



ISSN: 3060-4613



MAKTABGACHA
VA MAKTAB
TA'LIMI VAZIRLIGI



O'zbekiston
Milliy Pedagogika
Universiteti



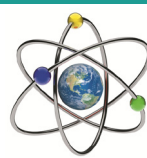
№10(1)
2026

- 13.00.00 Pedagogika fanlari
- 13.00.01 Pedagogika nazariyasi. Pedagogik ta'limotlar tarixi
- 13.00.02 Ta'lim va tarbiya nazariyasi va metodikasi (sohalar bo'yicha)
- 13.00.03 Maxsus pedagogika
- 13.00.04 Jismoniy tarbiya va sport mashg'ulotlari nazariyasi va metodikasi
- 13.00.05 Kasb-hunar ta'limi nazariyasi va metodikasi
- 13.00.06 Elektron ta'lim nazariyasi va metodikasi (ta'lim sohaları va bosqichlari bo'yicha)
- 13.00.07 Ta'limda menejment
- 13.00.08 Maktabgacha ta'lim va tarbiya nazariyasi va metodikasi
- 13.00.09 Ijtimoiy pedagogika
- 07.00.00 Tarix fanlari
- 19.00.00 Psixologiya fanlari
- 01.00.00 Fizika-matematika fanlari
- 02.00.00 Kimyo fanlari
- 03.00.00 Biologiya fanlari
- 09.00.00 Falsafa fanlari
- 10.00.00 Filologiya fanlari
- 11.00.00 Geografiya fanlari

M

MAKTABGACHA VA MAKTAB TA'LIMI

Pedagogika, psixologiya fanlariga ixtisoslashgan ilmiy jurnal



MAKTABGACHA VA MAKTAB TA'LIMI



Elektron nashr. 282 sahifa,
2-oktyabr, 2026-yil.

BOSH MUHARRIR:

Karimova E'zoza Gapirjanovna – Sotsiologiya fanlari nomzodi, dotsent

BOSH MUHARRIR O'RINBOSARI:

Ibragimova Gulsanam Ne'matovna – Pedagogika fanlari doktori, professor

TAHRIRIYAT KENGASHI A'ZOLARI

Ibragimov X.I. – pedagogika fanlari doktori, akademik
Shoumarov G'.B. – psixologiya fanlari doktori, akademik
Qirg'izboyev A.K. – Tarix fanlari doktori, professor
Jamoldinova O.R. – pedagogika fanlari doktori, professor
Sharipov Sh.S. – pedagogika fanlari doktori, professor
Shermuhhammadov B.Sh. – pedagogika fanlari doktori, professor
Ma'murov B.B. – pedagogika fanlari doktori, professor
Madraximova F.R. – pedagogika fanlari doktori, professor
Kalonov M.B. – iqtisodiyot fanlari doktori, professor
Nabiyev D.X. – iqtisodiyot fanlari doktori, professor
Qo'ldoshev Q. M. – iqtisodiyot fanlari doktori, professor
Ikramxanova F.I. – filologiya fanlari doktori, professor
Ismagilova F.S. – psixologiya fanlari doktori, professor (Rossiya)
Stoyuxina N.Yu. – psixologiya fanlari nomzodi, dotsent (Rossiya)
Magauova A.S. – pedagogika fanlari doktori, professor (Qozog'iston)
Rejep O'zyurek – psixologiya fanlari doktori, professor (Turkiya)
Woosyu Cha – Koreya milliy ta'lim universiteti rektori (Koreya)
Polonnikov A.A. – psixologiya fanlari nomzodi, dotsent (Belarus)
Mizayeva F. O. – Pedagogika fanlari doktori, dotsent
Baybayeva M.X. – pedagogika fanlari doktori, professor
Muxsiyeva A.T. – pedagogika fanlari doktori, professor
Aliyev B. – falsafa fanlari doktori, professor
Abdullayeva N. Sh. – Pedagogika fanlari doktori (DSc), professor
Doniyorov S. M. – "Yangi O'zbekiston" va "Pravda Vostoka" gazetalari tahririyati DM bosh muharriri, O'zbekiston Respublikasida xizmat ko'rsatgan jurnalist, filologiya fanlari bo'yicha falsafa doktori (PhD), dotsent
G'afurov D. O. – falsafa fanlari doktori (PhD)
Shomurodov R.T. – iqtisodiyot fanlari nomzodi (PhD), dotsent
Jalilova S.X. – psixologiya fanlari nomzodi (PhD), dotsent
Bafayev M.M. – psixologiya fanlari bo'yicha falsafa doktori (PhD), dotsent
Usmonova D.I. – Samarqand iqtisodiyot va servis instituti dotsenti
Saifnazarov I. – falsafa fanlari doktori, professor
Raxmonova Z. A. – filologiya fanlari bo'yicha falsafa doktori (PhD)
Nematov Sh.E. – pedagogika fanlari nomzodi (PhD)
Tillashayxova X.A. – psixologiya fanlari nomzodi (PhD), dotsent
Yuldasheva F.I. – pedagogika fanlari bo'yicha falsafa doktori (PhD), dotsent
Yuldasheva D.B. – filologiya fanlari bo'yicha falsafa (PhD) doktori, dotsent
Tangriyev A. T. – Toshkent davlat iqtisodiyot universiteti kafedra professori
Ashurov R. R. – psixologiya fanlari bo'yicha falsafa doktori (PhD), dotsent
Panjiyev M. A. – Qashqadaryo viloyati Maktabgacha va maktab ta'limi boshqarmasi boshlig'ining birinchi o'rinbosari
Xudayberganov N. A. – Xorazm Ma'mun akademiyasi Tabiiy fanlar bo'limining katta ilmiy xodimi, biologiya fanlari bo'yicha falsafa doktori (PhD)
Vaxobov Anvar Abdusattor o'g'li – Pedagogika fanlari bo'yicha falsafa doktori, dotsent

Muassis: "Tadbirkor va ishbilarmon" MChJ

Hamkorlarimiz: O'zbekiston Respublikasi Maktabgacha va maktab ta'limi vazirligi, O'zbekiston milliy pedagogika universiteti

EDITOR-IN-CHIEF:

Karimova E'zoza Gapirzhanovna – Minister of Perschool and School Education of the Republic of Uzbekistan

DEPUTY EDITOR-IN-CHIEF:

Ibragimova Gulsanam Ne'matovna – Doctor of Pedagogical Sciences, Professor

EDITORIAL BOARD MEMBERS:

Ibragimov X.I. – Doctor of Pedagogical Sciences, Academician

Shoumarov G'. B. – Doctor of Psychological Sciences, Academician

Qirg'izboyev A. K. – Doctor of Historical Sciences, Professor

Jamoldinova O.R. – Doctor of Pedagogical Sciences, Professor

Sharipov Sh.S. – Doctor of Pedagogical Sciences, Professor

Shermuhhammadov B.Sh. – Doctor of Pedagogical Sciences, Professor

Ma'murov B.B. – Doctor of Pedagogical Sciences, Professor

Madraximova F.R. – Doctor of Pedagogical Sciences, Professor

Kalonov M.B. – Doctor of Economic Sciences, Professor

Nabiyev D.X. – Doctor of Economic Sciences, Professor

Koldoshev K. M. – Doctor of Economic Sciences, Professor

Ikramxanova F.I. – Doctor of Philological Sciences, Professor

Ismagilova F.S. – Doctor of Psychological Sciences, Professor (Russia)

Stoyuxina N.Yu. – Candidate of Psychological Sciences (PhD), Associate Professor (Russia)

Magauova A.S. – Doctor of Pedagogical Sciences, Professor (Kazakhstan)

Rejep O'zyurek – Doctor of Psychological Sciences, Professor (Turkey)

Wookyu Cha – President of the National University of Education, Korea (South Korea)

Polonnikov A.A. – Candidate of Psychological Sciences (PhD), Associate Professor (Belarus)

Mizayeva F. O. – Doctor of Pedagogical Sciences, Professor

Baybayeva M.X. – Doctor of Pedagogical Sciences, Professor

Muxsiyeva A.T. – Doctor of Pedagogical Sciences, Professor

Aliyev B. – Doctor of philosophy, professor

Abdullayeva N. Sh. – Doctor of Pedagogical Sciences (DSc), Professor

Doniyorov S. M. – Editor-in-Chief of the DM Editorial Office of the newspapers “Yangi O'zbekiston” and “Pravda Vostoka”, Honored Journalist of the Republic of Uzbekistan, Doctor of Philosophy (PhD) in Philology, Associate Professor

Gafurov D. O. – Doctor of Philosophy (PhD)

Shomurodov R.T. – Candidate of Economic Sciences (PhD), Associate Professor

Jalilova S.X. – Candidate of Psychological Sciences (PhD), Associate Professor

Bafayev M.M. – Doctor of Philosophy in Psychological Sciences (PhD), Associate Professor

Usmonova D.I. – Associate Professor, Samarkand Institute of Economics and Service

Saifnazarov I. – Doctor of philosophy, professor

Rakhmonova Z.A. – Doctor of Philosophy (PhD) in Philological Sciences, Associate Professor

Nematov Sh.E. – Candidate of Pedagogical Sciences (PhD)

Tillashayxova X.A. – Candidate of Psychological Sciences (PhD), Associate Professor

Yuldasheva F.I. – Doctor of Philosophy in Pedagogical Sciences (PhD), Associate Professor

Yuldasheva D.B. – Doctor of Philosophy (PhD) in Philological Sciences, Associate Professor

Tangriyev A.T. – is a professor of Tashkent State University of Economics

Ashurov R. R. – Doctor of Philosophy (PhD) in Psychology, Associate Professor

Panjiyev M. A. – First Deputy Head of the Department of Preschool and School Education of the Kashkadarya Region

Khudaiberganov N. A. – Senior Researcher of the Department of Natural Sciences of the Khorezm Mamun

Academy, Doctor of Philosophy (PhD) in Biological Sciences

Vakhobov Anvar Abdusattor oğlu – Doctor of Philosophy in Pedagogical Sciences, Associate Professor

“Maktabgacha va maktab ta'limi” jurnali O'zbekiston Respublikasi Oliy attestatsiya komissiyasining quyidagi qarorlariga asosan pedagogika va psixologiya fanlari bo'yicha falsafa doktori (PhD) hamda fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiyalaridagi asosiy ilmiy natijalarni chop etish uchun milliy ilmiy nashrlar ro'yxatiga kiritilgan:

Pedagogika fanlari bo'yicha: OAK Kengashi tavsiyasi (26.08.2024-y., №11-05-4381/01) asosida:

- Ekspert kengashi (29.10.2024-y., №10)
- Rayosat qarori (31.10.2024-y., №363/5)

Psixologiya fanlari bo'yicha: Toshkent davlat pedagogika universiteti murojaatiga asosan OAK tavsiyasi (24.04.2025-y., №11-05-2566/01):

- Ekspert kengashi (25.05.2025-y., №10)
- Rayosat qarori (08.05.2025-y., №370/5)

“Maktabgacha va maktab ta'limi”
jurnali

26.09.2023-yildan

O'zbekiston Respublikasi Prezidenti
Administratsiyasi huzuridagi Axborot
va ommaviy kommunikatsiyalar
agentligi tomonidan **№C-5669363**
reyestr raqami tartibi bo'yicha
ro'yxatdan o'tkazilgan.

Litsenziya raqami: **№136361**

MUNDARIJA

Raqamli oliy ta'lim boshqaruvida kommunikativ kreativlikni rivojlantirishning tashkiliy pedagogik mexanizmi	10
Mamatova Xilola Muxiddinovna	
Ta'lim jarayonida sun'iy intellektdan foydalanishning ijobiy va salbiy jihatlari: tizimli tahlil va strategik yondashuvlar	14
Zamira Nishanova	
Maktabgacha yoshdagi bolalarda muloqot va ijodiy faoliyat orqali o'zini namoyon etish ko'nikmalarini shakllantirish	17
Abdug'afforova Navbahoroy Botirjon qizi	
Raqamli sivilizatsiya davrida oliy ta'limda axborot-analitik kompetensiyani shakllantirishning nazariy-pedagogik asoslari	22
Oybek Davlatov	
Maktabgacha yoshdagi bolalarning kommunikativ kompetensiyalarini rivojlantirishda o'yin texnologiyalarining ahamiyati	27
Bekmuhammedova Muhlisa Abdusattarovna	
Jaloliddin Rumiy asarlarida komil inson va shaxsiy kamolot masalalari	33
Qodirova Madina Odiljon qizi	
Interfaol ta'lim texnologiyalari asosida o'quvchilarning ijodiy va tanqidiy fikrlash kompetensiyalarini rivojlantirishning pedagogik imkoniyatlari	37
Ergasheva Shaxnoza Baxtiyor qizi	
Rhythmic Gymnastics: a Harmony of Art and Sport	42
Hurshida Murzoyeva	
Sport to'garaklari mashg'ulotlarining talaba-yoshlar jismoniy rivojlanishiga ta'siri	46
Jabborova Sevara Asroqul qizi	
Inklyuziv ta'lim sub'yektlari hamkorligini boshqarishning innovatsion pedagogik mexanizmlari	51
Kuranboyeva Ozoda Xakimovna	
Maktabdan tashqari ta'lim muhitining o'quvchilar ijtimoiy-emotsional rivojlanishiga ta'sir etuvchi pedagogik omillari	57
Mexriniso Nosirova	
Yangi O'zbekistonda maktabgacha va maktab ta'limi tizimi uchun zamonaviy pedagog kadrlar tayyorlashning rivojlanish istiqbollari	61
Sodiqova Yulduzxon Maxamat qizi	
Integrating Culture in Language Teaching and Developing Intercultural Tolerance: Pedagogical Approaches and Classroom Implications	65
Tairova Ruzakhon Umidbek qizi	
Boshlang'ich sinf o'quvchilarida darsga ijobiy munosabatlarni shakllantirish va rivojlantirish texnologiyalari	71
Tursunbayeva Yulduz	
Sun'iy intellekt asosidagi adaptiv ta'lim muhitining pedagogik-psixologik nazariy asoslari	75
Xakimov Utkirbek Ilhomjon o'g'li	
O'zbek xalq cholg'ulari orkestri ijrochiligi san'atining rivojlanish tarixi va uning pedagogik ta'limdagi o'rni ...	79
Miskinova Sabina Ziyadxon qizi	
Преодоление узбекско-русской лингвокультурной интерференции при текстоворждении студентов-филологов: методический аспект	85
Тахирова Маъсуда Абдужаббор кизи	
Инновационные методы преподавания русского языка в вузах Узбекистана: развитие коммуникативной и речевой компетенции учащихся	91
Шокаримова Камола Абдуллаевна	



Boshlang'ich sinf o'quvchilarida yumshoq ko'nikmalarni shakllantirishning nazariy-metodologik asoslari....	98
<i>Yakubjonov Xasan Qahramonovich</i>	
Bo'lajak boshlang'ich sinf o'qituvchilarining ilmiy-ijodiy kompetensiyalarini rivojlantirish texnologiyalari	103
<i>Umarov Xusniddin Egamboyevich</i>	
Sport maktabi o'quvchilarida yengil atletika intervalli mashg'ulotlari yordamida chidamlilikni oshirish	107
<i>Shuxrat Xudoyberdiyev Mamayusupovich, Abduqahhorov Nuriddin Xabibulloyevich</i>	
O'zbek va jahon adabiyoti matnlarini qiyosiy tahlil qilishda sun'iy intellekt va kompyuter lingvistikasining o'rni	112
<i>Sodikova Makhliyo</i>	
Milliy va umuminsoniy qadriyatlarga asoslangan harbiy tarbiya metodlari	117
<i>A'zamov Inom Salimjonovich</i>	
Boshlang'ich ta'lim o'qituvchilarining inklyuziv ta'lim sharoitida o'quvchilar bilan ishlashda STEAM yondashuvidan foydalanish metodikasini takomillashtirish	121
<i>Abidjanova Vasila To'likinovna</i>	
Kommunikativ topshiriqlar asosida ingliz tilida gapirish ko'nikmasini rivojlantirishning metodik imkoniyatlari	129
<i>Prenova Gulsara Barlikbayevna</i>	
Bo'lajak boshlang'ich sinf o'qituvchilarining pedagogik nutqiy kompetentligini rivojlantirishning nazariy asoslari	134
<i>Meyliqulova Manzura Abdullayevna</i>	
Kasbiy ta'lim sifatini boshqarishda sun'iy intellekt texnologiyalaridan foydalanish	138
<i>Ne'matova Zarina Furqat qizi</i>	
Kimyoning asosiy tushuncha va qonunlarini aksiomatik asosda talqin qilishning ilmiy-pedagogik ahamiyati	141
<i>Kamolova Nargiza Ibragimovna</i>	
Моделирование процесса формирования патриотизма у курсантов на основе культурного наследия казахской и узбекской интеллигенции	147
<i>Малдыбаев Куандик Бекбауович</i>	
Bo'lajak tarjimonlarda mumtoz asarlardagi leksik-frazeologik birliklarni tarjima qilish kompetensiyasini shakllantirish ("Firdavs ul-iqbol" asari misolida)	152
<i>Borasulova Dilnoza Dilmurodjon qizi</i>	
WHY I+1 is Not the ZPD: Reassessing Input-Based and Sociocultural Accounts of EFL Development	155
<i>Sattarova Zulxumor Nodirbek kizi</i>	
Лингвометодическая компетенция учителя русского языка: генезис понятия и структура в современной педагогической науке	159
<i>Икрамова Мунира Маруфжановна</i>	
Kasbiy axloq va pedagogik madaniyat masalasini tadqiq etishda ilmiy tadqiqot metodlarining o'rni va vazifalari	163
<i>Kazakov Shavkat Norqobilovich</i>	
Kichik maktab yoshidagi bolalar nutq faoliyatida ijtimoiy muhitning psixologik roli	168
<i>Shadiyev Feruz Djaxonovich</i>	
Intellektida nuqsoni bo'lgan kichik maktab yoshidagi o'quvchilarning so'z boyligini boyitishda o'yin texnologiyasidan foydalanish	173
<i>Xolmuxamedova Zilola Azimovna</i>	
PIRLS xalqaro baholash dasturida matn bilan ishlashning asosiy turlari va ularni boshlang'ich ta'limda qo'llash metodikasi	178
<i>Yusupova Gulshiraxan Zinatdinovna</i>	
Ona tili ta'limida sun'iy intellekt vositalaridan foydalanish orqali o'qituvchilarning kasbiy mahoratini oshirish	183
<i>Toshmatova Yulduz Olloqulovna</i>	
Kompetentlikka asoslangan ta'limda pedagogning kasbiy madaniyatining o'rni	190
<i>Xojamov Ilhomjon Rahmonaliyevich</i>	

Boshlang'ich ta'limda darslik va o'quv lug'atlari integratsiyasi asosida o'quvchilarning lingvistik-kommunikativ kompetensiyalarini rivojlantirish metodikasini takomillashtirish	195
Ergasheva Tuyg'unoy Shavkatjon qizi	
Bo'lajak musiqa ta'limi o'qituvchilarining raqamli-pedagogik kompetensiyasini rivojlantirishning integrativ-metodik modelini takomillashtirish	201
Ataboyeva Muxlisa Azimjon qizi	
Raqamli ta'lim nazariyasi va metodikasi asosida oliy ta'lim pedagoglarining kasbiy faoliyatini takomillashtirish.....	209
Choriyeva Lobar Fayzullo qizi, Safarov Azamat Rasul o'g'li	
Inklyuziv ta'limda o'quvchilarning individual ehtiyojlarini aniqlash va ularga mos pedagogik yondashuvlarni tashkil etish	214
Avazova Ozoda Ashurali qizi	
Integrativ yondashuv asosida talabalarning kreativ faoliyatini rivojlantirishning pedagogik texnologiyalari ...	219
Abdusabirova Laylo Azizjon qizi	
Boshlang'ich sinf o'quvchilarining imloviy savodxonligini oshirish usullari	224
Ashurova Zilolaxon Baxodir qizi	
Boshlang'ich ta'limda Jeopardylabs platformasi asosidagi o'yin-topshiriqlarnig dars samaradorligidagi ahamiyati.....	229
Atayeva Baxrinisa Abduvaxidovna	
Tasviriy san'at talabalarida perseptiv qobiliyatlarni rivojlantirishning pedagogik-psixologik asoslari	233
Bazarova Muyassarxon Maxamadyusubovna	
What Trainee English Teachers Notice in AI-Generated Assessment Tools: Fault Detection in Two Tesol Cohorts	238
Dildora Kdirbayeva	
Zamonaviy didaktik vositalar asosida bo'lajak boshlang'ich sinf o'qituvchilarining kasbiy kometensiyasini rivojlantirish mexanizmi.....	246
Ergasheva Ma'mura Suyunovna	
Maktabgacha ta'lim tashkilotlarida "Ilk qadam" milliy o'quv dasturi asosida bolalarning har tomonlama rivojlanishini ta'minlashning pedagogik asoslari.....	250
Ergasheva Umrinisso Komilovna	
Aqli zaif bolalarda kommunikativ ko'nikmalarni shakllantirishning diagnostik mezonlari va ko'rsatkichlari ...	254
Hamraqulova Durdona Dilshod qizi	
Parent Preparation for Meaningful Participation in Inclusive Education.....	258
Muazzam Bakhtierovna Abdurakhmanova	
Oliy ta'limning nofilologik yo'nalishlarida til ta'limi darsliklari mazmunini bosqichma-bosqich loyihalash	262
Muxamedova Muhlisa Avazjon qizi	
Interyer dizayn ta'limida sun'iy intellekt: bo'lajak o'qituvchilarning kasbiy kompetensiyasini shakllantirish metodikasi.....	268
Rustamov Kamoliddin Ziyodulla o'g'li	
Differences Between Professional and Amateur Boxing	273
Shuhrat Dilmuradov	
Лингводидактический потенциал интерактивного обучения в развитии речевой компетенции пятиклассников.....	277
Ергашева Эльвира Владимировна	

WHAT TRAINEE ENGLISH TEACHERS NOTICE IN AI-GENERATED ASSESSMENT TOOLS: FAULT DETECTION IN TWO TESOL COHORTS

Dildora Kdirbayeva

Department of foreign languages and general studies,
IT park university, Tashkent, Uzbekistan
E-mail: kdirbayevadi@gmail.com

ORCID: 0009-0004-6836-043X

Abstract: A teacher can now describe a practice activity to a language model and have a working assessment tool a few minutes later. Between that file and the learner, there is one check: the teacher deciding whether to use it. This study examined what trainee English teachers notice in descriptions of such tools. Forty-nine students in two graduate TESOL cohorts ($n = 24$ and $n = 25$) each judged seven tools carrying a planted fault. The faults were drawn from an argument-based account of validity and divided into those visible in the content and those residing in the relation between what a tool computes and what it claims to measure. Detection was higher for content faults than for construct faults in both cohorts: 62.5% against 35.4% in the first, and 83.0% against 52.0% in the second. The second cohort detected more than the first on six of the seven faults. The two clearest cohort differences were an incorrect answer key (58.3% against 88.0%) and a four-item claim to a CEFR level. Counting linking words and reporting the count as coherence remained the least frequently detected fault. The pattern supports the claim that language-trained noticing catches language errors and often misses measurement claims. It does not support the stronger claim that construct faults are nearly invisible.

Key words: teacher noticing, language assessment, generative artificial intelligence, validity, assessment literacy, teacher education, CEFR.

Annotatsiya: Bugungi kunda o'qituvchi amaliy mashg'ulotni til modeliga tavsiflab berishi va bir necha daqiqadan so'ng foydalanishga tayyor baholash vositasiga ega bo'lishi mumkin. Ushbu fayl bilan o'quvchi o'rtasida faqat bitta nazorat bosqichi mavjud: o'qituvchining undan foydalanish yoki foydalanmaslik to'g'risidagi qarori. Mazkur tadqiqotda ingliz tili bo'yicha bo'lajak o'qituvchilar bunday vositalarning tavsiflarini ko'rib chiqishda nimalarga e'tibor berishi o'rganildi. TESOL bo'yicha magistratura dasturining ikki guruhidagi jami 49 nafar talaba ($n = 24$ va $n = 25$) tarkibiga ataylab xato kiritilgan yettita vositaning har birini baholadi. Xatolar validlikka argumentga asoslangan yondashuv asosida ajratib olindi va ikki guruhga bo'lindi: mazmunning o'zida ko'rinadigan xatolar hamda vosita hisoblaydigan ko'rsatkich bilan u o'lchashni da'vo qilayotgan konstrukt o'rtasidagi munosabatga taalluqli xatolar. Har ikkala guruhda ham mazmun bilan bog'liq xatolarni aniqlash darajasi konstrukt bilan bog'liq xatolarni aniqlash darajasidan yuqori bo'lgan: birinchi guruhda mos ravishda 62.5% ga nisbatan 35.4% va ikkinchi guruhda 83.0% ga nisbatan 52.0%. Ikkinchi guruh yettita xatoning oltitasini birinchi guruhga qaraganda yuqoriroq darajada aniqlagan. Guruhlar o'rtasidagi eng yaqqol farqlar noto'g'ri javoblar kalitini aniqlashda (58.3% ga nisbatan 88.0%) va atigi to'rtta topshiriq asosida CEFR darajasini belgilash haqidagi da'voni baholashda kuzatilgan. Bog'lovchi so'zlarni sanash va ularning sonini matnning yaxlitligi ko'rsatkichi sifatida taqdim etish eng kam aniqlangan xato bo'lib qolgan. Olingan natijalar til bo'yicha tayyorgarlikka ega bo'lgan mutaxassislar til xatolarini aniqlashga moyil ekanligi, biroq o'lchashga oid da'volarni ko'pincha ilg'amay qolishini ko'rsatadigan fikrni qo'llab-quvvatlaydi. Shu bilan birga, natijalar konstrukt bilan bog'liq xatolarni deyarli aniqlab bo'lmaydi, degan kuchliroq da'voni tasdiqlamaydi.

Kalit so'zlar: o'qituvchi e'tibori, tilni baholash, generativ sun'iy intellekt, validlik, baholash savodxonligi, o'qituvchi ta'limi, CEFR.



Аннотация: В настоящее время учитель может описать практическое задание языковой модели и уже через несколько минут получить готовый инструмент оценивания. Между этим файлом и обучающимся существует лишь одна проверка: решение учителя о том, следует ли использовать данный инструмент. В настоящем исследовании изучалось, на что обращают внимание будущие преподаватели английского языка при ознакомлении с описаниями таких инструментов. Сорок девять студентов двух групп магистерской программы TESOL ($n = 24$ и $n = 25$) оценили по семь инструментов, каждый из которых содержал преднамеренно внесённую ошибку. Ошибки были выделены на основе аргументированного подхода к валидности и разделены на две группы: ошибки, непосредственно обнаруживаемые в содержании, и ошибки, связанные с соотношением между тем, что вычисляет инструмент, и тем, что, согласно заявлению, он измеряет. В обеих группах участников выявляемость ошибок содержания была выше, чем ошибок конструкта: 62.5% против 35.4% в первой группе и 83.0% против 52.0% во второй. Вторая группа выявила больше ошибок, чем первая, по шести из семи типов ошибок. Наиболее заметные различия между группами наблюдались при выявлении неправильного ключа ответов (58.3% против 88.0%) и необоснованного определения уровня CEFR на основе четырёх заданий. Подсчёт связующих слов и представление полученного количества как показателя связности оставались наименее часто выявляемой ошибкой. Полученные результаты подтверждают предположение о том, что подготовка в области языка способствует выявлению языковых ошибок, но часто не позволяет распознавать утверждения, касающиеся измерения. Вместе с тем результаты не подтверждают более сильное утверждение о том, что ошибки конструкта практически невозможно обнаружить.

Ключевые слова: внимание учителя, оценивание языковых навыков, генеративный искусственный интеллект, валидность, оценочная грамотность, педагогическое образование, CEFR.

INTRODUCTION

A teacher who wants a practice activity for tomorrow can describe it to a language model and have a working tool a few minutes later. The tool takes an answer, checks it, and returns a result. Nothing in that exchange asks what the result means, and the tool does not say where it stops being trustworthy. The only check between the generated file and the learner is the teacher, who looks at the screen and decides whether to use it. That check has rarely been examined. The consequences fall on learners. Black and Wiliam (1998) showed that feedback helps when it tells a learner what to do next, whereas a bare mark often does not.

A tool that reports a coherence score it never computed does something worse than fail to help. It tells the learner something false about their writing, and it does so with the authority of a number. In Uzbekistan, the question is not abstract. National policy is pushing artificial intelligence into education quickly (Republic of Uzbekistan, 2024), and teachers at every level are being urged to use these tools. This encouragement has not been matched by guidance on how to judge what the tools produce. Nothing in initial teacher preparation currently addresses this judgement.

LITERATURE REVIEW

Research on generative artificial intelligence in language assessment has grown quickly. Chuang and Yan (2025), reviewing 77 studies, identify automated scoring and feedback, item generation, classroom-based assessment, and academic integrity as the main strands, and argue that responsible use depends on sustained attention to validity. The most direct evidence of what these systems do when asked to assess comes from Anbreen et al. (2026), who compared a human rater with ChatGPT-5 on 50 IELTS essays. The scores agreed closely. The reasoning that produced them did not. A newer body of literature looks at educators who build tools rather than choose them. Sarkar and Drosos (2025) provide the first empirical account of the practice they call *vibe coding*: cycles of prompting and rapid inspection, with review compressed to whatever the user happens to read.

Chow and Ng (2025) report clinical applications built by a medical educator with no formal programming training. Song et al. (2026) propose a design framework for schoolteachers creating such tools and note how little is known about how teachers should go about it. In language education, the available work has looked at learners rather than teachers (Woo et al., 2025).

What a teacher perceives in an instructional situation is itself a developed object of study. Research on teacher noticing, following van Es and Sherin (2008), separates attending to relevant features from interpreting what they mean and has shown that the two are distinct and that both can be taught.

König et al. (2022), in a systematic review, report that most of this work is cognitive-psychological and divides noticing into attending, interpreting, and responding. Almost all of it, though, takes classroom events as its object. Noticing an artefact - a material that a teacher has to judge before use - has barely been studied.

Validity theory supplies the terms in which such an artefact can be judged. Messick (1989) argued that validity attaches to interpretations and uses, not to instruments. Kane (2013) recast validation as the evaluation

of an argument whose linked inferences each rest on a warrant that can fail. Chapelle et al. (2008) demonstrated the approach for TOEFL.

Fulcher (2012) argued that assessment literacy has to reach classroom teachers, whose ordinary decisions carry real consequences.

Within Uzbekistan, Baxtiyorova (2025) has examined how artificial intelligence might support the development of learners' English and treats the teacher as a user of such systems rather than as their reviewer. These literatures have not yet been brought together. Validity theory describes what an assessment instrument must support and assumes an institution with the resources to establish it. The noticing literature describes how teachers perceive and has looked at lessons rather than at materials. The tool-creation literature describes how artefacts are made and does not ask what they claim. No study has established what teachers actually detect in a generated assessment tool or whether detection varies by the kind of fault.

The aim was to establish what trainee English teachers notice in AI-generated assessment tools. The study set out to measure detection rates for faults of different kinds, using a taxonomy drawn from validity theory, and to test whether detection differs between faults visible in the content of a tool and faults that reside in what the tool claims to measure. A comparison of two cohorts was added once the task had been run twice. The study moves teacher noticing from classroom events to instructional artefacts. It uses a fault taxonomy derived from argument-based validity, so that faults are distinguished by the inference they defeat rather than by how they look. The predictions were written before the counts were examined, which makes the comparison in Section 4 a comparison with what the study expected, not with what it found convenient to conclude. If detection is uneven across fault types, the results indicate where teacher education should concentrate and distinguish instruction that would be useful from instruction that merely covers faults teachers already catch. For institutions adopting artificial intelligence under national policy, the findings provide a small evidence base for guidance that is currently being written without one.

RESEARCH METHODOLOGY

The study used a scenario-based judgement task administered as an online questionnaire. Participants read descriptions of AI-generated assessment tools, each with a known fault, and reported what they would change and which problems they identified. The scenario-based approach, standard in teacher noticing research (König et al., 2022), fixes the stimulus so that differences in response can be attributed to the participant rather than to the material. The principal comparison is within the set of scenarios: the same respondents judged every tool. A second comparison, between cohorts, became available because the task was administered twice. Participants were students in the graduate TESOL programme at Webster University in Tashkent. The task was completed by two cohorts, labelled here as Cohort 1t (n = 24) and Cohort 3t (n = 25). All students in each cohort were invited. Participation was voluntary. The programme admits both candidates preparing to enter the profession and practising teachers upgrading their qualifications.

Teaching status, experience, and prior study of language assessment were recorded as background variables. Those breakdowns are not included in the counts used for this report, so the cohorts are compared as wholes. The sample is a convenience sample from one programme at one institution. Eight scenarios were written, each describing an interactive assessment tool of a kind that current generative systems produce when asked for a classroom practice activity. Each description stated what the tool does, what the learner sees on completion, and, where the fault was in the tool's internal operation, how the tool arrived at its result. Descriptions ranged from 60 to 120 words. The scenarios are given in Appendix A. Textual descriptions were used rather than working artefacts so that every participant saw the same information and detection could not depend on layout or colour. Seven scenarios each contained one planted fault. One contained none and served as a control. Faults were chosen to represent distinct inferences in an argument-based account of validity (Kane, 2013). Council of Europe (2020) level claims are the reference point for the scenario that assigns a CEFR band. Scenario 3 contained no planted fault and served as a control.

Table 1: Fault taxonomy and distribution across scenarios

Scenario	Planted fault	Inference defeated	Locus
1.	Answer key incorrect for the item as written	Generation	Visible content
2.	Connective count reported as a coherence score	Extrapolation	Internal operation
3.	None (control); not in the counts below	–	–
4.	Lexis far above stated level; ungrammatical example	Generation	Visible content
5.	Typed word count reported as spoken fluency	Extrapolation	Internal operation



6.	Score reported with no indication of what was wrong	Use	Visible content
7.	Four items used to assign a CEFR level	Generalization	Visible content
8.	Exact string matching rejects correct responses	Scoring	Internal operation

The questionnaire was administered through Google Forms. Each scenario was followed by an open question asking whether the participant would use the tool and what they would change, a fixed list of nine candidate problems, and a five-point confidence rating. The open question elicited spontaneous noticing, while the list elicited cued noticing. Detection of the planted fault is reported below as a single rate for each scenario. Detection is reported as a count and percentage for each cohort and pooled across the 49 participants. Visible-content detection is the mean of Scenarios 1, 4, 6, and 7, calculated over judgements rather than participants. Construct-relation detection is the mean of Scenarios 2 and 5. Scenario 8 is held out of both classes.

Cohorts were compared scenario by scenario using Fisher's exact test on the 2×2 table of detected and missed cases. The alpha level was .05. Seven tests were run, so a Bonferroni threshold of 0.007 is also reported. Cohort comparisons use Fisher's exact test because the available figures are cohort totals. Class-level percentages pool judgements within participants and are interpreted descriptively. Participants were students of the author. The activity was presented as part of normal coursework, so declining to participate in the research carried no academic cost. Consent to use responses as research data was sought separately from completion of the class task. The instrument collected no names or email addresses.

ANALYSIS AND RESULTS

Forty-nine participants completed the task, 24 in Cohort 1t and 25 in Cohort 3t. Table 2 presents detection of the planted fault for each scenario. Percentages are rounded to one decimal place.

Table 2: Detection of the planted fault, by scenario and cohort

Scenario	Fault	Cohort 1t	Cohort 3t	Pooled	Fisher p
4	Level and grammar	20/24 (83.3%)	19/25 (76.0%)	39/49 (79.6%)	0.725
1	Incorrect answer key	14/24 (58.3%)	22/25 (88.0%)	36/49 (73.5%)	0.025
6	Score, no direction	15/24 (62.5%)	21/25 (84.0%)	36/49 (73.5%)	0.114
8	String matching	13/24 (54.2%)	19/25 (76.0%)	32/49 (65.3%)	0.140
7	Four items, CEFR	11/24 (45.8%)	21/25 (84.0%)	32/49 (65.3%)	0.007
5	Words as fluency	10/24 (41.7%)	14/25 (56.0%)	24/49 (49.0%)	0.396
2	Links as coherence	7/24 (29.2%)	12/25 (48.0%)	19/49 (38.8%)	0.244

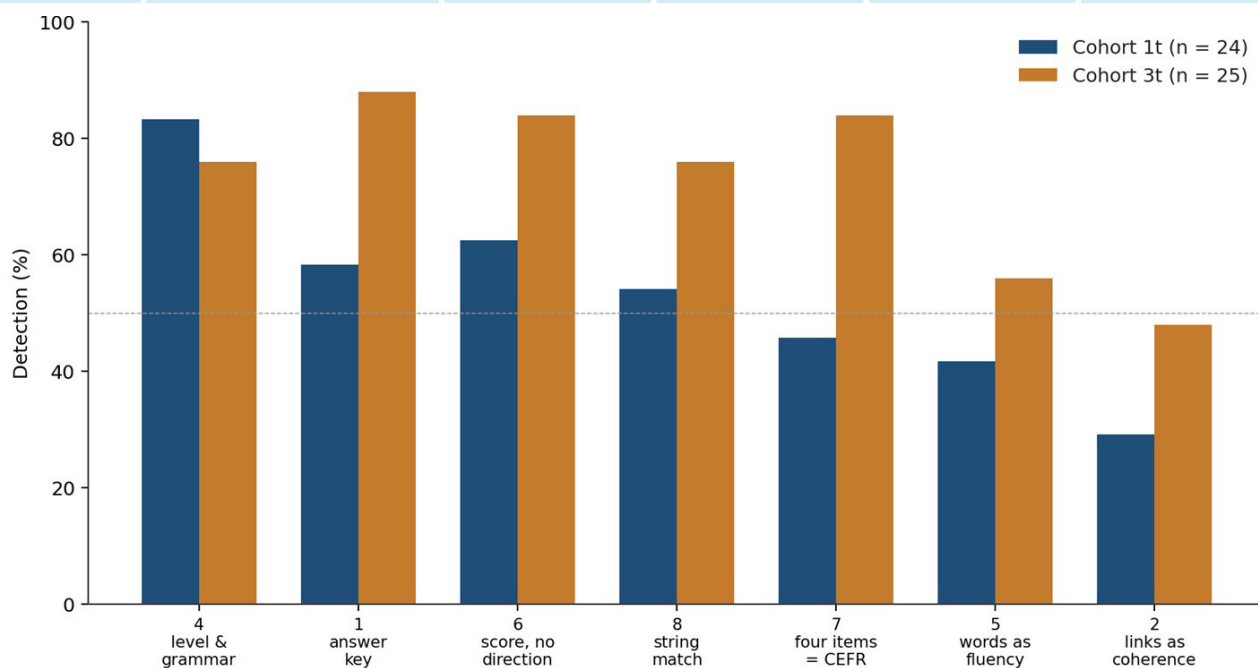


Figure 1: Detection of the planted fault, by scenario and cohort

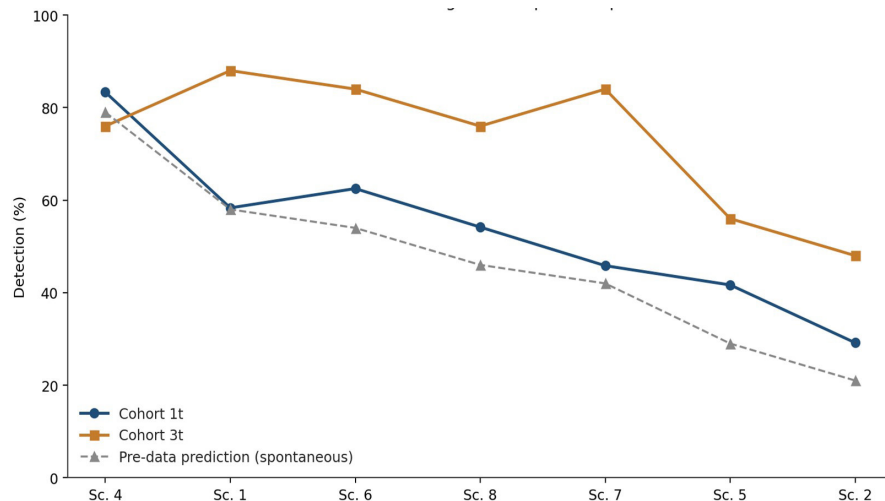


Figure 2: Observed detection against the spontaneous rate predicted before data collection. The prediction was written for a single cohort of 24.

In Cohort 1t, the order closely matched the prediction. The ungrammatical, off-level example was detected by 20 of 24 participants. The coherence score built from a count of linking words was detected by 7 of 24. The other scenarios fell between those two, in nearly the expected order. Cohort 3t did not reproduce that order. Detection was high for the incorrect answer key (22/25), the bare score (21/25), and the four-item CEFR claim (21/25). The off-level example, which had been the easiest fault in Cohort 1t, was detected by 19 of 25 participants, slightly less often than in the first cohort. The two construct faults remained the least frequently detected: 14 of 25 for the fluency score built from a word count and 12 of 25 for the coherence score built from linking words. Two cohort differences were unlikely under a null hypothesis of no difference. Cohort 3t detected the incorrect answer key more often (88.0% against 58.3%, $p = .025$) and the four-item CEFR claim more often (84.0% against 45.8%, $p = .007$). Only the second of these survives a Bonferroni adjustment. The other five comparisons were not significant. Cohort 3t was higher on five of them and lower on one, the level-and-grammar fault, for which both cohorts were already above 75%.

Table 3: Detection by fault class

Fault class	Scenarios	Cohort 1t	Cohort 3t	Pooled
Visible content	1, 4, 6, 7	60/96 (62.5%)	83/100 (83.0%)	143/196 (73.0%)
Construct relation	2, 5	17/48 (35.4%)	26/50 (52.0%)	43/98 (43.9%)
Difference	—	27.1 points	31.0 points	29.1 points

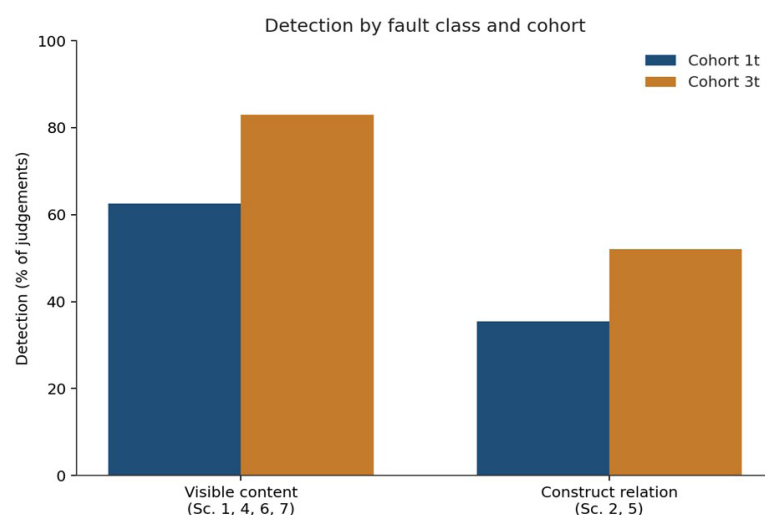


Figure 3: Detection by fault class and cohort. Visible-content faults were Scenarios 1, 4, 6 and 7. Construct-relation faults were Scenarios 2 and 5.

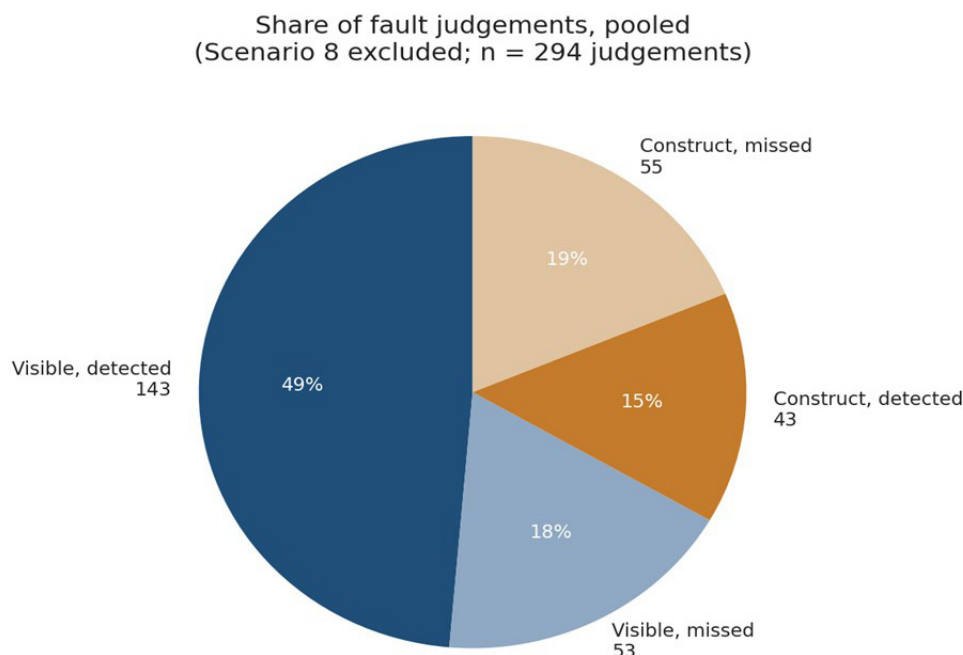


Figure 4: Share of pooled fault judgements, Scenario 8 excluded (294 judgements). Missed construct-relation judgements and missed visible-content judgements are similar in number. Detected visible-content judgements are the largest slice.

The class difference ran in the same direction in both cohorts. Visible-content faults were detected more often than construct-relation faults, by 27.1 percentage points in Cohort 1t and by 31.0 percentage points in Cohort 3t. The gap is close to the 33-point difference predicted before data collection. The absolute rates are not. Construct-relation detection was predicted at 25%. It was 35.4% in the first cohort and 52.0% in the second. Scenario 8, held out because a correct answer marked wrong would become apparent in use, sat between the two classes in Cohort 1t (54.2%) and with the visible faults in Cohort 3t (76.0%). The central prediction was that trainee teachers would see these tools in the same way they see language. A wrong verb form would be caught because it is a wrong verb form. A score that names something the tool never measured would pass because nothing in it is wrong to a reader attending to the words. The first half of that prediction is supported. In both cohorts, the off-level, ungrammatical example was among the most frequently detected faults, and the coherence score built from a count of linking words was the least frequently detected. The gap between the visible-content class and the construct-relation class was 27 percentage points and 31 percentage points, against a predicted 33.

The second half is only partly supported. The study had predicted that Scenarios 2 and 5 would be detected by about a quarter of a cohort and had set 50% spontaneous detection on both as a disconfirmation criterion. Cohort 1t stayed below that line (41.7% and 29.2%). Cohort 3t crossed it on the fluency scenario (56.0%) and came close on coherence (48.0%). Because the counts are not split into spontaneous and cued detection, the disconfirmation rule cannot be applied cleanly. What can be said is that construct faults were harder to detect, not that they were missed by almost everyone.

König et al. (2022) describe noticing as separating attending from interpreting. The pattern here is consistent with a problem of attending: the measurement claim is a feature that current preparation does not train teachers to see. It is not consistent with a claim that the feature is invisible. In the second cohort, about half the participants did ask, or at least did mark, the question the scenarios were designed to conceal. Cohort 3t was not a replication of Cohort 1t. It was higher on six scenarios, and the two differences that reached an unadjusted .05 were both faults that require the reader to check a claim against the items, rather than simply notice a bad sentence. The incorrect key asks whether lived can be correct after since 2019.

The CEFR scenario asks whether four grammar items can support a level claim (Council of Europe, 2020). Cohort 1t detected those faults at 58.3% and 45.8%, respectively. Cohort 3t detected them at 88.0% and 84.0%, respectively. The language fault, by contrast, was already caught by both cohorts.

The data do not say why. Cohort 3t may have had more assessment coursework, more teaching experience, or a different reading of the task. Background variables were collected and are not included in this file, so the difference is reported, not explained. It is large enough that a single-cohort study could have misdescribed the population. The CEFR result, in particular, should not be quoted from Cohort 1t alone.

Anbreen et al. (2026) found agreement between human and model scores and divergence in the reasoning that produced them. The scenarios here reverse the direction of that problem. The output looks acceptable, and the reasoning is the part a teacher has to notice.

Chow and Ng (2025) and Song et al. (2026) show that educators without programming experience can produce functioning applications. Functioning and validity are different properties. A framework for teacher-created tools that stops at whether the tool runs is incomplete. The question "What does this tool claim about a learner?" belongs in it.

With respect to Chuang and Yan's (2025) four strands, this study sits outside them. Their classroom-assessment strand concerns AI systems used for assessment. The artefacts here are assessment instruments produced by AI and owned by the teacher. When a commercial system misjudges a learner, the vendor holds the validity argument, however thin. When a teacher's generated tool misjudges a learner, nobody does. Within Uzbekistan, the findings complement Baxtiyorova (2025).

Her work asks whether AI-supported methods can improve learner outcomes. The present study asks a prior question about the teacher's capacity to judge the instruments involved. Competence frameworks for Uzbek English teachers need an assessment dimension alongside the technological one. The cohort difference suggests that some groups of trainees already have more of that dimension than others and that a single training recommendation would be too coarse.

CONCLUSION

Generative artificial intelligence has made every English teacher a potential author of assessment instruments. The only review such an instrument usually receives is the glance its author gives it before use. In two TESOL cohorts, that glance caught language faults more often than measurement claims. The gap was about 30 percentage points in both cohorts. It was not a gap between seeing and not seeing. In the second cohort, roughly half the participants detected a fluency score built from a word count, and just under half detected a coherence score built from linking words. The second cohort also caught an over-claimed CEFR level far more often than the first. The implication for teacher education is specific, and it is not uniform. Instruction in using artificial intelligence or in writing better prompts would not have been the missing piece for the fault these cohorts most often missed. The missing piece is the question of what evidence would justify a claim about a learner. That is an assessment question. The cohort difference indicates that this capacity cannot be assumed to be equally absent in every group of trainees.

References:

1. Anbreen, T., Özturan, T., Shrestha, P. N., & Maqsood, A. (2026). Human vs. generative artificial intelligence in writing assessment: Investigating feedback alignment, score validity, and teacher agency. *Educational Assessment, Evaluation and Accountability*. Advance online publication. <https://doi.org/10.1007/s11092-026-09486-z>
2. Baxtiyorova, D. I. (2025). Talabalarning ingliz tili ko'nikmalarini rivojlantirishda sun'iy intellektdan foydalanish metodi-kasini takomillashtirish [Abstract of doctoral dissertation, Bukhara State University]. https://buxdu.uz/media/aftoreferat/Durdona_Avtoreferati_20.08.2025docx.pdf
3. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
4. Chow, M., & Ng, O. (2025). From technology adopters to creators: Leveraging AI-assisted vibe coding to transform clinical teaching and learning. *Medical Teacher*, 47(12), 1927–1929. <https://doi.org/10.1080/0142159X.2025.2488353>
5. Chuang, P.-L., & Yan, X. (2025). Language assessment in the era of generative artificial intelligence: Opportunities, challenges, and future directions. *System*, 134, Article 103846. <https://doi.org/10.1016/j.system.2025.103846>
6. Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>
7. Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
8. Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
9. König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., & Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review*, 36, Article 100453. <https://doi.org/10.1016/j.edurev.2022.100453>
10. Republic of Uzbekistan. (2024). Resolution of the President of the Republic of Uzbekistan No. PP-358 of October 14, 2024, on approval of the Strategy for the Development of Artificial Intelligence Technologies to 2030. <https://lex.uz/ru/docs/7158606>



11. Sarkar, A., & Drosos, I. (2025). Vibe coding: Programming through conversation with artificial intelligence. In Proceedings of the 36th Annual Conference of the Psychology of Programming Interest Group (PPIG 2025). <https://arxiv.org/abs/2506.23253>
12. Song, Y., Choi, S., Kim, J., Kim, Y., Weisberg, L., & Moon, J. (2026). A guiding framework for K-12 teachers in creating AI-powered learning technologies through vibe coding [Preprint]. arXiv. <https://arxiv.org/abs/2607.05406>
13. van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, 24(2), 244–276. <https://doi.org/10.1016/j.tate.2006.11.005>
14. Woo, D. J., Guo, K., & Yu, Y. (2025). A vibe coding learning design to enhance EFL students' talking to, through, and about AI [Preprint]. arXiv. <https://arxiv.org/abs/2509.08854>

- 
- 13.00.00 Pedagogika fanlari
 - 13.00.01 Pedagogika nazariyasi. Pedagogik ta'limotlar tarixi
 - 13.00.02 Ta'lim va tarbiya nazariyasi va metodikasi (sohalar bo'yicha)
 - 13.00.03 Maxsus pedagogika
 - 13.00.04 Jismoniy tarbiya va sport mashg'ulotlari nazariyasi va metodikasi
 - 13.00.05 Kasb-hunar ta'limi nazariyasi va metodikasi
 - 13.00.06 Elektron ta'lim nazariyasi va metodikasi (ta'lim sohaları va bosqichlari bo'yicha)
 - 13.00.07 Ta'limda menejment
 - 13.00.08 Maktabgacha ta'lim va tarbiya nazariyasi va metodikasi
 - 13.00.09 Ijtimoiy pedagogika
 - 07.00.00 Tarix fanlari
 - 19.00.00 Psixologiya fanlari
 - 01.00.00 Fizika-matematika fanlari
 - 02.00.00 Kimyo fanlari
 - 03.00.00 Biologiya fanlari
 - 09.00.00 Falsafa fanlari
 - 10.00.00 Filologiya fanlari
 - 11.00.00 Geografiya fanlari



AKTABGACHA VA AKTAB TA'LIMI

Mas'ul muharrir: Ramzidin Ashurov

Ingliz tili muharriri: Murod Xoliyorov

Musahhih: Alibek Zokirov

Sahifalovchi va dizayner: Iskandar Islomov

2026. №10(1)

© Materiallar ko'chirib bosilganda "Maktabgacha va maktab ta'limi" jurnali manba sifatida ko'rsatilishi shart. Jurnalda bosilgan material va reklamalardagi dalillarning aniqligiga mualliflar ma'sul. Tahririyat fikri har vaqt ham mualliflar fikriga mos kelamasligi mumkin. Tahririyatga yuborilgan materiallar qaytarilmaydi.

"Maktabgacha va maktab ta'limi" jurnali 26.09.2023-yildan O'zbekiston Respublikasi Prezidenti Adminstratsiyasi huzuridagi Axborot va ommaviy kommunikatsiyalar agentligi tomonidan №C-5669363 reyestr raqami tartibi bo'yicha ro'yxatdan o'tkazilgan.
Litsenziya raqami: № 136361.

Manzilimiz: Toshkent shahar, Yunusobod tumani
19-mavze, 17-uy.