

# Вопросы оценивания результатов компьютерного тестирования

Юрий Мохир,  
Иван Ребрик

*Кафедра патологической физиологии  
Карагандинского государственного  
медицинского университета  
yurimokh@mail.ru*

**Статья посвящена разработке и обоснованию критериев оценивания результатов компьютерного тестирования. Рассмотрены различные варианты дифференцированного оценивания в зависимости от вида и значимости тестовых заданий и с учётом вероятности отгадывания правильного ответа. Все рассмотренные варианты оценивания могут быть реализованы с помощью системы MyTestXPro последующей обработкой и анализом результатов тестирования.**

*Ключевые слова:* компьютерное тестирование, виды и методы педагогического контроля, система оценивания результатов тестирования, виды тестовых заданий, супертест

Важнейшей составляющей процесса обучения является контроль знаний. Известно, что к одному из самых эффективных способов контроля знаний относится тестирование — метод объективного и стандартизированного педагогического измерения (педагогической диагностики) с помощью специально разработанных тестовых заданий. Результаты тестирования легко поддаются количественной оценке, статистической обработке и сравнительному анализу. В полной мере возможности тестирования могут быть реализованы лишь в компьютерном тестировании.

Компьютерное тестирование — способ проведения тестирования, в ходе которого тестовые задания предъявляются на мониторе компьютера. Это диалог испытуемого и компьютера без посредников (промежуточных звеньев), который завершается выводом результата на монитор.

При выполнении надлежащих требований компьютерное тестирование обеспечивает достижение следующих целей:

- одновременный контроль большого количества студентов с соблюдением принципа индивидуального контроля;
- оперативную проверку полноты усвоения определённой части или всего учебного материала;
- объективность, надёжность и валидность результатов оценки знаний;
- экономию времени;
- получение количественных показателей успеваемости, которые можно использовать в дальнейшем в целях совершенствования методики и организации преподавания.

Число тестовых заданий в тесте может варьировать в широком диапазоне в зависимости от целей тестирования<sup>1</sup>. С прагматической точки зре-

<sup>1</sup> Аванесов В.С. Супертест // Педагогические измерения, 2014. — № 2. — С. 3–13.

ния полезнее создать тест, имеющий сравнительно меньшее число заданий, но обладающий достоинствами, присущим тестам с большим числом заданий. Вот почему в тест желательно отобрать минимально достаточное количество заданий, которое позволяет сравнительно точно определить уровень и структуру подготовленности. Достаточность числа тестовых заданий определяется по результатам обязательной эмпирической апробации теста и применением методов многомерного статистического анализа.

### Требования к проведению компьютерного тестирования

1. Наличие структурированной базы качественных тестовых заданий различных видов, с различным количеством ответов, обеспечивающих:

- значимость проверяемых знаний в общей системе знаний и их полноту;
- репрезентативность содержания учебной дисциплины;
- соответствие уровню современного состояния науки.

2. Наличие нескольких параллельных по содержанию и сходных по трудности вариантов тестовых заданий. Наличие таких заданий помогает исключить искажение результатов, даёт возможность объективного сопоставления результатов тестирования.

3. Наличие краткой инструкции перед каждым заданием (при компьютерном тестировании), поскольку в тесте могут быть использованы различные виды тестовых заданий.

4. Наличие компьютерной автоматизированной системы, обеспечивающей:

- случайный выбор тестовых заданий из определённых групп и подгрупп;
- случайный порядок чередования предлагаемых ответов;

- оценивание результатов тестирования на основе апробированной системы критериев в зависимости от вида и значимости тестовых заданий;

- вывод результатов по окончании тестирования на монитор.

5. Определённый регламент времени.

6. Соблюдение условий тестирования, исключающих списывание, подсказки, возможность сдачи теста одним испытуемым вместо другого.

### Бланковое тестирование

Бланковое тестирование — достаточно распространённый способ проведения тестирования с помощью печатных бланков. Его называют также тестированием «на бумажных носителях». Испытуемые вносят ответы на особые бланки. Затем информация с бланков обрабатывается либо вручную, с помощью особых трафаретов, либо сканируется и поступает в электронную память компьютера.

На наш взгляд, бланковое тестирование — морально устаревший вид тестирования. Его существенными недостатками являются фиксированные варианты тестовых заданий, фиксированный порядок чередования предлагаемых ответов, наличие промежуточных звеньев между тестируемым и введением результатов в компьютер, что определяет отсроченность объявления результатов, т.е. наличие недостатков, которые в значительной мере могут исказить результаты тестирования.

### Система оценивания

Широкое использование компьютерного тестирования в учебном процессе делает весьма актуальной разработку стройной, гибкой системы дифференцированного оценива-

ния результатов. Однако в литературе имеются лишь отдельные работы, посвящённые дифференцированному оцениванию только некоторых видов тестовых заданий в тестовой форме<sup>2-5</sup>.

В данной работе предпринята попытка разработки системного дифференцированного подхода к оцениванию результатов в зависимости от различных видов и значимости тестовых заданий, с учётом вероятности угадывания правильного ответа.

## **Задания с выбором одного правильного ответа**

Эта форма тесно связана с такими основными законами логики, как закон противоречия и закон исключенного третьего. Первый требует, что нельзя одно и то же одновременно утверждать и отрицать, а закон исключенного третьего требует, чтобы задание вместе с одним из ответов превращалось либо в истинное высказывание, либо в ложное. Третьего здесь не дано.

Задания с выбором одного правильного ответа из нескольких предлагаемых — наиболее распространённый вид тестовых заданий. Число возможных ответов может быть любым, но в практике обычно применяются от 4 до 6 ответов.

Тут многое зависит от привлекательности неправильных ответов,

называемых дистракторами. Необходимо стремиться к тому, чтобы каждый дистрактор был одинаково привлекательным и равномерно использовался испытуемыми, выбирающими неправильный ответ. Если дистрактор неправдоподобен, то он перестаёт выполнять свою функцию. Механическое увеличение вариантов ответов, без учёта качества дистракторов, может привести к созданию логически некорректного задания.

С учётом снижения вероятности угадывания правильного ответа при увеличении количества ответов предлагается следующая шкала оценивания:

- правильный ответ при 3-х предлагаемых вариантах ответа — 0,6 балла;
- правильный ответ при 4-х предлагаемых вариантах ответа — 0,8 балла;
- правильный ответ при 5-и и более предлагаемых вариантах ответа — 1 балл;
- неправильный ответ — 0.

## **Задания с выбором одного наиболее правильного ответа**

Этот вид заданий формирует у испытуемых умение дифференцировать меру правильности ответов.

Наиболее правильный ответ предлагается оценивать 1 баллом, а остальные ответы — 0.

<sup>2</sup> Аванесов В.С. Применение заданий в тестовой форме и квантованных учебных текстов в новых образовательных технологиях / В.С. Аванесов // Педагогические измерения. — 2012. — № 2. — С. 75–91.

<sup>3</sup> Афонина М.В. Педагогический тест: требования к разработке и использованию (учебно-методическая разработка) / М.В. Афонина // Сборник материалов для работников образования по проекту «Развитие заочной распределенной многопрофильной школы Алтайского края» — Барнаул: Изд-во БГПУ, 2006. — С. 44–73.

<sup>4</sup> Агаева Ф. Контроль знаний тестовыми формами, посредством новой компьютерной программы / Агаева Ф., Исакова А. // Педагогические измерения, 2014. — № 2. — С. 64–70.

<sup>5</sup> Рыбанов А.А. Повышение качества и достоверности компьютерной оценки знаний в автоматизированных системах контроля и обучения // Педагогические измерения, 2014. — № 2. — С. 14–28.

### **Задания на установление правильной последовательности**

Цель этого вида заданий — формирование алгоритмического мышления, алгоритмических знаний, умений и навыков. Алгоритмическое мышление можно определить как интеллектуальную способность, проявляющуюся в определении наилучшей последовательности действий при решении учебных и практических задач. Характерные примеры проявления такого мышления — успешное определение правильной последовательности той или иной медицинской операции, последовательности лечения больных, выполнение различных действий в правильной последовательности, разработка самой эффективной программы для ЭВМ и т.п.

В патологической физиологии задания этого вида представляют интерес для установления правильной последовательности (причинно-следственных отношений) отдельных звеньев патогенеза различных заболеваний.

Оценивание таких заданий в литературе предлагается вести по-разному. Например, традиционно — всё правильно — один балл, наличие хотя бы одной ошибки — ноль баллов. Или даются два балла за правильный ответ и ноль — за ошибочный.

В данной работе правильный ответ предлагается оценивать числом баллов, равным количеству компонентов. При наличии 3-х компонентов правильный ответ предлагается оценивать 3 баллами, при наличии 4-х компонентов — 4 баллами и т. д. При наличии большего числа компонентов (более 5) возможно дифференцированное оценивание за частично правильный ответ. Наличие ошибки в начале последовательности оценивается 0 баллов. При наличии ошибки в середине последовательности с максимальной оценки снимаются 2 балла, при наличии ошибки в конце последовательности с максимальной оценки снимается 1 балл.

Упрощённый вариант реализации заданий на установление правильной последовательности: испытуемому предлагается несколько фиксированных ответов, из которых лишь один является правильным. По сути дела, это является одним из вариантов заданий с выбором одного правильного ответа из нескольких предлагаемых с соответствующим оцениванием (см. выше).

**Задания с выбором нескольких правильных ответов**

Сложность этих заданий выше, чем у заданий с выбором одного правильного ответа, и угадать правильный ответ сложнее. Испытуемый должен не только выбрать правильный ответ, но и определить количество правильных ответов, поскольку их число не указывается.

Количество предлагаемых ответов может достигать 10, а количество правильных ответов — 5-и и более баллов. Оптимальным соотношением правильных ответов и дистракторов, на наш взгляд, является ~ 1: (2–3). При меньшем соотношении вероятность угадывания правильных ответов возрастает.

### **Задания на установление соответствия**

Это вид задания, в котором испытуемому предлагается установить соответствие между элементами двух множеств, причём второй ряд содержит больше элементов, чем первый.

Каждую правильную пару предлагается оценивать 1 баллом, а за каждую неправильную пару — снимать 0,5 балла. При этом минимальная оценка за тестовое задание должна быть 0.

## Задания открытой формы

В заданиях открытой формы предусмотрен ввод цифр, символов и текста с клавиатуры. Оценивание заданий с вводом текста представляет проблему, т.к. грамматическая ошибка расценивается компьютером как смысловая.

В то же время оценивание заданий с введением символов или цифр представляется вполне реальной. Примером таких заданий, в частности, является указание диапазона колебаний в норме того или иного биохимического показателя и т.д. Правильный ответ в таких заданиях предлагается оценивать 1 баллом, а неправильный — 0.

## Итоговая оценка

По завершении тестирования общий балл рассчитывается путём суммирования результатов всех тестовых заданий. Итоговая оценка представляет собой отношение общего балла к максимально возможной сумме баллов, выраженное в процентах. Можно предусмотреть перевод процентов в принятую в некоторых странах буквенную оценку (от А до F), в 10-балльную, 12-балльную или 4-балльную оценку (отлично — 5, хорошо — 4, удовлетворительно — 3, неудовлетворительно — 2).

Предлагаемая система оценивания результатов компьютерного тестирования является рамочной и подлежит обязательной эмпирической проверке, с учётом специфики учебной дисциплины (предмета), с последующей корректировкой. Система не привязана к какому-либо определённому количеству тестовых заданий. Она позволяет сопоставлять баллы, полученные за тестовые задания различных видов.

Все рассмотренные варианты оценивания могут быть реализованы

с помощью компьютерной системы тестирования знаний MyTestXPro<sup>6</sup>, с последующей обработкой и анализом результатов тестирования. С помощью модуля тестирования можно организовать как локальное, так и сетевое тестирования. При сетевом тестировании результаты могут быть переданы по сети в модуль «Журнал», а могут быть отправлены по электронной почте или на веб-сервер в Интернет методом POST.

## Заключение

Компьютерное тестирование имеет ряд преимуществ по сравнению с использованием традиционных форм контроля. Автоматизированные тесты обеспечивают возможность быстрого и объективного оценивания качества знаний, способствуя повышению рентабельности образования за счёт экономии времени преподавателей.

Компьютерное тестирование является одной из наиболее технологичных форм проведения автоматизированного контроля, с управляемыми параметрами качества. В этом смысле ни одна из известных форм контроля знаний учащихся не может сравниться с тестированием. Компьютерное тестирование может быть с успехом использовано при различных видах педагогического контроля — базового уровня, текущего, тематического, рубежного и итогового.

Современный уровень компьютерного тестирования позволяет в настоящее время использовать в основном тестовые задания с выбором одного или нескольких правильных ответов. Существует мнение, что задания такой формы можно использовать лишь для проверки знаний фактологического материала и репродуктивного уровня умений. На наш взгляд, это не так. Использование различных видов тестовых зада-

<sup>6</sup> Инструкция пользователя программы MyTestXPro. Компьютерное тестирование знаний. — Режим доступа: <http://mytest.klyaksa.net>

ний с выбором одного или нескольких правильных ответов способствует развитию т.н. Дж. Гилфлордом конвергентного и дивергентного мышления. Задания на установление правильной последовательности формируют алгоритмическое мышление. Таким образом, есть все основания утверждать, что компьютерное тестирование способствует реализации творческого потенциала личности.

Тем не менее, наряду с компьютерным тестированием, в учебном процессе используются и рутинные методы педагогического контроля — устный и письменный контроль, хотя их существенными недостатками являются субъективность оценок и невоспроизводи-

мость результатов. Вместе с тем устный контроль способствует выработке быстрой реакции, развивает связную речь, а письменный контроль способствует развитию логического мышления, связности изложения мыслей.

Несмотря на широкое применение компьютерных технологий, ведущая роль в учебном процессе остается за преподавателем. Он является организатором учебного процесса, консультантом, разработчиком тестовых заданий, экспертом. Только опытный преподаватель в полной мере может оценить сферу творческих способностей обучающихся, их познавательную деятельность и творческую активность.