

Раздел 4. Компьютерное тестирование и обработка результатов. Интерпретация результатов тестирования

Текст лекций

Цели:

- провести сравнительную характеристику форм тестовых заданий;
- раскрыть специфику компьютерного тестирования;
- пояснить психологические и эмоциональные реакции учащихся на компьютерное тестирование;
- рассмотреть инновационные формы тестовых заданий при компьютерном тестировании;
- провести сравнительную характеристику инновационных форм заданий при компьютерном тестировании.

4.1. Сравнительная характеристика форм тестовых заданий

В процессе разработки теста у автора всегда возникает вопрос: «Остановиться на какой-нибудь одной форме заданий или совместить различные формы в одном тесте?» Выбор автора во многом должен определяться спецификой содержания учебной дисциплины, целями создания и применения теста. Немало в данном случае зависит от технологии проверки, сбора и обработки эмпирических данных, от технического и материального обеспечения процесса применения теста.

Организовать компьютеризованный сбор и анализ результатов выполнения теста легко, если тест состоит лишь из заданий с выбором ответов. Результаты выполнения заданий с конструируемыми ответами требуют ручной обработки и привлечения экспертов, а следовательно - дополнительных материальных затрат и времени на проверку. Обилие форм в тесте затрудняет работу ученика и значительно усложняет статистическую обработку эмпирических результатов выполнения теста.

К сожалению, требование моноформности не всегда выполнимо, поскольку не все знания и умения ученика можно проверить с помощью моноформного теста. В связи с этим часто приходится идти на совмещение форм, что при прочих равных условиях всегда негативно отражается на точности измерений, обеспечиваемой тестом.

Выбор оптимальной формы предтестовых заданий обычно связан со спецификой содержания теста. При этом приходится учитывать достоинства и недостатки каждой из форм (табл. 4) и принимать определенное компромиссное решение в процессе такого выбора.

Таблица 4

Достоинства и недостатки различных форм тестовых заданий

Форма тестовых заданий	Достоинства	Недостатки
Задания с двумя ответами	Благодаря краткости позволяют охватить большой объем материала, легко разрабатываются (только один дистрактор), результаты выполнения обрабатываются быстро с высокой объективностью	Стимулируют механическое запоминание, поощряют угадывание, требуют увеличения количества заданий и, соответственно, времени тестирования для компенсации эффекта угадывания

Форма тестовых заданий	Достоинства	Недостатки
Задания с выбором из четырех-пяти ответов	Годятся для самых различных предметов, в силу краткости формулировок в тесте можно охватить большой объем содержания, обеспечивают возможность автоматизированной проверки и высокую объективность оценок учащихся, позволяют провести развернутый статистический анализ своих характеристик, скорректировать их и значительно повысить надежность педагогических измерений	Требуют значительной работы авторов при подборе дистракторов, не годятся для проверки продуктивного уровня деятельности и когнитивных умений
Задания с конструируемыми регламентированными ответами	Просты в разработке, исключено угадывание, частично годятся для автоматизированной проверки	Проверяют в основном знание фактологического материала или понятийного аппарата, иногда (в гуманитарных предметах) приводят к неоднозначным правильным и частично правильным ответам
Задания со свободно конструируемыми ответами	Позволяют оценивать сложные учебные достижения, в том числе творческий уровень деятельности, легко формулируются, исключают угадывание	Требуют длительной дорогостоящей процедуры проверки, значительного времени выполнения, не позволяют охватить значительный объем содержания предмета, снижают надежность педагогических измерений
Задания на соответствие	Просты в разработке, идеально подходят для оценивания ассоциативных знаний и проведения текущего контроля, уменьшают эффект угадывания	В основном используются лишь для проверки репродуктивного уровня деятельности и алгоритмических умений, громоздки по форме представления

4.2. Специфика компьютерного тестирования и его формы

Общие представления о компьютерном тестировании. С начала XXI в. в образовании при проведении тестирования стали широко применяться компьютеры. В педагогических инновациях появилось отдельное направление — компьютерное тестирование, при котором предъявление тестов, оценивание результатов учащихся и выдача им результатов осуществляется с помощью ПК.

Этап генерации тестов технологически может протекать по-разному, в том числе путем ввода в компьютер бланковых тестов. На сегодняшний день по компьютерному тестированию имеются многочисленные публикации, разработаны программно-инструментальные средства для генерации и предъявления тестов.

Когда необходимо обращаться к компьютерному тестированию. Хотя компьютерное тестирование значительно облегчает работу учителя при предъявлении и оценивании результатов выполнения тестов, его распространение во многом не более чем дань моде, все негативные последствия которого до сих пор не выявлены в полной мере. Выбор компьютерного формата экзамена должен основываться на более важных и обоснованных предпосылках, чем просто увлечение инновациями, поскольку он порождает множество проблем и ставит учащихся в неравные условия. Обращаться к компьютерному тестированию следует в тех случаях, когда есть настоятельная потребность в отказе от традиционных бланковых тестов.

Например, компьютерное тестирование необходимо при проведении ЕГЭ в труднодоступных районах России. Сбор выпускников школ отдаленных районов в обозначенное время проведения ЕГЭ становится настолько сложным и дорогостоящим мероприятием, что обойтись без компьютерного тестирования и современных средств коммуникации просто невозможно. Компьютерное тестирование целесообразно также применять при проведении экзаменов для детей с ограниченными возможностями, имеющих серьезные нарушения зрения или слуха. С помощью ПК можно использовать большие по размерам шрифты, аудиозаписи, дополнительные устройства для ввода данных тестирования и другие приспособления, компенсирующие на экзаменах потенциальное отставание детей с ограниченными возможностями.

Формы осуществления компьютерного тестирования. Компьютерное тестирование может проводиться в различных формах, различающихся по технологии объединения заданий в тест (рис. 17). Часть из них пока не получили специального названия в литературе по тестовой проблематике.

Первая форма - самая простая. Готовый тест, стандартизованный или предназначенный для текущего контроля, вводится в специальную оболочку, функции которой могут различаться по степени полноты. Обычно при итоговом тестировании оболочка позволяет предъявлять задания на экране, оценивать результаты их выполнения, формировать матрицу результатов тестирования, обрабатывать ее и шкалировать первичные баллы испытуемых путем перевода в одну из стандартных шкал для выдачи каждому испытуемому тестового балла и протокола его оценок по заданиям теста.

Вторая форма компьютерного тестирования предполагает автоматизированную генерацию вариантов теста, осуществляемую с помощью инструментальных средств. Варианты создаются перед экзаменом или непосредственно во время его проведения из банка калиброванных тестовых заданий с устойчивыми статистическими характеристиками. Калибровка достигается благодаря длительной предварительной работе по формированию банка, параметры заданий которого получают на репрезентативной выборке учащихся, как правило, на протяжении 3 - 4 лет с помощью бланковых тестов. Содержательная валидность и параллельность вариантов обеспечиваются за счет строго регламентированного отбора заданий каждого варианта в соответствии со спецификацией теста.

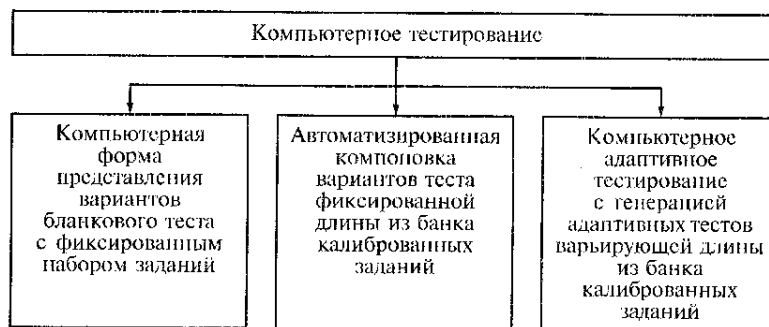


Рис. Формы компьютерного тестирования

Третья форма - компьютерное адаптивное тестирование - базируется на специальных адаптивных тестах. В основе идей адаптивности лежат соображения о том, что учащемуся бесполезно давать задания теста, которые он выполнит наверняка правильно без малейших затруднений или гарантированно не справится с ними в силу высокой трудности. Поэтому предлагается оптимизировать трудность заданий, адаптируя ее к уровню подготовленности каждого испытуемого, и сократить за счет исключения части заданий длину теста.

Достоинства и недостатки компьютерного тестирования. Компьютерное тестирование имеет определенные преимущества по сравнению с традиционным бланковым тестированием, которые проявляются особенно заметно при массовых проверках, например при проведении национальных экзаменов типа ЕГЭ. Предъявление вариантов теста на компьютере позволяет сэкономить средства, расходуемые обычно на печать и транспортировку бланковых тестов.

Благодаря компьютерному тестированию можно повысить информационную безопасность и предотвратить рассекречивание теста за счет высокой скорости передачи информации и специальной защиты электронных файлов. Упрощается также процедура подсчета результирующих баллов в тех случаях, когда тест содержит только задания с выбором ответов.

Другие преимущества компьютерного тестирования проявляются в текущем контроле, при самоконтроле и самоподготовке учащихся; благодаря компьютеру можно незамедлительно выдать тестовый балл и принять неотложные меры по коррекции усвоения нового материала на основе анализа протоколов по результатам выполнения корректирующих и диагностических тестов. Возможности педагогического контроля при компьютерном тестировании значительно увеличиваются за счет расширения спектра измеряемых умений и навыков в инновационных типах тестовых заданий, использующих многообразные возможности компьютера при включении аудио- и видеофайлов, интерактивности, динамической постановки проблем с помощью мультимедийных средств и др.

Благодаря компьютерному тестированию повышаются информационные возможности процесса контроля, появляется возможность сбора дополнительных данных о динамике прохождения теста отдельными учащимися и для осуществления дифференциации пропущенных и не достигнутых заданий теста.

Помимо неоспоримых достоинств компьютерное тестирование имеет ряд недостатков, которые представлены на рис. ***.

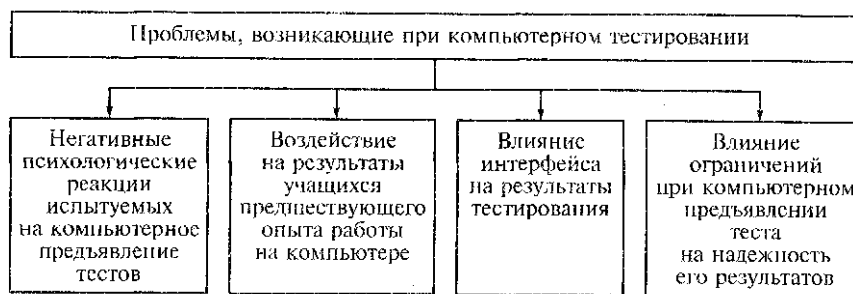


Рис. 18. Проблемы, возникающие при компьютерном тестировании

Типичные психологические и эмоциональные реакции учащихся на компьютерное тестирование. Обычно психологические и эмоциональные реакции учащихся на компьютерное тестирование носят позитивный характер. Учащимся нравится незамедлительная выдача тестовых баллов, протокола тестирования с результатами по каждому заданию, а также сам инновационный характер контроля в том случае, когда привлекаются современные гипермедийные технологии для выдачи теста. Динамическое мультимедийное сопровождение заданий на компьютере, объединенное программными средствами для представления в интерактивном режиме, по мнению учащихся, обеспечивает более точную оценку знаний и умений, сильнее мотивирует к выполнению заданий по сравнению с бланковыми тестами. Удобно также то, что вместо заполнения специальных форм для ответов можно просто выбрать ответ мышью. Если тестирование проходит в адаптивном режиме, то сокращаются время проведения экзамена и длина теста.

Негативные реакции обычно вызывают различные ограничения, которые иногда накладываются при выдаче заданий в компьютерном тестировании. Например, фиксируется либо порядок предъявления заданий, либо максимально возможное время выполнения каждого задания, после истечения которого независимо от желания испытуемого появляется следующее задание теста. В адаптивном тестировании учащиеся бывают недовольны тем, что они не имеют возможности пропустить очередное задание, просмотреть весь тест до начала работы над ним и изменить ответы на предыдущие задания. Иногда школьники возражают против компьютерного тестирования из-за трудностей, которые возникают при выполнении и записи математических вычислений и т.д.

Воздействие на выполнение теста **предшествующего** уровня компьютерного **опыта**. Результаты зарубежных исследований показали, что опыт работы на компьютерах, имеющийся у школьников, во многих случаях значительно влияет на валидность результатов выполнения теста. Если в тест включены задания без инноваций с выбором ответов, то влияние опыта работы с компьютером на результаты тестирования незначительно, поскольку от учащихся в таких заданиях не требуется никаких сложных действий при выполнении теста. При предъявлении на экране инновационных типов заданий, широко использующих средства компьютерной графики и другие новшества, влияние предшествующего компьютерного опыта на тестовый балл становится очень значительным. Таким образом, при компьютерном тестировании необходимо учитывать уровень компьютерного опыта учащихся, для которых предназначается тест.

Для снижения влияния опыта работы с компьютером на тестовые баллы рекомендуется включать в оболочки для компьютерного тестирования специальные инструкции и тренировочные упражнения для каждой инновационной формы заданий. Необходимо также предварительно ознакомить учащихся с интерфейсом, провести репетиционное тестирование и выделить в самостоятельные группы учащихся, не имеющих достаточного опыта работы с ПК, для того чтобы дополнительно обучить их или дать им бланковый тест.

Влияние интерфейса пользователя на результаты компьютерного тестирования. Интерфейс пользователя включает доступные учащемуся функции и возможности движения по заданиям теста, элементы размещения информации на экране, а также общий визуальный стиль представления информации. Хороший интерфейс пользователя должен обладать ясностью и корректностью логической последовательности взаимодействия с экзаменуемым, отражая общие принципы дизайна графической информации. Чем более продуман интерфейс, тем меньше внимания учащийся на него обращает, сосредоточивая все свои усилия на выполнении заданий теста.

4.4. Инновационные формы тестовых заданий при компьютерном тестировании

Цели разработки инновационных заданий в компьютерном тестировании. Инновационные задания, использующие возможности компьютерного тестирования, на сегодняшний день являются наиболее перспективным направлением развития автоматизации педагогических измерений. Основной причиной этого является большой потенциал инновационных заданий для повышения информативности педагогических измерений и увеличения содержательной валидности тестов.

Основная цель разработки инновационных заданий для компьютерного тестирования состоит в оценивании тех когнитивных умений, функциональной грамотности и коммуникативных умений, которые остаются не выявленными при традиционном контроле или использовании бланковых тестов.

Предметом оценивания при инновациях может быть уровень аналитико-синтетической деятельности обучаемого, скорость обобщения новой информации, гибкость мыслительного процесса и многие другие показатели умственной деятельности, сформировавшиеся в процессе обучения и не поддающиеся оцениванию с помощью обычных тестов.

Возможности инновационных заданий в компьютерном тестировании. В использовании инновационных заданий можно выделить два аспекта: дидактический и психолого-педагогический. Первый предполагает развернутую содержательную интерпретацию результатов тестирования в контексте освоенных на момент предъявления теста когнитивных, учебных и общеучебных умений, а второй позволяет оценить уровень развития мыслительных процессов у учащегося и выявить особенности усвоения им новых знаний. Большинство инновационных заданий, разработанных к настоящему времени, обеспечивают совершенствование измерений в обоих направлениях. Таким образом, инновационные задания позволяют расширить возможности самого педагогического измерения за счет получения результатов в новых, недоступных ранее направлениях оценивания качества подготовленности учащихся. Например, для оценивания уровня сформированное™ функциональной грамотности экзаменуемым можно предложить отрывок текста, в котором есть ошибки, а затем попросить идентифицировать их и исправить путем перепечатывания разделов текста.

Инновационные задания способствуют сокращению влияния случайного угадывания за счет увеличения числа возможных ответов без нарастания громоздкости заданий теста. Например, при оценивании понимания прочитанного текста можно попросить учащегося выбрать ключевое предложение в тексте и указать на него щелчком мыши. Таким образом, каждое предложение в текстовом отрывке становится опцией для выбора вместо 4-5 ответов в традиционных заданиях с готовыми ответами. Для совершенствования * формы заданий используют сложный рисунок, динамические элементы, включая изображения, мультипликацию или видео; тем самым сокращается время чтения условия. Расширение возможностей тестирования происходит при включении звука, что позволяет вести диалог с учащимся, оценивать фонети-

ческие особенности его произношения при тестировании по иностранному языку, проверять правильность интерпретации различных звуков.

Основные направления инноваций при разработке заданий. Инновации при разработке заданий для компьютерного тестирования охватывают пять связанных между собой направлений. К ним относятся: форма задания, действия испытуемого при ответе, уровень использования мультимедийных технологий, уровень интерактивности и методика подсчета баллов.

Нововведения в форме задания включают визуальный и звуковой информационные ряды или их сочетание. Визуальная информация может носить реалистический (фото, кино) и синтезированный (рисунок, анимация) характер. Тип информации в сочетании с тестовой формой определяет формат ответа, выбираемого или создаваемого экзаменуемым. При использовании фотографий или рисунков информация, содержащаяся в тестовых заданиях, носит статический характер. Кино, отражающее реальный мир, и анимация вносят динамику в выполнение теста.

Действия учащегося при ответе на задания зависят от тех инновационных средств, которые включены в тест. При включении в задания звуковой информации, предполагающей голосовой ответ учащегося, для ответа используются клавиатура, мышь или микрофон. Значительное место при ответах отводится интерактивным процессам. Интерактивный режим работы учащихся при компьютерном тестировании означает поочередную выдачу аудиовизуальной информации, при которой каждое новое высказывание со стороны учащегося или компьютера строится с учетом предыдущей информации с той и другой стороны. При организации интерактивного режима в компьютерном тестировании используется в основном экранное меню, в котором учащийся для ответа на тестовые задания выбирает, создает или перемещает объекты — компоненты ответа. Реже в интерактивном режиме применяют голосовой ввод ответа.

В целом уровень интерактивности, обеспеченный в компьютерном тестировании, характеризует степень, в которой определенная форма задания реагирует или отвечает на ввод информации со стороны экзаменуемого. Этот уровень варьируется от простейшего случая, когда совершается один шаг, до сложных, многоступенчатых заданий с разветвлением после каждого очередного ответа ученика.

Сравнительная характеристика инновационных форм заданий при компьютерном тестировании для различных целей совершенствования педагогического измерения приведена в табл. 5.

Проблемы, возникающие при использовании заданий повышенной трудности в компьютерном тестировании. Задания повышенной трудности всегда требуют больше времени для ответов вне зависимости от того, предъявляются ли они с помощью компьютерного моделирования виртуальной реальности, имеют ли форму лабораторной работы, эссе или используют мультимедийные технологии. Из-за временных затрат число сложных заданий должно быть незначительно — не более 10 — 15 %, в отдельных случаях — 20 — 25%. Многообразие звуковых и зрительных образов в компьютерном тестировании приводит к возникновению у школьников усталости, поэтому при включении в тест даже небольшого количества трудных инновационных заданий приходится значительно уменьшать длину теста, что негативно сказывается на содержательной валидности, надежности и информационной безопасности педагогического измерения.

Несмотря на преимущества инновационных форм заданий, предъявляемых с помощью компьютера, к ним нужно относиться с осторожностью, тщательно анализировать их адекватность целям измерения и уместность в тесте. Обычно инновационные задания высокой трудности выделяют в отдельный блок и помещают в конце теста. Их выполнение не должно отнимать времени у наиболее слабых учащихся, которые, скорее всего, не дойдут до конца теста.

4.5. Инновационные формы тестовых заданий

Таблица 5

Сравнительная характеристика инновационных форм заданий при компьютерном тестировании

Цель совершенствования педагогического измерения	Характеристика формы ответа	Основные направления инноваций	Характеристика трудности задания
Снизить эффект угадывания	Ответ числовой (или текстовый), конструируемый учащимся, ввод с клавиатуры или голосовой через микрофон	Использование формы задания с конструируемым ответом	Обычно высокая
Повысить содержательную валидность	Ответ выбирается мышью на графическом изображении, используется обычное меню или гипертекст	Использование аудиовизуального ряда. Включение мультимедиа без интерактивности	Низкая или средняя
Обеспечить повышение конструктивной и содержательной валидности	Ответ выбирается мышью на графическом изображении, запрашивается дополнительная информация, используется гипертекст	Использование мультимедиа для моделирования естественной окружающей среды и действий пользователя в ней. Представление объектов с помощью анимации вне режима интерактивности	Средняя или высокая
Расширить возможность измерения интеллектуальных умений, когнитивных навыков	Ответ осуществляется перемещением объектов на экране и конструируется учащимся, используется клавиатура, левая и правая кнопки мыши. Возможен интерактив	Использование формы задания с конструируемым ответом и интерактивом простейшего уровня	Средняя или высокая
Обеспечить возможность оценивания творческих и практических умений	При конструировании ответа учащимся обязательно используется двухступенчатый или многоступенчатый разветвляющийся ин-	Использование формы задания с конструируемым ответом и интерактивом сложного уровня	Средняя или высокая

Цель совершенствования педагогического измерения	Характеристика формы ответа	Основные направления инноваций	Характеристика трудности задания
	терактивный переход к различным этапам выполнения задания		
Обеспечить повышение конструктивной и содержательной валидности; расширить охват содержания; реализовать возможность измерения коммуникативных и интеллектуальных умений, когнитивных навыков	Ответ моделируется учащимся пошагово с использованием многоступенчатого разветвляющегося интерактивного перехода к различным этапам выполнения задания и виртуальной реальности	Действия испытуемого при ответе	Высокая

Подсчет баллов учащихся. Если в компьютерном тестировании не используются мультимедийные и интерактивные технологии, то подсчет первичных баллов учащихся проводится традиционно путем суммирования оценок по отдельным заданиям. Привлечение мультимедийных технологий приводит к многомерности результатов выполнения теста, поскольку оценивание целого спектра творческих, коммуникативных, общепредметных и других умений с помощью инновационных форм заданий всегда связано с несколькими переменными измерения. Появление интерактивности еще больше усложняет процедуру подсчета баллов учащихся, она становится зависимой от ответа экзаменуемого на каждом шаге выполнения заданий теста и требует полигамических оценок.

Проверка результатов выполнения заданий с конструируемым регламентированным ответом осуществляется путем сравнения ответа экзаменуемого с эталоном, хранящимся в памяти компьютера, и включает различные синонимы правильного ответа с приемлемыми орфографическими ошибками.

Намного сложнее автоматизированный подсчет баллов в заданиях со свободно конструируемым ответом (типа эссе) в гуманитарных дисциплинах. На сегодняшний день зарубежными тестологами разработаны специальные программы для автоматизированной проверки эссе. Критерии оценивания в этих программах довольно разнообразны: от рассмотрения поверхностных характеристик эссе типа длины и степени полноты ответа до сложных случаев анализа с использованием достижений компьютерной лингвистики. Обычно все эти различные автоматизированные программы подсчета баллов требуют участия экспертов только на момент начала работы, когда квалифицированным педагогам необходимо «обучить» компьютерную программу оцениванию любых развернутых ответов.