



УДК 37.013.2

ОПТИМИЗАЦИЯ ОБРАЗОВАТЕЛЬНОЙ ОЦЕНКИ: ПРАКТИЧНОСТЬ
КОМПЬЮТЕРНОГО АДАПТИВНОГО ТЕСТИРОВАНИЯ (КАТ)
С ИСПОЛЬЗОВАНИЕМ ТЕОРИИ ОТВЕТОВ НА ЗАДАНИЯ

Слюсарь В.В.

*кандидат технических наук, доцент, доцент Института системной и
программной инженерии и информационных технологий (СПИИТЕХ),
Национальный исследовательский университет «Московский институт
электронной техники»*

г. Зеленоград

vslyusar@mail.ru

Аннотация. Целью данного исследования является разработка системы компьютерных адаптивных тестов (КАТ) с использованием подхода теории ответов на задания. Данное исследование является частью разработки веб-системы с использованием метода исследований и разработок (R&D), применяя модель Four-D (4-D). По своей сути эта система похожа на компьютерный тест. Однако критическое отличие заключается в ее способности рандомизировать и предоставлять вопросы, соответствующие уровню подготовки тестируемого, используя алгоритм теории ответов на задания (Items Response Theory, IRT). В системе используется модель 3-PL из IRT, учитывающая уровень сложности вопросов, дискриминационную способность вопросов, а также вероятность угадывания или вмешательства в вопросы. Экзаменационная система случайным образом распределяет вопросы между студентами на основе их ответов на предыдущие вопросы, гарантируя, что каждый испытуемый получит уникальную последовательность вопросов. Экзамен завершается, когда тестируемый точно оценивает свои способности, либо получены ответы на все

вопросы. Результатом данного исследования является система компьютерных адаптивных тестов (КАТ), основанная на ИРТ, которая может быть использована для оценки результатов обучения студентов. Данное исследование было проведено в институте системной и программной инженерии и информационных технологий НИУ МИЭТ, в качестве исследовательской выборки использовались 90 студентов. Оценка практичности этой системы получила очень высокие баллы, что свидетельствует о том, что система компьютерного адаптивного тестирования (КАТ), основанная на теории ответов на задания, считается очень практичной и эффективной в достижении поставленных целей измерения.

Ключевые слова: компьютерный адаптивный тест; теория ответов на задания; исследование и разработка; модель Four-D; оценка образования.

OPTIMIZATION OF EDUCATIONAL ASSESSMENT:
THE PRACTICALITY OF COMPUTER ADAPTIVE TESTING (CAT)
USING TASK RESPONSE THEORY

Slyusar V. V.

*Candidate of Technical Sciences, Associate Professor, Associate Professor at the
Institute of System and Software Engineering and Information Technology
(SPINTECH), National Research University "Moscow Institute of Electronic
Technology"*

Zelenograd

vslyusar@mail.ru

Annotation. The purpose of this study is to develop a computer adaptive test (CAT) system using an item response theory approach. This research is part of the development of a web-based system using the research and development (R&D) method, applying the Four-D (4-D) model. At its core, this system is similar to a computer test. However, the critical difference is its ability to randomize and present questions appropriate to the test taker's skill level using an Items Response Theory (IRT) algorithm. The system uses the 3-PL model from question answering theory, which takes into account the difficulty level of questions, the discriminatory power of

questions, and the likelihood of guessing or tampering with questions. The examination system randomly assigns questions to students based on their answers to previous questions, ensuring that each test taker receives a unique sequence of questions. The exam ends when the test taker accurately assesses his abilities, or when all questions are answered. The result of this study is a computer adaptive test (CAT) system based on item response theory that can be used to assess student learning outcomes. This study was conducted at the Institute of System and Software Engineering and Information Technologies of the National Research University MIET, using 90 students as a research sample. The usability assessment of this system received very high scores, indicating that the computer adaptive testing (CAT) system, based on item response theory, is considered to be very practical and effective in achieving its measurement objectives.

Keywords: computer adaptive test; Item Response Theory; research and development; Four-D model; educational assessment.

Введение.

Образование тесно связано с оценкой или анализом результатов обучения студентов (Госс, 2022), (Suhendi, Ramdhani, Irwansyah, 2018). Одной из форм оценки в образовании является тестирование, например, ежедневные контрольные работы, промежуточные и выпускные экзамены, которые направлены на оценку достижения студентами целей обучения (Левент, Эрток, 2020). Между тем, оценка в образовании — это оценка роста и прогресса студентов в достижении целей и ценностей, установленных учебным планом (Sun, Zhang, Li, Zhao, Dong, 2020), (Чан, Лук, 2022).

Оценка обучения состоит из трех элементов, которые невозможно разделить: оценка, измерение и используемые инструменты (Нуртанто, Холифах, Масек,

Судира, Самсудин, 2021), (Мерта, Махфуд, 2022), (Gamage, Wijesuriya, Ekanayake, Rennie, Lambert, Gunawardhana, 2020). Для оценки требуется измерение, а для измерения - измерительный инструмент (Кларк, Уотсон, 2019), (Eloranta, Kaltiala, Lindberg, Kaivosoja, Peltonen, 2022). Оценка направлена на определение результатов процесса обучения, а измерение — это оценка, полученная по результатам тестирования. Тесты являются измерительными инструментами используемых при оценке результатов обучения.

В эпоху развития информационных технологий перед образовательным сектором встают все более сложные задачи по оценке успеваемости и эффективности обучения студентов (Алам, 2022), (Рахматулла, Муляса, Сиахрани, Понгпалилу, Путри, 2022), (Винсент-Ланкрин, Ван дер Влис, 2020). Постоянный прогресс в области доступа к технологиям и обилие данных об образовании открыли новые возможности для совершенствования методов оценки понимания учеников (Oktrilani, Delianti, Fajri, Samala, 2023), (Хусна, Хусна, 2023). Однако вместе с этим развитием необходимо эффективно решать не менее важные вопросы тестирования и оценки образования.

Постановка задачи.

Одна из фундаментальных проблем в оценке образования - неспособность традиционных методов оценки учесть различия в уровнях понимания учащихся (Mystakidis, Christopoulos, Pellas, 2022), (Мартинес-Хименес, Руис-Хименес, 2020), (Сингх, Стил, Сингх, 2021). На традиционном экзамене всем студентам предлагаются вопросы одинакового уровня сложности. Это может стать серьезной проблемой, поскольку у каждого студента свой уровень понимания

(Гарднер, 2021), (Путра, Лиривати, Тахрим, Сяфрудин, Аслан, 2020). Некоторые студенты могут хорошо понимать определенные концепции, в то время как другим может потребоваться дополнительная помощь. В результате результаты экзамена часто неточно отражают реальные способности каждого студента.

Кроме того, использование времени и ресурсов на тестирование становится менее эффективным, поскольку учащиеся могут столкнуться с вопросами, не соответствующими их способностям (Нгуен, Кейзман, Хамстон, 2020). Это также может нарушить общий процесс обучения, поскольку студенты могут испытывать разочарование, когда сталкиваются со слишком сложными вопросами, или скуку, когда сталкиваются со слишком легкими вопросами. Эта проблема не только ставит учащихся в невыгодное положение, но и негативно влияет на принятие решений в сфере образования, разработку учебных программ и оценку работы преподавателей. Поэтому существует острая необходимость в решении этой проблемы и совершенствовании методов оценки образования.

В этом контексте использование системы компьютерного адаптивного тестирования (КАТ), интегрированной с подходом Item Response Theory (IRT), предлагает интересное решение. Компьютерное адаптивное тестирование (КАТ) позволяет сделать экзамены более адаптивными, при этом уровень сложности вопросов динамически регулируется в зависимости от успеваемости студентов во время экзамена, (Есингельдинов, Рахымбаева, Жаппарова, Турсынова, 2023), (Boussakuk, Bouchboua, Ghazi, Bekkali, Fattah, 2021). Студенты, которые отвечают правильно, сталкиваются с более сложными вопросами, а те, кто отвечает неправильно, сталкиваются с более простыми вопросами. Это

позволяет более точно оценить способности отдельных студентов. С другой стороны, подход Item Response Theory позволяет провести более глубокий анализ каждого экзаменационного задания (Pliakos, Joo, Park, Cornillie, Vens, Van den Noortgate, 2019), (Bürkner, 2019). Это помогает выявить сильные и слабые стороны экзаменационных вопросов, создавая более качественные и релевантные тесты, основанные на изучаемом материале (Кларк, Уотсон, 2019).

Вопросы и цель исследования.

В статье более подробно описано, как Интеграция компьютерного адаптивного тестирования (КАТ) с подходом Item Response Theory может повысить эффективность оценки образования. Исследователь изучит практическое применение этого метода в различных образовательных контекстах, включая классные комнаты, высшие учебные заведения и стандартизированное тестирование. Исследователь описывает его преимущества в решении существующих проблем оценки, включая повышение точности измерения способностей учащихся, увеличение эффективности использования времени и повышение качества результатов экзаменов. Обсуждается практическая целесообразность внедрения компьютерного адаптивного тестирования (КАТ) с использованием подхода Item Response Theory (IRT) по оценке преподавателей и студентов, использующих эту систему для оценки учебного процесса.

Методы исследования.

В данной статье используется метод, основанный на исследовании и разработке нескольких типов моделей: модель разработки 4-D (Four D)

используется при разработке продукта и состоит из четырех основных этапов: определение, проектирование, разработка и тестирование (Sugiyono, 2021). Данные метод и модель были выбраны для создания продукта в виде системы компьютерного адаптивного тестирования (КАТ) с подходом, основанным на теории ответов на задания.

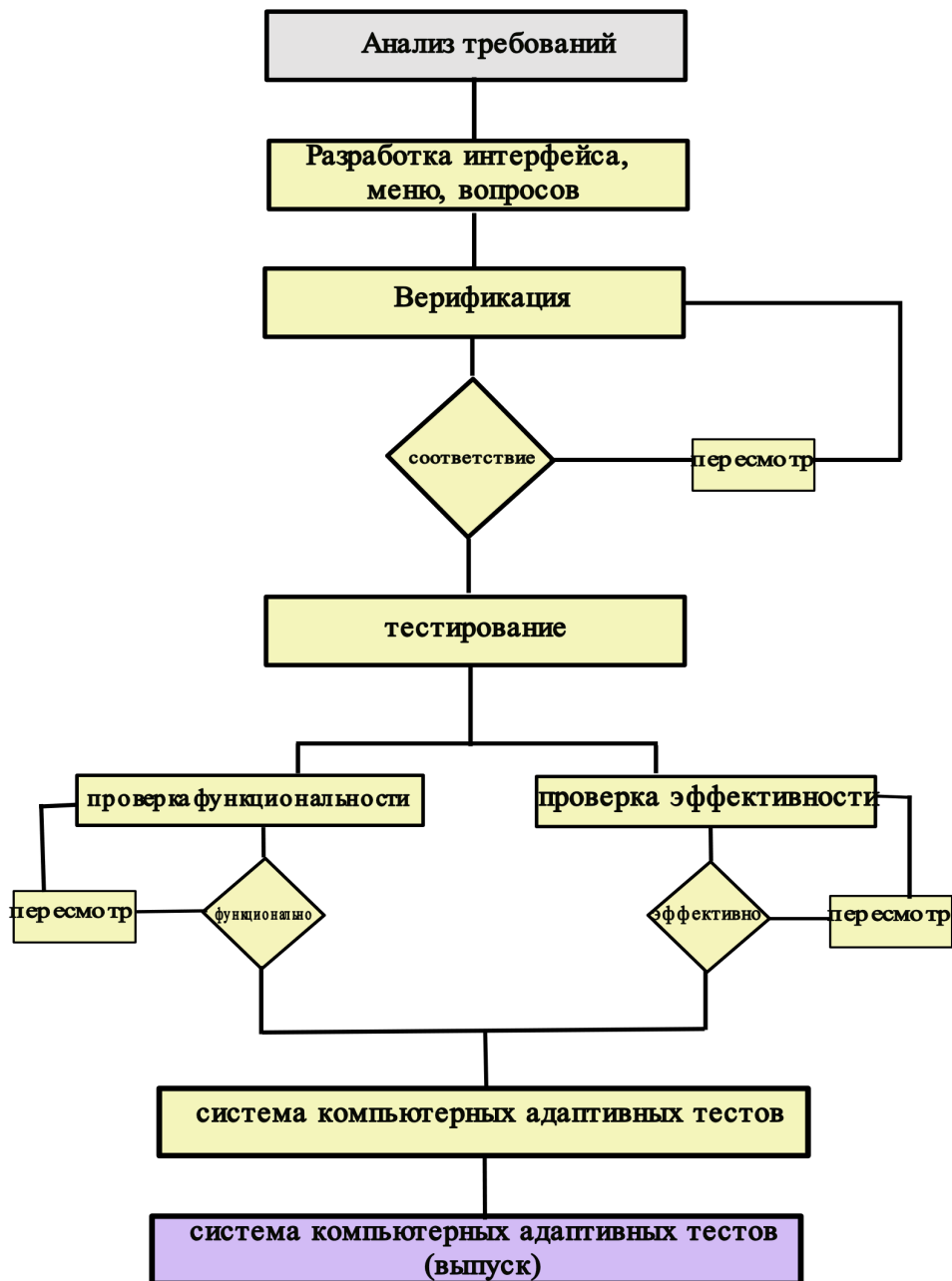


Рисунок 1-Пример модели 4-D для разработки системы компьютерного

тестирования

Процедура разработки инструмента оценки для системы компьютерных адаптивных тестов (КАТ) с использованием метода теории ответа на задания состоит из нескольких этапов:

Определение

Этот этап начинается с первоначального анализа для определения основных вопросов, которые необходимо решить в процессе разработки инструмента оценки КАТ с использованием IRT-подхода. Затем проводится анализ потребностей, чтобы определить требования, которым должен отвечать разрабатываемый продукт.

Анализ студентов помогает понять их характеристики, и проводится анализ концепций для определения концепции, необходимые для разработки инструмента. Наконец, цели анализируются для определения показателей достижения при разработке системы.

Проектирование

На основе первоначального определения выбирается подходящая среда: веб-система с подходом IRT. Затем выполняется первоначальное проектирование, включая разработку требований к системе. Варианты использования системы показаны на рисунке 2. Процесс включает в себя выбор технологии и инфраструктуры, которые поддерживают внедрение веб-системы с подходом IRT. Кроме того, первоначальный проект включает в себя шаги, необходимые для удовлетворения ранее выявленных потребностей. Рисунок 2 иллюстрирует основные сценарии использования системы, которые необходимо разработать, обеспечивая четкое понимание того, как система будет использоваться и взаимодействовать с пользователями, что важно для обеспечения соответствия дизайна системы первоначальным целям, установленным на предыдущем этапе.

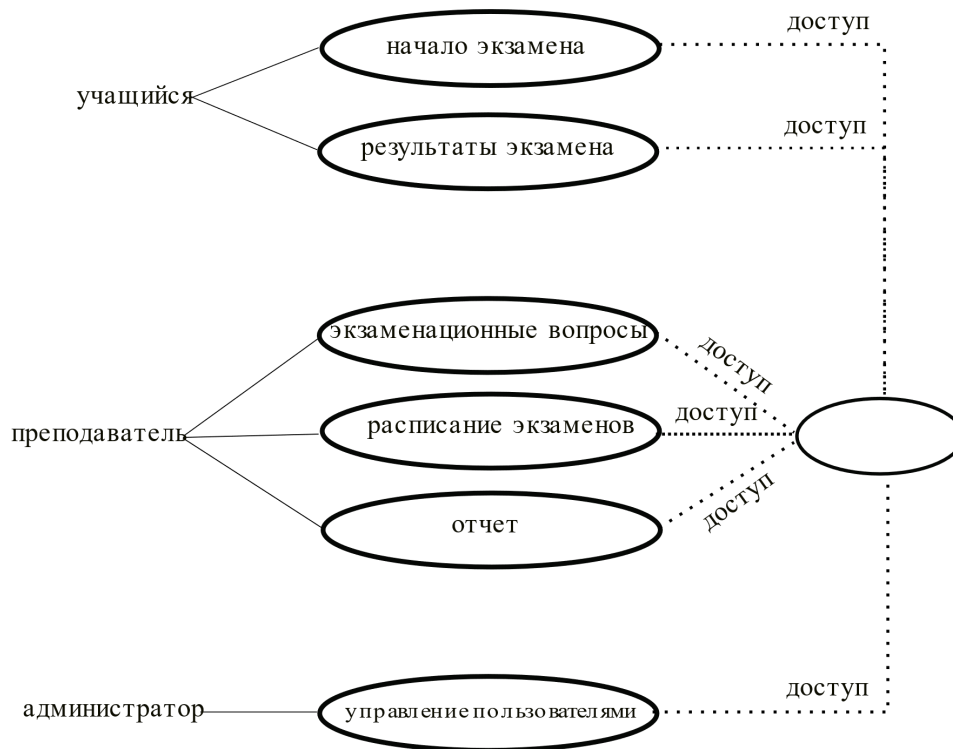


Рисунок 2-Диаграмма вариантов использования

На приведенной выше схеме использования показаны три уровня пользователей: участники тестирования/студенты, преподаватели и администраторы. Участники тестирования/студенты могут сдавать экзамены и просматривать их результаты. Преподаватели могут получить доступ к экзаменационным вопросам, расписанию экзаменов и отчетам о результатах экзаменов.

Разработка

Фаза разработки включает создание системного продукта на основе первоначального дизайна. Этот продукт будет тестироваться экспертами в соответствующей области знаний, и дизайн продукта может быть улучшен на основе отзывов. Метод проверки практичности системы компьютерных адаптивных тестов (КАТ) с использованием подхода Item Response Theory использует количественные методы, сбор данных в исследовании

осуществляется с помощью анкетирования. Аспекты, оцениваемые экспертами, представлены в таблице 1.

Таблица 1-Инструментарий оценки практичности

№	Объект оценки	Количество единиц	Форма оценивания
1	Содержание и вид	1-4	Опросник, шкала Ликерта
2	Цели и правила	5-14	Опросник, шкала Ликерта
3	Полезность	15-17	Опросник, шкала Ликерта

В таблице 1 приведен инструмент оценки или анкета, используемые в исследовании или оценке. Оцениваются три аспекта: Содержание и цель, представление и правила процесса, а также полезность. Каждый элемент оценивается с помощью разного количества утверждений или пунктов. Содержание и цель оцениваются с помощью четырех пунктов, презентация и правила процесса - с помощью десяти пунктов, а полезность - с помощью трех пунктов. Инструментом для сбора данных в данном исследовании послужил опросник со шкалой Ликерта.

Опросник со шкалой Ликерта — это широко используемый в исследованиях инструмент для измерения уровня согласия или несогласия респондентов с определенными утверждениями с помощью заранее определенной шкалы, например, шкалы от 1 до 5. Этот инструмент позволяет исследователям изучить восприятие и взгляды респондентов на оцениваемые аспекты и получить более глубокое представление о том, как респонденты оценивают содержание, представление, процесс и полезность оцениваемой системы.

Дальнейшее распространение и использование

Разработанный продукт будет выложен в открытый доступ для использования более широким кругом пользователей. Пользователям предлагается оставлять отзывы и предложения для дальнейшего улучшения качества продукта. Кроме того, продукт будет реализован в качестве пилотного проекта в НИУ МИЭТ в рамках данного исследования. Вся последовательность этих этапов образует комплексную процедуру проектирования и разработки

инструмента оценки компьютерно-адаптивного теста (КАТ) с использованием подхода теории ответов на задания.

Система компьютерных адаптивных тестов (КАТ), основанная на IRT, включает три типа пользователей с разными ролями: Студенты (либо Участники экзамена), Преподаватели и Администраторы. Перед началом работы с экзаменационной системой каждому необходимо зарегистрировать учетную запись и войти в систему (достаточно стандартная форма «имя пользователя – пароль» на отдельной странице).

Пользователи с ролью «Преподаватель» имеют права доступа к задачам по управлению экзаменами, таким как редактирование и управление коллекцией экзаменационных вопросов (в том числе изменение существующих вопросов, добавление новых и удаление устаревших либо неактуальных), установка временных параметров экзамена, печать экзаменационных билетов и карточек учеников, просмотр результатов экзаменов. Предполагается также добавить аналитико-статистический модуль для помощи преподавателю в анализе результатов индивидуальных и групповых экзаменов, что в конечном итоге способствует повышению эффективности преподавания.

Имея возможность создавать и редактировать экзаменационные вопросы, преподаватели могут обеспечить вариативность и актуальность экзамена, что способствует более точной оценке понимания студентами предмета. Другими словами, преподаватели могут разрабатывать экзамены, отвечающие как современным потребностям студентов, так и образовательным стандартам.

Преподаватели также могут просматривать результаты экзаменов учащихся с более подробной информацией, чем при обычных экзаменах. Данные о результатах экзамена отображают значения Theta, полученные каждым студентом на основе его ответов на каждый вопрос. В этих данных также представлены экзаменационные баллы студентов, основанные на полученных ими значениях Theta.

В IRT каждый вопрос обладает определенными характеристиками, такими как уровень сложности и способность к дискриминации. Система выбирает

вопросы, основываясь на ответах, тестируемых на предыдущие вопросы. Если испытуемый отвечает правильно, следующий вопрос будет более сложным. И наоборот, при неправильном ответе следующий вопрос будет легче. Это позволяет системе оценивать способности испытуемых с большей точностью и эффективностью, чем обычные тесты, в которых вопросы имеют фиксированный уровень сложности.

Использование системы КАТ, основанной на IRT, делает измерение способностей или знаний тестируемых более точным и справедливым. Каждый участник экзамена получает набор вопросов, соответствующий его способностям, и ни один из двух участников не получает одинаковой последовательности вопросов. Студенты начнут экзамен с вопросов умеренной сложности, а последующие вопросы будут корректироваться в зависимости от способности каждого студента правильно на них ответить. Представление вопросов происходит в несколько этапов:

1. Участники тестирования начинают работать над вопросами.
2. Система проверит, отвечал ли участник на вопросы ранее; если нет, система отобразит вопросы со средним весом 2.
3. Если да, то система считает наибольшее и наименьшее значения тета, чтобы определить, достигли ли они максимального (3) или минимального (-3) предела. Если достигли, система решает, что экзамен закончен.
4. Если нет, система считывает последний ответ участника, правильный (1) или неправильный (0). Если он точен, система увеличит вес вопроса на +1, если ошибочен - уменьшит на -1.
5. После получения веса система проверит, не осталось ли вопросов, соответствующих весу участника и не получивших ответа. Если подходящих вопросов больше нет, система решит, что экзамен закончен.

Система отобразит следующий вопрос в соответствии с весом, если есть совпадающие вопросы.

Компьютерный адаптивный тест (КАТ), основанный на IRT, отличается от обычного процесса подсчета баллов на экзамене. В этой системе подсчет баллов

начинается с вычисления значения Theta по следующим правилам.

1. Система проверяет ответы участника.
2. Является ли оценка правильного ответа участника = 1?
3. Если да, то θ после ответа вычисляется методом байесовской оценки
Если нет, то θ после ответа = θ исходного.

Далее происходит подсчет баллов студентов на основе их ответов и значений Theta, полученных самими студентами, с учетом следующих положений.

1. Система проверяет ответы всех участников, чтобы убедиться, что все ответы учеников неверны.
2. Если да, то итоговое θ = наименьшее θ , достигнутое любым участником.
3. Если нет, то итоговое θ = наибольшее θ , достигнутое любым участником.
4. После подсчета итоговых θ и баллов на экране отображаются итоговые θ и баллы.

После того как студенты сдадут экзамен, на экране появится информация, связанная с процессом сдачи экзамена. Эта информация включает в себя уровень сложности вопросов, на которые отвечали студенты, индекс дискриминации каждого вопроса, правильные или неправильные ответы на каждый вопрос, значение Theta студента после ответа на каждый вопрос, а также итоговый балл студента. Вся эта информация очень важна для анализа успеваемости студентов и понимания степени усвоения ими экзаменационного материала. С помощью этой информации преподаватели или специалисты по оценке могут обеспечить более точную обратную связь с учащимися и разработать более эффективные стратегии обучения, чтобы помочь учащимся улучшить понимание конкретных предметов.

Оценка полезности КАТ.

Анкета оценки полезности, состоит из 17 пунктов, разделенных на три различных аспекта:

4 пункта используются для измерения аспекта "Содержание и цель", 10 пунктов анкеты - для измерения аспекта "Внешний вид и процесс работы с правилами" и 3 пункта анкеты - для измерения аспекта "Полезность". В

контексте данного исследования данные, полученные от 90 студентов, которые использовали эту систему, затем были обработаны с помощью статистического программного обеспечения. Результаты оценки практичности этих 90 студентов можно увидеть в распределении баллов опросника практичности, представленном в таблице 2.

Таблица 2-Частотное распределение баллов

Интервальный класс	Частота	Процент (%)
71 - 72	1	1%
72 - 73	2	2%
73 - 75	5	6%
75 - 76	4	4%
76 - 78	19	21%
78 - 79	12	13%
79 - 81	47	52%
Всего	90	100%

В Таблице 2 представлено распределение частот, показывающее распределение данных по интервальным классам вместе с их частотами. Эта таблица состоит из семи интервальных классов, используемых для группировки данных, со значениями от 71 до 81 и различной шириной интервалов. Каждый интервальный класс имеет определенную частоту, отражающую количество вхождений данных в этот класс. Например, в классе интервалов 71-72 в эту категорию попадает одна точка данных.

Кроме того, в таблице также представлена информация о процентном соотношении, указывающая на вклад каждого класса интервалов в общем объеме данных. В качестве примера можно привести класс интервалов 79-81 с самой высокой частотой, равной 47, на долю которого приходится 52 % всего набора данных. Эта таблица дает наглядное представление о распределении данных и относительном вкладе каждого класса интервалов в набор данных. Эта информация очень важна для статистического анализа и помогает глубже понять

распределение оцениваемых данных. Таким образом, эта таблица служит важнейшим инструментом для понимания характеристик оцениваемых практических данных.

В заключение, оценки полезности можно свести к следующей таблице 3.

Таблица 3-Оценка полезности по трем аспектам

Объект	%	Категории
Содержание и цели	91	Очень полезно
Полезность внешнего вида и технологических правил	90	Очень полезно
Полезность	90	Очень полезно

Результаты исследования.

Исходя из таблицы 3, можно сделать вывод, что инструмент, использованный в оценке, имеет очень высокий уровень практичности. По трем аспектам, а именно: содержание и цели, внешний вид и правила, а также полезность инструмента, все они получили высокие проценты, а именно 91 % для содержания и целей и 90 % для внешнего вида и полезности. Отметка "Очень практично" указывает на то, что система компьютерных адаптивных тестов (КАТ), основанная на IRT, считается очень практичной и полезной для достижения поставленных целей измерения. Это указывает на то, что инструмент соответствует потребностям пользователей и может быть оптимально использован в соответствующих контекстах.

Заключение

Система компьютерных адаптивных тестов (КАТ), основанная на теории IRT, включает трех основных пользователей: ученики/тестируемые, учителя и администраторы. Каждый пользователь должен иметь учетную запись и войти в систему. После успешного входа в систему пользователи получают другой интерфейс, соответствующий их ролям и обязанностям. Система КАТ на основе IRT позволяет проводить адаптивное тестирование, при котором уровень сложности вопросов регулируется в зависимости от способностей тестируемых.

Каждый участник получает свою последовательность вопросов, и тест завершается, когда участник точно оценивает свои способности или когда он ответил на все имеющиеся вопросы. Результаты теста представлены в виде значений Тета, которые описывают способности студентов.

Результаты оценки практичности свидетельствуют о том, что эта система эффективна и полезна. Все оцениваемые аспекты, включая содержание и цели, дизайн пользовательского интерфейса и процессы управления, а также полезность системы, получили очень высокие баллы. Таким образом, система компьютерных адаптивных тестов (КАТ), основанная на теории реакции элементов, является эффективной и результативной в достижении поставленных целей измерения. Этот вывод свидетельствует о том, что система соответствует потребностям пользователей и может быть оптимально использована в соответствующих контекстах.

Литература

1. Алам, А. (2022) Использование адаптивного обучения и интеллектуальных роботов-репетиторов для виртуальных классов и умных кампусов: реформирование образования в эпоху искусственного интеллекта. *Advanced Computing and Intelligent Technologies. Proceedings of ICACIT*, Springer, 2022, 395.
2. Винсент-Ланкрин, С., Ван дер Влис, Р. (2020), Надежный искусственный интеллект (ИИ) в образовании: Обещания и проблемы.
3. Гарднер, Н. (2021) Дисциплинированный ум: Что должны понимать все студенты. Simon & Schuster.
4. Госс, Х. (2022) Оценка результатов обучения студентов в высшем образовании и академических библиотеках: обзор литературы. *J. Acad. Librariansh.*, 2022, 48 (2), 102485.
5. Есингельдинов, А., Рахымбаева, З., Жаппарова, А., Турсынова Л. (2023) Использование метода дифференцированного оценивания на уроках математики: компьютерное адаптивное тестирование для отслеживания прогресса учащихся. *Glob. J. Eng. Educ.*, 2023, 25 (1).

6. Кларк, Л., Уотсон, Д. (2019) Конструирование валидности: Новые разработки в создании объективных измерительных инструментов. *Psychol. Assess.*, 2019, 31 (12), 1412.
7. Левент, У., Эрток, Ш. (2020) Мнения студентов о подходе на основе заданий как формативном оценивании в сравнении с подходом на основе экзаменов как суммарном оценивании в образовании. *Sak. Univ. J. Educ.*, 2020, 10 (2), 226- 250.
8. Мартинес-Хименес, Р., Руис-Хименес, М.К. (2020) Повышение удовлетворенности студентов и эффективности обучения с помощью инвертированного класса. *Int. J. Manag. Educ.*, 2020, 18 (3), 100422.
9. Нгуен, Дж. Г., Кейзман, К. Дж., Хамстон Дж. Дж. (2020) Минимизация обмана при онлайн-оценке во время пандемии COVID-19. *J. Chem. Educ.*, 2020, 97 (9), 3429-3435.
10. Путра, Р., Лиривати, Ф.Й., Тахрим, Т., Сяфрудин, С., Аслан, А. (2020) Опыт учащихся, обучающихся на дому, во время политики закрытия школ Ковид-19 в Индонезии. *J. Iqra*, 2020, 5 (2).
11. Нуртанто, М., Холифах, Н., Масек, А., Судира, П., Самсудин, А. (2021) Основные проблемы при составлении плана урока для преподавателя профессионального обучения. *Int. J. Eval. Res. Educ.*, 2021, 10 (1), 345-354.
12. Рахматулла, С., Муляса, Э., Сиахрани, С., Понгпалилу, Ф., Путри, Р.Э. (2022) Цифровая эра 4.0: Вклад в образование и психологию студентов. *Лингвист. Cult. Rev.*, 2022, 6, (S3), 89-107.
13. Сингх, Дж., Стил, К., Сингх, Л. (2021) Сочетание лучшего из онлайн- и очного обучения: Гибридный и смешанный подход к обучению для COVID-19, мира после вакцинации и пандемии. *Educ. Technol. Syst.*, 2021, 50 (2), 140-171.
14. Чан, С.К, Лук, Л.Ю. (2022) Убеждения академиков в отношении целостного развития и оценки компетенций: Тематическое исследование в инженерном образовании. *Stud. Educ. Eval.*, 2022, 72, 101102.
15. Ченг, С.-К., Ченг, В.-П., Хуанг, И.-М. (2021) Реализация компьютерного адаптивного тестирования с автоматической настройкой

индекса сложности предметов на платформе адаптивного обучения английскому языку, *J. Internet Technol.*, 2021, 22 (7), 1599-1607.

16. Boussakuk, M., Bouchboua, A., Ghazi, M.El, Bekkali, M. El, Fattah, M. (2021) Design of Computerized Adaptive Testing Module into Our Dynamic Adaptive Hypermedia System. *Int. J. Emerg. Technol. Learn. IJET*, 2021, 16 (18), 113-128.

17. Bürkner, P.-C. (2019) Bayesian item response modeling in R with brms and Stan. ArXiv Prepr. ArXiv190509501.

18. Eloranta, S.J., Kaltiala, R., Lindberg, N., Kaivosoja, M., Peltonen, K. (2022) Validating measurement tools for mentalization, emotion regulation difficulties and identity diffusion among Finnish adolescents. *Nord. Psychol.*, 2022, (1), 30-52.

19. Gamage, K.A., Wijesuriya, D.I., Ekanayake, S.Y., Rennie, A.E., Lambert, C.G., Gunawardhana, N. (2020) Online delivery of teaching and laboratory practices: Continuity of University Programmes during COVID-19 Pandemic. *Educ. Sci.*, 2020, 10 (10), 291.

20. Mystakidis, S., Christopoulos, A., Pellas, N. (2022) A systematic mapping review of augmented reality appliKATions to support STEM learning in higher eduKATion. *Educ. Inf. Technol.*, 2022, 27 (2), 1883-1927.

21. Oktrilani, R., Delianti, V.I., Fajri, B.R., Samala, A.D., (2023) Rancang Bangun Media Pembelajaran Berbasis Augmented Reality pada Materi Sistem Pernapasan Mata Pelajaran Biologi Kelas XI MIPA Tingkat SMA. *JAVIT: Jurnal Vokasi Informatika*, 6, 79-86.

22. Pliakos, K., Joo, S.-H., Park, J.Y., Cornillie, F., Vens, C., W. Van den Noortgate, (2019) Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. *Comput. Educ.*, 2019, 137, 91-103

23. Sugiyono. A. (2013) Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D. Бандунг, Индонезия: Alfabeta.

24. Suhendi, H.Y., Ramdhani, M.A., Irwansyah, F.S. (2018) VerifiKATion concept of assesment for physics eduKATion student learning outcome, *Int. J. Eng. Technol. UEA*, 20187 (3.21), 321-325.

25. Sun, R., Zhang, H., Li, J., Zhao, J., Dong, P. (2020) Assessment-for-learning teaching mode based on interactive teaching approach in college English. *Int. J. Emerg. Technol. Learn.* 15 (21), 24.