

# Методика

## Разработка критериев оценивания результатов компьютерного тестирования

**Юрий Мохир**,  
кандидат медицинских наук, доцент,  
кафедра патологической физиологии  
Карагандинского государственного  
медицинского университета  
[yurimokh@mail.ru](mailto: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 угадывания правильного ответа. Все рассмотренные варианты оценивания могут быть реализованы с помощью системы MyTestXPro, с последующей обработкой и анализом результатов тестирования.

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

### Введение

Компьютерное тестирование при выполнении надлежащих требований 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 усовершенствования методики и организации преподавания. Компьютерное тестирование может быть с успехом использовано при различных видах педагогического контроля — контроля базового уровня, текущего, тематического, рубежного и итогового<sup>1</sup>.

Широкое использование компьютерного тестирования в учебном процессе делает весьма актуальной разработку стройной, гибкой системы дифференцированного оценивания результатов. Однако в литературе имеются лишь отдельные работы, посвящённые дифференцированному оцениванию, и только некоторых форм тестовых заданий<sup>2-5</sup>.

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

## Понятие о вероятности

Вероятность угадывания правильного ответа в тестовых заданиях с выбором одного или нескольких правильных ответов — это мера возможности угадать правильный от-

вет, при предположении полного незнания правильного ответа учебного материала, представленного в задании. Вероятность правильного ответа на задание обычно представляется как функция от уровня знаний испытуемого и от уровня трудности задания.

Вероятностью случайного события  $A$  называется отношение числа  $m$  несовместимых равновероятных элементарных событий, составляющих событие  $A$ , к числу всех возможных элементарных событий  $n$ :

$$P(A) = \frac{m}{n}, \quad (1)$$

где:  $P$  — вероятность событий, событие  $A$  — правильно угаданный ответ;  $m$  — число случаев, благоприятствующих этому событию, т.е. число правильно угаданных ответов;  $n$  — число всех случаев, т.е. все предлагаемые ответы.

## Задания с выбором одного правильного ответа

Задания с выбором одного правильного ответа из нескольких предлагаемых — одна из наиболее распространённых форм тестовых заданий.

Тестовые задания с выбором одного правильного ответа соответствуют альтернативной форме<sup>3</sup>. Эта форма тесно связана с такими основными законами логики, как закон противоречия и закон исключённого третьего. Первый требует, что нельзя одно и то же одновременно утверждать и отрицать, а за-

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

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

<sup>3</sup> Рыбанов А.А. Повышение качества и достоверности компьютерной оценки знаний в автоматизированных системах контроля и обучения // Педагогические измерения. 2014. № 2. С. 14–28.

<sup>4</sup> Guilford J.P. The Model of Mind: The Nature of Human Intelligence / Mc Graw Hill, N-Y, 1967. 538 p.

<sup>5</sup> Инструкция пользователя программы MyTestXPro. Компьютерное тестирование знаний. Режим доступа: <http://mytest.klyaksa.net>

кон исключённого третьего – чтобы задание вместе с одним из ответов превращалось либо в истинное высказывание, либо в ложное. Третьего здесь не дано.

Оптимальным количеством ответов является 4–5. 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. Увеличение вариантов ответов может привести к созданию некорректного задания. Как правило, трудно найти более 4–5 интересных, оригинальных дистракторов, которые будут одинаково привлекательны для выбора. Если дистракторов меньше (соответственно меньше вариантов ответов), то увеличивается вероятность угадывания правильного ответа.

С математической точки зрения такая вероятность обратно пропорциональна количеству предлагаемых ответов.

При  $m = 1$  (1) превращается в (2).

$$P(A) = \frac{m}{n} = \frac{1}{n}. \quad (2)$$

Вероятность отгадывания одного правильного ответа из нескольких предлагаемых ответов и рекомендуемые баллы за правильный ответ представлены в табл. 1.

Неправильный ответ оценивается 0 баллов.

### Задания с выбором одного наиболее правильного ответа

Основанием для введения этого вида заданий является целесообраз-

ность формирования у учащихся не только умения различать правильные ответы от неправильных, но и умения дифференцировать меру правильности ответов. Тестовые задания данного вида не нашли широкого применения на практике. Эти задания представлены здесь исключительно из соображений полноты рассматриваемых видов тестовых заданий.

### 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 (если не указано их количество).

Оптимальным соотношением правильных ответов и дистракторов, на наш взгляд, является ~ 1: (2–3).

С математической точки зрения, вероятность отгадывания полного правильного ответа, состоящего из  $k$  элементов, обратно пропорциональна числу сочетаний из  $n$  по  $k$ , где  $n$  – количество предлагаемых ответов:

$$P(A) = \frac{1}{C_n^k}. \quad (3)$$

В свою очередь, число сочетаний из  $n$  по  $k$  равно биномиальному коэффициенту:

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (4)$$

Подставляя (4) в (3), получаем:

$$P(A) = \frac{1}{C_n^k} = \frac{k!(n-k)!}{n!}. \quad (5)$$

Таблица 1

Число правильных ответов	1 из 2	1 из 3	1 из 4	1 из 5
Вероятность отгадывания	0,50	0,33	0,25	0,20
Рекомендуемые баллы за правильный ответ	0,5	0,6	0,8	1,0

Вероятность угадывания полного правильного ответа, состоящего из  $k$  элементов, из  $n$  предлагаемых ответов, и рекомендуемые баллы

оценивания при наличии подсказки (указание количества правильных ответов) представлены в табл. 2–5.

Таблица 2

**Вероятность угадывания полного правильного ответа, состоящего из двух ответов ( $k = 2$ ), из  $n$  предлагаемых ответов, и рекомендуемые баллы оценивания**

Число правильных ответов	2 из 4	2 из 5	2 из 6	2 из 7	2 из 8
Число сочетаний	6	10	15	21	28
Вероятность отгадывания	0,17	0,10	0,07	0,05	0,04
Рекомендуемые баллы за полный правильный ответ	1,0	2,0	3,0	4,0	5,0
Цена каждого правильного ответа	0,5	1,0	1,5	2,0	2,5

Таблица 3

**Вероятность отгадывания полного правильного ответа, состоящего из трёх ответов ( $k = 3$ ), из  $n$  предлагаемых ответов, и рекомендуемые баллы оценивания**

Число правильных ответов	3 из 6	3 из 7	3 из 8	3 из 9	3 из 10
Число сочетаний	20	35	56	84	120
Вероятность отгадывания	0,05	0,03	0,02	0,01	0,01
Рекомендуемые баллы за полный правильный ответ	4,0	6,0	10,0	18,0	21,0
Цена каждого правильного ответа	1,3	2,0	3,3	6,0	7,0

Таблица 4

**Вероятность отгадывания полного правильного ответа, состоящего из четырёх ответов ( $k = 4$ ), из  $n$  предлагаемых ответов, и рекомендуемые баллы оценивания**

Число правильных ответов	4 из 8	4 из 9	4 из 10
Число сочетаний	70	126	210
Вероятность отгадывания	0,01	0,008	0,005
Рекомендуемые баллы за полный правильный ответ	18,0	20,0	32,0
Цена каждого правильного ответа	4,5	5,0	8,0

Таблица 5

**Вероятность отгадывания полного правильного ответа, состоящего из пяти правильных ответов ( $k = 5$ ), из 10 ( $n = 10$ ) предлагаемых ответов, и рекомендуемые баллы оценивания**

Число правильных ответов	5 из 10
Число сочетаний	252
Вероятность отгадывания	0,004
Рекомендуемые баллы за полный правильный ответ	40,0
Цена каждого правильного ответа	8,0

Неполный правильный ответ рекомендуется оценивать в соответствии с ценой каждого правильного ответа. Отсутствие хотя бы одного правильного ответа рекомендуется оценивать 0 баллов.

Использование штрафных баллов при неправильных ответах, на наш взгляд, представляется дискуссионным. К данному вопросу планируется вернуться в одной из следующих работ. По-видимому, штрафные баллы целесообразно использовать в тестовых заданиях с выбором нескольких правильных ответов при отсутствии подсказки (указания количества правильных ответов).

### Задания на установление соответствия

Это вид заданий, в котором испытуемому предлагается установить соответствие между элементами двух множеств. Сложность этого вида заданий может варьировать в широком диапазоне, в зависимости от количества элементов в каждом из двух предлагаемых множеств от количества правильных ответов (комплементарных пар), от наличия или отсутствия подсказки (указания количества правильных ответов). Оценка различных вариантов сводится к рассмотренным выше ситуациям.

### 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 мышле-

ния — успешное выполнение различных заданий за короткое время, 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.

С математической точки зрения, вероятность отгадывания одного правильного ответа (правильной последовательности) зависит от числа элементов и обратно пропорциональна числу перестановок. Перестановками из  $n$  элементов называются соединения, каждое из которых содержит все  $n$  элементов, отличающихся только порядком расположения элементов. Число всех возможных перестановок, которые можно образовать из  $n$  элементов, обозначается символом  $P_n$ .

$$P(A) = \frac{1}{P_n}. \quad (6)$$

Число всех перестановок порядка  $n$  равно числу размещений из  $n$  по  $n$ , то есть факториалу:

$$P_n = n! \quad (7)$$

Подставляя (7) в (6), получаем

$$P(A) = \frac{1}{n!}. \quad (8)$$

Вероятность отгадывания одного правильного ответа (правильной последовательности) из  $n$  предлагаемых компонентов и рекомендуемые баллы представлены в табл. 6.

Наличие ошибки в начале последовательности однозначно оценивается 0 баллов.

При наличии большого числа компонентов (более 5) возможно дифференцированное оценивание за частично правильный ответ. Могут быть обсуждены и рассмотрены и другие варианты дифференцированного оценивания тестовых заданий при наличии большого числа компонентов.

Таблица 6

Число компонентов	3	4	5	6
Число перестановок	6	24	120	720
Вероятность отгадывания	0,17	0,04	0,008	0,001
Рекомендуемые баллы за правильный ответ	1,0	5,0	40,0	200,0

Возможен упрощённый вариант реализации заданий на установление правильной последовательности: испытуемому предлагается несколько фиксированных вариантов ответов, из которых лишь один является правильным. По сути дела, это является одним из вариантов заданий с выбором одного правильного ответов из нескольких предлагаемых с соответствующим оцениванием (см. выше, табл. 1).

## Итоговая оценка

По завершении (окончании) тестирования общий балл рассчитывается путём суммирования результатов всех тестовых заданий.

Абсолютно необходимое условие – одинаковое количество равно оцениваемых тестовых заданий для всех тестируемых.

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).

## Заключение

Предлагаемая система оценивания результатов компьютерного тестирования является рамочной и подлежит обязательной эмпирической

проверке, с учётом специфики учебной дисциплины и с последующей корректировкой. Она не привязана к какому-либо виду тестовых заданий или какому-нибудь определённому количеству тестовых заданий. Предлагаемая система оценивания результатов позволяет сопоставлять баллы, полученные за тестовые задания различных форм.

Современный уровень компьютерного тестирования позволяет 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 совсем так.

Использование различных форм тестовых заданий с выбором одного или нескольких правильных ответов способствует развитию соответственно конвергентного и дивергентного мышления. Понятия «конвергентного» и «дивергентного» мышления в своё время были детально разработаны Дж. Гилфордом в работе «Природа человеческого интеллекта»<sup>6</sup>. Задания на установление правильной последовательности формируют алгоритмическое мышление. Таким образом, есть все основания утверждать, что компьютерное тестирование способствует реализации творческого потенциала личности.

Все рассмотренные варианты оценивания могут быть реализованы с помощью системы тестирования знаний MyTestXPro<sup>7</sup> с после-

<sup>6</sup> Guilford J.P. The Model of Mind: The Nature of Human Intelligence / Mc Graw Hill, N-Y, 1967. 538 pp.

<sup>7</sup> Инструкция пользователя программы MyTestXPro. Компьютерное тестирование знаний. Режим доступа: <http://mytest.klyaksa.net>

дующей обработкой и анализом результатов тестирования. С помощью модуля тестирования можно организовать как локальное, так и сетевое тестирование. При сетевом тестировании результаты тестирования могут быть переданы по сети в модуль Журнал, отправлены по электронной почте или на веб-сервер в Интернет методом POST.

Автор выражает благодарность кандидату физико-математических наук доценту кафедры высшей математики Карагандинского государственного политехнического университета Л.М. Мустафиной за обсуждение данной работы.