

Казарян А.С.
Петросян А.А.
Московский Политехнический Университет
Факультет Информационных Технологий
kazarian.info@gmail.com
a.a.petrosyan@mospolytech.ru

ЭМПИРИЧЕСКИЙ АНАЛИЗ ПРЕДПОЧТЕНИЙ СТУДЕНТОВ В СИСТЕМЕ ОЦЕНИВАНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ

Аннотация. В статье исследуется восприятие студентами системы онлайн-тестирования и их предпочтения относительно альтернативных форматов оценивания в контексте доступности искусственного интеллекта. В рамках исследования было проведено анонимное анкетирование 96 студентов 1–3 курсов Московского Политехнического Университета, обучающихся на Транспортном факультете, факультетах Урбанистики и городского хозяйства (УиГХ), Информационных технологий (ФИТ) и Экономики и управления (ЭиУ). Анализ собранных данных показал, что комплексный формат с устной защитой значительно превосходит по популярности традиционные формы контроля. Не обнаружено связи между критическим отношением к текущим тестам и предпочтением нового формата, что указывает на универсальный характер запроса. Обнаружена сильная отрицательная корреляция между оценкой объективности тестов и восприятием их как формальности. Полученные результаты свидетельствуют о необходимости перепроектирования оценочных процедур в сторону комплексных, диалоговых форм, сочетающих аналитическую работу с устной верификацией. Предлагаются практические рекомендации по внедрению таких форматов.

Ключевые слова: онлайн-тестирование, кризис оценивания, предпочтения студентов, комплексные форматы, устная защита, искусственный интеллект в образовании, эмпирическое исследование, педагогический дизайн.

Abstract. The article examines students' perceptions of the online testing system and their preferences regarding alternative assessment formats in the context of artificial intelligence accessibility. An anonymous survey was conducted among 96 first- to third-year bachelor's students of Moscow Polytechnic University, enrolled at the Faculty of Transport, the Faculty of Urban Studies and Municipal Management, the Faculty of Information Technology and the Faculty of Economics and Management. Analysis of the collected data revealed that the complex assignment with oral defense significantly outperforms traditional forms of assessment in popularity. No relationship was found between a critical attitude towards current tests and the preference for the new format, indicating a universal nature of the demand. A strong negative correlation was identified between the perceived objectivity of tests and their perception as a formality. The findings highlight the need for redesigning assessment procedures toward complex, dialogic formats that combine analytical work with oral verification. Practical recommendations for implementing such formats are proposed.

Keywords: online testing, assessment crisis, student preferences, complex assignments, oral defense, artificial intelligence in education, empirical study, instructional design.

ВВЕДЕНИЕ

Повсеместное внедрение систем дистанционного обучения и стремительное развитие генеративного искусственного интеллекта (ИИ) обнажили системные ограничения традиционного онлайн-тестирования. В условиях доступности ИИ и краудсорсинговых инструментов тесты, ориентированные на проверку репродуктивных знаний, всё чаще фиксируют не столько освоенный материал, сколько умение эффективно обращаться к внешним источникам, что ставит под сомнение валидность ключевого инструмента педагогического измерения. Кризис имеет не техническую, а институциональную природу, укоренённую в более глубоких трансформациях высшего образования [1]. Интеграция технологий искусственного интеллекта в образовательный процесс признаётся неизбежной, что смещает исследовательский фокус с поиска способов запрета на выявление моделей педагогически осмысленного использования этих технологий [2]. Социологические опросы фиксируют неоднозначное отношение российского общества к ИИ: при общем высоком уровне информированности и умеренном доверии к технологии в сфере образования абсолютное большинство респондентов настаивает на его вспомогательной роли, причём молодёжь демонстрирует заметно более высокий уровень доверия [3].

Большинство современных исследований концентрируется на технических методах детекции недобросовестного поведения или этических аспектах использования ИИ [4]. При этом значительно меньше внимания уделяется тому, как сами студенты – ключевые участники образовательного процесса – воспринимают сложившийся кризис и какие альтернативные формы оценивания они готовы поддержать. Заполнение данного пробела необходимо для перехода от стратегии реагирования на последствия к содержательному перепроектированию оценочных процедур.

Цель настоящей работы состоит в эмпирическом выявлении того, как студенты университета оценивают текущую систему онлайн-тестирования и какие форматы итогового контроля они предпочитают.

Студентам предлагалось оценить четыре альтернативных формата итогового контроля: (1) очные контролируемые тесты на ПК; (2) традиционные письменные работы (по типу «Достаём двойные листочки»); (3) комплексные задания с возможностью использования любых источников, включая ИИ, и обязательной устной защитой; (4) устные опросы (собеседования).

В работе проверяются следующие **гипотезы**:

1. Комплексное задание с устной защитой получает более высокие оценки по сравнению с альтернативными форматами.
2. Существует положительная связь между восприятием тестов как формальности и предпочтением комплексных заданий.
3. Существует отрицательная связь между оценкой объективности тестов и восприятием их как формальности.

Научная новизна исследования заключается в смещении фокуса с контроля и предотвращения нарушений на восприятие системы оценивания самими студентами в

условиях широкой доступности искусственного интеллекта. В отличие от работ, сосредоточенных на методах детекции недобросовестного поведения, данная статья выявляет, какие форматы оценивания сами студенты считают адекватными современным образовательным целям, и связывает результаты с задачами педагогического проектирования. Обнаружено, что комплексные задания с устной защитой воспринимаются положительно большинством респондентов независимо от их отношения к тестам. Это обосновывает переход к диалоговым форматам как элемент осознанного перепроектирования оценочных процедур в логике педагогического дизайна.

Статья состоит из введения, описания методологии, представления результатов, их обсуждения, выводов и практических рекомендаций, а также указания ограничений исследования.

МЕТОДОЛОГИЯ

Для достижения поставленной цели было проведено количественное исследование методом анонимного анкетирования. Выбор анонимного формата обоснован необходимостью минимизировать эффект социальной желательности: отсутствие персональных данных позволило респондентам высказываться свободно, что особенно важно при изучении отношения к сложившейся системе оценивания.

В исследовании приняли участие 96 студентов бакалавриата 1–3 курсов Московского Политехнического Университета, обучающихся на Транспортном факультете, факультетах Урбанистики и городского хозяйства (УиГХ), Информационных технологий (ФИТ) и Экономики и управления (ЭиУ). Таким образом, выборка охватывает инженерно-технический и экономико-управленческий профили подготовки.

Выборка формировалась методом добровольного отклика в ходе очного взаимодействия в учебных аудиториях с предварительного разрешения преподавателей. Анкетирование проводилось на распечатанных бланках в течение 37 дней (с 3 февраля по 12 марта 2026 г.). Лаконичный формат анкеты — один печатный лист с тремя закрытыми вопросами и полем для опционального комментария на обороте — позволил органично встроить процедуру в ход учебного занятия, не нарушая образовательный процесс и не вызывая утомления респондентов.

Основным инструментом сбора данных выступила специально разработанная анкета (рисунок 1). Она комплексно охватывала основные форматы оценивания, используемые в высшем образовании: очные и электронные тесты, письменные работы, комплексные задания и устные опросы. Два первых вопроса — о восприятии объективности тестов в Moodle и об отношении к ним как к формальности — оценивались по шкале согласия из пяти градаций (от «полностью согласен» до «совсем не согласен»). Третий вопрос предлагал оценить отношение к каждому из четырёх форматов итогового контроля по пятибалльной шкале (от 1 — «крайне негативно» до 5 — «крайне позитивно»). Выбор пятибалльной шкалы обоснован её привычностью для студентов в образовательном контексте и достаточной дифференцирующей способностью для фиксации градаций от явного неприятия до полной поддержки. Закрытый перечень вопросов обеспечил стандартизацию ответов и возможность количественного анализа.

Ваш курс (1-4):		Ваш факультет (Например, ФИТ):	
1. Результаты моих тестов в Moodle (СДО, LMS) объективно отражают уровень моих знаний. Поставьте ✓ рядом с нужным ответом		2. При прохождении тестов в Moodle (СДО, LMS) у меня возникает чувство, что это просто формальность, необходимая для отчётности. Поставьте ✓ рядом с нужным ответом	
полностью согласен		полностью согласен	3. Как Вы относитесь к следующим формам оценки знаний? Поставьте оценку от 1 до 5 для каждого варианта, где 1 — Крайне негативно, 5 — Крайне позитивно: А) Очные тесты в аудитории: использование специального ПО на компьютерах, которое ограничивает доступ только к разрешенным преподавателем материалам и тесту
скорее согласен		скорее согласен	
затрудняюсь ответить		затрудняюсь ответить	Б) Письменные работы и тесты («Достаём двойные листочки»)
скорее не согласен		скорее не согласен	
совсем не согласен		совсем не согласен	В) Комплексные задания (эссе, анализ кейса) с возможностью использования любых источников, включая ИИ, но с обязательной последующей устной защитой, где нужно объяснить свою логику и аргументировать решения Г) Устные опросы (т.н. "Собеседования")
4. На обратной стороне этого листа Вы можете оставить обратную связь по этой теме. Спасибо за участие!			

Рисунок 1 – Бланк анкеты, использованный при сборе данных

Все заполненные бланки были проверены исследователем вручную на полноту заполнения: анкеты с пропусками или невалидными ответами не были включены в анализ. Данные были перенесены в электронную таблицу для последующего анализа.

Настоящее исследование имеет ряд ограничений, которые необходимо учитывать при интерпретации результатов. Во-первых, выборка ограничена студентами одного университета (Московский Политех), что не позволяет напрямую распространять выводы на всю популяцию студентов высшего образования. Во-вторых, участники отбирались на основе добровольного участия, что могло привести к смещению в сторону более заинтересованных и активных студентов. В-третьих, исследование основано на самооценках, поэтому отражает восприятие студентов, а не их реальное поведение в процессе обучения. Кроме того, исследование имеет разовый (кросс-секционный) характер и не позволяет отследить изменения взглядов во времени.

АНАЛИЗ ДАННЫХ

Статистическая обработка данных выполнялась с использованием языка программирования Python (библиотеки Pandas, SciPy и Matplotlib). Все расчёты выполнены автоматически с помощью стандартных статистических функций библиотеки SciPy. Визуализация результатов выполнена с использованием библиотеки Matplotlib.

Для изучения взаимосвязей между переменными применялся коэффициент корреляции Пирсона (r). Значимость корреляции оценивалась по стандартному критерию: связь считалась статистически значимой при $p < 0,05$. Корреляционный анализ проводился для всех пар переменных, соответствующих выдвинутым гипотезам.

Для проверки гипотезы 1 о том, что комплексные задания оцениваются выше альтернативных форматов, выполнялось сравнение средних значений с помощью t-теста для связанных выборок. Для контроля множественных сравнений (три попарных сравнения: комплексные задания с тремя другими форматами оценивания) использовалась поправка Бонферрони: результат считался значимым при $p < 0,0167$.

РЕЗУЛЬТАТЫ

Комплексное задание получило наиболее высокий средний балл (рисунок 2). Средние оценки четырёх предложенных форматов итогового контроля составили: контролируемое тестирование – 2,88, письменная работа – 2,69, комплексное задание с устной защитой – 3,90, устный опрос – 3,42. Таким образом, гипотеза 1 подтверждается: студенты значимо выше оценивают комплексный формат с устной защитой по сравнению с другими формами контроля.

Корреляция между восприятием тестов как формальности и оценкой комплексных заданий оказалась близка к нулю и статистически незначима ($r = 0,006$; $p = 0,95$). Это означает, что предпочтение комплексных заданий не связано с тем, считает ли студент текущие тесты формальностью. Гипотеза 2 не находит эмпирического подтверждения.

Обнаружена сильная отрицательная корреляция (рисунок 3) между оценкой объективности тестов и восприятием их как формальности: $r = -0,59$; $p < 0,001$. Чем выше студент оценивает объективность тестов, тем реже он считает их формальностью. Данный результат подтверждает внутреннюю согласованность ответов респондентов: студенты, оценивающие объективность тестов выше, закономерно реже воспринимают их как формальность. Гипотеза 3 подтверждается.

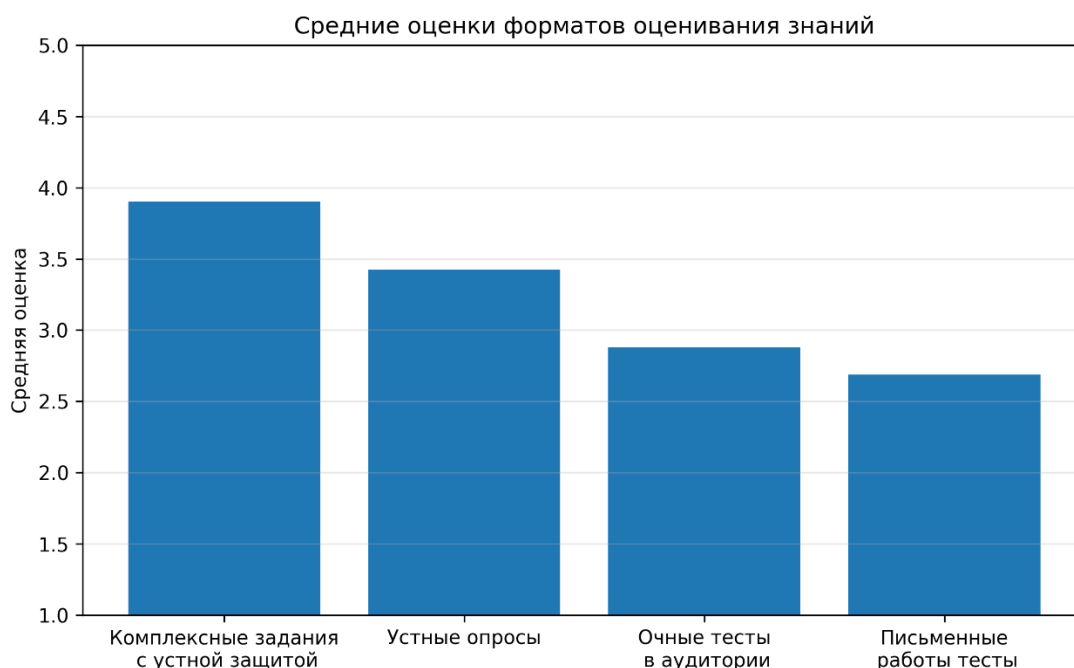


Рисунок 2 – Средние оценки форматов оценивания знаний

В открытых комментариях респонденты отмечали, что формат с защитой позволяет «проявить понимание, а не память». Студенты первого курса чаще связывали выбор комплексного задания со «справедливостью» оценки, тогда как старшекурсники акцентировали «полезность для формирования профессиональных навыков». Эти высказывания согласуются с количественными данными о предпочтении комплексных заданий.

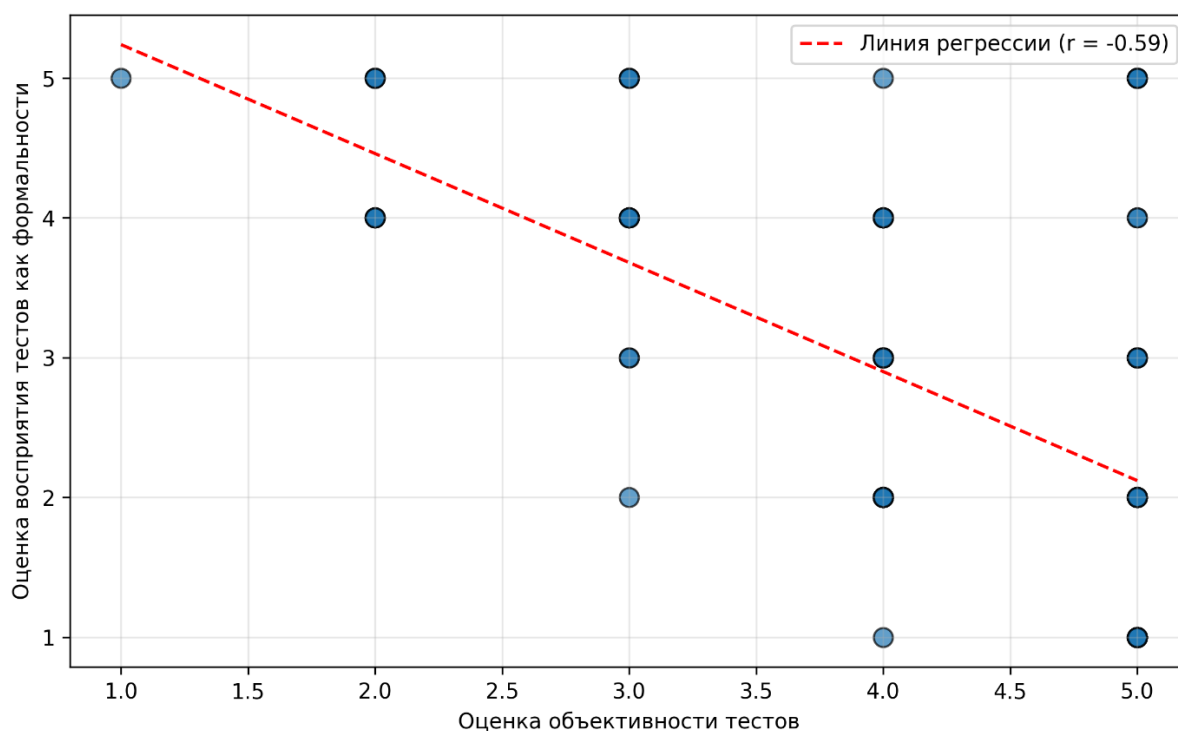


Рисунок 3 – Корреляция между оценкой объективности тестов и восприятием их как формальности

ОБСУЖДЕНИЕ

Полученные результаты позволяют интерпретировать выявленное предпочтение студентов как проявление педагогической и технологической устарелости существующего дизайна оценивания, а не как сугубо этическую проблему [5]. Подтверждение гипотезы 1 указывает на наличие устойчивого запроса на содержательные, требующие анализа и аргументации формы контроля.

Эффективность перехода к таким формам подтверждается исследованиями [6], показывающими влияние формирующего оценивания на развитие компетенций. Отсутствие значимой связи между восприятием тестов как формальности и предпочтением комплексных заданий означает, что запрос на новый формат не ограничивается группой студентов, критически настроенных к текущей системе, а носит универсальный характер. Комплексные задания воспринимаются студентами не столько как альтернатива традиционному тестированию, сколько как принципиально иная, более адекватная их представлениям об оценивании форма, привлекательная сама по себе. Данное наблюдение согласуется с выводами [7] о необходимости внедрения альтернативных методов оценивания, выходящих за рамки традиционных тестов.

Подтверждение гипотезы 3 демонстрирует внутреннюю согласованность студенческих оценок: те, кто верит в объективность тестов, закономерно не считают их пустой формальностью. Этот результат служит дополнительной проверкой валидности ответов респондентов и свидетельствует о том, что студенты подходят к оценке системы осмысленно, а не отвечают произвольно.

Данные выводы согласуются с результатами исследований [8], согласно которым аутентичные письменные задания не являются универсальным решением против использования ИИ и необходим сдвиг к оцениванию процессов через устные диалоговые

форматы, такие как защита, собеседование. Выявленный запрос на сочетание цифровых инструментов с устной верификацией находит своё теоретическое обоснование в концепции дополненного интеллекта (augmented intelligence) [9], которая предлагает модель синергичного партнёрства, где искусственный интеллект обрабатывает данные, а человек сохраняет за собой функции критического суждения.

Важным результатом представляется и то, что студенты, оценивающие комплексные задания выше, также выше оценивают устные опросы, что подтверждается положительной корреляцией между этими переменными. Данная связь, обнаруженная в ходе разведочного анализа, указывает на востребованность устной коммуникации как неотъемлемого компонента оценивания.

В то же время анализ средних оценок комплексного задания по факультетам и курсам позволяет предположить наличие профильной специфики, хотя различия между факультетами не достигают статистической значимости ($p = 0,50$) и требуют дополнительного изучения. Это же касается и курса обучения ($r = 0,13$; $p = 0,22$).

Включение в анкету контролируемого очного тестирования на ПК было обусловлено не только его ролью как альтернативы, но и его диагностическим потенциалом. В отличие от письменных работ на бумаге, требующих ручного ввода данных, и устных опросов, не оставляющих цифрового следа, контролируемый тест на ПК автоматически фиксирует детализированную информацию о процессе выполнения заданий — время ответа, последовательность действий, результаты по отдельным вопросам. Это создаёт основу для применения методов анализа образовательных данных (Educational Data Mining, EDM), позволяя выявлять типичные ошибки, отслеживать прогресс и адаптировать обучение. Таким образом, хотя студенты отдали предпочтение более открытому комплексному заданию, контролируемое компьютерное тестирование сохраняет свою ценность как инструмент сбора детализированных данных для анализа образовательного процесса. Следовательно, путь преодоления кризиса лежит в содержательном перепроектировании системы оценивания, ориентированном на комплексную деятельность. Такое перепроектирование способно не только повысить удовлетворённость студентов, но и генерировать качественные данные для анализа образовательных процессов EDM, смещая фокус с единовременного контроля на анализ индивидуального прогресса [10].

Подобные изменения требуют также развития оценочной самостоятельности студентов через систематическую обратную связь и вовлечение в процесс оценивания [11].

ЗАКЛЮЧЕНИЕ

Проведённое исследование позволило проверить три гипотезы. Гипотеза 1 о том, что комплексное задание с устной защитой оценивается студентами значимо выше альтернативных форматов, полностью подтвердилась. Гипотеза 2 о положительной связи между восприятием тестов как формальности и предпочтением комплексного задания не нашла эмпирического подтверждения — связь оказалась практически нулевой, что указывает на универсальный характер запроса на новый формат. Гипотеза 3 об отрицательной связи между оценкой объективности тестов и восприятием их как формальности подтвердилась, демонстрируя внутреннюю согласованность ответов.

Таким образом, кризис онлайн-тестирования носит системный, а не технический характер и вызван не внешними технологическими угрозами, а несоответствием устаревшего дизайна оценивания современным образовательным целям и ожиданиям студентов. Полученные данные указывают на необходимость перехода от контроля

репродуктивных знаний к оценке компетенций через комплексные, диалоговые формы с обязательной верификацией понимания. Стратегическая интеграция ИИ в учебный процесс, учитывающая выявленные предпочтения студентов, может способствовать преодолению кризиса [12].

На основании выводов сформулированы следующие практические рекомендации:

1. Целесообразна разработка методического каркаса для нового поколения заданий, основанного на принципах контекстуализации, аргументативной прозрачности и легализации цифровых инструментов с обязательной педагогической рефлексией.
2. Перепроектирование итогового контроля может быть дополнено интеграцией циклов формирующего оценивания, развивающего у студентов способность к самостоятельной оценке своей работы, что согласуется с современными тенденциями в образовании [13].
3. Рекомендуется инициировать пилотные проекты по внедрению комплексных форматов с устной верификацией для апробации предлагаемых альтернативных методов.

Направления будущих исследований включают проведение сравнительных исследований в нескольких университетах, лонгитюдный анализ влияния новых форматов на образовательные результаты, а также качественное изучение мотивации студентов к выбору комплексных заданий.

СПИСОК ЛИТЕРАТУРЫ

1. Leaton Gray S., Edsall D., Parapadakis D. AI-Based Digital Cheating At University, and the Case for New Ethical Pedagogies // *Journal of Academic Ethics*. 2025. Vol. 23. P. 2069–2086. <https://doi.org/10.1007/s10805-025-09642-y>.
2. Бермус А.Г. Искусственный интеллект в высшем образовании: синтез современных исследований (обзор) // *Высшее образование в России*. 2024. № 33(5). С. 9–31.
3. ВЦИОМ. Искусственный интеллект: доверие и опасения россиян. Аналитический обзор. 2024. URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/doverie-k-ii>.
4. Fawns T., Bearman M., Dawson P. et al. Authentic assessment: from panacea to criticality // *Assessment & Evaluation in Higher Education*. 2025. Vol. 50(3). P. 396–408. <https://doi.org/10.1080/02602938.2024.2404634>.
5. UNESCO. What's worth measuring in the future of assessment in the AI age. 2025. URL: <https://www.unesco.org/en/articles/whats-worth-measuring-future-assessment-ai-age>.
6. Bellido-García R.S., Venturo-Orbegoso C.O., Cruzata-Martínez A. et al. Involvement of the student in their learning: Effects of formative assessment on competency development // *Eurasia Journal of Mathematics, Science and Technology Education*. 2024. Vol. 20(5), em2440. <https://doi.org/10.29333/ejmste/14453>.
7. Хакимов Р.Р. Альтернативные методы оценки результатов образования — актуальность и проблемы реализации // *Проблемы педагогики*. 2019. №4 (43). URL: <https://cyberleninka.ru/article/n/alternativnye-metody-otsenki-rezultatov-obrazovaniya-aktualnost-i-problemy-realizatsii>.

8. Kofinas A.K., Tsay C.-H., Pike D. The impact of generative AI on academic integrity of authentic assessments within a higher education context // *British Journal of Educational Technology*. 2025. Vol. 56. P. 2522–2549. <https://doi.org/10.1111/bjet.13585>.
9. Kazarian A.S. Augmented intelligence: research on the effectiveness of collaboration between humans and artificial intelligence // *Материалы конференции*. 2025. P. 24–26. EDN VNVWBH.
10. Котюрова И.А. Применение образовательных данных в предметном обучении в вузе // *Вестник Российского университета дружбы народов. Серия: Информатизация образования*. 2024. Т. 21, № 2. С. 227–241. <https://doi.org/10.22363/2312-8631-2024-21-2-227-241>.
11. Solis Trujillo B.P., Velarde-Camaqui D., Gonzales Nuñez C.A. et al. The current landscape of formative assessment and feedback in graduate studies: a systematic literature review // *Frontiers in Education*. 2025. Vol. 10:1509983. <https://doi.org/10.3389/feduc.2025.1509983>.
12. Абалян Ж.А., Пивнева С.В. Потенциал и риски использования искусственного интеллекта в высшей школе при обучении иностранному языку в профессиональной сфере // *Международный научно-исследовательский журнал*. 2024. №3 (106). URL: <https://cyberleninka.ru/article/n/potentsial-i-riski-ispolzovaniya-iskusstvennogo-intellekta-v-vysshey-shkole-pri-obuchenii-inostrannomu-yazyku-v-professionalnoy>.
13. АНО «Цифровая экономика». Влияние искусственного интеллекта на образование. Аналитический отчет. 2024.