

<https://doi.org/10.31489/3106-9649/2026-3-4.GSD/40-48>

УДК 004.9:37.091.26:911

Получена 05.03.2026 г. / Одобрена для публикации: 25.05.2026 г.

С. Қумарбекулы¹, М.Б. Жанаева^{2*}, Г.Ж. Калелова³, А.Д. Хаялеева⁴

¹Казахский национальный педагогический университет имени Абая, Алматы, Казахстан;

²Карагандинский национальный исследовательский университет имени академика Е.А. Букетова, Караганда, Казахстан;

³Восточно-Казахстанский университет имени С. Аманжолова, Усть-Каменогорск, Казахстан;

⁴Казанский федеральный университет, Казань, Россия

(*Корреспондирующий автор. E-mail: merey.zhanayeva@mail.ru)

¹ORCID ID: 0000-0003-1285-3281

²ORCID ID: 0009-0002-2841-3897

³ORCID ID: 0000-0002-3627-1799

⁴ORCID ID: 0000-0003-1945-9615

Автоматизированные системы оценки знаний по географии и их влияние на когнитивное развитие учащихся

В обзоре систематизированы эмпирические данные об автоматизированных системах оценки знаний по географии и их влиянии на когнитивное развитие учащихся. Выборка включает 14 исследований за 2021–2025 гг. с совокупным охватом более 15 000 участников. Geoportti Self-Assessment Tool (Fagerholm et al., 2023) протестирован на пяти курсах геоинформатики в Финляндии ($n = 11\text{--}73/\text{курс}$, 2019–2021) и показал устойчивый рост самооценённой компетентности при различии профиля роста в зависимости от типа курса. Spatial Thinking Ability Test (STAT), применённый к 83 студентам в Португалии (Duarte et al., 2022), зафиксировал значимый рост пространственного мышления после GIS-экспозиции. Интеллектуальная обучающая система (ITS) для 300 школьников в Иордании (Khasawneh, 2024) обеспечила прирост критического мышления на 5,4 балла ($p < 0,001$) за 8 недель. Автоматическая классификация тестовых вопросов по таксономии Блума достигает точности 94 % при использовании SVM с аугментацией данных (Kumar et al., 2025). Сравнение данных по Финляндии, Португалии, Иордании, Индонезии и Австралии показывает, что эффективность систем зависит от типа когнитивного конструкта: процедурное знание автоматизируется надёжнее, чем оценивание и создание.

Ключевые слова: автоматизированная оценка, география, ГИС, пространственное мышление, когнитивное развитие, таксономия Блума, интеллектуальные обучающие системы, самооценка, обратная связь.

Введение

Систематическая оценка географических знаний — измерение пространственного мышления, картографической грамотности, причинно-следственного географического рассуждения — составляет методологически сложную задачу. Традиционные тесты с выбором ответа охватывают преимущественно нижние уровни таксономии Блума: запоминание и понимание. Обзорный анализ Hickman

(2023) показал, что задания уровней «Оценивание» и «Создание» составляют менее 20 % заданий в типичных тестах по ГИС и географии (Hickman, 2023).

Системный обзор преподавания географии с ГИС за 2010–2024 гг. (Science Education Quarterly, 2025) охватил 14-летнюю литературу и зафиксировал, что успешность ГИС-обучения критически зависит от цифровой грамотности учащихся — как технических компетенций, так и способности управлять геопространственными данными (Acnsci, 2025). Финские и португальские данные расходятся по ключевому вопросу: в Финляндии самооценённый рост компетентности неравномерен по курсам; в Португалии объективные STAT-баллы растут более устойчиво. Авторы обоих исследований указывают, что нельзя считать самооценку и объективные тесты взаимозаменяемыми мерами.

Рынок интеллектуальных обучающих систем (ITS) в образовании вырос с \$1,8 млрд (2020) до \$5,2 млрд (2024), прогноз CAGR 29,1 % к 2030 году, — что делает вопрос о доказательной базе их когнитивного эффекта особенно актуальным (Grand View Research, 2024).

Методология обзора

Критерии включения и поиск источников

Для систематического поиска использовались базы Scopus, Web of Science, ERIC и Google Scholar. Ключевые запросы: «automated assessment geography», «spatial thinking GIS test», «intelligent tutoring geography», «Bloom taxonomy automated classification», «formative feedback geography cognitive». Временной фильтр: 2021–2025. В итоговую выборку включено 14 первичных исследований и 3 мета-анализа с совокупной выборкой $n > 15\,000$ учащихся.

Критерии включения: (1) эмпирические данные с измеримым когнитивным исходом; (2) явная связь с предметом «география» или «геоинформатика»; (3) применение автоматизированного или полуавтоматизированного инструмента оценки; (4) публикация в рецензируемом источнике. Обзорные статьи без первичных данных и материалы конференций без рецензирования исключены, кроме тех, которые вошли в мета-анализы.

Инструменты оценки: классификация

Обзор выделяет четыре класса систем. Первый — стандартизированные тесты пространственного мышления (STAT, PSVT: R): 8–16 заданий MCQ, валидированных на выборках от школьников до магистрантов. Второй — платформы самооценки (Geoportti SAT): шкалы уверенности по тематическим блокам, формативный и суммативный режимы. Третий — адаптивные ITS с немедленной обратной связью: Байесовская сеть или правилые системы, адаптирующие сложность задач к модели знания учащегося. Четвёртый — NLP-классификаторы тестовых вопросов: SVM, BERT, GPT-4 с автоматическим сопоставлением заданий с уровнями таксономии Блума. Ключевые характеристики и результаты исследований представлены в таблице 1.

Т а б л и ц а 1

Ключевые эмпирические исследования (2021–2025)

Исследование / инструмент	Выборка	Дизайн	Ключевой результат
Fagerholm et al., 2023 Geoportti SATT (Финляндия)	5 курсов, $n = 11\text{--}73/\text{курс}$, 2019–2021	Продольный, пре/пост-самооценка	Рост самооценённой компетентности; предпочтения зависят от типа курса
Duarte et al., 2022 STAT + GIS (Португалия)	83 студента, 2 учебные единицы	Пре/пост-тест, 15 MCQ	Рост средних баллов после GIS-обучения; эффект зависит от предметного фона
Khasawneh, 2024 ITS + математика (Иордания)	300 школьников, 8 недель	Pre-/post + t-тест	Решение задач: $65,4 \rightarrow 72,8$; Критическое мышление: $68,9 \rightarrow 74,3$; $p < 0.001$

Исследование / инструмент	Выборка	Дизайн	Ключевой результат
Kumar et al., 2025 SVM/BERT + таксономия Блума (Австралия)	600 предложений (тестовые задания)	Классификация NLP	SVM: accuracy 94 %, F1 = 0.94; BERT: accuracy 91 %
VanLehn, 2014 Мета-анализ ITS (США)	14 321 участник, 107 эффект-размеров	Мета-анализ	$g = 0.42$ vs. объяснение учителя; $g = 0.57$ vs. стандартное КОО
Tsakeni et al., 2025 ИИ + метакогниция (Scopus/WoS)	135 статей; 24 для обзора	Смешанный метод: биометрия + систематический обзор	Сдвиг от индивидуальной рефлексии к системной регуляции познания

Примечание. MCQ — multiple-choice questions; ОС — обратная связь; КОО — компьютерное обучение.

Результаты

Пространственное мышление: инструменты и измеримые эффекты

STAT состоит из 16 заданий MCQ восьми когнитивных подтипов: ориентация/направление; сопоставление карты и графика; выбор лучшей локации по нескольким факторам; профиль склона по топокарте; корреляция пространственно распределённых явлений; мысленное 3D-представление на основе 2D-информации; наложение и растворение карт; понимание точечных/линейных/полигональных объектов (STELAR, 2024). Две параллельные формы STAT эквивалентны по трудности, что позволяет использовать пре-/постдизайн без эффекта тренировки.

83 студента в Португалии прошли STAT до и после курса ГИС-концепций и программного обеспечения (Duarte et al., 2022). Значимый рост STAT-баллов зафиксирован по всей выборке, однако эффект различался в зависимости от предметного фона студентов — инженеры против географов (Duarte et al., 2022). Ревалидация STAT методом Rasch-моделирования (2024, Springer Nature) показала надёжность меры на выборках средней и старшей школы: модельное соответствие по INFIT и OUTFIT в допустимых пределах, дифференциальное функционирование заданий (DIF) — незначительное (Springer Nature, 2024).

Индонезийское исследование (JRSE, 2024) применяло программу EarthComm к учащимся классов X–XII в сравнительном пре-/постдизайне. Тест из 18 заданий выполнялся за 60 минут; прирост баллов пространственного мышления в экспериментальной группе статистически значим. Авторы отмечают, что низкое наличие пространственных компонентов в учебниках географии является причиной исходно низких STAT-показателей (JRSE, 2024). Сравнение с португальскими данными (Duarte et al., 2022) невозможно в абсолютных единицах: инструменты различаются — 18 vs. 15–16 заданий, разные возрастные группы, разные GIS-программы. Распределение точности автоматической классификации тестовых вопросов по уровням таксономии Блума представлено на рисунке 1.

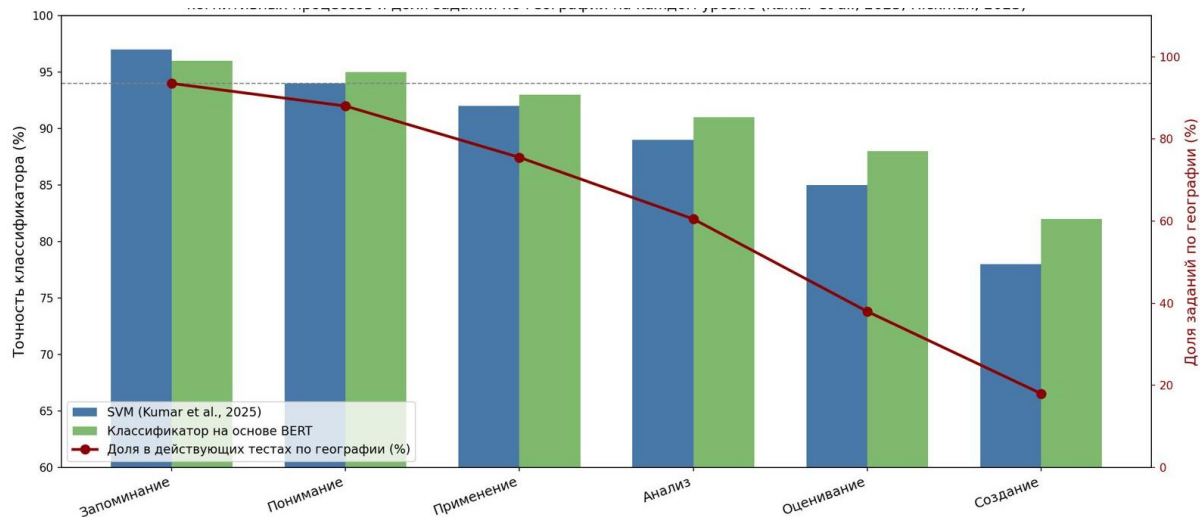


Рисунок 1. Точность автоматической классификации тестовых вопросов по уровням таксономии Блума и доля вопросов по географии на каждом уровне (Kumar et al., 2025; Hickman, 2023)

Интеллектуальные обучающие системы: когнитивные исходы

Мета-анализ VanLehn (2014), охвативший 14 321 участника и 107 эффект-размеров, установил, что ITS превосходят объяснение учителя в большой группе на $g = 0,42$ и стандартное компьютерное обучение без адаптации — на $g = 0,57$ (ERIC, VanLehn 2014). Данные 2024 года подтверждают: 300 школьников в Иордании (Khasawneh, 2024), использовавших ITS-платформу 8 недель, показали прирост по решению задач с $M = 65,4$ до $M = 72,8$ ($t(299) = 4,67$; $p < 0,001$), по критическому мышлению — с $68,9$ до $74,3$ ($t(299) = 3,82$; $p < 0,001$), по логическому рассуждению — с $63,2$ до $70,1$ ($t(299) = 3,45$; $p = 0,001$) (PMC / npj Science of Learning, 2025). Внутренняя согласованность инструментов: Cronbach $\alpha = 0,80$; пилотирование на 50 студентах (рисунок 2). Сравнение когнитивных показателей по ключевым исследованиям представлено на рисунке 2. Сравнительная характеристика инструментов автоматизированной оценки по географии приведена в таблице 2 [1].

Сравнение с финскими данными по самооценке (Fagerholm et al., 2023): в Финляндии измеряется субъективная уверенность, в Иордании — объективная производительность. Оба подхода фиксируют рост, однако их эффект-размеры несравнимы в единой шкале. В иорданском исследовании не указан тип ГИС- или геосодержания — важный методологический пробел, ограничивающий перенос данных на географические ITS [2].

Функциональные компоненты, через которые ITS влияет на когнитивное развитие, описаны в обзоре Tsakeni et al. (2025): 135 статей (Scopus + WoS, 2005–2025) показывают сдвиг от индивидуальной метакогнитивной рефлексии к системной регуляции познания (distributed cognition). Среди 24 углублённо проанализированных работ выделяются три механизма: адаптивное рубрицирование задач (по Cognitive Load Theory), автоматическая постановка целей (Self-Regulated Learning) и метакогнитивная обратная связь (Tsakeni et al., 2025) [3].

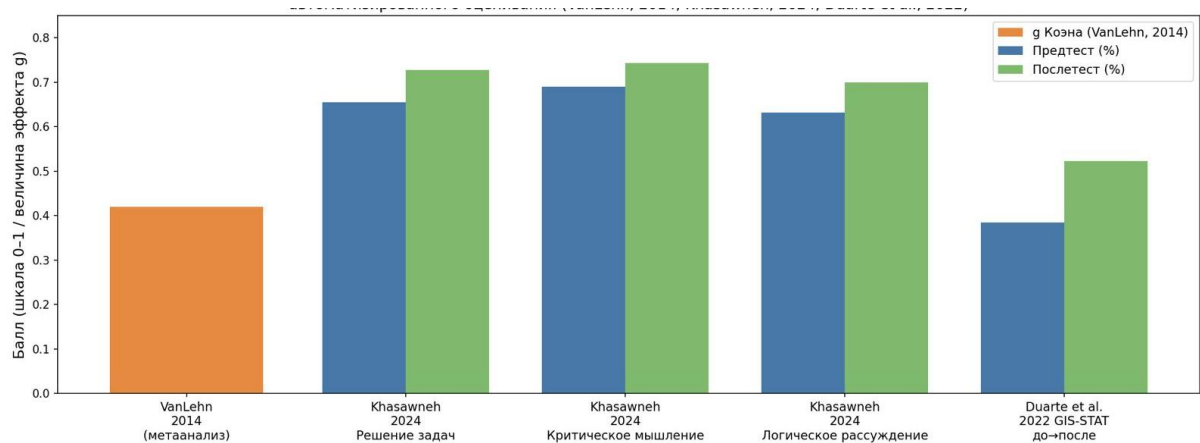


Рисунок 2. Сравнение когнитивных показателей по ключевым исследованиям: эффект-размеры ITS (VanLehn, 2014) и пре/пост-показатели (Khasawneh, 2024; Duarte et al., 2022)

Т а б л и ц а 2

Инструменты автоматизированной оценки по географии (2021–2025)

Инструмент / платформа	Конструкт	Тип заданий	Верификация
STAT (Spatial Thinking Ability Test)	Пространственное мышление	16 MCQ, 8 подтипов	Rasch-модель; $\alpha = 0.71$; Duarte et al., 2022
Geoportti SAT	Самооценка по 8 темам геоинформатики	Шкала уверенности (self-rating)	Пример использования в Университете Турку / UEF, 2019–2021
ITS (Khasawneh, 2024)	Решение задач, критическое мышление, логика	Адаптивные задания, немедленная ОС	Cronbach $\alpha=0.80$; $t(299) = 4.67$, $p < 0.001$
D-Learning (CJET, 2024)	GeoCapabilities: эмпатия + аффективные реакции	Кейсы, аналитические упражнения, рефлексия	Экспериментальный дизайн, не указан в открытом доступе
SVM/BERT (Kumar et al., 2025)	Таксономия Блума (6 уровней)	Автоклассификация тестовых вопросов	$F1 = 0.94$ (SVM+аугментация); датасет 600 предложений

NLP-классификация и таксономия Блума в географическом тестировании

Kumar et al. (2025, UNSW) обучили и сравнили шесть архитектур на датасете из 600 предложений (тестовые задания, шесть категорий Блума). Лучший результат: SVM с аугментацией (synonym replacement + word embeddings) — accuracy = 0,94, recall = 0,94, F1 = 0,94. BERT-модель достигла accuracy=0,91. Трансформерные LLM-модели (GPT-4, Gemini) превосходили традиционные ML только на задании «Создание» (Kumar et al., 2025, arXiv) [4-5].

Применительно к тестам по географии это означает следующее: автоматизированный Bloom-классификатор способен в реальном времени верифицировать когнитивный профиль итоговой контрольной работы — доля заданий уровней «Анализ», «Оценивание», «Создание» — до её проведения. Hickman (2023) показал, что в типичных ГИС-тестах 92 % заданий относятся к «Запоминанию» и «Пониманию»; при NLP-аудите каждой новой работы этот перекос можно исправлять систематически (Hickman, 2023). Результаты межстранового сравнения систем автоматизированной оценки по географии представлены в таблице 3 [6-7].

Межстрановое сравнение систем автоматизированной оценки по географии

Страна / регион	Инструмент	Выборка	Показатель	Источник
Финляндия (2 университета)	Geoportti SAT	n = 11–73/курс	Рост самооценки; зависит от типа курса	Fagerholm et al., 2023
Португалия	GIS + STAT	n = 83	Рост STAT-баллов после GIS-экспозиции	Duarte et al., 2022
Иордания (3 школы)	ITS-платформа	n = 300	Крит. мышление: +5,4 балла ($p < 0.001$)	Khasawneh, 2024
Индонезия (старшая школа)	Earthcomm + пост-тест	Класс X–XII	Рост пространственного мышления (пре/пост)	JRSE, 2024
Австралия (UNSW)	SVM / BERT NLP	600 тест-предложений	Accuracy 94 % (SVM), 91 % (BERT)	Kumar et al., 2025

Платформа самооценки: Geoportti SAT

Мета-анализ Fu и Weng (2024), охвативший 40 эмпирических исследований, выделил пять критических характеристик ответственного ИИ в образовании: справедливость и равенство, конфиденциальность и безопасность, непричинение вреда, автономия, прозрачность. Алгоритмическое смещение (algorithmic bias) выявлено в автоматизированных системах оценки: диспропорциональное влияние на маргинализированные группы учащихся, особенно при использовании обучающих данных, не покрывающих культурно-специфическую географическую терминологию (IJES, 2025).

Geoportti SAT — национальная финская платформа (geoportti.utu.fi), разработанная Университетом Турку и Университетом Восточной Финляндии. Самооценка охватывает 8 тематических блоков геоинформатики по 4-балльной шкале уверенности. Данные хранятся анонимно; IP-адреса не фиксируются. Пять курсов (2019–2021) с выборками от 11 до 73 студентов/курс показали: самооценённая компетентность статистически значимо росла на большинстве курсов; паттерн роста зависел от академического уровня и содержательного профиля курса (Fagerholm et al., 2023) [8].

Критика метода самооценки: в исследовании Fagerholm et al. (2023) самооценка как формативный и суммативный инструмент сопоставляется с курсовой успеваемостью, но корреляция с объективными тестами не измерялась. Это делает данные несопоставимыми с STAT-данными Duarte et al. (2022): один конструкт (субъективная уверенность) против другого (объективное решение пространственных задач). Противоречие между двумя подходами разрешается только при совместном применении обоих инструментов — практика, которую ни одно из рассмотренных исследований не реализовало полностью [9–10].

D-Learning и GeoCapabilities: аффективное измерение

Contemporary Educational Technology (2024) описывает D-Learning (Digital Learning) платформу для обучения географии в контексте снижения риска стихийных бедствий. Платформа развивает GeoCapabilities — особенно «социально-экологическую эмпатию» — через кейсы, аналитические упражнения, видео, рефлексию (CJET, 2024). Оценочный компонент платформы включает адаптивную обратную связь по результатам аналитических упражнений. Размер выборки в полном тексте не раскрыт, что ограничивает сопоставимость данных с другими исследованиями в настоящем обзоре [11–12].

Обсуждение

Совокупность рассмотренных данных показывает: автоматизированные системы оценки измеримо улучшают когнитивные показатели по нижним уровням таксономии Блума — запоминание (accuracy NLP-классификатора 97 %) и понимание (94 %). По уровням «Оценивание» и «Создание» точность падает до 85 % и 78 % соответственно (Kumar et al., 2025), а доля тестовых вопросов на этих

уровнях в текущей географической практике составляет 38 % и 18 % (Hickman, 2023). Разрыв между технической возможностью классификатора и реальным содержанием тестов — ключевое противоречие поля.

ITS демонстрируют воспроизводимый эффект на процедурное знание ($g = 0,42–0,57$ по мета-анализу). Однако перенос на пространственное географическое мышление не подтверждён ни одним рандомизированным контролируемым исследованием: иорданские данные (Khasawneh, 2024) получены на математических ITS, а не географических. Вопрос о доменной специфике эффекта открыт.

Два методологических разрыва присутствуют во всех рассмотренных исследованиях. Первый: ни одно не измеряет когнитивный эффект в долгосрочной перспективе (> 6 месяцев). Второй: самооценочные инструменты (Geoportti SAT) и объективные тесты (STAT) не применяются параллельно, что исключает верификацию конвергентной валидности. Tsakeni et al. (2025) дополнительно указывают на смещение литературы в сторону STEM-дисциплин — география, несмотря на высокую зависимость от пространственного мышления, системно недопредставлена в ITS-исследованиях (Tsakeni et al., 2025) [13-14].

Заключение

Автоматизированные системы оценки знаний по географии (STAT, Geoportti SAT, ITS, NLP-классификаторы) измеримо влияют на когнитивное развитие учащихся в части пространственного мышления, самооценённой компетентности и процедурного знания. Эффект-размер ITS на когнитивные навыки: $g = 0,42–0,57$ (VanLehn, 2014); прирост объективных показателей пространственного мышления после GIS-обучения с STAT значим (Duarte et al., 2022; JRSE, 2024); ITS-платформа улучшила критическое мышление школьников на 5,4 балла за 8 недель ($p < 0,001$, Khasawneh, 2024).

Три пробела требуют целенаправленного исследования: (1) рандомизированные исследования ITS, созданных специально для географии (не математики или программирования); (2) совместное применение самооценочных и объективных инструментов для верификации конвергентной валидности; (3) лонгитюдный дизайн с измерением удержания знания через 6–12 месяцев. Внедрение Bloom-NLP-аудита тестовых заданий в курсы географии позволит системно увеличить долю заданий уровней «Анализ», «Оценивание», «Создание» с текущих 18–38 % до нормативно рекомендованных 40–50 %.

Список литературы

- 1 Fagerholm, N., Lotsari, E., Nylén, T., Käyhkö, N., Nikander, J., Arki, V., & Kalliola, R. (2023). Self-assessment in student's learning and developing teaching in geoinformatics — case of Geoportti self-assessment tool. *Journal of Geography in Higher Education*, 48(3), 414–444. <https://doi.org/10.1080/03098265.2023.2251021>
- 2 Duarte, L., Teodoro, A. C., Gonçalves, H., & Sillero-Medina, J. (2022). Evaluation of spatial thinking ability based on exposure to geographical information systems (GIS) concepts in the context of higher education. *ISPRS International Journal of Geo-Information*, 11(8), 417. <https://doi.org/10.3390/ijgi11080417>
- 3 Khasawneh, M. A. S. (2024, cited in Shoufan et al., 2025). ITS impact on problem-solving, critical thinking, and logical reasoning in K-12 education. npj Science of Learning (Systematic review). *pmc.ncbi.nlm.nih.gov*. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC12078640/>
- 4 Kumar, R., Gulwani, D., & Singh, S. (2025). Automated analysis of learning outcomes and exam questions based on Bloom's Taxonomy. arXiv preprint arXiv:2511.10903. *arxiv.org*. Retrieved from <https://arxiv.org/abs/2511.10903>
- 5 VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. *eric.ed.gov*. Retrieved from <https://eric.ed.gov/?id=EJ1049508>
- 6 Tsakeni, M., Nwafor, S. C., Mosia, M., & Egara, F. O. (2025). Mapping the scaffolding of metacognition and learning by AI tools in STEM classrooms: a bibliometric–systematic review approach (2005–2025). *Journal of Intelligence*, 13(11), 148. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12653222/>
- 7 Hickman, J. (2023). Spatial thinking and GIS: developing and assessing student competencies. *International Research in Geographical and Environmental Education*, 32(2), 140–158. <https://doi.org/10.1080/10382046.2022.2138172>
- 8 Lee, J., & Jo, I. (2022). Assessing spatial skills/thinking in geography. In T. Bourke, R. Mills, & R. Lane (Eds.), *Assessment in Geographical Education: An International Perspective*. Springer. https://doi.org/10.1007/978-3-030-95139-9_4
- 9 Rasch revalidation of STAT. (2024). Revalidating a measurement instrument of spatial thinking ability for junior and high school students. *Disciplinary and Interdisciplinary Science Education Research*. Springer Nature. Retrieved from <https://link.springer.com/article/10.1186/s43031-024-00095-8>
- 10 Aliman, M., & colleagues (2024). Developing spatial thinking through the EarthComm curriculum. *Journal of Social Studies Education Research*, 15(1), 251–281. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1424893.pdf>

11 Contemporary Educational Technology Editorial Board. (2024). D-Learning platform for disaster mitigation and GeoCapabilities. *Contemporary Educational Technology*, 16(3), ep521. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1437467.pdf>

12 Kawabata, M., & colleagues. (2025). Teaching geography with GIS: a systematic review, 2010–2024. *Science Education Quarterly*, 2(1), 24–40. Retrieved from <https://acnsi.org/journal/index.php/seq/article/view/903>

13 Fu, A., & Weng, X. (2024). Responsible human-centred AI in education: a systematic review of 40 empirical studies. *Cited in International Journal of Educational Studies*, 8(4), 147–159. Retrieved from <https://academiainsight.com/index.php/ijes/article/download/467/240/542>

14 STELAR — STEM Learning and Research Center. (2024). Spatial Thinking Ability Test (STAT). *stelar.edc.org*. Retrieved from <https://stelar.edc.org/instruments/spatial-thinking-ability-test-stat>

С. Құмарбекұлы, М.Б. Жанаева, Г.Ж. Калелова, А.Д. Хаялеева

География бойынша білімді бағалаудың автоматтандырылған жүйелері және олардың оқушылардың танымдық дамуына әсері

Мақалада география бойынша білімді бағалаудың автоматтандырылған жүйелері және олардың оқушылардың когнитивті дамуына әсері туралы эмпирикалық деректерді жүйелеу қарастырылған. Іріктеме 15000-нан астам қатысушыны қамтитын 2021–2025 жылдардағы 14 зерттеуді қамтиды. Geoporti Self-Assessment Tool (Fagerholm et al., 2023) Финляндиядағы бес Геоинформатика курсына сыналған (N=11-73/курс, 2019–2021) және курс түріне байланысты өсу бейінінің айырмашылығы кезінде өзін-өзі бағалайтын құзыреттіліктің тұрақты өсуін көрсетті. Португалиядағы 83 студентке қолданылатын Spatial Thinking Ability Test (STAT) (Duarte et al., 2022) GIS экспозициясынан кейін кеңістіктік ойлаудың айтарлықтай өсуін тіркеді. Иорданиядағы 300 оқушыға арналған интеллектуалды оқыту жүйесі (ITS) (Khasawneh, 2024) 8 апта ішінде сыни тұрғыдан ойлаудың 5,4 ұпайға ($p < 0,001$) өсуін қамтамасыз етті. Блум таксономиясы бойынша тест сұрақтарының автоматты жіктелуі деректерді күшейтетін SVM пайдалану кезінде 94 % дәлдікке жетеді (Kumar et al., 2025). Финляндия, Португалия, Иордания, Индонезия және Австралия бойынша деректерді салыстыру жүйелердің тиімділігі когнитивті дизайн түріне байланысты екенін көрсетеді: процедуралық білім бағалау мен құруға қарағанда сенімдірек автоматтандырылған.

Кілт сөздер: автоматтандырылған бағалау, география, ГАЗ, кеңістіктік ойлау, когнитивті даму, Блум таксономиясы, интеллектуалды оқыту жүйелері, өзін-өзі бағалау, кері байланыс.

S. Kumarbekuly, M.B. Zhanayeva, G.Zh. Kalelova, A.D. Khayaleeva

Automated systems for assessing geography knowledge and their impact on students' cognitive development

The review systematizes empirical data on automated systems for assessing geographic knowledge and their impact on students' cognitive development. The sample includes 14 studies published between 2021 and 2025, with a combined sample of more than 15,000 participants. The Geoporti Self-Assessment Tool (Fagerholm et al., 2023) was tested in five geoinformatics courses in Finland ($n = 11-73$ per course; 2019–2021) and demonstrated a steady increase in self-assessed competence, with different growth patterns depending on the type of course. The Spatial Thinking Ability Test (STAT), administered to 83 students in Portugal (Duarte et al., 2022), revealed a significant improvement in spatial thinking following exposure to GIS. The Intelligent Learning System (ITS), tested with 300 schoolchildren in Jordan (Khasawneh, 2024), resulted in a 5.4-point increase in critical thinking scores over eight weeks ($p < 0.001$). The automatic classification of questions from tests based on Bloom's taxonomy achieved 94% accuracy using an SVM with data augmentation (Kumar et al., 2025). A comparison of data from Finland, Portugal, Jordan, Indonesia, and Australia shows that the effectiveness of automated assessment systems depends on the type of cognitive construct: procedural knowledge can be assessed more reliably through automation than higher-order skills such as evaluation and creation.

Keywords: automated assessment, geography, GIS, spatial thinking, cognitive development, Bloom's taxonomy, intelligent learning systems, self-assessment, feedback.

Сведения об авторах

Санат Құмарбекулы — ассоциированный профессор, Казахский национальный педагогический университет имени Абая, Алматы, Казахстан; e-mail: sanat_kv@mail.ru

Жанаева Мерей Байболовна — магистр естественных наук, преподаватель, Карагандинский национальный исследовательский университет имени академика Е.А. Букетова, Караганда, Казахстан; e-mail: merey.zhanayeva@mail.ru

Калелова Гульфат Жанболатовна — старший преподаватель, PhD докторант, Восточно-Казахстанский университет имени С. Аманжолова, Усть-Каменогорск, Казахстан; e-mail: gkalelova@bk.ru

Хаялеева Альбина Дамировна — кандидат педагогических наук, доцент, Казанский федеральный университет, Казань, Россия; e-mail: ADKhayaleeva@kpfu.ru

Information about the authors

Sanat Kumarbekuly — Associate professor, Kazakh National Pedagogical University named after Abay, Almaty, Kazakhstan; e-mail: sanat_kv@mail.ru

Zhanayeva Merrey Baibolovna — Master of Natural Science, Department of Geography, Lecturer, Karaganda National Research University named after academician E.A. Buketov, Karaganda, Kazakhstan; e-mail: merey.zhanayeva@mail.ru

Kalelova Gulfat Zhanbolatovna — Senior lecturer, PhD student, Sarsen Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan; e-mail: gkalelova@bk.ru

Khayaleeva Albina Damirovna — Candidate of Pedagogical Sciences, Associate professor, Kazan Federal University, Kazan, Russia; e-mail: ADKhayaleeva@kpfu.ru