olimi Fisherning 1864-yilda e’lon qilingan «Scale books» kitobida
B) Amerika olimi J.I.Royoshning «Orfografiya – bilimlarni tekshirish» jadvalida
C) nemis olimi Ebbengausning nuqtali bo‘lib qoldirilgan matnlarida
D) Avanesov V.S. tadqiqotlarida

Ushbu test topshirig‘ida A muqobil javobi boshqa muqobil javoblardan uzunligi bilan ajralib turadi va test topshiruvchi aynan shu javob to‘g‘ri ekanligini bilib olishi qiyin emas. Shu bilan birga, A muqobil javobida topshiriq o‘zagida ishlatilgan so‘zlarga ma’no jihatidan yaqin so‘zlarning mavjudligi («birinchi mar-

ta» – «dastlabki ma'lumotlar») ham bu muqobil javobning kalit ekanligini ko'rsatib turadi. Ushbu test topshirig'ining yana bir muammosi – A muqobil javobi grammatik jihatdan noto'g'ri tuzilgan, C muqobil javobi tushunarsiz bo'lib qolgan.

Yana bir misol:

Psixologik testlar deganda nimani tushunasiz?

- A) shaxsning psixologik xususiyatlarini, uning intellektual tomonlarini o'rganadi va aniqlaydi
- B) ta'lim-tarbiya jarayonida bilimlar darajasini aniqlaydi
- C) shaxsning ijtimoiy tomonlarini o'rganadi va aniqlaydi
- D) pedagogikada qo'llanadigan testlar

Ushbu test topshirig'ida ham A muqobil javobi boshqa muqobil javoblardan uzunligi bilan ajralib turadi. Shuningdek, hamma muqobil javoblar grammatik jihatdan bir xil shaklda emas (D javobi boshqalardan grammatik shakli jihatidan ajralib turadi).

Yana bir misol:

Maktabda o'zlashtirishni aniqlashga xizmat qiluvchi testlar dastlab qachon va qayerda paydo bo'lgan?

- A) XIX asrning oxirida Germaniyada
- B) XIX asrning oxirida Angliyada
- C) XIX asrning oxirida Fransiyada
- D) XIX asrning oxirida Turkiyada

Ushbu test topshirig'ida barcha muqobil javoblarda «XIX asrning oxirida» jumlasini qaytarilgan. Bu ma'lumotni testning o'zagida keltirish kerak («Maktabda o'zlashtirishni aniqlashga xizmat qiluvchi testlar dastlab XIX asrning oxirida ... paydo bo'lgan» shaklida).

Yana bir misol:

Iqtisodiyot nazariyasiga «ko'rinmas qo'l» iborasini kim olib kirgan?

- A) Leonardo da Vinchi B) Jan Renuar
C) Adam Smit D) Mark Tven

Ushbu test topshirig'ida distraktorlar (chalg'ituvchi javoblar) asosli tanlanmagan va ma'no jihatidan gomogen (bir turkumga mansub) emas: A, B, D variantlardagi shaxslar iqtisodchi emas.

13. Test topshirig'idan maqsad test topshiruvchini **chalg'itish, yiqitish** bo'lmasligi kerak. Test topshirig'i chalkash bo'lmasligi, test topshiruvchi uchun **tushunarli**, sodda va ravon tilda yozilishi lozim¹. Chalkash, tushunarsiz yozilgan test topshiriqlari o'z vazifasini bajarmaydi, ya'ni bilimlini bilimsizdan ajratib bermaydi, chunki tekshirilayotgan domenni yaxshi o'zlashtirgan test topshiruvchi ham bunday test topshiriqlariga javob bera olmaydi.

¹ Downing S. Selected-Response Item Formats in Test Development // Downing S. and Haladyna, T. (Eds) Handbook of Test Development. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006. – P.294.

4.2.2. «To‘g‘ri-noto‘g‘ri» shaklidagi test topshiriqlari

«To‘g‘ri-noto‘g‘ri» shaklidagi test topshiriqlarida bir yoki bir nechta **jumla (fikir, hukm)** beriladi va test topshiruvchi «to‘g‘ri» / «noto‘g‘ri», «ha» / «yo‘q», «qo‘shilaman» / «qo‘shilmayman» shaklidagi ikkita bir-biriga teskari javob variantidan birini tanlashi talab qilinadi. Bunday test topshiriqlari **dixotom** (qarama-qarshi xarakterga ega) tushunchalar haqidagi bilimlarni tekshirishda qo‘l keladi.

Quyida berilgan gaplarning to‘g‘ri yoki noto‘g‘riligini toping. Agar gap to‘g‘ri bo‘lsa, javoblar varaqasida A javob variantini, noto‘g‘ri bo‘lsa, B javob variantini belgilang.

1. O‘zbekiston Respublikasi Oliy Majlisi ikki palatadan iborat.

A) To‘g‘ri

B) Noto‘g‘ri

javob variantlari

topshiriq sharti

jumla (hukm)
shaklida ifodalangan test topshirig‘i

2. O‘zbekiston Respublikasi Oliy Majlisining barcha a‘zolari umumxalq saylovlarida saylanadi.

A) To‘g‘ri

B) Noto‘g‘ri

3. O‘zbekiston Respublikasi Oliy Majlisi O‘zbekiston Prezidenti va hukumati qarorlarining Konstitutsiyaga mosligini tekshiradi.

A) To‘g‘ri

B) Noto‘g‘ri

«To‘g‘ri-noto‘g‘ri» shaklidagi test topshirig‘i namunasi

«To‘g‘ri-noto‘g‘ri» shaklidagi test topshiriqlarini tuzishda quyidagi **qoidalar va tavsiyalarga** rioya qilmoq lozim:

1. Test topshirig‘i to‘g‘ri tuzilganligini tekshirib ko‘rish uchun jumlaning (hukmning) **teskarisini** tuzib ko‘rish kerak – jumlaning (hukmning) teskarisi mavjud bo‘lsagina topshiriq to‘g‘ri tuzilgan bo‘ladi¹.

2. Jumla (hukm) **yuzaki** tahlil etishga moyil yoki mavzuni bilmaydigan test topshiruvchini chalg‘ita olishi kerak².

3. Test topshirig‘idagi jumalarda **inkor** gaplarni ishlatmaslik maqsadga muvofiq³.

Misol:

Tavsiya qilinmaydi

Quyida berilgan hukmlarning to‘g‘ri yoki noto‘g‘riligini belgilang.

Abdulla Qahhor «O‘tkan kunlar»ni yozmagan.

A) To‘g‘ri B) Noto‘g‘ri

Tavsiya qilinadi

Quyida berilgan hukmlarning to‘g‘ri yoki noto‘g‘riligini belgilang.

Abdulla Qahhor «O‘tkan kunlar»ni yozgan.

A) To‘g‘ri B) Noto‘g‘ri

4. Bitta jumlada (hukmda) **bittadan** ortiq ma‘lumotni (tushunchani) singdirmaslik lozim⁴. Agar bitta jumlada

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.235.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.235.

³ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.236.

⁴ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.236.

(hukmda) bittadan ortiq ma'lumot (tushuncha) singdirilsa, test topshiruvchi ushbu ma'lumotlardan aynan qaysisining to'g'ri yoki noto'g'ri ekanligini topishni bilmay qoladi.

Misol:

Tavsiya qilinmaydi

Quyida berilgan hukmlarning to'g'ri yoki noto'g'riligini belgilang.

O'zbekiston Respublikasida davlat tili o'zbek tilidir, shuning uchun davlat idoralariga ishga kirishda davlat tilidan imtihon topshirilishi majburiy.

A) To'g'ri B) Noto'g'ri

Tavsiya qilinadi

Quyida berilgan hukmlarning to'g'ri yoki noto'g'riligini belgilang.

1. O'zbekiston Respublikasida davlat tili o'zbek tilidir.

A) To'g'ri B) Noto'g'ri

2. O'zbekiston Respublikasida davlat idoralariga ishga kirishda davlat tilidan imtihon topshirilishi majburiy.

A) To'g'ri B) Noto'g'ri

5. «To'g'ri» va «noto'g'ri» javob variantli topshiriqlar soni **nisbatan teng** (lekin aynan teng emas) bo'lishi maqsadga muvofiq¹.

6. To'g'ri va noto'g'ri jumlar **uzunligi** nisbatan bir xil bo'lishi lozim². Test tuzuvchilar ko'pincha to'g'ri jummalarni

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.237.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.238.

uzunroq tuzishadi va bu test topshiruvchiga to‘g‘ri javobni bilmasdan ham topish imkonini yaratadi.

Ba‘zan «to‘g‘ri-noto‘g‘ri» shaklidagi test topshiriqlari **bir nechta to‘g‘ri javobga ega** bo‘lgan muqobil javobli test topshirig‘i shaklida berilishi mumkin. Misol uchun:

«To‘g‘ri-noto‘g‘ri» shaklidagi test topshirig‘i

Quyidagi xususiyatlar tuyuq janriga xosmi? Agar tuyuq janriga xos bo‘lsa, A variantni, aks holda, B variantni belgilang.

1. Tuyuq aruz va barmoq vaznlarida yozilishi mumkin.

A) Ha B) Yo‘q

2. Tuyuq to‘rt misradan iborat bo‘ladi.

A) Ha B) Yo‘q

3. Tuyuqda shakldosh so‘zlar qofiya bo‘lib keladi.

A) Ha B) Yo‘q

4. Tuyuq aabb shaklida qofiyalanadi.

A) Ha B) Yo‘q

Bir nechta to‘g‘ri javobga ega bo‘lgan muqobil javobli test topshirig‘i

1. Quyidagi xususiyatlardan qaysilari tuyuq janriga xos? A – D javob variantlaridan ikkitasini tanlang.

A) Tuyuq aruz va barmoq vaznlarida yozilishi mumkin.

B) Tuyuq to‘rt misradan iborat bo‘ladi.

C) Tuyuqda shakldosh so‘zlar qofiya bo‘lib keladi.

D) Tuyuq aabb shaklida qofiyalanadi.

Bir nechta to‘g‘ri javobga ega bo‘lgan muqobil javobli test topshirig‘iga javob berish uchun ham test topshiruvchi har bir muqobil javobga xayolan «ha/yo‘q» yoki «to‘g‘ri/ noto‘g‘ri» shaklida javob berib, keyin javobi «ha» («to‘g‘ri») bo‘lgan muqobil javoblarni tanlashi kerak bo‘ladi. Bunday shakldagi test topshiriqlarini tuzishdan oldin ularni baholash mezonlarini aniqlashtirib olish talab etiladi: test topshiruvchi barcha «ha» («to‘g‘ri») javoblarni to‘g‘ri topsa, uning natijasi qanday hisoblanadi va qisman to‘g‘ri topsa, uning natijasi qanday hisoblanadi? Masalan, ingliz tilini bilish darajasini baholash uchun qo‘llaniladigan xalqaro IELTS (International English Language Testing System) testida bunday shakldagi test topshiriqlari qo‘llanilgan holatlarda har bitta «ha» («to‘g‘ri») javobi uchun alohida ball beriladi:

Questions 21 – 23: Choose three letters, A – F.

Which THREE things are the students required to submit to their professor?


- | | |
|-----------------------------|----------------------|
| A) a written summary | B) maps |
| C) a case study | D) charts and graphs |
| E) a list of resources used | F) a video |

IELTS testi topshirig‘i namunasi¹

¹Jakeman V. and McDowell C. Cambridge IELTS 10. Authentic Examination Papers from Cambridge ESOL. – Cambridge: Cambridge University Press, 2015. – P.15.

4.2.3. Moslashtirishni talab qiluvchi test topshiriqlari

Moslashtirishni talab qiluvchi test topshiriqlarida ikki **ro'yxat** beriladi va ular o'rtasidagi qandaydir **bog'liqlikka** asoslanib (bu bog'liqlik topshiriq shartida ko'rsatilishi lozim) birinchi ro'yxatdagi savollarga (tavsiflar, ta'riflar, nomlar va hokazo) ikkinchi ro'yxatdan mos javoblarni topish talab etiladi. Bunday test topshiriqlari **bog'liqliklar (assotsiatsiyalar)** haqidagi bilimlarni tekshirishda qo'l keladi. Moslashtirishni talab etuvchi test topshiriqlari yakka holda kela olmaydi: birinchi ro'yxatda nechta savol bo'lsa, shuncha test topshiriqlaridan iborat bo'ladi.

Quyida berilgan asarlarning (1 – 4) mualliflarini (A – E) toping. Har bir javob faqat bir marta ishlatilishi mumkin. Diqqat: A – E javoblardan biri ortiqcha! 

Asarlar:

1. «Abulfayzxon»
2. «Kecha va kunduz»
3. «Oygul bilan Baxtiyor»
4. «O'tkan kunlar»

} birinchi ro'yxat savollar

Mualliflar:

- A) Abdulla Qodiriy
- B) Botu
- C) Cho'lpon
- D) Fitrat
- E) Hamid Olimjon

} ikkinchi ro'yxat javoblar

Quyida berilgan ta'riflarga mos davlat boshqaruvi shaklini toping. Diqqat: A – C javob variantlaridan har biri bir marta ortiq ishlatilishi mumkin.

Ta'riflar:

1. Bunday davlat boshqaruvida davlat boshlig'i umumxalq saylovida saylanadi.
2. Bunday davlat boshqaruvida davlat boshlig'i lavozimi meros bo'lib o'tadi.
3. Bunday davlat boshqaruvida davlat boshlig'i qonun chiqaruvchi organ tomonidan saylanadi.
4. Bunday davlat boshqaruvida davlat boshlig'i, odatda, qonun chiqaruvchi organni tarqata oladi.

Davlat boshqaruvi shakllari:

- A) monarxiya B) parlament respublikasi
C) prezidentlik respublikasi

Moslashtirishni talab qiluvchi test topshiriqlari namunalari

Moslashtirishni talab etuvchi test topshiriqlarini tuzishda quyidagi **qoidalar va tavsiyalarga** rioya qilmoq lozim:

1. Ro'yxatlar nisbatan **qisqa** bo'lishi lozim – haddan tashqari uzun ro'yxatlar test topshiruvchini toliqtiradi¹.
2. Qisqaroq nomlardan iborat ro'yxatni **ikkinchi** ro'yxat (**javob variantlari**) sifatida berish lozim. Masalan, agar test topshiruvchilardan berilgan ta'riflarga mos terminlarni topish talab qilinsa, ta'riflarni birinchi ro'yxatga (savollarga), terminlarni ikkinchi ro'yxatga (javoblarga) kiritish maqsadga

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.238.

muvofiq. Bu test topshiruvchining topshiriqlar bilan ishlashini osonlashtiradi¹.

3. Ikkala ro‘yxatda keltirilgan nomlar yoki tavsiflar **gomon** (bir sohaga mansub) bo‘lishi lozim, agar birinchi ro‘yxatda jumlar yoki gaplar ishlatilgan bo‘lsa, ular **grammatik** jihatdan bir-biriga o‘xshash bo‘lishi kerak².

4. Ro‘yxatlar **nomlanishi** zarur³. Har bir ro‘yxatning boshida bu nimaning ro‘yxati ekanligini aniq ko‘rsatuvchi sarlavha beriladi.

5. Ro‘yxatlardagi savollar va javoblar soni **teng bo‘lmasligi** talab qilinadi⁴. Aks holda, test topshiruvchi ro‘yxatdagi ayrim savollarga bilmasdan ham to‘g‘ri javob topishi mumkin. Agar ikkinchi ro‘yxatdagi (javoblardagi) nomlar faqat bir martadan ishlatiladigan bo‘lsa, unda bu ro‘yxat birinchi ro‘yxatdan taxminan ikki baravar uzunroq bo‘lishi tavsiya etiladi⁵.

6. Ikkinchi ro‘yxatni (javob variantlarini) biron-bir **mantiqiy ketma-ketlikda** (masalan, alifbo tartibida yoki xronologik tartibda) joylashtirish darkor⁶.

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.238.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.238.

³ Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002. – С.195.

⁴ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.239.

⁵ Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002. – С.195.

⁶ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.239.

7. Test topshiriqlari shartida ikki ro‘yxat o‘rtasidagi **bog‘-liqlik** aniq ko‘rsatilishi, ikkinchi ro‘yxatdagi javoblar **qaytarilishi** mumkinligi yoki mumkin emasligi ham aniq ko‘rsatilishi lozim¹.

8. Ro‘yxatlar **haddan tashqari uzun** bo‘lmasligi zarur – ro‘yxatlardagi elementlar sonini 10 tadan oshirmaslik tavsiya qilinadi².

4.2.4. Ketma-ketlikni topishni talab qiluvchi test topshiriqlari

Ketma-ketlikni topishni talab qiluvchi test topshiriqlarida bir qancha elementlar (voqea-hodisalar, xatti-harakatlar, jarayonlar, tushunchalar, nomlar va hokazo) beriladi va test topshiruvchidan ushbu elementlarning **to‘g‘ri ketma-ketligini** topish talab etiladi.

Birinchi jahon urushining boshlanishiga olib kelgan voqealarning to‘g‘ri ketma-ketligini toping. Har bir voqeaning oldiga 1 dan 7 gacha tegishli raqam qo‘yib chiqing.

- Germaniya Fransiyaga urush e‘lon qildi
- Rossiya umumiy harbiy safarbarlikni boshladi
- Sarayevoda Frans Ferdinand o‘ldirildi
- Angliya Germaniyaga urush e‘lon qildi

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.239.

² Мамай С.П. Методика составления тестовых заданий. – Екатеринбург: Издательство УГППУ, 2001. – С.41.

- Germaniya Belgiyaga bostirib kirdi
- Avstriya-Vengriya Serbiyaga ultimatum e’lon qildi
- Germaniya Rossiyaga urush e’lon qildi

Ketma-ketlikni topishni talab qiluvchi test topshirig’i namunasi

Bunday shakldagi topshiriqlar **xronologik, texnologik yoki algoritmik** ketma-ketlik haqidagi bilimlarni tekshirishda qo‘l keladi. Masalan, tarix fanidan tarixiy voqealarning ketma-ketligini topish, adabiyotdan asar syujeti voqealarini topish, aniq fanlardan masalani yechish ketma-ketligini topish, biologiyadan organizmlar evolyutsiyasidagi jarayonlar ketma-ketligini topish va shu kabi bilimlarni tekshirishda ishlatilishi mumkin.

4.2.5. Yopiq test topshiriqlari yordamida yuqori darajadagi kognitiv ko‘nikmalarni tekshirish

Aksar hollarda yopiq test topshiriqlari quyi darajadagi kognitiv ko‘nikmalarni (asosan bilish, ya’ni ma’lumotni eslab qolish, ba’zan tushunish va qo‘llash) tekshirish uchun xizmat qiladi, deya tanqid ostiga olinadi. Shu bilan birga, agar quyidagi tamoyillarga rioya qilinsa, yopiq test topshiriqlari yordamida ham yuqori darajadagi kognitiv ko‘nikmalarni baholash mumkin¹:

¹ Brookhart S.M. How To Assess Higher-Order Thinking Skills in Your Classroom. – Alexandria, VA: ASCD, 2010. – P.17.

1. Test topshiruvchini o‘ylashga majbur qiladigan **kirish material** (matn, xarita, grafik, jadval, rasm va hokazo) berilishi.

Bunday material test topshiruvchini tekshiriladigan yuqori darajadagi kognitiv ko‘nikmalarni ishga solishga majbur qilishi lozim. Bitta kirish material bilan bog‘liq bir nechta test topshirig‘i berilishi mumkin (bitta kirish material asosida berilgan bir qancha test topshiriqlari ketma-ketligi ingliz tilidagi ilmiy adabiyotlarda «testlet» deb ataladi).

Malika tovuqxonada eshigidagi qulfnı burab eshikni ochdi, hovliga rang-barang bulutlar singari otilib chiqayotgan tovuqlarga qarab ohista jilmaydi. Ular qanotlarini qoqqan Malikaning uydan olib chiqqan ovqat qoldiqlarini cho‘qiy boshlashdi. Odatdagidek, malla tovuq eng yaxshi bo‘laklarni olar va o‘zini xo‘jayinlarcha tutib, shovqin qilgan qanotlarini qoqqan holda o‘ziga xalaqit berishga uringan har qanday tovuqni cho‘qishga tushardi.

1. Hikoyaning boshida Malika nima qilayotgan edi?

- A) tovuqni quvlayotgan edi.
- B) tovuqlarga don berayotgan edi.
- C) tuxum qidirayotgan edi.
- D) patlar to‘playotgan edi.

2. Muallif qanday qilib senga malla tovuqning fe‘l-atvorini ko‘rsatib bermoqda?

- A) Malla tovuq qanday ko‘rinishda ekanligini tavsiflayapti.
- B) Malla tovuqning yoqtiradigan ovqati haqida hikoya qilayapti.

- C) Malla tovuq qayerda yashashi haqida hikoya qilyapti.
D) Malla tovuq o'zini qanday tutayotgani haqida hikoya qilyapti.

Testlet namunasi¹

Test topshiriqlari bilan ishlatiladigan kirish materiali tekshirilayotgan konstruktga bog'liq bo'lishi kerak². Aks holda, bunday material testdan chiqariladigan xulosalarning validligiga salbiy ta'sir ko'rsatadi.

Misol:

Rasmda ko'rsatilgan qadimgi yunon faylasufiga tegishli fikrni toping³.

- A) «Insonda hech qanday ehtiyojlar bo'lmasligi shart»
B) «Hamma narsa oqadi, hamma narsa o'zgaruvchandir»



- C) «Menga tayanch nuqtasini topib bersangiz, Yerni ag'darib yuboraman!»

¹ O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Ta'lim sifatini nazorat qilish davlat inspeksiyasi Ta'lim sifatini baholash bo'yicha xalqaro tadqiqotlarni amalga oshirish milliy markazi: O'quvchilarning matematik, tabiiy-ilmiy fanlar hamda o'qish savodxonligini baholashga mo'ljallangan topshiriqlar to'plami. – B.46 // <http://markaz.tdi.uz>

² Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013. – P.80.

³ Davlat test markazi: Tarix fanidan namunaviy savollar (2019). Davlat test markazining rasmiy saytidan (www.dtm.uz) yuklab olingan.

D) «Hech narsani bilmasligimni bilganim uchun ham meni donishmand deb hisoblashsa kerak»

Yuqoridagi test topshirig‘ida test topshiruvchi oldin rasmda qadimgi yunon faylasufi Diogen tasvirlanganini topishi, so‘ngra Diogenga tegishli fikrni topishi kerak. Diogenning aynan shu (fransuz rassomi Jan Leon Jerom qalamiga mansub) rasmini ko‘rmagan yoki ko‘rib eslab qolmagan test topshiruvchi berilgan fikrlardan aynan qaysisi Diogenga tegishli ekanligini bilsa ham, test topshirig‘ini bajara olmaydi. Demak, testda konstruktga bog‘liq bo‘lmagan o‘zgaruvchi mavjud.

Kirish materiali to‘g‘ri va o‘z o‘rnida ishlatilsa, test topshiriqlarining faqat xotirani tekshirishga qaratilishi yoki tekshirilayotgan konstruktga xotira omili ta‘sir qilishining oldi olinadi.

Misol:

Qaysi javobda Navoiyning «Deyin» radifli g‘azali haqida ma’lumot berilgan?¹

A) G‘azalda «soch», «chin», «shikan» timsollari yordamida tasavvufiy ma’no ifoda etiladi.

B) Radd ul-matla’ san’ati asosida birinchi misra g‘azalning so‘nggi satri sifatida yana bir karra kuchaytiriladi.

C) G‘azalda shoir «ko‘hkan», ya’ni tosh yo‘nuvchi timsoli orqali Farhodga ishora qiladi.

D) Matlada tardi aks san’atining ajoyib namunasi yaratilgan.

¹ Davlat test markazi. Ona tili va adabiyot. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.92.

Test topshiruvchida gʻazalni tahlil qila olish koʻnikmasi shakllangan boʻlsa ham, agar u soʻralgan gʻazalni yod olmagan boʻlsa, ushbu test topshirigʻiga javob bera olmaydi, yaʼni tekshirilayotgan konstruktga (gʻazalni tahlil qila olish koʻnikmasiga) bogʻliq boʻlmagan oʻzgaruvchi (xotira) testning natijasiga taʼsir qiladi. Bunday muammoning oldini olish uchun gʻazal matnini keltirib, soʻng shu gʻazalni tahlil etishga qaratilgan test topshiriqlarini berish mumkin.

2. Test topshiruvchi uchun **yangi, notanish** boʻlgan materialning ishlatilishi.

Test topshirigʻi test topshiruvchi uchun tanish (dars jarayonida ishlatilgan, oʻquvchining qoʻlidagi darslikda berilgan va hokazo) materialga asoslansa, test topshiruvchi baholanayotgan koʻnikma shakllanganligi sababli emas, materialni eslab qolganligi sababli topshiriqqa toʻgʻri javob berishi mumkin. Buning oldini olish uchun taʼlim jarayonida shakllantirilgan koʻnikmalar yangi material asosida tekshirilishi lozim. Masalan, taʼlim jarayonida Navoiy gʻazaliyotini tushunish koʻnikmasi «Qaro koʻzum...» gʻazali misolida shakllantirilgan boʻlsa, testda ayni koʻnikmani baholash boshqa gʻazal vositasida amalga oshirilishi maqsadga muvofiq.

3. Test topshirigʻining qiyinchilik darajasi va test topshirigʻida talab qilinadigan kognitiv koʻnikma darajasini **farqlash**.

Kognitiv ko'nikmalarning quyi darajasini tekshirishga qaratilgan topshiriqlar ham, yuqori darajasini tekshirishga qaratilgan topshiriqlar ham oson yoki qiyin bo'lishi mumkin. Test topshiriqlarini tuzayotganda kognitiv ko'nikmalarning quyi darajasini tekshirishga qaratilgan topshiriqlarni avtomatik tarzda «oson», yuqori darajasini tekshirishga qaratilgan topshiriqlarni esa «qiyin» deb ajratish testning ishonchliligiga va testdan chiqariladigan xulosalarning validligiga salbiy ta'sir ko'rsatadi. Test topshirig'ining qiyinchilik darajasi statistik tahlil yordamida aniqlanadi.

Asosiy xulosa:

Yopiq test topshiriqlarining bir qancha turlari mavjud. Testdan chiqariladigan xulosalar valid bo'lishi uchun test topshiriqlarini tuzish qoidalariga amal qilish va har bir topshiriq turini o'z o'rnida qo'llash lozim.

Har qanday test topshirig'ida quyidagilarga e'tibor berish talab etiladi:

- test topshiruvchi uchun yo'riqnomalar (ko'rsatmalar) aniq va lo'nda bo'lishi kerak;
- test topshirig'ining mazmuni tushunarli bo'lishi, imkon qadar qisqa, lo'nda, ortiqcha ma'lumotlarsiz ifodalinishi zarur;
- tekshirilayotgan domenni bilmaydigan test topshiruvchiga ham to'g'ri javobni topish imkoni yaratilmasligi darakor.

4.3. Ochiq test topshiriqlarini yaratish va baholash

Yopiq test topshiriqlarida test topshiruvchi berilgan muqobil javoblardan (javob variantlaridan) birini tanlashi talab qilinsa, **ochiq** test topshiriqlarida javobni o‘zi **yaratishi** talab qilinadi. Qanday javob yaratilishiga qarab ochiq test topshiriqlari quyidagi turlarga bo‘linadi:

- qisqa javobni talab qiladigan test topshirig‘i;
- kengaytirilgan javobni talab qiladigan test topshirig‘i;
- amaliy test topshirig‘i.

Ochiq test topshiriqlarining turlari

Qisqa javobni talab qiladigan test topshirig‘i

Kengaytirilgan javobni talab qiladigan test topshirig‘i

Amaliy test topshirig‘i (biron-bir mahsulot yaratilishi yoki biron-bir amal bajarib ko‘rsatilishi talab qilinadigan test topshirig‘i)

Misollar

Berilgan hukmni to‘ldiring: Mumtoz she‘riyatda to‘rt misradan iborat, tajnis san‘atiga asoslangan janr ... deyiladi.

XX asrning oxirgi 10 yilligida jahon siyosiy xaritasida ro‘y bergan yirik o‘zgarishlar va ularning sabablarini izohlang.

Microsoft Excel dasturida oxirgi 2 oyda so‘mning dollarga nisbatan rasmiy kursi o‘zgarishlari dinamikasini ko‘rsatadigan grafik tuzing.

Ochiq test topshiriqlari turlari

Qisqa javobni talab qiladigan test topshirig‘i so‘roq gap shaklida yoki to‘liqsiz gap shaklida bo‘lishi mumkin. Bunday test topshiriqlari quyi darajadagi kognitiv ko‘nikmalarni (bilish, tushunish, qo‘llash) tekshirish imkonini beradi¹.

Qisqa javobni talab etadigan test topshiriqlarini tuzishda quyidagi **qoidalar va tavsiyalarga** rioya qilmoq lozim:

1. Bunday test topshiriqlarini to‘liqsiz gap shaklidan ko‘ra **so‘roq gap** shaklida ifodalash samaraliroq² sanaladi. Misol uchun, quyidagi ikki test topshirig‘idan o‘ng tarafdagisini qo‘llash tavsiya qilinadi:

<i>To‘liqsiz gap shaklidagi test topshirig‘i</i>	<i>So‘roq gap shaklidagi test topshirig‘i</i>
Amir Temur davlatining poytaxti ... shahri bo‘lgan.	Amir Temur davlatining poytaxti qaysi shahar bo‘lgan?

2. Test topshirig‘i shunday tuzilishi kerakki, **qisqa, lo‘nda** javob (bir yoki bir nechtagacha so‘z, so‘z birikmasi, raqam, simvol, formula) talab qilsin³. Test topshiruvchiga yo‘riqnomada (ko‘rsatmada) javobning hajmini ko‘rsatish ham test topshiruvchini yo‘naltirib, topshiriqning samaradorligini oshiradi.

¹ Brookhart S.M. How To Design Questions and Tasks To Assess Student Thinking. – Alexandria, VA: ASCD, 2014. – P.105.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.264.

³ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.265.

Questions 26 – 30: Complete the notes below. Write NO MORE THAN TWO WORDS AND/OR A NUMBER for each answer.

City Art Museum

The main part of museum was built in **26** _____

The **27** _____ was built sixty years later.

Collections: modern art, works by **28** _____, sculpture, European art.

Classes: **29** _____ classes for adults

Arts and crafts workshops for children

Weekly **30** _____ in the fall and winter

IELTS test topshiriqlari namunasi¹

Yuqoridagi IELTS xalqaro imtihonidan olingan namunada test topshiruvchi uchun yo‘riqnomada javoblar maksimum ikkita so‘z va/yoki raqam shaklida bo‘lishi kerakligi ko‘rsatilgan.

3. Agar test topshirig‘i to‘liqsiz gap shaklida bo‘lsa, gapning to‘ldiriladigan qismi gapning **boshida** kelmasligi kerak².

¹ Jakeman V. and McDowell C. Cambridge IELTS 10. Authentic Examination Papers from Cambridge ESOL. – Cambridge: Cambridge University Press, 2015. – P.32.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.265.

<i>Tavsiya qilinmaydi</i>	<i>Tavsiya qilinadi</i>
... – bu mumtoz adabiyotda har bir bandi besh misradan iborat bo‘lgan lirik janr.	Mumtoz adabiyotda har bir bandi besh misradan iborat bo‘lgan lirik janr ... deyiladi.

4. Har bir test topshirig‘i faqat **bitta javobni** talab qilishi kerak¹.

<i>Tavsiya qilinmaydi</i>	<i>Tavsiya qilinadi</i>
Samarqandning Registon maydonidagi Sherdor madrasasi ... yilda ... tomonidan qurdirilgan.	Samarqandning Registon maydonidagi Sherdor madrasasi 1. Nechanchi yilda qurilgan? 2. Kim tomonidan qurdirilgan?

Qisqa javobni talab qiladigan test topshiriqlarining javoblari maxsus blankalarda (javoblar varaqasida) yozilsa, bunday blankalar optik qurilmalar tomonidan qayta ishlanib, so‘ng kompyuterda tekshirilishi mumkin. Bu esa test natijalarini hisoblashni osonlashtiradi. Shu bilan birga, bunday topshiriqlarda test tuzuvchi kutgandan boshqa javoblar (masalan, to‘g‘ri javobning sinonimlari) ham to‘g‘ri bo‘lib qolishi mumkinligi ehtimolini hisobga olib qo‘yish lozim.

Kengaytirilgan javobni talab qiladigan test topshirig‘i yuqori darajadagi kognitiv ko‘nikmalar, ko‘nikmalar inte-

¹ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.136.

gratsiyasi, mantiqiy fikrlash va fikrni asoslash, muammolarga yechim taklif qilish qobiliyati, shuningdek, kognitiv bo‘lmagan ko‘nikmalar (masalan, yozma yoki og‘zaki nutq) haqida asosli xulosalar chiqarish imkonini beradi. Yopiq test topshiriqlari yoki qisqa javobni talab qiladigan ochiq test topshiriqlari yordamida bunday ko‘nikmalarni baholash aksar hollarda imkonsizdir. Masalan, test topshiruvchining o‘z ona tilida yoki chet tilida rasmiy xat yozish ko‘nikmasini baholash uchun unga xat yozdirib ko‘rish darkor. Shu bilan birga, kengaytirilgan javobni talab qiladigan test topshiriqlarini baholash uchun vaqt kerak va bahoning subyektiv bo‘lish ehtimoli yuqori ekanini ham qayd etish joiz.

Kengaytirilgan javobni talab qiladigan test topshiriqlarini tuzishda quyidagi **qoidalar va tavsiyalarga** rioya etish lozim:

1. Test topshiruvchidan nima talab qilinayotganligi **aniq** ko‘rsatilishi zarur¹. Test topshirig‘i shunday tuzilishi kerakki, u test topshiruvchini ma’lum bir qolipga solsin, yo‘nalish ko‘rsatsin. Misol uchun, quyidagi jadvalda chap ustunda keltirilgan test topshiriqlari shu qadar mavhum ifodalanganki, tekshirilayotgan domenni yaxshi o‘zlashtirmagan test topshiruvchi ham istalgancha «olib qochishi» mumkin. Shu bilan birga, chap ustundagi test topshiriqlari mavhum bo‘lganligi sababli baholashda muammolar keltirib chiqaradi:

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.268.

<i>Tavsiya qilinmaydi</i>	<i>Tavsiya qilinadi</i>
XX asr – texnika mo‘jizalari asri.	XX asrning, sizningcha, eng muhim bo‘lgan uch texnik kashfiyotini sanang va bu kashfiyotlar insonlar hayotini qanday o‘zgartirganini tasvirlang.
O‘zbekiston mustaqillik ostonasida.	1980-yillarning ikkinchi yarmida O‘zbekiston SSRda siyosiy mustaqillikka erishish yo‘lida qo‘yilgan qadamlarni va bu qadamlarning natijalarini ko‘rsating.
«O‘tkan kunlar»da Zaynab obrazi.	Abdulla Qodiriyning «O‘tkan kunlar» romanida Zaynab obrazidagi qarama-qarshiliklar nimalarda namoyon bo‘ladi? Fikringizni asardan misollar bilan asoslang.
G‘azal mulkinging sultoni.	Alisher Navoiyning o‘zbek mumtoz adabiyotida g‘azal janrining rivojiga qo‘shgan hissasi nimalarda namoyon bo‘ladi?
Spartak qo‘zg‘oloni.	Spartak qo‘zg‘oloniga zamin yaratgan eng muhim ijtimoiy omillarni ko‘rsating.

2. Har bir topshiriqqa beriladigan **maksimal ball** va topshiriqni bajarish uchun **ajratilgan vaqtni** ko‘rsating¹. Bu test topshiruvchiga yo‘l-yo‘riq vazifasini bajaradi hamda testda vaqtini va kuchini to‘g‘ri taqsimlashga yordam beradi.

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.269.

3. Bitta uzun (bir necha betdan iborat) javob talab qiladigan test topshirig‘idan ko‘ra **bir nechta** nisbatan qisqaroq javob talab qiladigan test topshirig‘i samaraliroq hisoblanadi¹. Bu tekshirilayotgan domenni to‘liqroq qamrab olish imkonini beradi va test natijalari asosida chiqariladigan xulosalarning validligini oshiradi.

4. Test topshiruvchilarga bir nechta test topshirig‘idan bit-tasini **tanlash** imkonini berish tavsiya etilmaydi². Test topshiruvchilarga tanlash imkonining berilishi ularning natijalarini solishtirish asosida chiqariladigan xulosalar validligiga salbiy ta’sir qiladi.

5. Test topshirig‘i shartini «**kim, nima, qachon, qayer-da?**» so‘roqlari bilan boshlash tavsiya etilmaydi³. Bunday so‘roqlar test topshiruvchidan yuqori darajadagi kognitiv ko‘nikmalarni ishga solishni talab qilmaydi va kengaytirilgan javobni talab qiladigan test topshiriqlarini qo‘llashning asosiy maqsadini yo‘qqa chiqaradi.

6. Topshiriq shartida javob **hajmiga** chegaralar qo‘yish tavsiya etiladi⁴.

Kengaytirilgan javobni talab qiladigan test topshiriqlari bilan birga ularni **baholash mezonlari** va **baholash sxemasini**

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.269.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.269.

³ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.139.

⁴ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.139.

ham oldindan tayyorlab qo'yish tavsiya etiladi. Bunda ikki xil baholash sxemasi qo'llanilishi mumkin: **murakkab baholash sxemasi** va **yig'ma baholash sxemasi**.

Murakkab baholash sxemasi – test topshiruvchining ochiq test topshirig'iga javobini uning umumiy sifatidan kelib chiqqan holda baholash uchun qo'llaniladigan sxema.

Murakkab baholash sxemasida baholar (masalan, 1 dan 4 gacha) beriladi va har bir baho uchun mezonlar tavsifi (deskriptorlar) keltiriladi. Baholovchi test topshiruvchi bergan javobning umumiy sifatidan kelib chiqqan holda uning qaysi bahoga mosligini aniqlaydi va tegishli bahoni qo'yadi.

4	3	2	1
Asarning asosiy va ikkilamchi g'oyalari ajratilgan, fikrini asardan missollar bilan asoslagan, asarda ko'tarilgan muammolarga munosabat bildirgan va munosabatini asoslagan	Asarning asosiy va ikkilamchi g'oyalarini ajratgan, asarda ko'tarilgan muammolarga munosabat bildirgan	Asarning asosiy g'oyasini ajratgan, asarda ko'tarilgan muammoga munosabat bildirgan	Asarning asosiy g'oyasini ajratishga harakat qilgan

Badiiy asar tahlilini baholash uchun murakkab baholash sxemasi namunasi

Yig‘ma baholash sxemasi – test topshiruvchining ochiq test topshirig‘iga javobini har bir mezon bo‘yicha alohida baholash uchun qo‘llaniladigan sxema.

Yig‘ma baholash sxemasida har bir baholash mezoni alohida beriladi va bu mezon bo‘yicha baholarning tavsifi keltiriladi. Bunda turli mezonlarning yakuniy bahodagi ulushi turlicha bo‘lishi ham mumkin. Baholovchi test topshiruvchi bergan javobni har bir mezon bo‘yicha alohida baholaydi, so‘ng yig‘ilgan ballarni qo‘shib umumiy bahoni chiqaradi. Kerakli o‘rinlarda umumiy baho boshqa shkalaga (masalan, 5 ballik tizimga) keltirilishi mumkin.

Murakkab baholash sxemasi qo‘llashda osonroq bo‘lishi va kamroq vaqt talab qilishi mumkin, chunki bunday baholash sxemasini ishlatganda test topshiruvchining javobi u qoldirgan **umumiy taassurot** asosida baholanadi. Yig‘ma baholash sxemasini qo‘llash ko‘proq vaqt oladi, chunki baholovchi **har bir mezon** asosida test topshiruvchining javobini baholash uchun javobni bir necha marotaba o‘qishi (eshitishi) kerak bo‘lishi mumkin. Shu bilan birga, bunday baholash sxemasi test topshiruvchi javobining har bir jihatiga alohida e‘tibor qaratish imkonini beradi va test topshiruvchiga baho bo‘yicha tushuntirish («*feedback*») berishni (masalan, apellyatsiya jarayonida) osonlashtiradi.

Til (ona tili, chet tili) bo‘yicha ko‘nikmalarni tekshirayorganda grammatik, leksik, orfografik, uslubiy xatolar test topshiruvchining bahosiga ta’sir qilishi tabiiy, chunki grammatika, leksi-

ka, orfografiya, uslubiyat til domenining bir qismi hisoblanadi. Boshqa fanlarda (masalan, tarix, iqtisodiyot, huquq va hokazo) baholash mezonlari va baholash sxemasida test topshiruvchilarning bunday xatolari bahoga ta'sir qilishi yoki qilmasligi masalasi oldindan hal etilishi va bu haqda test topshiruvchilarga ma'lum qilinishi lozim. Shuningdek, agar test topshiriqlari **bahsli masalalar va muammolar** asosida tuzilgan bo'lsa (masalan, «Oliy ta'lim barcha xohlovchilar uchun bepul bo'lishi kerak. Bu fikrga qay darajada qo'shilasiz? Fikringizni asoslang» singari), test topshiruvchilarning bunday topshiriqlarga javoblarini baholashda ularning **qanday xulosaga** kelganliklarini emas, balki bu xulosaga qanday kelganliklarini (o'z xulosalarini qay darajada **asoslay** olishlarini) baholash kerak¹.

Qaysi baholash sxemasi ishlatilishidan qat'iy nazar:

- baholovchilar baholash sxemasini qo'llash bo'yicha maxsus treningdan o'tishlari lozim;
- baholovchilarning ishi doimiy monitoring qilib borilishi zarur;
- baholovchilar orasidagi ishonchlilik va baholovchining ichki ishonchliligi tahlil qilib borilishi darkor;
- agar test ham yopiq, ham ochiq test topshiriqlaridan iborat bo'lsa, test topshiruvchilarning yopiq test topshiriqlaridan olgan ballari va ochiq test topshiriqlaridan olgan ballari o'rtasidagi korrelyatsiya tahlil etilishi kerak.

¹ Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.163.

	4	3	2	1
Mavzu-ning ochib berilishi	Qo'yilgan savol to'la ochib berilgan	Qo'yilgan savol qisman ochib berilgan	Qo'yilgan savolning faqat bir qismi ochib berilgan	Qo'yilgan savol ochib berilmagan
Uslub	Tanlangan uslub topshiriqqa to'la mos	Tanlangan uslub topshiriqqa mos , ayrim o'rinlarda uslubdan chekinishlar kuzatilishi mumkin	Tanlangan uslub topshiriqqa qisman mos , ko'p o'rinlarda uslubdan chekinishlar kuzatiladi	Tanlangan uslub topshiriqqa mos emas
Grammatika	Grammatik xatolar uchramaydi, turli grammatik strukturalardan o'z o'rnida foydalanilgan	Ayrim grammatik xatolar uchraydi , lekin ular mazmunga ta'sir qilmaydi	Ayrim o'rinlardagi grammatik xatolar mazmuni tushunishni qiyinlashtiradi	Grammatik qoidalariga rioya qilinmagan
So'z boyligi	So'z tanlashda xatoliklar uchramaydi, turli leksik vositalardan o'z o'rnida foydalanilgan	So'z tanlashda ayrim xatoliklar uchraydi , lekin ular mazmunga ta'sir qilmaydi	So'z boyligi cheklangan , so'z tanlashdagi xatoliklar mazmuni tushunishni qiyinlashtiradi	So'z boyligi juda cheklangan

Matn yaxlitligi	Matn izchil , mantiqiy bog‘langan, bog‘lovchi vositalardan o‘z o‘rnida foydalanilgan	Matn izchil , ayrim o‘rinlarda mantiqiy bog‘lanish buzilgan, bog‘lovchi vositalarni ishlatishda ayrim xatoliklar kuzatiladi	Aksar hollarda mantiqiy bog‘lanish buzilgan, bog‘lovchi vositalar noo‘rin qo‘llanilgan	Matn uzuq-yuluq, bir-biriga bog‘lanmagan gaplardan iborat
-----------------	---	--	--	--

Chet tilida yozma ishni baholash uchun yig‘ma baholash sxemasi namunasi

Amaliy test topshirig‘i ham yuqori darajadagi kognitiv ko‘nikmalar, ko‘nikmalar integratsiyasi, mantiqiy fikrlash va fikrni asoslash, muammolarga yechim taklif etish, murakkab vaziyatlarda qaror qabul qilish qobiliyati, shuningdek, kognitiv bo‘lmagan ko‘nikmalar (masalan, amaliy ko‘nikmalar, ijtimoiy ko‘nikmalar) haqida asosli xulosalar chiqarish imkonini beradi. Aksar mutaxassislar amaliy test topshiriqlarida boshqa turdagi test topshiriqlariga nisbatan **autentiklik** xususiyati – real hayotiy vaziyatlarga yaqinligi ko‘proq deb hisoblashadi¹. Masalan, test topshiruvchining transport vositasini boshqarish ko‘nikmasini baholash uchun

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.283.

unga transport vositasini berib, shahar ko‘chalarida boshqarishini kuzatish baholash jarayonini real hayotiy vaziyatga yaqinlashtiradi.

Autentiklik – test jarayonining test orqali baholanayotgan domenning (konstruktning) real hayotiy vaziyatlarda qo‘llanilishiga yaqinligi.

Amaliy test topshiriqlari orqali baholanadigan ko‘nikmalar **muhim** va **o‘rgatilishi mumkin** bo‘lishi (tug‘ma qobiliyatga bog‘liq bo‘lmasligi) lozim¹. Shuningdek, amaliy test topshiriqlari orqali **jarayon** (test topshuruvchi test topshirig‘ini qanday bajarganligi), **natija** (test topshiruvchi yaratgan mahsulot), yoki **ham jarayon, ham natija** baholanishi mumkin². Bunday test topshiriqlarini baholashda, odatda, yig‘ma baholash sxemasi qo‘llaniladi, chunki topshiriq ko‘p qamrovli ekanligi sababli test topshiruvchining javobini bir qancha mezonlar asosida baholashga ehtiyoj tug‘iladi. Shuningdek, amaliy ko‘nikmalarni tekshirishga qaratilgan topshiriqlarni baholash uchun har bir elementi «ha/yo‘q» (yoki 1/0) shaklida baholanadigan **ro‘yxat shaklidagi baholash sxemasi** («*checklists*») ham ishlatilishi mumkin.

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.285.

² Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.184.

Ha	Yo‘q	
		Kompyuterni yoqa oladi
		Kompyuterni to‘g‘ri o‘chira oladi
		Kerakli dasturni topib, ishga tushira oladi
		Bajargan ishini saqlab, dasturdan chiqib keta oladi
		Ishini saqlagan joyini topa oladi

Kompyuterda ishlash ko‘nikmasini baholash uchun ro‘yxat shaklidagi baholash sxemasi namunasi

Amaliy test topshiriqlarini yaratish va qo‘llashda quyidagilarga e‘tibor berish tavsiya qilinadi¹:

– yakuniy natijaga qo‘yilgan talab **aniq** bo‘lishi kerak: test topshiruvchilar ulardan nima talab qilinayotganligini, qanday mahsulot kutilayotganligini aniq tasavvur qila olishlari lozim, biroq qo‘yilgan vazifani bajarish uchun ancha o‘ylash talab etilishi darkor;

– topshiriq shunday tuzilgan bo‘lishi kerakki, test topshiruvchilarning topshiriqni bajarish jarayonini kuzatish yoki natijasini baholash orqali ularda kognitiv yoki kognitiv bo‘lmagan **ko‘nikmalarning** qay **darajada** shakllanganligi haqida **asosli xulosalar** chiqarish imkonini bersin;

– topshiriq **har tomonlama baholash** imkonini berishi kerak, masalan, tabiiy fanlarda test topshiruvchining labora-

¹ Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.192.

toriya ishini kuzatish bilan birga uning bu ish bo'yicha tushuntirishini eshitish va yozma hisobotini o'qish orqali baholash mumkin;

– topshiriq shunday tuzilishi kerakki, uning yagona yechimi bo'lmasin, shu bilan birga, yechim algoritmik (oddiy ketma-ketlikka asoslangan) bo'lmasin, ya'ni topshiriqning yechimi **ko'p qirrali** va **har xil tarafdin baholanadigan** bo'lishi lozim;

– topshiriqni bajarish jarayonida test topshiruvchi **yechimga kelish uchun turli kognitiv strategiyalarni ishga solishga**, mustaqil qarorlar qabul qilishga majbur bo'lishi darkor;

– test topshirig'ini bajarish uchun qo'yiladigan **cheklovlar** (beriladigan vaqt; qo'shimcha resurslardan va texnik vositalardan foydalanish, shuningdek, o'zgalardan yordam olish imkoni berilishi yoki berilmasligi va hokazo) oldindan aniq belgilab olinishi, bu cheklovlar real hayotiy vaziyatda shu kabi topshiriqni bajaradigan shaxslar duch keladigan cheklovlarga imkon qadar yaqin bo'lishi zarur.

Umuman, ochiq test topshiriqlarini tuzishda quyidagi ishlarni amalga oshirish tavsiya etiladi¹:

1. Tekshirilishi lozim bo'lgan domen va konsruktlarni **aniq** belgilab olish.

2. Tekshirilishi lozim bo'lgan kognitiv ko'nikma darajasiga **mos** topshiriq turini tanlash va yozish. Quyi darajadagi kog-

¹ Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013. – P.213.

nitiv ko‘nikmalarni baholash uchun ochiq test topshiriqlarini ishlatish maqsadga muvofiq emas.

3. Bir testga kiritilgan turli test topshiriqlari tekshirilishi lozim bo‘lgan domen va konstruktni adekvat darajada **qamrab** olishiga ishonch hosil qilish.

4. Test topshiruvchilar uchun yo‘riqnomalarni va test topshiriqlarini o‘qib chiqib, har qanday xatolar (grammatik, leksik, uslubiy) va noaniqliklarni bartaraf qilish, yo‘riqnomalar va test topshiriqlarining test topshiruvchilar uchun **tushunarli** ekanligiga ishonch hosil qilish.

5. Baholash **mezonlari** va baholash **sxemasini** tuzish.

6. Test topshiriqlari va test jarayonini **aprobatsiyadan** o‘tkazish.

7. Aprobatsiya natijalari asosida yo‘riqnomalar va test topshiriqlarini **tahrir qilish**, konstruktg bog‘liq bo‘lmagan omillarni **bartaraf etish**.

Umuman, yopiq va ochiq test topshiriqlarining kerakli turini tanlash, topshiriqlarni tuzish va testni yaratish jarayonida quyidagi savollarga javob berilishi lozim¹:

1. Topshiriq testning maqsadi, test topshiruvchilar kontingenti va test spetsifikatsiyasiga **mosmi?**

2. Test tekshirilishi lozim bo‘lgan domen va konstruktlarni adekvat darajada **qamrab olganmi?**

3. Test topshiriqlari va test uchun yo‘riqnomalar **aniq, tushunarli va to‘liqmi?**

¹Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013. – P.210.

4. Test topshiriqlarining o‘zi **aniq, tushunarli va to‘liqmi?**

Test topshiriqlari shaklini tanlash hamda test topshiriqlarini tuzish va qo‘llashda testni tuzuvchilar yuqoridagi savollarga javob bera olishi, shuningdek, test topshiriqlarini yaratish qoidalariga rioya etishi test natijalaridan chiqariladigan xulosalarning validligiga hissa qo‘shadi.

Asosiy xulosa:

Ochiq test topshiriqlarining ham bir qancha turlari mavjud. Ulardan qisqa javobni talab qiladigan test topshiriqlari quyi darajadagi kognitiv ko‘nikmalarni (bilish, tushunish, qo‘llash) tekshirish imkonini beradi. Kengaytirilgan javobni talab qiladigan test topshiriqlari va amaliy test topshiriqlari esa yuqori darajadagi kognitiv ko‘nikmalar, ko‘nikmalar integratsiyasi, mantiqiy fikrlash va fikrini asoslash, muammolarga yechim taklif etish, murakkab vaziyatlarda qaror qabul qilish qobiliyati, shuningdek, kognitiv bo‘lmagan ko‘nikmalarni (masalan, yozma va og‘zaki nutq yoki amaliy va ijtimoiy ko‘nikmalar) tekshirish imkonini beradi. Testdan chiqariladigan xulosalar valid bo‘lishi uchun test topshiriqlarini tuzish qoidalariga amal qilish va har bir topshiriq turini o‘z o‘rnida qo‘llash lozim.

Kengaytirilgan javobni talab qiladigan test topshiriqlari va amaliy test topshiriqlarini baholash uchun murakkab baholash sxemasi yoki yig‘ma baholash shuningdek, ro‘yxat shaklidagi baholash sxemasi qo‘llanilishi mumkin

4.4. Chet tilidan test topshiriqlarining o'ziga xos xususiyatlari (ingliz tili bo'yicha testlar misolida)

Chet tilidan test topshiriqlarining maqsadi test topshiruvchilarning til bilan bog'liq bilim, ko'nikma va malakalari haqida xulosa chiqarishdir. Bu xulosa asosida test topshiruvchini ta'lim chet tilida olib boriladigan o'quv muassasasiga qabul qilish yoki qabul qilmaslik, chet elga malaka oshirishga yuborish yoki yubormaslik, ishga qabul qilish yoki qabul qilmaslik singari turli qarorlar qabul qilinishi mumkin. Xulosa asosli bo'lishi uchun til bilan bog'liq bilim, ko'nikma va malakalar nimalardan tashkil topishini, ya'ni tekshirilishi lozim bo'lgan domen va konstruktarni aniqlashtirib olish zarur.

Til muloqot vositasidir. **Kommunikativ kompetensiya** (tilda muloqot qila olish kompetensiyasi) til haqidagi bilimlar (so'z boyligi, fonetika, grammatika, stilistika haqidagi bilimlar) va tildan turli vaziyatlarda foydalana olish ko'nikmasini o'z ichiga oladi¹. Shu bilan birga, muloqot samarali bo'lishida mavjud bilimlar va hayotiy tajriba («*background knowledge*») ham muhim rol o'ynaydi, chunki insonlar nimadir haqida muloqot qiladilar va shu nimadir haqida tasavvur bo'lmasa, muloqot jarayoni qiyinlashadi. Shundan kelib chiqqan holda chet tili testlari **tilga doir bilimlarni** (masalan, so'z boyligini, grammatika qoidalarini bilishni) va **til ko'nikmalarini** (pas-

¹ Douglas D. Understanding Language Testing. – London-New York: Routledge, 2010. – P.30.

siv ko‘nikmalar – o‘qib tushunish va tinglab tushunish hamda aktiv ko‘nikmalar – gapirish va yozish) tekshirishga qaratilgan bo‘lishi mumkin.

Chet tili testlari **diskret test topshiriqlaridan** iborat bo‘lishi yoki **kommunikativ test** shaklida tuzilgan bo‘lishi mumkin. **Diskret** test topshiriqlari tilning konkret bir elementini (masalan, test topshiruvchining biron-bir fonema, so‘z yoki grammatik qoidani bilishini) tekshirishga qaratilgan bo‘ladi. **Kommunikativ** testlar esa test topshiruvchi tilni real hayotiy vaziyatlarda qo‘llay olishini tekshiradi.

Diskret test topshirig‘i – tilning konkret bir elementini tekshirishga qaratilgan test topshirig‘i.

Kommunikativ test – test topshiruvchining tilni real hayotiy vaziyatlarda qo‘llay olishini tekshirishga qaratilgan test.

Kommunikativ testlar quyidagi **talablarga** javob berishi talab etiladi¹:

- mazmunli muloqotga asoslanishi;
- autentik (real hayotiy) vaziyatga qurilishi;
- test topshiruvchiga notanish til materialini asosida tuzilishi;
- test topshiruvchidan chet tilida nimadir yaratishni talab qilishi;
- barcha til ko‘nikmalarini qamrab olishi lozim.

¹ Brown J.D. Testing in Language Programs: A Comprehensive Guide to English Language Assessment, Upper Saddle River. – N.J.: Prentice Hall Regents, 2005. – P.21.

Choose the correct answer.

When I entered the room, John ... a TV.

A) watched B) was watching C) is watching D) watches

Diskret test topshirig'i namunasi

Part 2

Write an answer to one of the questions 2 – 5 in this part.

Write your answer in 120 – 180 words in an appropriate style on the next page, putting the question number in the box.

4. You are attending a summer language course and have been asked to report on a local leisure facility (e.g. cinema, sports hall, etc.) for the benefit of students attending the next course.

Write your report describing the facility and what it has to offer, and commenting on its good and bad points.

Kommunikativ testdan parcha¹

So‘z boyligini (*vocabulary*) tekshirishga qaratilgan test topshiriqlari quyidagi ko‘nikmalarni tekshirishi mumkin²:

- so‘zni tanish (test topshiruvchi so‘zni ko‘rganda uning ma’nosini tushunadimi?);
- so‘zni eslash (test topshiruvchi kerakli so‘zni eslay oladimi?);

¹ Carne P., Hashemi L. and Thomas B. Cambridge Practice Tests for First Certificate 1, Self-Study Edition. – Cambridge: Cambridge University Press, 1996. – P.45.

² Read J. Assessing Vocabulary. – Cambridge: Cambridge University Press, 2000. – P.155.

- soʻzni tushunish (test topshiruvchi soʻzni biron-bir kontekstda eshitganda yoki oʻqiganda tushunadimi?);
- soʻzni qoʻllash (test topshiruvchi gapirganda yoki yozganda kerakli soʻzni ishlata oladimi?).

Dastlabki ikki koʻnikma diskret test topshiriqlari yordamida tekshirilishi mumkin boʻlsa, keyingi ikki koʻnikma kommunikativ test yordamida tekshirilishi lozim.

1. *Choose the correct answer. loathe means ...*

- A) dislike intensely B) become seriously ill
C) search carefully D) look very angry

2. *Write the correct word to complete the sentence.*

Because of the snow, the football match was ... until the following week.

Soʻz boyligini tekshirishga qaratilgan diskret test topshiriqlari namunasi – birinchi topshiriq soʻzni tanishni, ikkinchi topshiriq soʻzni eslashni tekshiradi

Grammatikani (*grammar*) tekshirishga qaratilgan test topshiriqlari quyidagi koʻnikmalarni tekshirishi mumkin¹:

– grammatik shakllarni bilish (test topshiruvchi kerakli grammatik shaklni taniy oladimi?) – bu koʻnikma diskret test topshiriqlari yordamida tekshirilishi mumkin;

– grammatik shakllarni qoʻllash (test topshiruvchi grammatik jihatdan toʻgʻri va kommunikativ vaziyatga mos kela-

¹ Purpura J.E. *Assessing Grammar*. – Cambridge: Cambridge University Press, 2004. – P.89.

digan gaplarni tuza oladimi?) – bu ko‘nikma kommunikativ test doirasida tekshiriladi.

Passiv til ko‘nikmalarini (o‘qib tushunish va tinglab tushunish) tekshirishga qaratilgan testlar o‘qilgan yoki eshitilgan matnning umumiy ma’nosini ilg‘ash, konkret ma’lumotlarni tushunish, ochiq aytilmagan fikrlarni anglash (inferensiya) singari ko‘nikmalarni tekshiradi. **Aktiv til ko‘nikmalarini** (gapirish va yozish) tekshirishga qaratilgan testlar esa test topshiruvchining tilga doir bilimlarni ishga solgan holda berilgan kommunikativ vaziyatga mos og‘zaki yoki yozma nutq hosil qilish ko‘nikmasini tekshiradi. **Kombinatsiyalashgan testlar** esa test topshiruvchini bir paytning o‘zida **bir nechta parametrlar** bo‘yicha baholash imkonini beradi. Masalan, test topshiruvchidan matnni o‘q‘ib, uning qisqacha mazmunini (*summary*) yozishni talab qiladigan test topshirig‘i ham o‘qib tushunish, ham yozish ko‘nikmalarini tekshiradi.

Chet tili testlarida standart test topshiriqlaridan tashqari **qo‘shimcha topshiriq turlari** ham ishlatilishi mumkin:

1. «Cloze» testi (inglizcha «*closure*» – «*yopilish*» so‘zidan). Test topshiruvchiga matn beriladi, bu matnda ayrim so‘zlar tushirib qoldirilgan bo‘ladi. Test topshiruvchidan tushirib qoldirilgan so‘zlarni topish talab qilinadi. Test yopiq shaklda (muqobil javoblar berilgan) yoki ochiq shaklda (javoblarsiz) bo‘lishi mumkin. Matndagi tushirib qoldirilgan so‘zlar mexanik tarzda (har nechanchidir, masalan, har oltinchi so‘z)

yoki ma'lum bir maqsadga qaratilgan holda (masalan, olmoshlar yoki bog'lovchi so'zlar) tanlanishi mumkin. Bunday testlar tilga doir bilimlarni yoki passiv til ko'nikmalarini tekshirish maqsadida qo'llanilishi mumkin.

Read the text below and choose which answer A, B, C or D best fits each space.

The Bongou people are an ethnic minority in a (Q1) ... British colony in Africa. They have their own language, culture and traditions, all of which are under (Q2) ... from the government's policies and the (Q3) ... of the population from the countryside to cities.

- Q1. A) last B) past C) previous D) former
Q2. A) pressure B) force C) weight D) strength
Q3. A) progress B) movement C) ramble D) motion

Yopiq shakldagi «cloze» testi namunasi

For questions 1 – 8, read the text below and think of the word which best fits each gap. Use only one word in each gap. There is an example at the beginning (0).

Example: (0) AS

Motorbike Stunt Rider

I work (0) ... a motorbike stunt rider – that is, I do tricks on my motorbike at shows. The Le Mans racetrack in France was (Q1) ... I first saw some guys doing motorbike stunts. I'd never seen anyone riding a motorbike using just the back wheel

before and I was (Q2) ... impressed I went straight home and taught (Q3) ... to do the same. It wasn't very long before I began to earn my living at shows performing my own motorbike stunts.

Ochiq shakldagi «cloze» testi namunasi

2. Transformatsiya test topshirig'i. Test topshiruvchiga gap beriladi va bu gapning ma'nosini saqlab qolgan holda shaklini o'zgartirish talab etiladi. Grammatik strukturalarni qo'llay olish ko'nikmasini tekshirish uchun ishlatiladi. Bunday test topshiriqlari, odatda, qisqa javobni talab etuvchi ochiq test topshiriqlari shaklida beriladi – berilgan gapning transformatsiya qilingan shaklini (ikkinchi gapni) to'ldirish kerak.

Complete the second sentence so that it has a similar meaning to the first sentence, using the word given. Do not change the word given. You must use between two and five words, including the word given.

So that Susan would be fit for the skiing, she went to the gym three times a week.

order

Susan went to the gym three times a week fit for the skiing.

Transformatsiya test topshirig'i namunasi

3. Xatoni topish (to‘g‘rilash). Test topshiruvchi berilgan gapda (matnda) xatolarni topishi yoki topib, to‘g‘rilashi talab qilinadi. Odatda, grammatik strukturalarni yoki so‘zlarni qo‘llay olish ko‘nikmasini tekshirish uchun ishlatiladi. Bunday test topshiriqlari yopiq shaklda (xatoni topish) yoki qisqa javobni talab qiluvchi ochiq test topshiriqlari shaklida (xatoni to‘g‘rilab yozish) bo‘lishi mumkin.

There are four underlined sections in the text. Select the one which is wrong.

The minimum wage bill provides for a gradual reduction of the minimum wage from \$5,15 an hour to \$7,25 an hour over two years.

A) bill

B) reduction

C) minimum

D) over two years

Xatoni topish test topshirig‘i namunasi

4. «To‘g‘ri – noto‘g‘ri – berilmagan» shaklidagi test topshiriqlari. O‘qish uchun matn va bir nechta gaplar beriladi. Test topshiruvchi har bir gapning o‘qilgan matn mazmuniga mos yoki mos emasligini topishi lozim. Ayrim gaplar esa matnda berilmagan bo‘ladi. Yozma matnni tushunish ko‘nikmasini tekshirish uchun ishlatiladi. Ayrim davlatlarning imtihon tizimlarida (masalan, Rossiyaning Yagona davlat imtihonida) bunday shakldagi test topshiriqlari tinglab tushunish ko‘nikmasini tekshirish uchun ham qo‘llanilsa-da,

aksar mutaxassislar tinglab tushunishni tekshirishda bunday topshiriqlarni **ishlatmaslikni** tavsiya etadilar.

Questions 1 – 4

Do the following statements agree with the information given in the passage? Write:

TRUE if the statement agrees with the information

FALSE if the statement contradicts the information

NOT GIVEN if there is no information on this

1. Insects are often very different in their shape and colour.
2. Insects are a vital part of the environment.
3. Insects share many of the same characteristics as other animals.

INSECTS

Despite the dazzling diversity of shape and colour among insects, they all share three fundamental characteristics in common. They are made up of three component parts. All are invertebrates, that is, they have no backbone. And, finally, all have six legs.

«To ‘g‘ri – noto ‘g‘ri – berilmagan» shaklidagi test topshiriqlari namunasi

Chet tilidan test topshiriqlarini yaratishda test topshiriqlari uchun umumiy bo‘lgan qoida va tavsiyalardan tashqari quyidagi qo‘shimcha **qoidalar va tavsiyalarga** rioya qilmoq lozim:

1. Test topshirig‘ida ishlatilgan til materiali (leksik va grammatik birliklar) test topshiruvchilarning **darajasiga** mos bo‘lishi kerak. Misol uchun, agar test Umumyevropa shkalasi (CEFR – Common European Framework of Reference for Languages) bo‘yicha B2 darajasidagi test topshiruvchilar uchun mo‘ljallangan bo‘lsa, test topshirig‘idagi so‘zlar va grammatik konstruksiyalar shu daraja uchun tanish bo‘lishi lozim¹. Chet tilida o‘qib tushunish muammolari bo‘yicha tadqiqot olib borgan aksar olimlarning fikricha, o‘qilgan matnni tushunish uchun undagi leksik birliklarning kamida 95 foizi tanish bo‘lishi talab qilinadi².

Misol:

Read the text and answer questions 1 – 10.

<...> Godbole is not very old, perhaps in his mid-forties. He has a balding head and a round face dominated by a handlebar moustache. He walks with heavy steps and his overfed stomach droops over his khaki trousers.<...>

Q7. Inspector Godbole is portrayed as ...

A) hairless B) elderly C) stout D) active

B2 daraja uchun mo‘ljallangan testdan olingan ushbu test topshirig‘ida kalitda ishlatilgan «*stout*» so‘zi test topshiruv-

¹ Brown J.D. Testing in Language Programs: A Comprehensive Guide to English Language Assessment, Upper Saddle River. – N.J.: Prentice Hall Regents, 2005. – P.44.

² Qarang: Schmitt R., Jiang X. and Grabe W. The Percentage of Words Known in a Text and Reading Comprehension // The Modern Language Journal. 2011. №95. – P.26 – 43.

chilar darajasidan yuqori bo‘lganligi uchun aprobatsiyada test topshiruvchilarning faqat 17 foizi bu topshiriqqa to‘g‘ri javob bergan.

Yana bir misol:

Read the text and answer questions 11 – 20.

<...> I just wanted a piece of my life that didn’t involve my little sister. When I complained to my mother, she’d just smile and say, «One day you’ll want **her** around». Sure.

Q17. The word «her» in Paragraph 4 refers to ...

A) mother B) sibling C) group mate D) girl friend

B1 daraja uchun mo‘ljallangan testdan olingan ushbu test topshirig‘ida ham kalitda ishlatilgan «*sibling*» so‘zi test topshiruvchilar darajasidan yuqori.

2. Test topshiruvchilar test topshirig‘ining maqsadini tushunib olishlari uchun yetarli **kontekst** berilishi lozim¹. Masalan, test topshirig‘ida to‘ldirilishi kerak bo‘lgan gap berilsa, gapning to‘ldiriladigan qismi gap boshida kelmasligi tavsiya qilinadi.

Misol:

Choose the correct answer.

... nail won’t be enough for this job. I need several.²

A) the B) an C) – D) a

¹ Brown J.D. Testing in Language Programs: A Comprehensive Guide to English Language Assessment, Upper Saddle River. – N.J.: Prentice Hall Regents, 2005. – P.52.

² Davlat test markazi. Ingliz tili. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.8.

Ushbu diskret test topshirig‘ida gapning to‘ldirilishi kerak bo‘lgan qismi gap boshida berilganligi uchun test topshiruvchi topshiriqni qayta o‘qigandan keyingina undan nima kutilayotganini tushunadi.

Shuningdek, yetarli kontekstning berilmasligi (ayniqsa, diskret test topshiriqlarida) bir nechta javobning to‘g‘ri bo‘lib qolishiga olib kelishi mumkin.

Misol:

Choose the correct answer.

Usually the cows in this farm ... milked twice a day.

A) were B) is C) have been D) are

*Mazkur misol «Test topshiriqlari to‘plami:
amaliy qo‘llanma»dan¹ olingan*

Ushbu test topshirig‘ida kontekst yetarli bo‘lmagani uchun berilgan gap hozirgi zamondagi ish-harakat sifatida ham, o‘tgan zamondagi ish-harakat sifatida ham talqin etilishi mumkin, natijada A va D muqobil javoblari to‘g‘ri javob sifatida qabul qilinishi mumkin.

3. Tekshirilayotgan mavzuni bilmaydigan test topshiruvchiga to‘g‘ri javobni topishga yordam beradigan grammatik yoki boshqa shakldagi «**yordamlar**» bo‘lmasligi zarur.²

¹ Davlat test markazi. Ingliz tili. 2019-yil. Test topshiriqlari to‘plami: amaliy qo‘llanma. – Toshkent: Davr-press, 2019. – B.19.

² Brown J.D. Testing in Language Programs: A Comprehensive Guide to English Language Assessment, Upper Saddle River. – N.J.: Prentice Hall Regents, 2005. – P.48.

Misol:

Choose the correct answer.

According to the passage, the fruit that Adam ate was an ...

A) pear B) banana C) apple D) papaya

Ushbu test topshirig‘i matni tushunishni tekshirishga qaratilgan bo‘lsa ham, ingliz tilida noaniq artiklning «*an*» shakli unli tovushlar bilan boshlanadigan so‘zlardan oldin ishlatilishini bilgan test topshiruvchi bu topshiriqqa matni o‘qimasdan turib ham javob berish mumkin: berilgan muqobil javoblar ichidan faqat «*apple*» so‘zi unli tovush bilan boshlangan.

4. Matnlar asosida tushunishni tekshirishga qaratilgan test topshiriqlari tuzishda quyidagilarga e‘tibor berish lozim¹:

– test topshiriqlari matnda ma‘lumotlarning berilish ketma-ketligida joylashishi shart: birinchi topshiriq matnning boshida keltirilgan ma‘lumotni tushunganlikni tekshiradi, ikkinchi topshiriq – undan keyingi ma‘lumotni va hokazo (moslashtirishni talab etuvchi test topshiriqlarida bu qoidaga rioya qilinmaydi);

– matndagi ayni bir ma‘lumotni tushunganlikni tekshirish uchun bittadan ortiq test topshirig‘i ishlatilmasligi zarur;

¹ Green A. Exploring Language Assessment and Testing: Language in Action. – NY: Routledge, 2014. – P.113.

– test topshiriqlarining tili (ularda qoʻllanilgan leksik va grammatik birliklar) matnning tiliga nisbatan soddaroq boʻlishi darkor;

– topshiriqlar shunday tuzilishi lozimki, test topshiruvchi matnni oʻqimasdan (eshitmasdan) turib (masalan, oldindan bilgan maʼlumoti asosida yoki mantiqiy fikrlash yoʻli bilan) unga javob bera olmasin;

– test topshiriqlariga javob berish uchun test topshiruvchidan maxsus bilim (masalan, madaniy kontekstga oid bilimlar) talab qilinmasligi kerak;

– topshiriqlardagi maʼlumotlar matndan aynan (soʻzmasoʻz) olinmay, parafraza qilinishi (boshqacha soʻzlar bilan ifodalanishi) shart, aks holda, test topshiruvchi matnni tushunmasa ham, shunchaki bir xil soʻzlarni topish orqali topshiriqlarni bajara oladi;

– muqobil javobli test topshiriqlarida muqobil javoblar bir-birini inkor qilmasligi darkor (masalan, «A. The river is narrow. B. The river is wide. C. The river is pretty» shaklida boʻlishi mumkin emas – bunda A va B muqobil javoblari bir-birini inkor qiladi);

– tinglab tushunish koʻnikmasi tekshirilayotganda test topshiruvchi vaziyatga moslashib olishi uchun yetarlicha vaqt berilishi zarur – buning uchun tekshirilayotgan maʼlumotlar eshitilayotgan matnning dastlabki soniyalaridan boshlanmasligi talab qilinadi;

– tinglab tushunish ko‘nikmasi tekshirilayotganda tekshirilayotgan ma’lumotlar matnda bir-biriga juda yaqin joylashmasligi talab qilinadi, aks holda, test topshiruvchi bir test topshirig‘iga javob bergandan keyin diqqatini keyingi topshiriqqa qaratishga ulgurmaydi.

Asosiy xulosa:

Chet tilidan test topshiriqlari tilga doir bilimlarni va til ko‘nikmalarini diskret test topshiriqlari yoki kommunikativ test yordamida tekshirishi mumkin. Bunda standart test topshiriqlaridan tashqari bir qancha qo‘shimcha topshiriq turlari ham («cloze» testi, transformatsiya test topshirig‘i, xatoni topish yoki to‘g‘rilash, «to‘g‘ri – noto‘g‘ri – berilmagan» shaklidagi test topshiriqlari) ishlatilishi mumkin. Chet tilidan test topshiriqlarini yaratishda umumiy qoida va tavsiyalardan tashqari bir qancha o‘ziga xos talablarni ham hisobga olish talab etiladi.

5. TEST TOPSHIRIQLARINING STATISTIK TAHLILI

5.1. Statistika tahlilning ahamiyati

Testning maqsadi hamda test orqali tekshiriladigan domen va konstruktlar aniqlab olingandan keyin test spetsifikatsiyasiga muvofiq test topshiriqlari tuziladi. Bu topshiriqlar mutaxassislar tomonidan ekspertiza qilinishi va aprotatsiyadan o'tishi shart. Aprotatsiya test va testga kiritilgan test topshiriqlarining sifati haqida **empirik ma'lumotlarni** yig'ish imkonini beradi. Bu ma'lumotlar asosida test topshiriqlari statistik tahlil qilinib, test topshiriqlari bilan bog'liq muammolar bartaraf etilishi lozim. Bu **testning ishonchliligini** hamda test yordamida chiqariladigan **xulosalarning validlik darajasini** oshirishga xizmat qiladi.

Test topshiriqlarining statistik tahlili testning sifatini oshirishda quyidagicha yordam beradi¹:

- haddan tashqari qiyin va haddan tashqari oson savollarni aniqlab beradi;
- tanlangan test uslubi qay darajada samarali ishlayotganini ko'rsatib beradi;

¹Green R. Statistical Analyses for Language Testers. – NY: Palgrave Macmillan, 2013. – P.25.

- test topshiriqlari test topshiruvchilarning ichidan o‘lchanayotgan domen va konstruktlarni **egallaganlarni** bu domen va konstruktlarni **egallamaganlardan** (bilimlilarni bilimsizlardan) qay darajada ajratayotganini aniqlab beradi;
- har bir test topshirig‘i testning ichki barqarorligiga qay darajada hissa qo‘shayotganini aniqlab beradi.

Statistik tahlil yordamida qaysi test topshiriqlari sifatli tuzilgan va testda ishlatsa bo‘ladi, qaysilarini qayta ishlash va qaysilaridan voz kechish kerakligi aniqlab olinadi. Statistik tahlil nafaqat **test topshiriqlari aprotasiyasidan** keyin, balki **haqiqiy imtihondan** keyin ham o‘tkazilishi lozim, chunki statistik tahlil orqali olingan ma’lumot test yordamida chiqariladigan xulosalarga ta’sir qilishi mumkin.

Asosiy xulosa:

Statistik tahlil test topshiriqlarining sifati haqida obyektiv ma’lumotlar olish imkonini beradi. Test topshiriqlarining sifat ko‘rsatkichlari testning ishonchliligi hamda test yordamida chiqariladigan xulosalarning validligiga bevosita ta’sir qiladi.

5.2. Test topshiriqlarining sifat ko'rsatkichlari

Test topshiriqlari statistik tahlilida quyidagi asosiy ko'rsatkichlarga e'tibor berish lozim:

- test topshirig'ining qiyinlik darajasi;
- test topshirig'ining diskriminatsiya indeksi;
- test topshiruvchilar javoblarining distraktorlar bo'yicha taqsimoti;
- test topshirig'ining DIF ko'rsatkichi.

Statistik tahlilni maxsus dasturlar (masalan, Statistical Package for the Social Sciences – SPSS dasturi) yoki Microsoft Office tarkibiga kiruvchi Excel dasturi yordamida amalga oshirish mumkin.

5.2.1. Test topshirig'ining qiyinlik darajasi

Test topshirig'ining qiyinlik darajasi (p qiymat deb ham ataladi) quyidagicha hisoblanadi:

$$p = \frac{R}{T}$$

bunda: p = test topshirig'ining qiyinlik darajasi;

R = test topshirig'iga to'g'ri javob berganlar soni;

T = test topshirig'ini bajarganlarning umumiy soni.

Misol uchun, test topshirig'ini 300 kishi bajargan bo'lsa va shulardan 180 tasi to'g'ri javob bergan bo'lsa, test top-

shirig'ining qiyinlik darajasi (p qiymati) $180/300 = 0,60$ ga teng.

Test topshirig'ining qiyinlik darajasi 0 dan 1 gacha qiymatda o'lanadi, bu qiymat 1 ga qanchalik yaqin bo'lsa, test topshirig'i shunchalik oson hisoblanadi. Test topshirig'ining qiyinlik darajasini tahlil qilishda quyidagilarni hisobga olish lozim¹:

– qiyinlik darajasi 0,50 atrofida bo'lgan test topshiriqlari eng sifatli hisoblanadi, chunki test topshiriqlarining qiyinlik darajasi 0,50 bo'lganda testning dispersiyasi maksimal qiymatga ega bo'ladi;

– qiyinlik darajasi 0,30 va 0,70 oralig'ida bo'lgan test topshiriqlarini ham testda ishlatish mumkin;

– qiyinlik darajasi 0,20 va 0,80 oralig'ida bo'lgan test topshiriqlarini ishlatishda boshqa ko'rsatkichlarni (test topshirig'ining diskriminatsiya indeksi, test topshirig'ining butun test bilan korrelyatsiyasi) ham hisobga olish zarur;

– qiyinlik darajasi 0,20 dan past (juda qiyin) va 0,80 dan yuqori (juda oson) test topshiriqlarini ishlatish tavsiya etilmaydi.

Test topshirig'ining qiyinlik darajasi juda muhim ko'rsatkich bo'lib, testning ishonchliligini ko'rsatuvchi boshqa juda ko'plab parametrlari shu ko'rsatkichga bevosita bog'liq. Shu

¹ Green R. Statistical Analyses for Language Testers. – NY: Palgrave Macmillan, 2013. – P.27.

bilan birga, test topshirig‘ining qiyinlik darajasi nisbiy ko‘rsatkich ekanligini ham hisobga olish kerak: masalan, «Hisoblang: $7x8+14$ » shaklidagi test topshirig‘i 1-sinf o‘quvchilari uchun qiyin, yuqori sinf o‘quvchilari uchun oson bo‘lishi mumkin. Shuning uchun test topshiriqlarini aprotatsiyadan o‘tkazayotganda aprotatsiyada qatnashayotgan kontingent o‘z xususiyatlari jihatidan testni haqiqiy topshiradiganlar kontingentiga yaqin bo‘lishi darkor.

5.2.2. Test topshirig‘ining diskriminatsiya indeksi

Test topshirig‘ining diskriminatsiya indeksi – bu test topshirig‘ining bilimli (tekshirilayotgan domenni yaxshi o‘zlashtirgan yoki tekshirilayotgan konstrukt bo‘yicha yuqori ko‘rsatkichga ega bo‘lgan) test topshiruvchini bilimsiz (tekshirilayotgan domenni yaxshi o‘zlashtirmagan yoki tekshirilayotgan konstrukt bo‘yicha past ko‘rsatkichga ega bo‘lgan) test topshiruvchidan **ajratib berish** qobiliyatidir. Diskriminatsiya indeksi quyidagicha hisoblanadi:

$$D = p_h - p_l$$

bunda: D = test topshirig‘ining diskriminatsiya indeksi;

p_h = test topshirig‘ining yuqori guruh uchun p qiymati;

p_l = test topshirig‘ining quyi guruh uchun p qiymati.

Demak, test topshirig‘ining diskriminatsiya indeksini topish uchun oldin test topshirig‘ining yuqori va quyi guruhlari

uchun p qiymatini topish kerak bo‘ladi. Buning uchun quyidagilarni amalga oshirishimiz lozim:

– test topshiruvchilar testda olgan natijalariga qarab kamayib borish tartibida (ko‘proq ball olganlardan boshlab pastroq ball olganlarga qarab) joylashtiriladi;

– test topshiruvchilardan yuqori guruh va quyi guruh ajratib olinadi: yuqori guruh – eng ko‘p ball olganlar va quyi guruh – eng past ball olganlar, bunda har ikkala guruhda test topshiruvchilarning soni teng bo‘lishi shart (odatda, eng yuqoridan 27% va eng pastdan 27% olinadi, lekin test topshiruvchilarning umumiy soni kam bo‘lsa, teng ikkiga bo‘lish ham mumkin);

– har ikkala guruh uchun test topshirig‘ining p qiymati (qiyinlik darajasi) hisoblanadi: buning uchun shu guruhda test topshirig‘iga to‘g‘ri javob berganlar soni shu guruhda test topshirig‘ini bajarganlarning umumiy soniga bo‘linadi.

Misol uchun, 100 kishi test topshirdi. Test topshiruvchilarni testda olgan natijalariga qarab kamayib borish tartibida joylashtiramiz. Shunda eng yuqori natija ko‘rsatgan 27 ta test topshiruvchi yuqori guruhni, eng past natija ko‘rsatgan 27 ta test topshiruvchi quyi guruhni tashkil qiladi. Yuqori guruhda 27 ta test topshiruvchidan 25 tasi birinchi test topshirig‘iga to‘g‘ri javob bergan, demak, $p_h = 25/27 = 0,93$. Quyi guruhda 27 ta test topshiruvchidan 10 tasi xuddi shu test topshirig‘iga to‘g‘ri javob bergan, demak, $p_l = 10/27 = 0,37$. Demak, test topshirig‘ining diskriminatsiya indeksi $0,93 - 0,37 = 0,56$ ga teng.

Test topshirig‘ining diskriminatsiya indeksi -1 dan $+1$ gacha bo‘lgan qiymatlar oralig‘ida o‘lchanadi. Diskriminatsiya indeksini tahlil etishda quyidagilarni hisobga olish kerak¹:

– diskriminatsiya indeksi $0,40$ va undan yuqori bo‘lgan test topshiriqlari eng sifatli hisoblanadi;

– diskriminatsiya indeksi $0,30$ va $0,39$ oralig‘ida bo‘lgan test topshiriqlarini qayta ishlab, sifatini oshirish mumkin;

– diskriminatsiya indeksi $0,20$ va $0,29$ oralig‘ida bo‘lgan test topshiriqlarini qayta ishlash lozim;

– diskriminatsiya indeksi $0,19$ va undan past bo‘lgan test topshiriqlarini ishlatish tavsiya etilmaydi.

Test topshirig‘ining diskriminatsiya indeksi manfiy qiymatga ega bo‘lsa, bunday test topshirig‘iga bilimsizlar to‘g‘ri javob berib, bilimlilar noto‘g‘ri javob berayotgan bo‘ladi – bunday test topshirig‘ini testda ishlatib bo‘lmaydi.

Test topshirig‘ining diskriminatsiya indeksini hisoblashning boshqa usullari ham mavjud. Shulardan biri **nuqtali biseral korrelyatsiya** deb ataladi va quyidagi formula bo‘yicha hisoblanadi:

$$r_{pbis} = \frac{\bar{X}_1 - \bar{X}}{S_x} \sqrt{\frac{p}{q}}$$

bunda: r_{pbis} = test topshirig‘ining nuqtali biseral korrelyatsiya koeffitsiyenti;

¹ Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.324.

\bar{X}_1 = test topshirig'iga to'g'ri javob berganlarning o'rtacha bali;

\bar{X} = barcha test topshiruvchilarning o'rtacha bali;

S_x = testning standart og'ishi;

p = test topshirig'ining qiyinlik darajasi;

$q = 1 - p$

Bu formula test topshirig'ining bajarilishi va testning umumiy natijasi o'rtasida qay darajada bog'liqlik borligini ko'rish hamda test topshirig'ining diskriminatsiya indeksini aniqroq hisoblash imkonini beradi. Test topshirig'ining nuqtali biseral korrelyatsiya koeffitsiyenti ham -1 dan $+1$ gacha bo'lgan qiymatlar oralig'ida o'lchanadi.

Shuni hisobga olish lozimki, test topshirig'ining diskriminatsiya indeksi uning qiyinlik darajasiga (p qiymatiga) bevosita bog'liq: haddan tashqari oson va haddan tashqari qiyin test topshiriqlarining diskriminatsiya indeksi ham talab darajasida bo'lmaydi.

5.2.3. Test topshiruvchilar javoblarining distraktorlar bo'yicha taqsimoti

Test topshirig'ining sifatini tahlil etishda test topshiruvchilar **javoblarining distraktorlar** (chalg'ituvchi javoblar) **bo'yicha taqsimotini** ham hisobga olish lozim. Bunda quyidagi ko'rsatkichlarga e'tibor berish kerak:

1. Distraktorlarning **tanlanishi** test topshirig‘iga noto‘g‘ri javob berganlar orasida **taxminan teng** bo‘lishi zarur.

Misol uchun, test topshirig‘ining qiyinlik darajasi 0,60 ga teng, ya’ni test topshiruvchilarning 60 foizi to‘g‘ri javob berishdi. Agar bu topshiriqda muqobil javoblar soni 4 ta bo‘lsa, demak, distraktorlar (chalg‘ituvchi javoblar) soni 3 taga teng. Test topshirig‘iga noto‘g‘ri javob bergan 40% test topshiruvchilar distraktorlar orasida taxminan teng taqsimlanishi shart:

A (distraktor) – 17% test topshiruvchilar tanladi

B (kalit) – 60% test topshiruvchilar tanladi

C (distraktor) – 12% test topshiruvchilar tanladi

D (distraktor) – 11% test topshiruvchilar tanladi

Yuqoridagi misolda barcha chalg‘ituvchi javoblar o‘z vazifasini bajaryapti – tekshirilayotgan domenni yaxshi o‘zlashtirmagan test topshiruvchini chalg‘ityapti. Agar biron-bir distraktorni test topshirig‘iga noto‘g‘ri javob berganlarning 5 foizdan kami tanlagan bo‘lsa, bu distraktor o‘z vazifasini bajarmayotgan hisoblanadi va uni o‘zgartirish darkor¹.

2. Kalitning (to‘g‘ri javobning) diskriminatsiya indeksi **musbat** bo‘lishi kerak, ya’ni yuqori guruhdagi test topshiruvchilar (test natijalariga ko‘ra eng ko‘p ball olganlar) kalitni ko‘proq tanlashi lozim. **Distraktorlarning** (noto‘g‘ri javoblarning) diskriminatsiya indeksi esa **manfiy** bo‘lishi zarur, ya’ni quyì guruhdagi test topshiruvchilar (test natijalariga

¹ Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения. Изд. 5-е, пер. и доп. – Москва: Академия, 2013. – С.194.

ko‘ra eng past ball olganlar) distraktorlarni ko‘proq tanlashi darkor.

Misol uchun, quyidagi vaziyatni ko‘raylik. Deylik, 30 kishidan test olindi. Test topshiruvchilarni olgan ballari ketma-ketligida joylashtirib, 15 kishini yuqori guruhga, 15 kishini quyi guruhga ajratdik. Birinchi test topshirig‘iga test topshiruvchilarning javobi quyidagicha taqsimlandi:

A (distraktor) – yuqori guruhdan 3 kishi, quyi guruhdan 0 kishi tanladi.

B (kalit) – yuqori guruhdan 4 kishi, quyi guruhdan 10 kishi tanladi.

C (distraktor) – yuqori guruhdan 3 kishi, quyi guruhdan 2 kishi tanladi.

D (distraktor) – yuqori guruhdan 5 kishi, quyi guruhdan 3 kishi tanladi.

Ushbu test topshirig‘ida kalitni quyi guruhdagi test topshiruvchilar, distraktorlarni esa yuqori guruhdagi test topshiruvchilar ko‘proq tanlashgan. Test topshirig‘ining qiyinlik darajasi $14/30 = 0,47$ ga teng, diskriminatsiya indeksi esa $(4 - 10) / 15 = -0,40$ ga teng. Demak, test topshirig‘i sifatsiz tuzilgan va uni testda ishlatib bo‘lmaydi.

Test topshirig‘iga javoblarning distraktorlar bo‘yicha taqsimotini tahlil etish test topshirig‘i bilan bog‘liq boshqa muammolarni ham aniqlashda yordam beradi¹:

¹ Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.231.

1. Yuqorida aytilganidek, test topshirig'ida to'g'ri javobni (kalitni) tanlaganlar ulushi istalgan distraktorni (chalg'ituvchi javobni) tanlaganlar ulushidan ko'proq bo'lishi kerak. Agar yuqori guruhga kiruvchi test topshiruvchilarning ko'pchiligi to'g'ri javobni (kalitni) emas, distraktorni (chalg'ituvchi javobni) tanlagan bo'lsa, bu test topshirig'ining **kaliti noto'g'ri kiritilgan** bo'lishi mumkin. Misol uchun, tizimda test topshirig'ining kaliti B deb kiritilgan bo'lsa, yuqori guruhga kiruvchi test topshiruvchilarning ko'pchiligi D javob variantini tanlagan bo'lsa, kalitning haqiqatda D emas, B ekanligini yana bir bor tekshirib ishonch hosil qilish lozim.

2. Agar yuqori guruhga kiruvchi test topshiruvchilarning javoblari barcha muqobil javoblar (kalit va distraktorlar) bo'yicha taxminan teng taqsimlangan bo'lsa, bu test topshirig'iga bilimli test topshiruvchilar ham bilib emas, **tavakkal qilib** javob berayotganini bildiradi. Buning sabablari har xil bo'lishi mumkin, masalan, test topshirig'ining qiyinlik darajasi juda yuqori bo'lishi, test topshirig'i tekshirilayotgan domenga kirmaydigan material asosida tuzilgan bo'lishi yoki test topshirig'i o'ta xususiy ma'lumotni bilishni tekshirayotgan bo'lishi mumkin. Bunday test topshirig'ini qayta ishlash yoki testdan chiqarib tashlash talab etiladi.

3. Agar yuqori guruhga kiruvchi test topshiruvchilarning javoblari taqsimotida kalit va bir distraktorning ulushi teng yoki taxminan teng bo'lsa (distraktorlardan birini tanlaganlarning ulushi kalitni tanlaganlarning ulushiga teng yoki

juda yaqin bo'lsa), test topshirig'i yoki uning javobi **bahsli** bo'lishi mumkin. Bunday test topshirig'ini diqqat bilan o'rganib, tahlil va tahrir qilish yoki testdan chiqarib tashlash lozim bo'ladi.

5.2.4. Test topshirig'ining DIF ko'rsatkichi

Test topshirig'ining **DIF ko'rsatkichi** (inglizcha: «*differential item functioning*» – «test topshirig'ining farqlovchi funksiyasi») – test topshiruvchilarning ayrim guruhlarini natijalarining tekshirilayotgan konstruktga bog'liq bo'lmagan holda boshqalarnikidan yaqqol ajralib turishi (inglizcha «*bias*»). Masalan, test topshiruvchilarga matematikadan test berilsa va testga kiritilgan topshiriqlardan biri futbolga oid tushunchalar asosida qurilgan bo'lsa, bunday topshiriqqa test topshiruvchilar orasida qizlarga nisbatan o'g'il bolalar ko'proq to'g'ri javob berishi ehtimoli yuqori, chunki aksar qizlar futbolga oid tushunchalarni tushunmasliklari ularning natijalariga ta'sir qilishi mumkin. Test topshirig'ida bunday tekshirilayotgan konstruktga bog'liq bo'lmagan faktorlarning mavjud bo'lishi test asosida chiqariladigan xulosalarning validligiga salbiy ta'sir qiladi. Shuning uchun ham test topshirig'ida DIF ko'rsatkichi 0 ga teng bo'lishi kerak, ya'ni tekshirilayotgan konstruktga bog'liq bo'lmagan faktorlar test topshiruvchilarning biron-bir guruhi natijalariga aslo ta'sir qilmasligi shart.

Misol uchun, o‘qish savodxonligini tekshiradigan test oldik. Test natijalarini tahlil qilganimizda test topshiruvchilar orasida qizlar o‘g‘il bolalarga nisbatan, yaxshiroq natijalar qayd etishganini ko‘rdik. Agar bir xil ball olgan qiz bola test topshiruvchilar va o‘g‘il bola test topshiruvchilarning qaysidir bir test topshirig‘iga bergan javobini tahlil qilsak va aynan shu test topshirig‘ida umumiy holatga teskari holatni kuzatsak (test topshiruvchilar ichida o‘g‘il bolalar qizlarga nisbatan ko‘proq to‘g‘ri javob bergan bo‘lsa), bu hol mazkur topshiriqda DIF ko‘rsatkichi 0 dan yuqori ekanligini ko‘rsatadi.

Shu o‘rinda «DIF va «ta’sir» (inglizcha: «impact») tushunchalarining farqiga aniqlik kiritish lozim. Ta’sir – bu test topshiruvchilarning test orqali o‘lchanayotgan konstruktning egalashda o‘ziga xos farqlari. Masalan, odatda, SAT Matematika testlarida kelib chiqishi osiyolik bo‘lgan test topshiruvchilar boshqalarga nisbatan yuqori natijalarga erishadilar. Ya’ni **ta’sir** – bu test topshiruvchilarning turli guruhlarida orasidagi **barqaror** farqlar. **DIF** esa test topshiruvchilar turli guruhlarining test topshirig‘iga bergan javoblaridagi **kutilmagan** farqlardir¹.

Albatta, test topshiruvchilarning turli guruhlarida natijasi orasida farqning **mavjudligi** testda DIF mavjudligini anglat-

¹ Dorans N.J. and Holland P.W. DIF Detection and Description: Mantel-Haenszel and Standardization, ETS Research Report RR-92-10. – Princeton, NJ: Educational Testing Service, 1992. – P.5.

maydi. Misol uchun, boy oilalardan chiqqan test topshiruvchilar guruhi SAT testida kambag'al oilalardan chiqqan test topshiruvchilar guruhiga nisbatan yuqoriroq natijalar qayd etishi boy oilalarda farzandlarini testga yaxshiroq tayyorlash uchun imkoniyatlar ko'pligi bilan izohlanishi mumkin¹. Shuning uchun ham bunday farqlarning sabablari tahlil etilishi va testda DIF yo'qligiga ishonch hosil qilinishi lozim.

DIF ko'rsatkichini hisoblashning oson yo'llaridan biri quyidagicha²:

1. Test topshirganlarni qaysidir jihati (masalan, jinsi, millati, ijtimoiy kelib chiqishi va hokazo) bo'yicha guruhlariga bo'lamiz.

2. Guruhlarning testdagi umumiy natijalarini solishtiramiz. Misol uchun, bir guruhning testda qayd etgan umumiy natijasi ikkinchi guruhning testda qayd etgan umumiy natijasiga nisbatan 15% yuqori bo'ldi.

3. Guruhlarning test topshirig'ini bajarishdagi natijalarini solishtiramiz. Bu natijalar orasidagi farq guruhlarning testdagi testda aniqlangan umumiy natijalari orasidagi farqqa yaqin bo'lishi lozim. Yuqoridagi misolimizda bir guruhning testda aniqlangan umumiy natijasi ikkinchi guruhning testda aniqlangan umumiy natijasidan 15% yuqori bo'lganligi uchun ko'rib chiqilayotgan test topshirig'ida ham shunga ya-

¹ Koretz D.M. *Measuring Up: What Educational Testing Really Tells Us.* – Cambridge, MA: Harvard University Press, 2008. – P.264.

² Popham W.J. *Modern Educational Measurement: Practical Guidelines for Educational Leaders.* 3rd ed. – Boston: Allyn and Bacon, 2000. – P.156.

qin farqni kuzatishimiz kerak. Aks holda, bu topshiriqni tahlil qilib qayta ishlash yoki undan voz kechish lozim.

Yuqori ta'sirga ega testlardagi test topshiriqlari uchun DIF ko'rsatkichini hisoblashning matematik modellari ishlatiladi. Shunday modellardan biri **Mantel-Haenszel (MH) protsedurasi** deb ataladi¹. Bu protsedura bo'yicha:

1. Test topshiruvchilarni qaysidir jihati (masalan, jinsi, millati, ijtimoiy kelib chiqishi va hokazo) bo'yicha ikki guruhga bo'lamiz.

2. Har ikkala guruhning test topshirig'iga to'g'ri javob berishi ehtimolini hisoblaymiz. Buning uchun:

– oldin birinchi guruhda test topshirig'iga to'g'ri javob berganlar sonini shu guruhdagi test topshiruvchilarning umumiy soniga bo'lamiz, so'ng test topshirig'iga noto'g'ri javob berganlar sonini shu guruhdagi test topshiruvchilarning umumiy soniga bo'lamiz, so'ng birinchi raqamni ikkinchi raqamga bo'lamiz;

– xuddi shu amallarni ikkinchi guruh uchun bajaramiz.

3. Birinchi guruhning test topshirig'iga to'g'ri javob berishi ehtimolini ikkinchi guruhning test topshirig'iga to'g'ri javob berishi ehtimoliga bo'lamiz.

4. Yuqoridagi amallarni test topshiruvchilarning turli jihati bo'yicha guruhlari uchun bajaramiz. Natijalarning o'r-

¹ Karami H. An Introduction to Differential Item Functioning // The International Journal of Educational and Psychological Assessment. September 2012. Vol. 11 (2). – P.66.

tacha arifmetik qiymatini hisoblaymiz. Bu bilan α_{MH} ni (Mantel-Haenszel bo'yicha alfani) hisoblaymiz.

5. Test topshirig'ining Mantel-Haenszel bo'yicha DIF ko'rsatkichi MH_{DIF} ni quyidagi formula bo'yicha hisoblaymiz:

$$MH_{DIF} = -2,35 \ln(\alpha_{MH})$$

Hosil bo'lgan natija 0 ga teng yoki 0 ga yaqin bo'lsa, test topshirig'ida DIF funksiyasi (tekshirilayotgan konstruktga bog'liq bo'lmagan holda farqlash funksiyasi) mavjud emas deb olishimiz mumkin.

Masalan, yirik test topshiruvchilar auditoriyasida test o'tkazdik. Bu auditoriyani turli jihati bo'yicha guruhlariga bo'lish mumkin: jinsi, millati, ijtimoiy kelib chiqishi va hokazo. Bir jihatini (masalan, millatini) asos qilib olib, test topshiruvchilarni ikki guruhga ajratamiz. Birinchi guruhdagi 20 ta test topshiruvchidan 14 tasi birinchi test topshirig'iga to'g'ri javob berdi, 6 tasi noto'g'ri javob berdi. Xuddi shu test topshirig'iga ikkinchi guruhdagi 20 ta test topshiruvchidan 8 tasi to'g'ri javob berdi, 12 tasi noto'g'ri javob berdi.

Birinchi guruhning test topshirig'iga to'g'ri javob berishi ehtimolini hisoblaymiz: $(14/20) / (6/20) = 0,7 / 0,3 = 2,33$.

Ikkinchi guruhning test topshirig'iga to'g'ri javob berishi ehtimolini hisoblaymiz: $(8/20) / (12/20) = 0,4 / 0,6 = 0,66$.

Birinchi raqamni ikkinchi raqamga bo'lamiz: $2,33 / 0,66 = 3,53$.

Demak, birinchi guruhimizning test topshirig'iga to'g'ri javob berishi ehtimoli ikkinchi guruhga nisbatan 3,53 baravar ko'p. Test topshiruvchilarni boshqa jihatlari bo'yicha guruhlab, har bir holatda yuqoridagi amallarni bajarashimiz kerak bo'ladi. Oxirida Mantel-Haenszel bo'yicha alfani va u asosida Mantel-Haenszel bo'yicha DIF ko'rsatkichini hisoblash lozim.

Test topshirig'ining DIF ko'rsatkichini hisoblashning matematik modellari qo'llanilganda maxsus dasturiy ta'minotdan (masalan, SPSS) foydalanish maqsadga muvofiqdir.

Asosiy xulosa:

Test topshiriqlarining asosiy sifat ko'rsatkichlariga quyidagilar kiradi: test topshirig'ining qiyinlik darajasi, test topshirig'ining diskriminatsiya indeksi, test topshiruvchilar javoblarining distraktorlar bo'yicha taqsimoti va test topshirig'ining DIF ko'rsatkichi. Test topshiriqlari aprobat-siyasidan keyin ushbu ko'rsatkichlarni hisoblab, tahlil etish va talab darajasida ekanligiga ishonch hosil qilish lozim.

5.3. Mezonga mo'ljallangan testlarda test topshiriqlari tahlili

Yuqorida ko'rib chiqilgan test topshirig'ining statistik tahlili usullari normaga mo'ljallangan testlar uchun qo'llaniladi. Normaga mo'ljallangan testlar samarali bo'lishi uchun ballar taqsimoti xilma-xil bo'lishi talab etiladi. **Mezonga mo'ljallangan testlarda** esa (agar ta'lim jarayoni samarali tashkil qilingan bo'lsa) test topshiriqlari test topshiruvchilar uchun qiyin bo'lmasligi kutiladi, buning natijasida ballar taqsimoti ham katta bo'lmasligi kerak. Shuning uchun mezonga mo'ljallangan testlarda test topshiriqlarining sifatini tahlil etishda boshqacha usullar qo'llaniladi. Bunda ikki xil yondashuv mavjud.

Birinchi yondashuv: Test aynan bir guruhda ikki marta – test orqali tekshirilayotgan domenni o'qitishdan (konstrukt-ni shakllantirishdan) **oldin** va undan **keyin** o'tkaziladi. Ideal holatda birinchi marta test o'tkazilganda ko'pchilik test topshiruvchilar past natija ko'rsatishi, ikkinchi marta o'tkazilganda esa natija yuqori bo'lishi kerak. Birinchi martada aksar test topshiriqlarining qiyinlik darajasi (p qiymati) yuqori, ikkinchi martada esa past bo'lishi kutiladi. Test topshirig'ining **diskriminatsiya indeksi** quyidagi formula bo'yicha hisoblanadi:

$$D = -\rho_{post} - \rho_{pre}$$

bunda: D = test topshirig'ining diskriminatsiya indeksi;

ρ_{post} = test topshirig‘ining domen o‘qitilishidan keyingi p qiymati;

ρ_{pre} = test topshirig‘ining domen o‘qitilishidan oldingi p qiymati.

Misol uchun, ayni bir testni bir guruhda ikki marta (domen o‘qitilishidan oldin va keyin) o‘tkazdik. Birinchi marta test olganimizda birinchi test topshirig‘iga 25 nafar test topshiruvchidan 6 nafari to‘g‘ri javob berdi. Demak, $\rho_{pre} = 6 / 25 = 0,24$. Ikkinchi marta test olganimizda xuddi shu test topshirig‘iga 21 kishi to‘g‘ri javob berdi. Demak, $\rho_{post} = 21 / 25 = 0,84$. Bu test topshirig‘ining diskriminatsiya indeksi $0,84 - 0,24 = 0,60$.

Test topshirig‘ining diskriminatsiya indeksi -1 dan $+1$ gacha qiymatda o‘lchanadi va sifatli test topshirig‘ining diskriminatsiya indeksi $+1$ ga yaqin (hech bo‘lmaganda $0,40$ dan yuqori) bo‘lishi tavsiya qilinadi¹.

Test topshiruvchilar javoblarining distraktorlar bo‘yicha taqsimotini tahlil etishda ham ikki marta o‘tkazilgan test natijalari asos qilib olinadi. **Kalitning** (to‘g‘ri javobning) diskriminatsiya indeksi **musbat** bo‘lishi kerak, ya‘ni ikkinchi marta (o‘qitishdan keyin) test o‘tkazilganda kalitni tanlaganlar soni ko‘proq bo‘lishi zarur. **Distraktorlarning** (noto‘g‘ri javoblarning) diskriminatsiya indeksi esa **manfiy**

¹ Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley, 2013. – P.235.

bo'lishi lozim, ya'ni birinchi marta (o'qitishdan oldin) test o'tkazilganda distraktorlarni tanlaganlar soni ko'proq bo'lishi darkor.

Misol uchun, testni birinchi marta o'tkazganimizda birinchi test topshirig'iga javoblar quyidagicha taqsimlandi (test topshiruvchilar soni 25 ga teng):

A (distraktor): 9 kishi tanladi

B (distraktor): 7 kishi tanladi

C (distraktor): 3 kishi tanladi

D (kalit): 6 kishi tanladi

Ikkinchi marta test o'tkazganimizda xuddi shu test topshirig'iga javoblar quyidagicha taqsimlandi:

A (distraktor): 1 kishi tanladi B (distraktor): 1 kishi tanladi

C (distraktor): 2 kishi tanladi D (kalit): 21 kishi tanladi

Ko'rib turganimizdek, birinchi marta test o'tkazilganda distraktorlarni test topshiruvchilarning ko'pchiligi tanlagan, ikkinchi marta test o'tkazilganda esa kalitni test topshiruvchilarning ko'pchiligi tanlagan. Bu test topshirig'ining sifatli ekanligini ko'rsatadi.

Shuningdek, test topshiriqlarini tahlil etishda testni birinchi marta topshirganda to'g'ri javob berganlar foizi bilan testni ikkinchi marta topshirganda to'g'ri javob berganlar foizini **solishtirish** mumkin. Bunda birinchi holat bilan ikkinchi holat orasidagi **farq** hisoblanadi. Bu farq **musbat** bo'lishi kerak. Agar farq bo'lmasa yoki manfiy bo'lsa, demak, test topshirig'ining sifatida muammo bo'lishi mumkin.

Misol uchun, quyidagi holatni ko‘rib chiqaylik:

Test topshirig‘i	Testni birinchi marta topshirganda to‘g‘ri javob berganlar foizi	Testni ikkinchi marta topshirganda to‘g‘ri javob berganlar foizi	Farq
1	16	79	+63
2	10	82	+72
3	75	75	0
4	27	91	+64
5	67	53	-14

Ushbu holatda uchinchi va beshinchi test topshiriqlari qayta ko‘rib chiqilishi kerak.

Test topshiriqlarini tahlil etishdagi yana bir usul birinchi va ikkinchi marta test o‘tkazilganda test topshiruvchilar **to‘g‘ri javob bergan test topshiriqlari foizini** tahlil etishga asoslanadi. Buning uchun:

1. Har bir test topshiruvchining nechta test topshirig‘iga birinchi marta noto‘g‘ri javob berib, ikkinchi marta to‘g‘ri javob bergani hisoblanadi.

2. Barcha bir test topshiruvchilarning nechta test topshirig‘iga birinchi marta noto‘g‘ri javob berib, ikkinchi marta to‘g‘ri javob bergani yig‘indisi olinadi va bu raqam test topshiruvchilar soniga bo‘linadi.

3. Hosil bo‘lgan raqam testdagi jami test topshiriqlari soniga bo‘linadi.

4. Natija 100 ga ko'paytiriladi.

Test topshiruvchilar birinchi marta noto'g'ri javob berib, ikkinchi marta to'g'ri javob bergan test topshiriqlari ulushi qancha yuqori bo'lsa, test shunchalik sifatli hisoblanadi.

Misol uchun, besh kishidan iborat test topshiruvchilar guruhida ayni bir test ikki marta (o'qitishdan oldin va o'qitishdan keyin) olindi. Har bir test topshiruvchining nechta test topshirig'iga birinchi marta noto'g'ri javob berib, ikkinchi marta to'g'ri javob berganini hisoblaymiz:

Test topshiruvchi A: 18 Test topshiruvchi D: 20

Test topshiruvchi B: 15 Test topshiruvchi E: 13

Test topshiruvchi C: 22

Sonlarni qo'shamiz va test topshiruvchilar soniga bo'lamiz:
 $(18 + 15 + 22 + 20 + 23) / 5 = 88 / 5 = 17,6$.

Hosil bo'lgan sonni test topshiriqlari soniga bo'lamiz va 100 ga ko'paytiramiz:

$$(17,6 / 25) \times 100 = 0,70 \times 100 = 70.$$

Demak, testimizdagi 70% test topshiriqlariga test topshiruvchilar birinchi marta noto'g'ri javob berib, ikkinchi marta to'g'ri javob berishgan.

Ko'rib turganimizdek, bu yondashuv ayni bir testni ayni bir guruhda ikki marta o'tkazishni talab qiladi. Bunda ayrim muammolar yuzaga kelishi mumkin. Masalan, agar testlar orasidagi vaqt qisqa bo'lsa, test topshiruvchilar test topshiriqlarini eslab qolishi, bu ularning ikkinchi marta test topshirgandagi natijalariga ta'sir qilishi mumkin. Shuningdek, o'qitish-

dan oldin test o'tkazilishi o'qitish jarayoniga ta'sir qilishi – test topshiruvchilar o'qitish jarayonida aynan nimalarga e'tibor berish kerakligini bilib olishlari mumkin. Bunday muammolarning oldini olish uchun **ikkinchi yondashuv** qo'llanilishi mumkin. Bu yondashuvning yuqorida tasvirlangan birinchi yondashuvdan farqi – test ayni bir guruhda ikki marta o'tkazilmaydi, balki **ikkita har xil guruhda** o'tkaziladi. Birinchi guruh sifatida test orqali tekshirilayotgan domen o'qitilmagan (konstrukt shakllantirilmagan) guruh va ikkinchi guruh sifatida domen o'qitilgan (konstrukt shakllantirilgan) guruh olinadi. Hisob-kitoblar esa yuqorida tasvirlangan shakllarda amalga oshiriladi. Shuni hisobga olish lozimki, tanlangan har ikkala guruh intellektual qobiliyati va boshqa xususiyatlari jihatidan bir xil bo'lishi talab qilinadi.

Asosiy xulosa:

Mezonga asoslangan test topshiriqlarining sifatini tahlil etishda test topshiruvchilarning test orqali tekshiriladigan domen o'qitilishidan oldingi va keyingi natijalari asos qilib olinadi. Ikkinchi holatda (o'qitishdan keyin) test topshirig'iga test topshiruvchilarning ko'pchiligi to'g'ri javob berishi lozim.

6. TESTNI O‘TKAZISH VA UNING NATIJALARINI QAYTA ISHLASH. O‘TISH BALLARINI BELGILASH

6.1. Testni o‘tkazish

Test topshiriqlari yaratilib, tahlil etilib, eng maqbullari asosida test shakllantirilgandan keyingi muhim jarayon **testni o‘tkazish** va **uning natijalarini tahlil qilishdir**. Testni shakllantirish, o‘tkazish va natijalarini tahlil qilishda test natijalaridan chiqariladigan xulosalarning **validligi** va testning **ishonchliligi** bilan bir qatorda **adolat** va **shaffoflik** tamoyillariga ham amal qilinishi lozim. Aslida, adolat va shaffoflik tamoyillariga rioya etilishi ham test natijalari asosida chiqariladigan xulosalarning validligiga hissa qo‘shadi.

Adolat – testning mazmuni, konteksti (o‘tkazilish shart-sharoitlari) va testdan kutilayotgan natijalar barcha test topshiruvchilarga o‘z qobiliyatlarini birdek namoyon etishlariga sharoit yaratib berishi.

Testda adolat tamoyili ikki muhim **qoidaga** asoslanadi¹:

– test tizimining o‘zida test topshiruvchilarga nisbatan **nohaqlik** bo‘lmasligi kerak;

¹ Isaacs T. et al. Key Concepts in Educational Assessment. – London: SAGE, 2013. – P.57.

– test topshiruvchilarda mavjud test tizimlaridan foydalanish uchun **teng imkoniyatlar** bo‘lishi lozim.

Test tizimining o‘zida test topshiruvchilarga nisbatan **nohaqlik** bo‘lmasligi uchun:

– test topshiriqlarining **DIF ko‘rsatkichlari** tahlil etiladi va test topshiruvchilarning ayrim guruhleri natijalariga konstruktga bog‘liq bo‘lmagan omillar ta’sir ko‘rsatmayotganligiga (barcha test topshiruvchilarning natijalari faqat tekshirilayotgan konstruktga bog‘liq ekanligiga) ishonch hosil qilinadi;

– testni o‘tkazish jarayonida barcha test topshiruvchilarga **bir xilda qulay shart-sharoitlar** yaratilishi lozim va bu shart-sharoitlar konstruktga bog‘liq bo‘lmagan o‘zgaruvchilarning (masalan, shovqin, xona harorati, test topshiruvchilarning tashqaridan yordam olishi yoki ko‘chirishi va hokazo) test natijalariga ta’sir qilmasligini ta’minlashi kerak – buning uchun test topshirish jarayoni bilan bog‘liq barcha qonun-qoidalar **standartlashtirilishi**, test jarayonini o‘tkazish va nazorat qilishga jalb etilganlar esa bu qonun-qoidalar bilan tanish bo‘lishi hamda ularga rioya qilishi talab etiladi¹. Test natijalarini hisoblash va e’lon qilish jarayonida barcha test topshiruvchilarga bir xil yondashuv qo‘llanilishi, baholash mezonlarining barcha uchun bir xilda tatbiq qilinishi ham adolat tamoyilini ta’minlash uchun muhim hisoblanadi.

¹ American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. – Washington D.C.: American Educational Research Association, 2014. – P.114.

Test topshiruvchilarda mavjud test tizimlaridan foydalanishda **teng imkoniyatlar** bo'lishi uchun test topshiruvchining individual xususiyatlari (masalan, jinsi, ijtimoiy kelib chiqishi, jismoniy imkoniyatlari) uning test topshirishiga ta'sir o'tkazmasligi kerak. Masalan, ko'rish imkoniyatlari cheklangan test topshiruvchilar uchun katta harflar bilan yoki Brayl alifbosida chop etilgan test kitoblaridan foydalanish imkoni yoki jismoniy harakati cheklangan test topshiruvchilar uchun test o'tkaziladigan binolarga kirish imkoniyatlari yaratilishi lozim. Bunday holatlar uchun kimga, qay tartibda murojaat qilinishi, qanday asoslarga ko'ra test topshiruvchilarga ularning ehtiyojlarini hisobga olgan holda sharoit yaratilishi haqidagi ma'lumotlar oldindan e'lon qilinishi zarur¹.

Test bilan bog'liq yana bir muhim tamoyil shaffoflikdir.

Shaffoflik – test jarayonining barcha qatnashchilari (test topshiruvchilar, testni o'tkazishga mas'ullar, testni baholaydiganlar, test natijalarini ishlatadiganlar) testda nima talab qilinishi va test topshiruvchilarning ishlari qanday baholanishi haqida yetarlicha ma'lumotga ega bo'lishi.

Shaffoflik ta'minlanishi uchun test qanday domen va konstruktarni tekshirishi, bu tekshirish qay tarzda amalga

¹ American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. – Washington D.C.: American Educational Research Association, 2014. – P.115.

oshishi, test topshiruvchilar qanday mezonlar asosida baholanishi haqidagi ma'lumotlar oldindan e'lon qilinishi va bu ma'lumot bilan tanishish uchun barchada imkoniyat bo'lishi shart. Odatda, testlarni (ayniqsa, yuqori ta'sir kuchiga ega testlarni) tayyorlashga mas'ul tashkilotlar bunday ma'lumotlarni o'z rasmiy saytlarida e'lon qiladilar va doimiy ravishda yangilab boradilar. Shuningdek, test kunida ham test topshiruvchilarga testni topshirish jarayoni bilan bog'liq yetarlicha ma'lumot berilishi lozim¹. Test topshiruvchilarga beriladigan test materiallari quyidagilarni o'z ichiga olishi kerak²:

- testning maqsadi;
- test uchun ajratilgan vaqt – agar test bir necha qismlardan iborat bo'lsa, testning har bir qismi uchun ajratilgan vaqt, shuningdek, ochiq test topshiriqlari uchun sarflanishi lozim bo'lgan vaqt ham ko'rsatilishi maqsadga muvofiq;
- test topshiriqlariga qanday javob berish va javoblarni qayerga qayd qilish bo'yicha aniq yo'riqnomalar (masalan, javoblarni alohida javoblar varaqasiga ko'chirish lozim bo'lsa, bu haqda aniq ma'lumot berilishi lozim) berilishi;

¹ American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. – Washington D.C.: American Educational Research Association, 2014. – P.115.

² Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000. – P.359.

– agar turli test topshiriqlari turlicha baholanadigan bo‘lsa (masalan, qaysidir test topshiriqlariga boshqalariga nisbatan ko‘proq ball beriladigan bo‘lsa), bu ham ko‘rsatilishi darkor.

Testni shakllantirish, o‘tkazish va natijalarini tahlil etishda **amaliy** masalalarni – vaqt va resurslar bilan bog‘liq masalalarni hisobga olish ham lozim. Bu masalalar quyidagilarni o‘z ichiga oladi¹:

– testni o‘tkazish bilan bog‘liq barcha masalalar (masalan, test o‘tkaziladigan joy, vaqt, test o‘tkazish qoidalari) hal qilinganmi?

– test uchun ajratilgan vaqt ichida test topshiruvchilar barcha test topshiriqlariga javob bera olishadimi?

– testni o‘tkazish jarayonida muammo tug‘dirishi mumkin bo‘lgan to‘siqlar bartaraf etilganmi?

– test uchun barcha resurslar (test materiallari, lozim bo‘lgan texnika va boshqalar) tayyormi?

– test bilan bog‘liq xarajatlar belgilangan byudjetdan chiqib ketmaydimi?

– test topshiruvchilarni baholash va testning natijalarini qayta ishlash haddan tashqari ko‘p vaqt va resurs talab qilmaydimi?

– test natijalarini e‘lon qilish usullari usullari (masalan, rasmiy sayt orqali, elektron pochta orqali va hokazo) va muddatlari aniq belgilanganmi?

¹Brown D. and Abeywickrama P. Language Assessment, Principles and Classroom Practices. White Plains. – NY: Pearson Education, 2010. – P.31.

Ushbu masalalarni hal etishda test natijalari asosida chiqariladigan xulosalarning validligiga alohida e'tibor qaratilishi – qabul qilinadigan har qanday qarorlar **validlikka putur yetkazmasligi** lozim.

Asosiy xulosa:

Testni shakllantirish, o'tkazish va natijalarini tahlil etishda adolat va shaffoflik tamoyillariga rioya qilish lozim. Buning uchun barcha test topshiruvchilarga testda o'z qobiliyatlarini namoyish qilish uchun birdek sharoit yaratish, konstruktga bog'liq bo'lmagan omillarning test topshiruvchilar natijalariga ta'sirini yo'qotish, test jarayonining barcha ishtirokchilari testda nima talab etilishi va test topshiruvchilarning ishlari qanday baholanishi haqida oldindan ma'lumotga ega bo'lishlarini ta'minlash kerak. Shu bilan birga, amaliy (vaqt va resurslar bilan bog'liq) masalalar oldindan hal qilinishi zarur.

6.2. Test natijalarini yagona shkalaga keltirish

Tasavvur qiling, biz ikki shahardagi havo haroratini solishtirib, qaysi shaharda havo issiqroq, qaysisida sovuqroq ekanligini bilmoqchimiz. Qohirada havo harorati 35°C ni (Selsiy shkalasi bo'yicha 35 darajani), Nyu Yorkda esa 50°F ni (Farengeyt shkalasi bo'yicha 50 darajani) tashkil etadi. Agar biz har ikki shahardagi havo haroratini umumiy shkalaga keltirmasak, qayerda issiq va qayerda sovuq ekanligi haqida to'g'ri xulosa chiqara olmaymiz.

Shunga o'xshash, test topshiruvchilarning natijalari ham tahlil va e'lon qilinishidan oldin **yagona shkalaga** keltirilishi lozim. Aks holda, test natijalarini tushunish va test topshiruvchilarni bir-birlari bilan yoki normativ guruhlar bilan solishtirib asosli xulosalar chiqarish imkonsiz bo'ladi. Masalan, adabiyot fanidan test olindi. Aliyev 30 ta test topshirig'idan 15 tasini to'g'ri bajardi. Bu yaxshi natijami yoki yomon? Agar test qiyin bo'lsa va ko'pchilik bu testda 30 ta test topshirig'idan 10 tasini to'g'ri yechgan bo'lsa, bu holda 30 tadan 15 tasi to'g'ri bajarilishi yaxshi natijadir. Agar test oson bo'lsa va ko'pchilik bu testda 30 ta test topshirig'idan 25 tasini to'g'ri yechgan bo'lsa, bu holda 30 tadan 15 tasi to'g'ri bajarilishi yaxshi natija emas.

Demak, test natijalarini talqin qilish uchun bu natijalar normalar bilan (agar test normaga mo'ljallangan bo'lsa) yoki

mezonlar bilan (agar test mezonga mo'ljallangan bo'lsa) solishtirilishi lozim.

Normalar – test topshiruvchilarning maxsus tanlab olingan yirik guruhi (normativ guruh) testni bajarganda qayd etgan natijalari asosida belgilangan ko'rsatkichlar.

Masalan, adabiyot bo'yicha tuzilgan testimizni normativ guruhga yehtirsak va ularning test natijalarini (o'rtacha ballar, standart og'ish va hokazo) normalar sifatida qabul qilsak, keyinchalik boshqa test topshiruvchilarning shu testdan qayd etgan natijalarini normalar bilan solishtirishimiz mumkin.

Normalar quyidagi **talablarga** javob berishi lozim¹:

– test topshiruvchilarning tayyorgarlik darajasi turlicha bo'lgan guruhleri orasidagi **farqlarni** hisobga olishi kerak, masalan, matematika testi uchun belgilangan normalar matematika chuqurlashtirib o'tiladigan maktablar va oddiy maktablar uchun bir xil bo'lishi mumkin emas;

– test topshiruvchilarning **representativ** tanlanmasi asosida olingan bo'lishi zarur, ya'ni normativ guruh xususiyatlari jihatidan (yoshi, tayyorgarlik darajasi va hokazo) shu testni topshiradiganlar guruhi bilan bir xil bo'lishi va kamida 200 – 300 kishidan tashkil topishi talab etiladi;

¹ Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002. – С.385.

– test mo‘ljallangan guruh uchun **mos** bo‘lishi, ya’ni test topshiruvchilarning real kontingentiga asoslanib real sharoitlarda olingan bo‘lishi shart.

Normalar belgilanishi uchun tanlangan normativ guruh **qatlamlashtirilgan** bo‘lishi ham talab qilinadi, ya’ni testni topshiradigan haqiqiy kontingentning turli qatlamlarini (masalan, yoshi, jinsi, ijtimoiy kelib chiqishi bo‘yicha) aks ettirishi kerak. Xususan, biz tuzgan test turli yoshdagilar uchun mo‘ljallangan bo‘lsa, normativ guruh ham turli yoshdagilardan iborat bo‘lishi lozim.

Misol uchun, oliy ta’lim muassasasiga kirish imtihonlarida qo‘llanilishi mo‘ljallangan test uchun normalar belgilamoqchi bo‘lsak, kamida 200 – 300 kishidan iborat normativ guruh tuzamiz. Bu guruhning yoshi, tayyorgarlik darajasi va boshqa jihatlari testimizni topshiradigan abituriyentlarniki bilan bir xil bo‘lishi kerak. Bu guruh haqiqiy imtihon sharoitida testni yechadi. Shundan so‘ng bu guruhning test natijalari tahlil etiladi va bu natijalar asosida normalar belgilanadi.

Normalar belgilanishi bilan bir qatorda test natijalari e‘lon qilinishidan oldin boshqa **omillar** ham hisobga olinishi mumkin. Agar testda bir paytning o‘zida parallel test formalari ishlatilsa, unda test natijalarini bir-biri bilan o‘zaro yoki normalar asosida solishtirish qiyin bo‘ladi. Masalan, Aliyev va Valiyev oliy ta’lim muassasasiga kirish uchun test topshirishdi. Ali-

yev 90 ta test topshirig‘idan 57 tasini to‘g‘ri ishladi, Valiyevda esa 61 ta to‘g‘ri javob bo‘ldi. Yuzaki qaraganda, Valiyev yaxshiroq natija ko‘rsatgan degan xulosaga kelishimiz mumkin. Biroq Valiyevga tushgan test varianti Aliyev yechgan test variantidan qiyinroq bo‘lishi ham mumkin, shu sababli Valiyevning natijasi ball hisobida Aliyevning natijasiga qaraganda pastroq bo‘lsa ham, sifat jihatidan unikidan yaxshiroq bo‘lishi mumkin.

Test turli vaqtlarda o‘tkazilsa va testning xavfsizligini ta‘minlash maqsadida har safar yangi parallel test formalari ishlatilsa, bu ham solishtirishni qiyinlashtiradi. Masalan, Aliyev chet tilidan imtihonni yanvar oyida topshirganda 90 ta test topshirig‘idan 60 tasini to‘g‘ri yechgan edi. U o‘z natijalaridan qoniqmay imtihonni mart oyida qayta topshirdi va bu safar 64 ta test topshirig‘ini to‘g‘ri yechdi. Yuzaki qaraganda, biz uning chet tilidan bilim darajasi yaxshilangan deb hisoblashimiz mumkin, lekin uning yanvar oyida yechgan test formasi (varianti) bilan mart oyida yechgan test formasining qiyinlik darajasi bir xilmi?

Shunday muammolarni bartaraf etish uchun ham test natijalari, odatda, **xom (boshlang‘ich) ballarda** emas, **persentillarda** yoki yagona shkalaga keltirilgan **standart ballarda** e‘lon qilinadi.

■ **Xom (boshlang‘ich) ball** – test topshiruvchining to‘g‘ri yechgan test topshiriqlari asosida hisoblangan ball.

Yuqoridagi misolimizda agar har bir to'g'ri ishlangan test topshirig'iga 1 baldan bersak, Aliyevning xom bali 57 ga, Valiyevniki esa 61 ga teng.

Persentil – test topshiruvchi bilan bir xil yoki undan past ball olganlar necha foizni tashkil etishi.

Test topshiruvchi mansub bo'lgan persentil qanchalik **yuqori** bo'lsa, uning natijasi boshqalarga nisbatan shunchalik **yaxshiroq** bo'ladi. Masalan, Aliyev 67-persentilga kirsam, bu u bilan bir xil yoki undan past ball olganlar 67% ni tashkil etadi deganidir. Valiyev 45-persentilga kirsam, u bilan bir xil yoki undan past ball olganlar 45% ni tashkil etadi.

Test natijalari bo'yicha 50-persentildan yuqori guruhga kiradiganlar testni o'rtachadan yaxshi yechgan, 50-persentildan quyi guruhga kiradiganlar testni o'rtachadan past yechgan hisoblanadi. 25-persentil **birinchi kvartil**, 75-persentil esa **uchinchi kvartil** deb ataladi.

Test topshiruvchining qaysi persentilga mansubligi quyidagi formula bo'yicha hisoblanadi:

$$pr = \frac{n(x \leq X)}{N} \times 100$$

bunda: pr = persentil;

X = test topshiruvchining bali;

$n(x \leq X)$ = bali X ga teng yoki undan past bo'lgan test topshiruvchilar soni;

N = barcha test topshiruvchilar soni.

Misol uchun, 20 kishi test topshirdi. 3 nafar test topshiruvchi 5 ball, 8 nafar test topshiruvchi 4 ball, 6 nafar test topshiruvchi 3 ball va 3 nafar test topshiruvchi 2 ball olishdi. Aliyev 4 ball oldi. Uning qaysi persentilga mansubligini topamiz. Buning uchun oldin test topshiruvchilarning natijalarini kamayish tartibida joylashtiramiz:

5, 5, 5, 4, 4, 4, 4, 4, 4, 4, 3, 3, 3, 3, 3, 3, 2, 2, 2

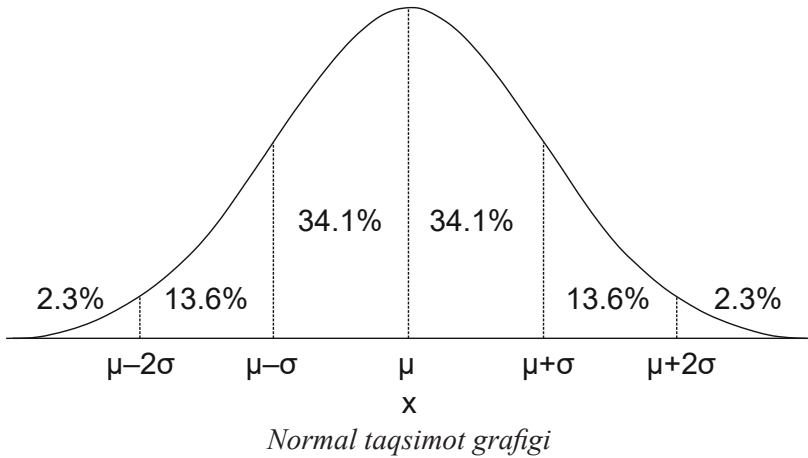
So'ng necha kishi 4 ball yoki undan past natija ko'rsatganini hisoblaymiz: 17. Demak, Aliyev 85-persentilga kiradi: $(17 / 20) \times 100 = 85$. Bu testda Aliyevga teng yoki undan past natija ko'rsatganlar 85 foizni tashkil qiladi deganidir.

Persentillar ayni bir test formasini yechgan test topshiruvchilarni o'zaro solishtirish uchun qulaydir. **Parallel test formalarini** yechgan test topshiruvchilarning natijalarini solishtirish uchun esa xom ballarni standart ballarga aylantirishimiz kerak.

Standart ball – xom balning test natijalari o'rtacha arifmetik qiymatidan qancha standart og'ishga farq qilishini ko'rsatuvchi ball.

Standart balning mohiyatini tushunish uchun normal taqsimot grafigini eslashimiz lozim. Normal taqsimot grafigida test topshiruvchilarning natijalari jami test topshiruvchilarning o'rtacha baliga nisbatan qancha standart og'ishga farq qilishi aks etadi. Standart ballar esa test top-

shiruvchilarning natijalarini o‘rtacha natijalar bilan solishtirish imkonini beradi.



Xom balni standart ball shaklida aks ettirishning bir qancha yo‘llari mavjud.

1) z ball

z ball test topshiruvchi olgan ball test natijalarining o‘rtacha qiymatidan qancha standart og‘ishga farq qilishini ko‘rsatadi. Masalan, agar z ball +1,5 ga teng bo‘lsa, bu test topshiruvchining natijasi o‘rtacha baldan 1,5 standart og‘ishga yuqori ekanligini bildiradi. Agar z ball $-2,5$ ga teng bo‘lsa, test topshiruvchining natijasi o‘rtacha baldan 2,5 standart og‘ishga past demakdir.

Xom ball z balga quyidagi formula bo‘yicha o‘giriladi:

$$z = \frac{X - \bar{X}}{s}$$

bunda: $z = z$ ball;

X = xom ball;

\bar{X} = test natijalari yig'indisining o'rta arifmetik ko'rsatkichi;

s = test natijalari to'plamining standart og'ishi.

Misol uchun, test topshiruvchilar guruhidan test olindi. Guruhning o'rtacha bali (test natijalari yig'indisining o'rta arifmetik ko'rsatkichi) 32,8 ni tashkil qildi. Test natijalari to'plamining standart og'ishi 2,2 ga teng. Agar test topshiruvchi 35 ball olgan bo'lsa, uning z bali +1 ga $((35 - 32,8) / 2,2 = 2,2 / 2,2 = 1)$ teng. Demak, test topshiruvchining natijasi o'rtacha natijadan 1 standart og'ishga yuqori.

Normal taqsimot grafigida test topshiruvchilarning ballari asosan o'rtacha baldan -3 va $+3$ standart og'ishgacha farq qilishi mumkin (maksimal farq -5 va $+5$ standart og'ishgacha). Shu tufayli z ballar ham asosan -3 dan $+3$ gacha qiymatda bo'ladi (maksimal z ballar -5 va $+5$ oralig'ida bo'lishi mumkin). z ball test topshiruvchining normal taqsimot grafigida joylashuvini tasavvur qilish imkonini beradi. Masalan, bali o'rtacha baldan 1 standart og'ishga yuqori bo'lgan test topshiruvchining natijasi jami test topshiruvchilarning 84,2 foizidan yuqori bo'ladi.

Ammo z ballar testologiyadan xabarsiz bo‘lganlar uchun tushunilishi qiyin ekanini aytish joiz. Masalan, mutaxassis bo‘lmagan shaxslar testni yechib, –1 ball olish mumkinligini tasavvur qila olmasligi mumkin. Shuni hisobga olib xom ballni standart ball shaklida aks ettirishning boshqa bir yo‘li – T ballni qo‘llash mumkin.

2) T ball

T ball (inglizcha «*transformed*» – «o‘zgartirilgan» so‘zidan olingan) z ballning o‘zgartirilgan shaklidir: z ball 10 ga ko‘paytiriladi va hosil bo‘lgan raqamga 50 qo‘shiladi:

$$T = 50 + 10z$$

yoki

$$T = 50 + 10 \frac{X - \bar{X}}{s}$$

Demak, T ballar 0 dan 100 gacha oraliqda bo‘lishi mumkin (ko‘pchilik ballar 20 – 80 oralig‘ida bo‘ladi). Bunda test natijalari yig‘indisining o‘rta arifmetik ko‘rsatkichi har doim 50 ga, standart og‘ish har doim 10 ga teng. Misol uchun, agar test topshiruvchining T bali 45 ga teng bo‘lsa, bu uning natijasi test topshiruvchilarning o‘rtacha natijasidan 0,5 standart og‘ishga past deganidir. Agar test topshiruvchining T bali 60 ga teng bo‘lsa, bu uning natijasi test topshiruvchilarning o‘rtacha natijasidan 1 standart og‘ishga yuqori deganidir.

3) Z₁ ball

Z₁ ball ham z ballning o‘zgartirilgan shaklidir. Bunda quyidagi formula qo‘llaniladi:

$$Z_1 = M + sz$$

bunda: $Z_1 = Z_1$ ball;

M = yangi o‘rta arifmetik qiymat;

s = yangi standart og‘ish;

z = z ball.

M va s uchun istalgan qiymatlar qo‘yilishi mumkin. Misol uchun, IQ testlari yangi o‘rta arifmetik qiymat (M) sifatida 100 ni, yangi standart og‘ish (s) sifatida 15 ni ishlatadi. Demak, IQ testlarida 130 ball natija test topshiruvchining IQ qobiliyati o‘rtachadan 2 standart og‘ishga yuqori ekanligini, 85 ball esa o‘rtachadan 1 standart og‘ishga past ekanligini ko‘rsatadi. Xalqaro PISA testlarida va SAT imtihonlarida $M = 500$, $s = 100$, demak, 400 ball o‘rtachadan 1 standart og‘ishga past, 600 ball esa o‘rtachadan 1 standart og‘ishga yuqori bo‘ladi.

z ball hamda uning asosida hisoblanadigan T ball yoki Z_1 ball normal taqsimot qonuni asosida test topshiruvchilarning natijalarini solishtirish imkonini beradi. Biroq test natijalari har doim ham normal taqsimot grafigiga tushmasligi mumkin. Bunday holatda xom ballar **normallashtirilgan standart ballarga** aylantiriladi. Buning uchun:

1) har bir xom ballning qaysi persentilga mansubligi aniqlanadi (persentilni aniqlash formulasi yuqorida keltirilgan edi);

2) maxsus jadval yordamida shu persentilga mos **normallashtirilgan z ball** aniqlanadi (bu jadvalning qisqartirilgan ko‘rinishi quyida keltirilgan, kengaytirilgan shaklini statistika bo‘yicha manbalardan topish mumkin);

3) normallashtirilgan z ball asosida T ball yoki Z_1 ball hisoblanadi (T ball va Z_1 ballni hisoblash formulalari yuqorida keltirilgan edi).

<i>Normal-lashgan z ball</i>	<i>Persentil</i>	<i>Normal-lashgan z ball</i>	<i>Persentil</i>	<i>Normal-lashgan z ball</i>	<i>Persentil</i>
3,0	99,9	1,0	84,1	-1,0	15,9
2,5	99,4	0,5	69,1	-1,5	6,7
2,0	97,7	0	50,0	-2,0	2,3
1,5	93,3	-0,5	30,9	-2,5	0,6
1,4	91,9	-0,7	24,2	-3,0	0,1

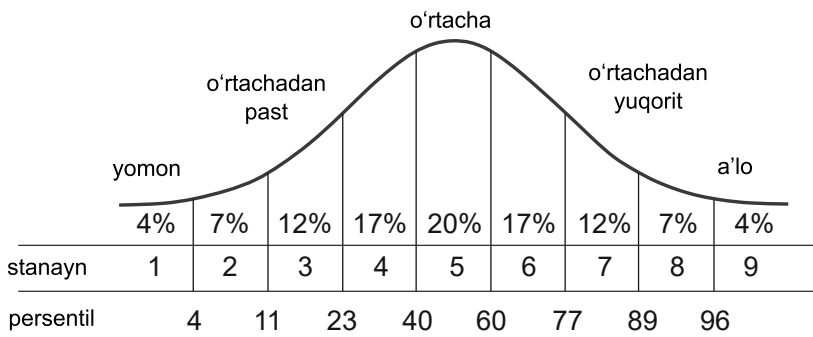
Persentillarning normallashtirilgan z ballarga moslik jadvali (qisqartirilgan)

Masalan, test topshiruvchining xom bali 92-persentilga mansub bo'ldi. Jadvaldan uning normallashtirilgan z bali 1,4 ga tengligini ko'ramiz. Test topshiruvchining T bali 64 ga teng ($50 + 1,4 \times 10 = 64$).

Normallashtirilgan standart ballar **stanayn** shaklida ham ifodalanishi mumkin.

Stanayn (inglizcha «standard» va «nine» – «standart» va «to'qqiz» so'zlaridan olingan) – test ballari taqsimotini 0,5 standart og'ish qadam bilan to'qqiz oraliqqa bo'lish natijasida hosil bo'lgan standart shkala.

Stanayn shaklida ifodalangan ballar 1 dan 9 gacha bo‘lishi mumkin, bunda test natijalari taqsimotining o‘rtacha bali 5 ga, standart og‘ishi taxminan 2 ga teng. Ayrim standartlashtirilgan testlar, masalan, xalqaro ingliz tili test tizimi (IELTS) imtihoni test natijalari 9 ballik tizim asosida e‘lon qilinadi.



Test topshiruvchilarning standart taqsimot grafigida stanaynlar bo‘yicha taqsimlanishi

Misol uchun, test topshiruvchining natijasi 33-persentilga mansub bo‘lsa, 9 ballik tizimda uning bali 4,5 ga teng bo‘ladi.

Test topshiruvchining natijalarini standart ballar shaklida yagona shkalada ifodalashning yuqorida ko‘rib chiqilgan usullari **klassik test nazariyasiga** asoslangan. Bunda testga kiritilgan test topshiriqlariga, ularning qiyinlik darajasidan qat’iy nazar, teng ball beriladi. **Zamonaviy test nazariyasi** («savol-javob» nazariyasi, inglizcha «Item-Response Theory», IRT) asosidagi yagona shkalaga keltirish usullari esa

Guttman shkalasi asosida tuzilgan testlar bo'yicha natijalarni solishtirish imkonini beradi.

Guttman shkalasi – test topshiriqlarining qiyinlik darajasi bo'yicha qat'iy ketma-ketlik asosida joylashuvi, bunda ma'lum bir test topshirig'iga to'g'ri javob bergan test topshiruvchi bundan oldingi barcha test topshiriqlariga to'g'ri javob bera olishi kerak deb hisoblanadi.

Misol uchun, matematikadan test tuzilsa va bu testga ko'paytirish bo'yicha test topshiriqlari quyidagi ketma-ketlikda joylashtirilsa:

1) $17 \times 2 = ?$ 2) $236 \times 12 = ?$ 3) $1437 \times 382 = ?$ 4) $57342 \times 7412 = ?$

To'rtinchi test topshirig'ini to'g'ri bajargan test topshiruvchi undan oldingi topshiriqlarni ham bajara olishi kerak, chunki ko'p xonali sonlarni bir-biriga ko'paytira olish uchun kichik sonlarni bir-biriga ko'paytirishni bilish zarur.

IRT nazariyasi asosida tuzilgan testlar bir o'lchamli bo'lishi, testga kiritilgan test topshiriqlari bir-biri bilan bog'liq bo'lmasligi va kuchli diskriminatsiya xususiyatiga ega bo'lishi, test natijalari normal taqsimot qonuniga mos kelishi talab etiladi.

IRT nazariyasi asosida test natijalarini yagona shkalaga keltirishning bir qancha usullari mavjud. Shulardan biri **Rash modeli** deb ataladi (*Rasch model*, daniyalik olim *Georg*

Rasch ismidan olingan). Bu modelda oldin maxsus formulalar yordamida har bir test topshirig‘ining logit qiymati hisoblana-
di. Logit qiymat, odatda, -3 dan $+3$ gacha qiymatda bo‘ladi, qiyinlik darajasi yuqoriroq test topshiriqlarining logit qiymati yuqoriroq bo‘ladi. Shundan so‘ng matematik modellashtirish yordamida test topshiruvchilarning xom ballari standart ballarga aylantiriladi. Bu model asosan yuqori ta‘sir kuchiga ega bo‘lgan testlarning natijalarini qayta ishlash uchun qo‘llaniladi.

Demak, standart ballar testning parallel formalarini (har xil variantlarini) yechgan test topshiruvchilarning natijalarini o‘zaro va normalar bilan solishtirish imkonini beradi. Masalan, testning har xil variantlarini yechgan Aliyev va Valiyevning xom ballari mos ravishda 57 va 61 bo‘lsa, biz ularning natijalarini bu ballar asosida solishtira olmaymiz. Ularning xom ballari standart ballarga aylantirilgandan keyin ularning natijalarini solishtirish mumkin bo‘ladi.

Asosiy xulosa:

Test topshiruvchining xom bali uning natijalarini boshqa test topshiruvchilarning natijalari bilan solishtirish imkonini bermaydi. Buning uchun test topshiruvchilarning natijalari standart ballarga aylantirilishi lozim.

6.3. O'tish ballarini belgilash texnologiyalari

Testlar o'tkazilib, ularning natijalari e'lon qilingandan keyin bu natijalar asosida ma'lum bir qarorlar qabul qilina-di. Masalan, ta'lim muassasasida o'quv yili oxirida yakuniy test olinsa, bu test natijalari asosida ta'lim oluvchilarni ke-yingi bosqichga (sinfga, kursga) o'tkazish yoki o'tkazmaslik to'g'risida qaror qabul qilinadi. Buning uchun esa o'tish bal-larini belgilash talab etiladi.

O'tish bali – test topshiruvchilarni natijalariga ko'ra ikki yoki undan ortiq kategoriyalarga bo'lish uchun ishlatila-digan ball.

Test topshiruvchilar nechta kategoriyaga ajratilishiga qarab bittadan ortiq o'tish bali ishlatilishi mumkin. Misol uchun, agar test topshiruvchilar ikki kategoriyaga («o'tdi» – «o'tmadi» sin-gari) ajratiladigan bo'lsa, bitta o'tish bali, agar test topshiruv-chilar uch kategoriyaga («yomon» – «o'rtacha» – «yaxshi» sin-gari) ajratiladigan bo'lsa, ikkita o'tish bali kerak bo'ladi.

Test orqali test topshiruvchilar qanday kategoriyalarga bo'linishi va bu kategoriyalarning tavsifi testni rejalashtirish jarayonida testning maqsadlari bilan birga aniqlashtirib olini-shi, validatsiya jarayonida bu kategoriyalarga bo'linish asosli ekanligi haqida dalillar yig'ilishi lozim¹.

¹ Cizek G.J. and Bunch M.B. Standard Setting. Thousand Oaks: SAGE, 2011. – P.13.

Normaga mo'ljallangan testlarda o'tish ballari test topshiruvchilar natijalariga ko'ra joylashtirilgandan keyin test topshiruvchilarning ma'lum bir foizi sifatida belgilanishi mumkin. Misol uchun, ta'lim muassasasiga kirish test sinovida test natijalariga ko'ra eng yuqori natija ko'rsatgan 30 foiz test topshiruvchilarni o'qishga qabul qilishimiz mumkin. Bunday vaziyatda biz test topshiruvchilarning ichidan natijalari **nisbatan** yaxshilarini ajratib olamiz.

Mezonda mo'ljallangan testlarda o'tish ballarini belgilashda test topshiruvchilar ma'lum bir kategoriyaga (masalan, o'tganlar, yaxshilar, o'rtachalar va hokazo) kirishi uchun **mezonlar** (qanday bilim va ko'nikmalarga ega bo'lishi lozimligi) belgilanadi hamda shu mezonlarga mos keluvchi ballar aniqlanadi.

Ayrim hollarda o'tish ballarini belgilashda har ikki yondashuvning elementlari qo'llanilishi mumkin: **mezonlar** asosida eng quyi ball belgilanadi, shu baldan yuqori natija qayd etganlarning ichidan **nisbatan** yaxshilari tanlanadi. O'zbekiston oliy ta'lim muassasalariga kirish test sinovlarida shu usul qo'llaniladi: belgilangan kvota doirasida eng yuqori ballar to'plagan abituriyentlar o'qishga qabul qilinadi, lekin bunda abituriyentning bali belgilangan eng quyi o'tish balidan past bo'lmasligi kerak.

O'tish ballarini belgilashning bir qancha keng tarqalgan usullari mavjud. Qaysi usul qo'llanilishidan qat'iy nazar, bunda quyidagi **bosqichlarga** rioya etilishi lozim¹:

¹ Hambleton R.K., Pitoniak M.J. and Copella J.M. Essential Steps in Setting Performance Standards on Educational Tests and Strategies for Assessing the Reliability of Results // Cizek G.J. (Ed.) Setting Performance Standards: Foundations, Methods, and Innovations. 2nd ed. – London-New York, Routledge, – P.53.

1. O'tish balini belgilash metodi tanlanadi va ekspertlarning birinchi yig'ilishiga tayyorgarlik ko'riladi.

2. Ekspertlar guruhi tanlanadi. Guruh, odatda, 15 – 20 kishidan iborat bo'ladi va bunda turli manfaatdor guruhlarning vakillari (masalan, o'qituvchilar, ta'lim dasturlari mualliflari, ta'limni boshqarish organi vakillari, ish beruvchilar vakillari singari) jalb qilinganligiga, guruh yoshi, jinsi va boshqa xususiyatlariga ko'ra turli kishilardan iborat ekanligiga alohida e'tibor qaratish lozim.

3. Test topshiruvchilarni bo'lish uchun tanlangan kategoriyalar tavsifi tayyorlanadi. Masalan, test topshiruvchilar «o'tdi» – «o'tmadi» kategoriyalariga bo'linsa, o'tganlar va o'tmaganlar orasidagi farqlar nimalarda mamoyon bo'lishi kerak?

4. Ekspertlar guruhi uchun o'tish balini belgilash metodini qo'llash bo'yicha trening o'tkaziladi.

5. Reytinglar (ekspertlarning turli kategoriya uchun bergan ballari) yig'iladi va ularning o'rta arifmetik ko'rsatkichi, medianasi, dispersiyasi singari statistik ko'rsatkichlari tahlil etiladi.

6. Reytinglar ekspertlar bilan muhokama qilinadi va test topshiruvchilarning natijalari bilan solishtiriladi.

7. Reytinglar ekspertlar tomonidan qaytadan belgilanadi va o'tish ballari belgilanadi.

8. O'tish ballarini belgilash texnologiyasiga (jarayonning qanday o'tganligiga) baho beriladi, validlik dalillari yig'iladi.

Agar test yopiq test topshiriqlaridan iborat bo'lsa, o'tish ballarini belgilash uchun, odatda, **Angoff** yoki **Bookmark** usuli, ochiq test topshiriqlaridan iborat testlar uchun, odatda, **Bookmark** usuli qo'llaniladi. Har ikkala usul uchun ham test topshiruvchilarning natijalari kerak bo'ladi,

Angoff usulida (amerikalik olim Angoff /*William Angoff*/ ismidan olingan) ekspertlarga oldin testni o'zlari yechib ko'rishlari taklif etiladi. Shundan so'ng:

– ekspertlar test topshiruvchilarni kategoriyalarga ajratish uchun tanlangan kategoriyalar tavsiflari bilan tanishishadi;

– ekspertlarga «belgilangan kategoriyalar bo'yicha minimum talabga javob beradigan test topshiruvchining shu test topshirig'iga to'g'ri javob berish ehtimoli necha foiz?» degan savol qo'yiladi;

– ekspertlar tomonidan har bir test topshirig'i tahlil qilindi va har bir ekspert yuqorida qo'yilgan savolga o'z javobini yozadi (bunda ekspertlar bir-birlari bilan gaplashishmaydi);

– ekspertlar qo'ygan foizlar yig'ib olinadi va bu foizlarning o'rta arifmetik qiymati yoki medianasi asosida reyting tuziladi, bu reyting har bir kategoriya uchun o'tish balini belgilaydi;

– ekspertlarga o'zlarining test natijalari statistik tahlili va haqiqiy test topshiruvchilarning natijalari statistik tahlili, shuningdek, har bir topshiriq uchun o'zlari oldin belgilagan to'g'ri javob berish ehtimoli foizi beriladi, bu ma'lumotlar ekspertlar orasida muhokama qilib olinadi;

– ekspertlar test topshiriqlarini va o‘zlari oldin belgilagan foizlarini qaytadan tahlil etishadi hamda kerakli o‘rinlarga o‘zgartirish kiritishadi;

– ekspertlar qo‘ygan foizlar yana yig‘ib olinadi va bu foizlarning o‘rta arifmetik qiymati yoki medianasi asosida har bir kategoriya uchun o‘tish ballari belgilanadi.

Bu jarayon test topshiruvchilar ikki («o‘tdi» – «o‘tmadi») yoki undan ortiq kategoriyaga bo‘linadigan testlar uchun qo‘llanilishi mumkin.

O‘tish ballarini belgilash uchun qo‘llaniladigan yana bir usul «**Bookmark**» (inglizcha «*bookmark*» – «*kitob uchun xatcho‘p*» so‘zidan olingan) deb ataladi. Bu usulni qo‘llash uchun, birinchi navbatda, testda qo‘llanilgan test topshiriqlari maxsus tartibda joylashtirilgan test kitobi tuziladi. Test kitobini tuzishda quyidagi qoidalarga rioya qilinadi¹:

– dixotom test topshiriqlari qiyinchiligi o‘tib borish tartibida joylashtiriladi;

– politom test topshiriqlari bir necha o‘rinlarda joylashtiriladi: masalan, test topshirig‘i 0, 1, 2 ballar bilan baholanadigan bo‘lsa, oldin test topshiruvchi bu test topshirig‘i uchun 1 ball olishi mumkin bo‘lgan o‘rinda keltiriladi, so‘ngra shu test topshirig‘i uchun 2 ball olishi mumkin bo‘lgan o‘rinda qaytariladi.

Shundan so‘ng quyidagi jarayonlar amalga oshiriladi:

¹ Lewis D. Bookmark Method // Frey, B. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. Thousand Oaks: SAGE, 2018. – P.216.

1. Ekspertlarga test kitobi beriladi va ular har bir test topshirig'ini, bu test topshirig'i orqali qanday konstrukt o'lchanayotganini, bu test topshirig'iga to'g'ri javob bergan test topshiruvchi qanday bilim va ko'nikmalarga egaligini muhokama qilishadi. Shuningdek, ular test topshiruvchilarni kategoriyalarga (masalan, o'tgan va o'tmaganlarga) bo'lish uchun tanlangan kategoriyalar tavsifi bilan tanishishadi.

2. Har bir ekspert mustaqil ravishda test topshiriqlarini yana bir bor ko'rib chiqib, har bir kategoriyadagi test topshiruvchilarning ma'lum bir foizi (odatda, 67%) to'g'ri javob berishi kerak bo'lgan test topshiriqlari joylashgan o'rinlarni belgilab chiqadi («xatcho'p qo'yib chiqadi»). Bu o'rinlar shunday tanlanadiki, shu test topshirig'idan oldingi barcha test topshiriqlariga ko'rib chiqilayotgan kategoriyaga mansub test topshiruvchilarning 67 foizdan ko'pi to'g'ri javob bera olishi, undan keyingi barcha test topshiriqlariga 67 foizdan kami to'g'ri javob bera olishi kerak.

Misol uchun, test topshiruvchilarni «o'tdi» – «o'tmadi» kategoriyalariga ajratmoqchi bo'lsak va ekspert test kitobidagi beshinchi test topshirig'ini belgilagan bo'lsa, uning fikricha, «o'tmadi» kategoriyasiga kiruvchi test topshiruvchilarning 67 foizdan ko'pi birinchidan to'rtinchigacha test topshiriqlariga to'g'ri javob bera oladi, oltinchi va undan keyingi test topshiriqlariga esa 67 foizdan kami to'g'ri javob bera oladi;

– ekspertlarga haqiqiy test topshiruvchilar natijalarining statistik tahlili beriladi;

– ekspertlar test topshiriqlarini va o‘zlari belgilagan «xatcho‘plar»ni tahlil qilishadi hamda lozim o‘rinlarga o‘zgartirish kiritishadi;

– ekspertlar belgilagan «xatcho‘plar» tahlil qilinadi va o‘rta arifmetik qiymati yoki medianasi asosida har bir kategoriya uchun o‘tish ballari belgilanadi.

Demak, har ikkala ko‘rib chiqilgan usulda ham ekspertlarning fikri, ularning tasniflanayotgan kategoriyalarga mansub test topshiruvchilarni qanday tasavvur etishlari muhim rol o‘ynaydi.

Asosiy xulosa:

Normaga asoslangan testlarda o‘tish balini belgilash uchun testda yuqori natija qayd etgan test topshiruvchilarning ma’lum bir foizi erishgan ballar asos qilib olinishi mumkin. Mezonga mo‘ljallangan testlarda esa test topshiriqlari ekspertlar tomonidan tahlil etilib, har bir kategoriya uchun o‘tish ballari belgilanadi. Ayrim hollarda aralash usul qo‘llanilishi mumkin: test topshiriqlari tahlili asosida minimal o‘tish bali belgilanadi va bu baldan yuqori ball to‘plagan test topshiruvchilarning ichida eng yuqori natija qayd etganlarining ma’lum bir foizi olinadi.

ILOVA

MICROSOFT EXCEL DASTURIDA TEST TOPSHIRIQLARI VA TEST BILAN BOG'LIQ STATISTIK KO'RSATKICHLARNI HISOBLASH NAMUNASI

Test topshiriqlari va test bilan bog'liq statistik tahlil asosan maxsus dasturlarda (masalan, Statistical Package for the Social Sciences – SPSS dasturida) amalga oshiriladi. Shu bilan birga, ko'p statistik ko'rsatkichlarni Microsoft Office tarkibiga kiruvchi Excel dasturi yordamida ham hisoblash mumkin.

Test topshiriqlari va butun testning statistik ko'rsatkichlarini Excel dasturida hisoblash uchun:

1. Jadval yaratamiz va unga **test natijalarini kiritamiz.** Bunda test topshiruvchi test topshirig'iga to'g'ri javob bergan bo'lsa, 1, noto'g'ri javob bergan bo'lsa, 0 raqami kiritiladi.

Test topshiruvchilar	1-test topshirig'i	2-test topshirig'i	3-test topshirig'i	4-test topshirig'i
00001	1	1	0	1
00002	1	0	1	1
00003	1	0	0	1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
5	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Jadvalning Microsoft excel dasturida ko'rinishi. Har bir qator – test topshiruvchi test topshiriqlariga bergan javoblari: 1 – to'g'ri javob, 0 – noto'g'ri javob.

2. Har bir qatorning **yig'indisini** (har bir test topshiruvchining jami olgan balini) СУММ funksiyasi yordamida hisoblaymiz.

3. Qatorlarni bir test topshiruvchilarning jami olgan ballarini **kamayib borishi tartibida saralaymiz**. Buning uchun test topshiruvchilarning jami ballari keltirilgan ustunni belgilab, «сортировка по убыванию» funksiyasini ishga solamiz.

4. Test topshirig'ining **qiyinlik darajasini** СРЗНАЧ funksiyasi yordamida hisoblaymiz, bunda СРЗНАЧ funksiyasining argumenti sifatida test topshirig'i uchun barcha test topshiruvchilarning javoblari keltirilgan kataklar diapazoni ko'rsatiladi.

5. Test topshirig'ining **diskriminatsiya indeksini** hisoblash uchun:

– test topshirig'ining yuqori guruh uchun qiyinlik darajasini СРЗНАЧ funksiyasi yordamida hisoblaymiz, bunda СРЗНАЧ funksiyasining argumenti sifatida yuqori guruhga kiruvchi test topshiruvchilarning javoblari kiritilgan kataklar diapazoni ko'rsatiladi;

– test topshirig'ining quyi guruh uchun qiyinlik darajasini СРЗНАЧ funksiyasi yordamida hisoblaymiz, bunda СРЗНАЧ funksiyasining argumenti sifatida quyi guruhga kiruvchi test topshiruvchilarning javoblari kiritilgan kataklar diapazoni ko'rsatiladi;

B2 Test Booklet v1 Paper 2 - Reading Analysis [Режим с

Файл Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Справка АВВУ

Вставить функцию Библиотека функций

Логические Текстовые Дата и время

Ссылки и массивы Математические Другие функции

Статистические Инженерные Аналитические Информационные Совместимость Идентификация

ГAMMA.OBP ГAMMA.RASP ГAMMA.РАСП ГAMMANЛОГ ГAMMANЛОГ.ТОЧН ГАУСС ГИПЕРГЕОМ.РАСП ДИСП.В ДИСПГ ДИСПА ДИСПРА ДОВЕРИТ. ДОВЕРИТ. Логические и текстовые значения игнорируются. КВАДРОТК КВАРТИЛЬ КВАРТИЛЬ.ИСКЛ КВПИРСОН КОВАРИАЦИЯ.В КОВАРИАЦИЯ.Г КОРРЕЛ

Диспетчер имен Создать из выделенного

Задать имя Использовать в формуле Создать из выделенного

Влияющие Зависимые Убрать стр

Вставить функцию

fx

	A	B	C	D	E
31	1000011	1	0	1	1
32	1000014	1	1	0	1
33	1000021	0	1	0	1
34	1000025	1	1	1	0
35	1000075	0	0	0	0
36	1000004	0	0	1	1
37	1000003	0	0	1	1
38	1000029	0	0	0	0
39	1000040	0	0	1	0
40	1000042	1	0	0	1
41	1000038	1	1	1	0
42	1000039	0	0	0	1
43	1000030	0	0	1	0
44	1000024	0	0	1	0
45	1000022	0	1	0	0
46	1000043	0	0	0	1
47					
48					
49	dispersiya	0,250	0,232	0,243	0,196
50					

Дополнительные сведения

ДИСПГ (число1;число2)
Вычисляет дисперсию для генеральной совокупности. Логические и текстовые значения игнорируются.

Вставить функцию...

fx

0,96 0,246 0,125

Test topshiriqlari dispersiyasini hisoblash

– birinchi raqamdan (yuqori guruh uchun test topshirig‘i-ning qiyinlik darajasidan) ikkinchi raqamni (quyi guruh uchun test topshirig‘ining qiyinlik darajasini) ayiramiz.

6. Test topshirig‘ining **dispersiyasini** hisoblash uchun ДИСП.Г funksiyasini («формулы» – «другие формулы» – «статистические формулы» – ДИСП.Г) ishlatamiz. Bunda funksiyaning argumenti sifatida test topshirig‘i uchun barcha test topshiruvchilarning javoblari keltirilgan kataklar diapazoni ko‘rsatiladi.

7. Butun testning **dispersiyasini** hisoblash uchun ham ДИСП.Г funksiyasini ishlatamiz. Bunda funksiyaning argumenti sifatida test topshiruvchilarning jami olgan ballari keltirilgan kataklar diapazoni ko‘rsatiladi.

8. Testning **asosiy tendensiyasi** bilan bog‘liq ko‘rsatkichlarni (o‘rtacha arifmetik, moda, mediana) hisoblash uchun СРЗНАЧ, МЕДИАНА («формулы» – «другие формулы» – «статистические формулы» – МЕДИАНА) va МОДА.ОДН («формулы» – «другие формулы» – «статистические формулы» – МОДА.ОДН) funksiyalarini ishlatamiz. Bunda funksiyaning argumenti sifatida test topshiruvchilarning jami olgan ballari keltirilgan kataklar diapazoni ko‘rsatiladi.

9. Test natijalari to‘plamining **standart og‘ishini** hisoblash uchun СТАНДОТКЛОН.Г («формулы» – «другие формулы» – «статистические формулы» – СТАНДОТКЛОН.Г) funksiyasini ishlatamiz. Bunda funksiyaning argu-

menti sifatida test topshiruvchilarning jami olgan ballari keltirilgan kataklar diapazoni ko'rsatiladi.

Har bir test topshirig'i dispersiyasi, test topshiriqlari dispersiyalari yig'indisi, butun testning dispersiyasi va test natijalari to'plamining standart og'ishi hisoblangandan keyin hosil bo'lgan qiymatlarni kerakli formulalarga qo'yib, **ishonchlilik ko'rsatkichlarini** hisoblash mumkin.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Alderson J. and Wall D. Does Washback Exist? // Applied Linguistics. 1993. 14:2. – P.115 – 129.
2. American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing. – Washington D.C.: American Educational Research Association. 2014.
3. Brookhart S.M. How To Assess Higher-Order Thinking Skills in Your Classroom. – Alexandria, VA: ASCD. 2010.
4. Brookhart S.M. How To Design Questions and Tasks To Assess Student Thinking. – Alexandria, VA: ASCD. 2014.
5. Brown D. and Abeywickrama P. Language Assessment, Principles and Classroom Practices. White Plains. – NY: Pearson Education, 2010.
6. Brown J.D. Testing in Language Programs: A Comprehensive Guide to English Language Assessment, Upper Saddle River. – N.J.: Prentice Hall Regents, 2005.
7. Cizek G.J. (Ed.) Setting Performance Standards: Foundations, Methods, and Innovations. 2nd ed. – London-New York: Routledge.
8. Cizek G.J. and Bunch M.B. Standard Setting, Thousand Oaks: SAGE. 2011.

9. Green A. *Exploring Language Assessment and Testing: Language in Action*. – New York, NY: Routledge, 2014.
10. Dorans N.J. and Holland P.W. *DIF Detection and Description: Mantel-Haenszel and Standardization*, ETS Research Report RR-92-10. – Princeton, NJ: Educational Testing Service. 1992.
11. Douglas D. *Understanding Language Testing*. – London-New York: Routledge, 2010.
12. Downing S. and Haladyna T. (Eds) *Handbook of Test Development*. – Mahwah, NJ: Lawrence Erlbaum Associates, 2006.
13. Fein M. *Test Development: Fundamentals for Certification and Evaluation*. – Alexandria, VA: American Society for Training and Development, 2012.
14. Frederiksen N. *The Real Test Bias*, ETS Research Report Series RR 81-40. – Princeton, NJ: Educational Testing Service, 1981.
15. Frey B. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, Thousand Oaks: SAGE. 2018.
16. Green A. *Watching for Washback: Observing the Influence of the International English Language Testing System Academic Writing Test in the Classroom // Language Assessment Quarterly*. 2006. 3:4. – P.333 – 368.
17. Green R. *Statistical Analyses for Language Testers*. – NY: Palgrave Macmillan, 2013.

18. Haladyna T. and Rodriguez M. Developing and Validating Test Items. – London-New York: Routledge, 2013.
19. Karami H. An Introduction to Differential Item Functioning // The International Journal of Educational and Psychological Assessment. September 2012. Vol. 11 (2).
20. Koretz D.M. Measuring Up: What Educational Testing Really Tells Us. – Cambridge, MA: Harvard University Press, 2008.
21. Kubiszyn T. and Borich G. Educational Testing and Measurement: Classroom Application and Practice. – Hoboken, NJ: Wiley. 2013.
22. Lord F. and Novick M. Statistical Theories of Mental Test Scores. – Reading, MA: Addison-Wesley Pub. Co, 1968.
23. McDonald R.P. Test Theory: A Unified Treatment. – Mahwah, NJ: Lawrence Erlbaum Associates, 1999.
24. Newton P.E., and Shaw S.D. Validity in Educational and Psychological Assessment. – London: Sage, 2014.
25. Ovendahl N.V. Testwise: Understanding Educational Assessment. Vol.1. – Plymouth: Rowman & Littlefield Education, 2011.
26. Petty G. Teaching Today: A Practical Guide, 4th ed. – Cheltenham: Nelson Thornes, 2009.
27. Popham W.J. Modern Educational Measurement: Practical Guidelines for Educational Leaders. 3rd ed. – Boston: Allyn and Bacon, 2000.

28. Purpura J.E. *Assessing Grammar*. – Cambridge: Cambridge University Press, 2004.
29. Raykov T. and Marcoulides G. A. *Introduction to Psychometric Theory*. – London-New York: Routledge, 2011.
30. Read J. *Assessing Vocabulary*. – Cambridge: Cambridge University Press, 2000.
31. Rowntree D. *Assessing Students: How Shall We Know Them*. – London: Kogan Page, 1987.
32. Schmitt R., Jiang X. and Grabe W. The Percentage of Words Known in a Text and Reading Comprehension // *The Modern Language Journal*. 2011. 95. – P.26 – 43.
33. Shohamy E., Donitsa-Schmidt S. and Ferman I. Test Impact Revisited: Washback Effect Over Time // *Language Testing*. 1996. 13.3. – P.298 – 317.
34. Stobart G. *Testing Times: The Uses and Abuses of Assessment*. – London-New York: Routledge, 2008.
35. Балыхина Т.М. *Словарь понятий и терминов тестологии*. – Москва: Русский язык, 2006.
36. Воробьева С.В. *Современные средства оценивания результатов обучения в общеобразовательной школе*. Изд. 2-е, пер. и доп. – Москва: Юрайт, 2019.
37. Звонников В.И., Чельшкова М.Б. *Современные средства оценивания результатов обучения*. Изд. 5-е, пер. и доп. – Москва: Академия, 2013.
38. Кирейцева А.Н. *Азбука тестирования*. – Санкт-Петербург: Златоуст, 2013.

39. Крокер Л., Алгина Д. Введение в классическую и современную теорию тестов. – Москва: Логос, 2010.
40. Майоров А.Н. Теория и практика создания тестов для системы образования: как выбирать, создавать и использовать тесты для целей образования. – Москва: Интеллект-центр, 2002.
41. Пермяков О.Е., Максимова О.А. Основы технологии комплексной экспертизы качества педагогических тестов. – Томск: ТОИПКРО, 2008.
42. Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – Москва: Логос, 2002.

KITOBDA QO‘LLANILGAN ASOSIY TUSHUNCHALAR VA ATAMALAR INDEKSI

Ushbu indeksga kitobda qo‘llanilgan asosiy tushunchalar va atamalar alifbo tartibida keltirilgan. Kitobxonlarga xorijiy manbalardan ham foydalanish qulay bo‘lishi uchun har bir atamaning ingliz va rus tillaridagi muqobillari keltirildi. Tushuncha va atamalardan keyingi raqamlar ushbu kitobda ularga ta’rif keltirilgan, mohiyati ochib berilgan boblarni bildiradi.

adaptiv test (*ing*: adaptive test, *rus*: адаптивный тест) 1.3

adolat (*ing*: fairness, *rus*: справедливость) 6.1

amaliy test topshirig‘i (*ing*: performance items, *rus*: практическое задание) 4.3

Angoff usuli (*ing*: Angoff method, *rus*: метод Ангоффа) 6.3

aprobatsiya (*ing*: field test, *rus*: апробация) 4.1

aralash test (*ing*: mixed test, *rus*: смешанный тест) 1.3

asosiy tendensiya (*ing*: central tendency, *rus*: центральная тенденция) 3.5

autentiklik (*ing*: authenticity, *rus*: аутентичность) 4.3

baholash (*ing*: assessment, *rus*: оценивание) 1.1

baholash nazariyasi (*ing*: theory of assessment, *rus*: теория оценивания) 1.1

baholovchilar orasidagi ishonchlilik (*ing*: inter-rater reliability, *rus*: согласованность между экспертами) 2.3

baholovchining ichki ishonchliligi (*ing*: intra-rater reliability, *rus*: надежность эксперта) 2.3

binominal taqsimot (*ing*: binominal distribution, *rus*: биномиальное распределение) 3.5

bir o'lchamli test (*ing*: unidimensional test, *rus*: одномерный тест) 2.2

Blum taksonomiyasi (*ing*: Bloom's taxonomy, *rus*: таксономия Блума) 4.1

«Bookmark» usuli (*ing*: Bookmark method, *rus*: метод «Bookmark») 6.3

diagnostik test (*ing*: diagnostic test, *rus*: диагностический тест) 1.3

DIF (test topshirig'ining farqlovchi funksiyasi) (*ing*: differential item functioning, *rus*: дифференциальное функционирование тестового задания) 5.2.4

diskret test topshirig'i (*ing*: discrete-point test item, *rus*: дискретное тестовое задание) 1.3

diskriminant dalillar (*ing*: discriminant evidences, *rus*: дискриминантные доказательства) 2.2

dispersiya (*ing*: variance, *rus*: дисперсия, разброс) 3.2

distraktor (*ing*: distractor, *rus*: дистрактор) 4.2

dixotom test topshirig'i (*ing*: dichotomous test item, *rus*: дихотомное тестовое задание) 3.2

domen (*ing*: domain, *rus*: домен) 1.1

egallash testi (*ing*: proficiency test, *rus*: тест владения) 1.3

heterogen test (*ing*: heterogeneous test, *rus*: гетерогенный тест) 1.3

gomogen test (*ing*: homogenous test, *rus*: гомогенный тест) 1.3

Guttman shkalasi (*ing*: Guttman scale, *rus*: шкала Гуттмана) 6.2

haqiqiy baho (*ing*: true score, *rus*: истинный балл) 3.1

«ikkita yarim» usuli (*ing*: split-half method, *rus*: метод расщепления теста) 3.2

ipsativ test (*ing*: ipsative test, *rus*: ипсативный тест) 1.3

ishonchlik (*ing*: reliability, *rus*: надежность) 3.1

kalit (*ing*: key, *rus*: ключ) 4.2

kengaytirilgan javobni talab qiladigan test topshirig‘i (*ing*: extended response / essay item, *rus*: тестовое задание с развернутым ответом) 4.3

ketma-ketlikni topishni talab qiluvchi test topshirig‘i (*ing*: sequencing item, *rus*: тестовое задание на установление последовательности) 4.2

kommunikativ kompetensiya (*ing*: communicative competence, *rus*: коммуникативная компетенция) 4.4

kommunikativ test (*ing*: communicative test, *rus*: коммуникативный тест) 4.4

konstrukt (*ing*: construct, *rus*: конструкт) 1.1

konstruktga bog‘liq bo‘lmagan o‘zgaruvchilar (*ing*: construct-irrelevant variances, *rus*: переменные, не связанные с конструктом) 2.3

konstruktiv validlik (*ing*: construct validity, *rus*: конструктивная /концептуальная/ валидность) 2.1

konstruktning yetarli darajada aks etmasligi (*ing*: construct underrepresentation, *rus*: недостаточный охват конструкта) 2.3

konvergent dalillar (*ing*: convergent evidences, *rus*: конвергентные доказательства) 2.2

korrelyatsiya (*ing*: correlation, *rus*: корреляция) 2.2

Kronbah alfasi (*ing*: Cronbach alpha, *rus*: альфа Кронбаха) 3.2

kuchli test (*ing*: power test, *rus*: тест мощности) 1.3

Kuder-Richardson formulasi (*ing*: Kuder-Richardson formula, *rus*: формула Кьюдера-Ричардсона) 3.2

ko'p o'lchamli test (*ing*: multidimensional test, *rus*: многомерный тест) 2.2

manfiy og'gan (chapga og'gan) (*ing*: negatively skewed, left skewed, *rus*: отрицательно скошенный, скошенный влево) 3.4

mazmuniy validlik (*ing*: content validity, *rus*: содержательная валидность) 2.1

mediana (*ing*: median, *rus*: медиана) 3.5

mezonga mo'ljallangan test (*ing*: criterion-oriented test, *rus*: критериально-ориентированный тест) 1.3

mezoniy (kriterial) validlik (*ing*: criterion validity, *rus*: критериальная валидность) 2.1

moda (*ing*: mode, *rus*: мода) 3.5

monomorf test (*ing*: monomorphic test, *рус*: мономорфный тест) 1.3

moslashtirishni talab qiluvchi test topshirig'i (*ing*: matching item, *рус*: тестовое задание на установление соответствия) 4.2

muqobil javobli test topshirig'i (*ing*: multiple-choice item, *рус*: тестовое задание на выбор правильного ответа / тестовое задание множественного выбора) 4.2

murakkab baholash sxemasi (*ing*: holistic rubric, *рус*: усложненная схема оценивания) 4.3

musbat og'gan (o'ngga og'gan) (*ing*: positively skewed, right skewed, *рус*: положительно скошенный, скошенный вправо) 3.4

muvaffaqiyat testi (*ing*: achievement test, *рус*: тест достижений) 1.3

normaga mo'ljalangan test (*ing*: norm-oriented test, *рус*: нормативно-ориентированный тест) 1.3

normal taqsimot (*ing*: normal distribution, *рус*: нормальное распределение) 3.3

normalar (*ing*: norms, *рус*: нормы) 6.2

normallashtirilgan standart ball (*ing*: normalised standard score, *рус*: нормализованный стандартный балл) 6.2

nuqtali biseral korrelyatsiya (*ing*: point-biseral korrelation, *рус*: точечно-бисеральная корреляция) 5.2.2

ochiq test topshirig'i (*ing*: constructed response item, *рус*: тестовое задание открытого типа) 1.3, 4.3

parallel test formalari (*ing*: parallel test forms, *rus*: параллельные формы теста) 2.2

«parallel test formalari» usuli (*ing*: parallel forms method, *rus*: метод параллельных форм) 3.2

pedagogik o'lovlar (*ing*: educational measurement, *rus*: педагогические измерения) 1.1

persentil (*ing*: percentile, *rus*: процентиль) 6.2

polimorf test (*ing*: polymorphic test, *rus*: полиморфный тест) 1.3

politom test topshirig'i (*ing*: polytomous test item, *rus*: политомическое тестовое задание) 3.2

prediktiv dalillar (*ing*: predictive evidences, *rus*: предиктивные доказательства) 2.2

prediktiv validlik (*ing*: predictive validity, *rus*: предиктивная /прогностическая/ валидность) 2.1

prognostik test (layoqatni aniqlash testi) (*ing*: aptitude test, *rus*: тест способностей, прогностический тест) 1.3

qisqa javobni talab qiladigan test topshirig'i (*ing*: short-answer item, *rus*: тестовое задание с кратким ответом) 4.3

Rash modeli (*ing*: Rasch model, *rus*: модель Раша) 6.2

solishtirma dalillar (*ing*: concurrent evidences, *rus*: сопоставительные доказательства) 2.2

solishtirma validlik (*ing*: concurrent validity, *rus*: сопоставительная /сопряженная/ предиктивная валидность) 2.1

stanayn (*ing*: stanine, *rus*: станайн) 6.2

standart ball (*ing*: standard score, *рус*: стандартный балл) 6.2
standart og'ish (*ing*: standard deviation, *рус*: стандартное отклонение) 3.3

statistik tahlil (*ing*: statistical analysis, *рус*: статистический анализ) 5.1

T ball (*ing*: T score, *рус*: T балл) 6.2

tasniflash barqarorligi (*ing*: classification consistency, *рус*: устойчивость классификации) 3.3

tasniflash xatosi (*ing*: classification error, *рус*: ошибка классификации) 3.3

tasodifiy xatolik (*ing*: random error, *рус*: случайная ошибка) 3.1

test (*ing*: test, *рус*: тест) 1.1

test spetsifikatsiyasi (tafsiloti) (*ing*: test specification, *рус*: спецификация теста) 2.2

test topshirig'i (*ing*: test item, *рус*: тестовое задание) 1.1

test topshirig'ining diskriminatsiya indeksi (*ing*: item's discrimination index, *рус*: дискриминативность / различительная способность тестового задания) 5.2.2, 5.3

test topshirig'ining qiyyinlik darajasi (*ing*: item's difficulty index, *рус*: сложность тестового задания) 5.2.1

test topshiruvchilar javoblarining distraktorlar bo'yicha taqsimoti

(*ing*: distribution of test takers' responses by distractors, *рус*: распределение ответов тестируемых по дистракторам) 5.2.3, 5.3

«test – qayta test» usuli (*ing*: test – retest method, *rus*: метод ретеста) 3.2

tezkor test (*ing*: speed test, *rus*: тест скорости) 1.3

tizimli (sistematik) xatolik (*ing*: systematic error, *rus*: систематическая ошибка) 3.1

«to‘g‘ri-noto‘g‘ri» shaklidagi test topshirig‘i (*ing*: «true-false» item, *rus*: тестовое задание «правда-ложь» / тестовое задание альтернативных ответов) 4.2

transformatsiya test topshirig‘i (*ing*: transformation test item, *rus*: тестовое задание на трансформацию) 6.4

validatsiya (*ing*: validation, *rus*: валидация) 2.1

validlik (*ing*: validity, *rus*: валидность) 2.1

validlikning dalillari (*ing*: evidences of validity, *rus*: доказательства валидности) 2.2

xom (boshlang‘ich) ball (*ing*: raw score, *rus*: сырой балл, первичный балл) 6.2

xulosalovchi baholash (*ing*: summative assessment, assessment of learning, *rus*: суммирующая оценка) 1.2

yig‘ma baholash sxemasi (*ing*: analytic rubric, *rus*: накопительная схема оценивания) 4.3

yolg‘on ijobiy (*ing*: false positive, *rus*: ложный положительный) 3.3

yolg‘on salbiy (*ing*: false negative, *rus*: ложный отрицательный) 3.3

yoriq test topshirig‘i (*ing*: selected response item, *rus*: тестовое задание закрытого типа) 1.3, 4.2

yuqori ta'sir kuchiga ega test (*ing*: high-stakes test, *rus*: тест с высокими ставками) 2.2.5

z ball (*ing*: z score, *rus*: z балл) 6.2

Z_1 ball (*ing*: Z_1 score, *rus*: Z_1 балл) 6.2

o'lchashning standart xatosi (*ing*: standard error of measurement, *rus*: стандартная ошибка измерения) 3.3

o'lchashning xatosi (*ing*: error of measurement, *rus*: ошибка измерения) 3.1

o'rtacha arifmetik (*ing*: mean, *rus*: среднее арифметическое) 3.5

o'tish bali (*ing*: cut score, *rus*: проходной балл, пороговый балл) 6.3

o'zak (*ing*: stem, *rus*: основная часть) 4.2

shaffoflik (*ing*: transparency, *rus*: прозрачность) 6.1

shakllantiruvchi baholash (*ing*: formative assessment, assessment for learning, *rus*: формирующая оценка) 1.2

chiziqli test (*ing*: linear test, *rus*: линейный тест) 1.3

«cloze» testi (*ing*: cloze test, *rus*: тест типа «cloze») 4.4

«washback» 2.2.5

Izoh va qaydlar uchun

Izoh va qaydlar uchun

Ilmiy-ommabop nashr

Komil Jalilov

BAHOLASH NAZARIYASI ASOSLARI

(TESTOLOGIYA, PEDAGOGIK O'LCHOVLAR)

Muharrir: Abdulla SHAROPOV

Badiiy muharrir: Bahridin BOZOROV

Texnik muharrir: Dilshod NAZAROV

Sahifalovchi: Inomjon O'SAROV

Musahhih: Dilorom MATKARIMOVA

Nashriyot litsenziyasi: AI №134, 27.04.2009

Terishga berildi: 15.07.2020-y.

Bosishga ruxsat etildi: 28.08.2020-y.

Ofset qog'oz. Qog'oz bichimi: 84x108 $\frac{1}{64}$.

Times garnituras. Ofset bosma.

Hisob-nashriyot t.: 15,6. Shartli b.t.: 12,6.

Adadi: 500 nusxa.

Buyurtma № 116

«Akademnashr» nashriyotida tayyorlandi va chop etildi.
100156, Toshkent shahri Chilonzor tumani 20^A-mavze 42-uy.

Tel.: (+99871) 217-16-77

e-mail: info@akademnashr.uz

web: www.akademnashr.uz