

МАСТЕР-КЛАСС

**Критериально-ориентированное
педагогическое тестирование**

**Модуль № 3. Оценка надежности
теста**

© РЭА им. Г.В. Плеханова
© Управление качеством учебного процесса

Оценка надежности теста (1)

- Надежность теста отражает точность педагогического измерения и устойчивость результатов тестирования к воздействию посторонних случайных факторов.
- Для оценки надежности теста необходимо иметь матрицу ответов группы студентов на тест.
- Для каждого студента путем суммирования числа правильных ответов определяется «сырой балл».

<i>№ испытуемого</i>	<i>Задания</i>								
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>
<i>1</i>	1	1	1	1	1	0	0	1	1
<i>2</i>	1	1	1	0	0	0	0	1	1
<i>3</i>	1	1	1	0	0	0	1	1	1
<i>4</i>	1	1	1	1	1	0	0	1	1
<i>5</i>	1	0	0	0	1	1	1	1	0
<i>6</i>	0	0	1	1	1	1	1	0	1
<i>7</i>	0	0	0	1	1	1	1	0	0
<i>8</i>	0	0	1	1	1	1	0	0	0
<i>9</i>	1	1	1	0	0	0	1	1	1
<i>10</i>	0	0	0	0	1	1	1	1	1
<i>11</i>	1	1	1	1	1	1	1	1	1
<i>12</i>	1	1	1	1	1	1	1	1	1

Оценка надежности теста (2)

- Пусть критерий зачета – 60% и более выполненных заданий. Тогда для теста из 26 заданий зачет будет получать студенты, выполнившие 16 и более заданий.
- На основании результатов тестирования составляется таблица сопряжения результатов тестирования «2x2», в которой a, b, c, d – доли испытуемых, получивших соответствующую аттестацию по результатам двух параллельных форм теста.
- На основе таблицы сопряжения вычисляются коэффициент надежности как фи-коэффициент корреляции и каппа-коэффициент надежности теста.

		<i>Параллельная форма теста - А</i>	
		Незачет	Зачет
<i>Параллельная форма теста - Б</i>	Незачет	a=0,5	b=0
	Зачет	c=0,05	d=0,45

Оценка надежности теста (3)

1. $\varphi = \frac{ad - bc}{\sqrt{(a + c)(b + d)(a + b)(c + d)}}$ - фи-коэффициент корреляции

2. $P = a + d$ - вероятность принятия согласованного решения по таблице сопряжения результатов критериально-ориентированного теста, представляющая собой сумму вероятностей принятия согласованных решений по отнесению испытуемых в каждую из групп.

$P_c = (c + d)(d + b) + (a + b)(c + a)$ – вероятность случайного согласования

$\kappa = (P - P_c) / (1 - P_c)$ – оценка надежности критериально-ориентированного теста, учитывающая случайную согласованность. Каппа-коэффициент надежности теста $\kappa = [-1; + 1]$

3. Равенство фи и каппа коэффициентов надежности, а также превышение этими величин значения 0,8, говорит о надежности проведенных тестовых испытаний.

Способы увеличения надежности теста

- Провести тщательный анализ каждого тестового задания по соответствию его формы и содержания основам тестовой теории.
- Увеличить количество заданий в тесте.
- Удалить из теста те задания, которые плохо различают подготовленных и неподготовленных учащихся.
- Использовать задания в закрытой форме с четырьмя и более вариантами ответов для уменьшения влияния случайного угадывания

Статистический анализ качества тестовых заданий. Коэффициент D_j

- По каждому заданию в тесте рассчитывается коэффициент различающей способности D_j путем выделения двух контрастных групп испытуемых (на основе анализа сырых баллов отделяются 27% слабых и 27% сильных студентов)
- P_U и P_L - доли студентов (процент от общего числа сдававших тест) в лучшей и слабой группе, правильно ответивших на данное j -е задание.
- Если $D_j \geq 0,4$ – задание вполне эффективно
- От 0,3 до 0,39 – удовлетворительное
- От 0,2 до 0,29 – задание следует проанализировать на пригодность использования в тесте
- $< 0,19$ – задание необходимо изъять или тщательно проанализировать и переработать. Оно практически не обладает дифференцирующей способностью.
- Если $D_j < 0$ – задание некачественное, т.к. лучшая группа студентов отвечает на него хуже, чем слабая.

$$D_j = p_U - p_L$$

Методы подсчета тестовых баллов – коррекция на угадывание (1)

Пусть испытуемый **А** дал 35 верных ответов, 10 неверных и 5 заданий пропустил в тесте с выбором из 5 вариантов ответа. Испытуемый **Б**: 36 верных и 14 неверных в том же тесте. Испытуемый **В**: 34 дает верных ответа и пропускает все остальные задания.

1. Расчет поправки на угадывание по алгоритму «Правильный плюс пропуск» осуществляется по формуле:

$$x_i = R + \frac{H}{k}$$

где x_i - скорректированный индивидуальный балл;

R – число правильных ответов;

H – число пропущенных заданий;

k – число вариантов ответов в задании.

- Испытуемый **А**

$$X_n = 35 + \frac{5}{5} = 35 + 1 = 36$$

- Испытуемый **Б**

$$X_n = 36 + \frac{0}{5} = 36 + 0 = 36$$

- Испытуемый **В**

$$X_n = 34 + \frac{16}{5} = 34 + 3,2 = 37,2$$

Методы подсчета тестовых баллов – коррекция на угадывание (2)

2. Расчет поправки на угадывание по алгоритму «Правильный минус неправильный» осуществляется по формуле:

$$x_i = R - \frac{w}{(k-1)}$$

где x_i - скорректированный на угадывание индивидуальный балл;

R – число правильных ответов;

N – число неправильных ответов;

k – число вариантов ответов в задании.

- Испытуемый **А**

$$X_n = 35 - \frac{10}{5-1} = 35 - \frac{10}{4} = 35 - 2,5 = 32,5$$

- Испытуемый **Б**

$$X_n = 36 - \frac{14}{5-1} = 36 - \frac{14}{4} = 36 - 3,5 = 32,5$$

- Испытуемый **В**

$$X_n = 34 - \frac{0}{5-1} = 34 - 0 = 34$$

Результаты теста – сырые баллы испытуемых

№ испыту емого	Задания										
	1	2	3	4	5	6	7	8	9	...	X_i
1	1	1	1	1	1	0	0	1	1		34
2	1	1	1	0	0	0	0	1	1		26
3	1	1	1	0	0	0	1	1	1		34
4	1	1	1	1	1	0	0	1	1		31
5	1	0	0	0	1	1	1	1	0		28
6	0	0	1	1	1	1	1	0	1		32
7	0	0	0	1	1	1	1	0	0		30
8	0	0	1	1	1	1	0	0	0		30
9	1	1	1	0	0	0	1	1	1		36
10	0	0	0	0	1	1	1	1	1		27
11	1	1	1	1	1	1	1	1	1		46
12	1	1	1	1	1	1	1	1	1		44
13	1	1	1	1	1	1	1	1	1		50

Методы подсчета тестовых баллов

- Традиционный - $x_i = \sum_{j=1}^n x_{ij}$, где x_i – общий индивидуальный балл i -го испытуемого, x_{ij} – индивидуальный балл i -го испытуемого при ответе на j -е задание теста.
- За правильный ответ испытуемый получает 1, за неправильный – 0.
- Однако, при ответе на закрытые вопросы (с вариантами ответа) существует различие между испытуемыми по степени желания пропускать в тесте те вопросы, на которые они не знают правильного ответа. Чтобы уравнивать испытуемых существуют алгоритмы коррекции на угадывание.