

Универсальная система оценивания разных видов тестовых заданий закрытого типа

**Мохир
Юрий Михайлович**

кандидат медицинских наук, доцент кафедры патологии,
Медицинский университет г. Караганды, Казахстан.
yurimokh@mail.ru

Ключевые слова: тестовые задания закрытого типа, вероятность угадывания, система оценивания.

Анализ результатов педагогического тестирования обычно проводится на основе классической теории тестов, на основе Item Response Theory или на основе модели Г. Раша¹. Несмотря на различия, указанные теории в качестве исходного материала используют измерения, выполненные в дихотомических и порядковых шкалах. Такой подход не позволяет дифференцированно оценивать разные виды тестовых заданий, а также оценивать неполные правильные ответы в заданиях с выбором нескольких правильных ответов.

В литературе имеются лишь отдельные работы, посвящённые дифференцированному оцениванию только некоторых видов тестовых заданий². Необходимо признать, что отсутствие дифференцированной системы оценивания разных видов тестовых заданий, а также оценки за неполный правильный ответ в заданиях с выбором нескольких правильных ответов, не позволяют в полной мере использовать все возможности компьютерного тестирования.

Правильный ответ в тестовом задании закрытого типа может явиться результатом осмысленного (осознанного) выбора, основанного на знаниях, или случайного выбора (угадывания). Вероятность угадывания правильного ответа — мера возможности угадать правильный ответ при предположении полного незнания учебного материала. В качестве теоретической оценки угадывания правильного ответа можно использовать величину, обратную от количества ответов в тестовом задании с выбором одного из нескольких предлагаемых: $P(A) = \frac{1}{n}$, где n — число предлагаемых ответов.

Например, в тестовом задании с четырьмя предлагаемыми ответами вероятность угадывания равна $1/4$, с пятью предлагаемыми ответами — $1/5$ и т.д. Чем выше вероятность угадывания правильного ответа, тем ниже информативность тестового задания. Чем ниже вероятность угадывания правильного ответа, тем выше информативность тестового задания. Высокая

¹ Аванесов В.С. Применение тестовых форм в Rasch Measurement / В.С. Аванесов // Педагогические измерения. — 2005. — № 4. — С. 3–20.

² 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;
Афонина М.В. Педагогический тест: требования к разработке и использованию (учебно-методическая разработка) / М.В. Афонина // Сб. матер. для работников образования по проекту «Развитие заочной распределенной многопрофильной школы Алтайского края» — Барнаул: Изд-во БГПУ. — 2006. — С. 44–73;

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

Таблица 1

Выбор одного правильного ответа ($k = 1$) из нескольких предлагаемых ($n = 2 \div 10$)

Число предлагаемых ответов n	2	3	4	5	6	7	8	9	10
Вероятность угадывания правильного ответа $P(A)$	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10
Рассчитанная рекомендуемая оценка за правильный ответ (УТЕ)	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8	2,0
Отсутствие правильного ответа (УТЕ)	0	0	0	0	0	0	0	0	0

вероятность угадывания снижает информативность и трудность тестового задания, и наоборот.

За условную тестологическую единицу (1,0 УТЕ) предлагается принять правильный ответ с вероятностью угадывания 1/5 (1 правильный ответ из 5 предлагаемых).

В таблице 1 представлены вероятности угадывания одного правильного ответа из $2 \div 10$ предлагаемых ответов. На основании этого рассчитаны соответствующие рекомендуемые оценки за правильный ответ.

Задания с выбором одного наиболее правильного ответа из нескольких предлагаемых и с выбором одного неправильного ответа из нескольких предлагаемых являются зеркальным отражением заданий с выбором одного правильного ответа из нескольких предлагаемых. Задания данных видов не нашли широкого применения на практике, но представлены здесь исключительно из соображений полноты рассматриваемых видов тестовых заданий. К тестовым заданиям данного вида в полной мере применимы подходы, рассмотренные выше (см. табл. 1).

В заданиях с выбором нескольких правильных ответов из нескольких предлагаемых с математической точки зрения вероятность угадывания полного правильного ответа обратно пропорциональна количеству числу возможных сочетаний, где k — число правильных ответов, а n — число предлагаемых ответов³.

В таблицах 2–4 представлены вероятности угадывания k правильных ответов

из n предлагаемых ответов на основе числа возможных сочетаний C_n^k ($k \leq n/2$). Исходя из этого рассчитаны соответствующие оценки за полный правильный ответ. Неполные правильные ответы предлагается оценивать пропорционально полному правильному ответу.

Альтернативный вариант оценивания тестовых заданий с выбором нескольких правильных ответов из предлагаемых: каждый правильный ответ оценивается аналогично выбору одного правильного ответа из нескольких предлагаемых ответов (см. таблицу 1), а сумма правильных ответов представляет собой оценку за полный правильный ответ⁴.

примеры применения данного подхода представлены в таблицах 5 и 6.

В заданиях на установление соответствия испытуемому предлагается установить соответствие между элементами двух множеств. Первое множество (объекты, субъекты) содержит m элементов. Второе множество (признаки, характеристики) содержит n элементов; m может быть равно или не равно n ($m = n$ или $m \neq n$).

Сложность этого вида тестовых заданий может варьировать в широком диапазоне в зависимости от числа элементов в предлагаемых множествах и от количества правильных ответов (комплементарных пар). Вероятность парного соответствия для каждого из m_1 первого множества ($m_1 < m$, т. е. могут быть задействованы не все элементы первого множества) составляет $1/n$, что соответствует определённой оценке (см. табл. 1).

³ Мохир Ю. Разработка критериев оценивания результатов компьютерного тестирования / Ю. Мохир // Педагогическая диагностика. — 2016. — № 6. — С. 66–72.

⁴ Мохир Ю. Система оценивания результатов компьютерного тестирования / Ю. Мохир // Педагогическая диагностика. — 2017. — № 6. — С. 44–53.

Таблица 2

**Выбор двух правильных ответов ($k = 2$) из нескольких предлагаемых ответов ($n = 4 \div 10$)
(на основе числа возможных сочетаний C_n^2)**

Число предлагаемых ответов n	4	5	6	7	8	9	10
Число возможных сочетаний C_n^2	6	10	15	21	28	36	45
Вероятность угадывания полного правильного ответа $P(A)$	1/6	1/10	1/15	1/21	1/28	1/36	1/45
Рассчитанная рекомендуемая оценка за полный правильный ответ (УТЕ)	1,2	2,0	3,0	4,2	5,6	7,2	9,0
Рассчитанная рекомендуемая оценка за неполный правильный ответ (УТЕ) — один из двух правильных ответов	0,6	1,0	1,5	2,1	2,8	3,6	4,5
Отсутствие правильного ответа (УТЕ)	0	0	0	0	0	0	0

Таблица 3

**Выбор трёх правильных ответов ($k = 3$) из нескольких предлагаемых ответов ($n = 6 \div 10$)
(на основе числа возможных сочетаний C_n^3)**

Число предлагаемых ответов n	6	7	8	9	10
Число возможных сочетаний C_n^3	20	35	56	84	120
Вероятность угадывания полного правильного ответа $P(A)$	1/20	1/35	1/56	1/84	1/120
Рассчитанная рекомендуемая оценка за полный правильный ответ (УТЕ)	4,0	7,0	11,2	16,8	24,0
Рассчитанная рекомендуемая оценка за неполный правильный ответ (УТЕ) — один из трёх правильных ответов два из трёх правильных ответов	1,3 2,6	2,3 4,6	3,7 7,5	5,6 11,2	8,0 16,0
Отсутствие правильного ответа (УТЕ)	0	0	0	0	0

Таблица 4

**Выбор четырёх правильных ответов ($k = 4$) из нескольких предлагаемых ответов ($n = 8 \div 10$)
(на основе числа возможных сочетаний C_n^4)**

Число предлагаемых ответов n	8	9	10
Число возможных сочетаний C_n^4	70	126	210
Вероятность угадывания полного правильного ответа $P(A)$	1/70	1/126	1/210
Рассчитанная рекомендуемая оценка за полный правильный ответ (УТЕ)	14,0	25,2	42,0
Рассчитанная рекомендуемая оценка за неполный правильный ответ (УТЕ) — один из четырёх правильных ответов два из четырёх правильных ответов три из четырёх правильных ответов	3,5 7,0 10,5	6,3 12,6 18,9	10,5 21,0 31,5
Отсутствие правильного ответа (УТЕ)	0	0	0

Таблица 5

**Выбор двух правильных ответов ($k = 2$) из нескольких предлагаемых ответов ($n = 4 \div 10$)
(на основе оценки каждого правильного ответа)**

Число предлагаемых ответов n	4	5	6	7	8	9	10
Вероятность угадывания одного правильного ответа $P(A)$	1/4	1/5	1/6	1/7	1/8	1/9	1/10
Рекомендуемая оценка за один (каждый) правильный ответ (УТЕ)	0,8	1,0	1,2	1,4	1,6	1,8	2,0
Рекомендуемая оценка за полный правильный ответ (УТЕ)	1,6	2,0	2,4	2,8	3,2	3,6	4,0
Отсутствие правильного ответа (УТЕ)	0	0	0	0	0	0	0

Таблица 6

**Выбор четырёх правильных ответов ($k = 4$) из нескольких предлагаемых ответов ($n = 8 \div 10$)
(на основе оценки каждого правильного ответа)**

Число предлагаемых ответов n	8	9	10
Вероятность угадывания одного правильного ответа $P(A)$	1/8	1/9	1/10
Рекомендуемая оценка за один (каждый) правильный ответ (УТЕ)	1,6	1,8	2,0
Рекомендуемая оценка за полный правильный ответ (УТЕ)	6,4	7,2	8,0
Отсутствие правильного ответа (УТЕ)	0	0	0

При наличии $m_j > n$ возможно повторное использование элементов второго множества для соответствующих (комплементарных) пар. Однако и в этом случае, вероятность парного соответствия для каждого из m_j элементов также будет составлять $1/n$.

Тестовые задания на установление соответствия по сути дела представляют собой совокупность нескольких ТЗ по выбору одного правильного ответа из нескольких предлагаемых. Общая оценка за задания на установление соответствия будет равна сумме оценок за каждую правильную пару в соответствии с вероятностью каждого правильного ответа.

Приведём пример: в задании нужно установить соответствие между элементами первого множества (А, В, С и D) и элементами второго множества (g, h, i и j). Предположим, что верный ответ: Ag, Bi, Ch, Dj. Тогда оценка за каждую правильную пару (выбор одного правильного ответа из 4 предлагаемых) составляет 0,8 УТЕ. Общая оценка за тестовое задание: $0,8 \times 4 = 3,2$ УТЕ.

Пусть в задании нужно установить соответствие между элементами первого множества (А, В, С и D) и элементами второго множества, состоящего из меньшего числа элементов (g, h и i). Предположим, что верный ответ: Ag, Bi, Ch, Di. Оценка за каждую правильную пару (выбор одного правильного ответа из 3 предлагаемых) составляет 0,6 УТЕ. Общая оценка за тестовое задание: $0,6 \times 4 = 2,4$ УТЕ.

Для заданий на **установление правильной последовательности** с математической точки зрения, вероятность угадывания правильной последовательности обратно пропорциональна числу перестановок P_n , где n — число компонентов: $P(A) = \frac{1}{P_n}$.

В таблице 7 представлены вероятности угадывания правильной последовательности из n предлагаемых компонентов на основе числа возможных перестановок и соответствующие оценки.

Альтернативный вариант оценивания тестовых заданий на установление правильной последовательности: испытуемому предлагается на выбор несколько

Таблица 7

**Установление правильной последовательности
(на основе числа возможных перестановок P_n)**

Число компонентов n	2	3	4	5	6
Число возможных перестановок P_n	2	6	24	120	720
Вероятность угадывания правильного ответа	1/2	1/6	1/24	1/120	1/720
Рассчитанная рекомендуемая оценка за правильный ответ (УТЕ)	0,4	1,2	4,8	24,0	144,0
Отсутствие правильного ответа (УТЕ)	0	0	0	0	0

фиксированных последовательностей, из которых лишь одна является правильной. По сути дела, это является одним из вариантов ТЗ с выбором одного правильного ответа из нескольких предлагаемых с соответствующим оцениванием (см. табл. 1).

* * *

Предлагаемая система оценивания разных видов тестовых заданий закрытого типа является рамочной и подлежит обязательной эмпирической проверке с учётом специфики учебной дисциплины с последующей корректировкой. Система не привязана к какому-либо виду тестовых заданий или какому-нибудь определённому количеству заданий. Предлагаемая система позволяет сопоставлять оценки, полученные за тестовые задания разных видов, а также оценивать неполные правильные ответы. Абсолютно необходимое условие — одинаковое количество равноценяемых заданий для всех тестируемых.

Предлагаемая система оценок, как показывает наш опыт, применима для оценки заданий с указанием числа правильных ответов. Отсутствие подсказки (указания числа правильных ответов) повышает сложность заданий. Поэтому целесообразно вводить повышающий коэффициент,

например, 1,2–1,3. На наш взгляд, повышающего коэффициента заслуживают также тестовые задания, представляющие особую значимость для какой-либо конкретной дисциплины.

Выбор того или иного варианта оценивания осуществляется в процессе конструирования теста. По завершении тестирования определяется сумма полученных результатов всех тестовых заданий. Итоговая оценка представляет собой отношение суммы полученных результатов к максимально возможной сумме результатов всех заданий, выраженное в процентах.

Предлагаемая система оценивания результатов тестирования может быть реализована с помощью компьютерной системы тестирования знаний MyTestXPro⁵ с последующей обработкой и анализом результатов тестирования. С помощью модуля тестирования можно организовать как локальное, так и сетевое тестирование. 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.

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