

Приложение 1  
к Методическим рекомендациям по  
проведению стандартизации экзаменационных  
заданий государственной итоговой аттестации  
(п.4.2.)

### Определение трудности тестовых заданий и их дискриминативности

Для проверки трудности тестовых заданий и их дискриминативности необходимо создать Матрицу результатов. Она представляет собой таблицу, в которой располагают результаты обучающихся, полученные при проведении пробного тестирования. Составлять и обрабатывать матрицу результатов следует в программе Excel.

#### ШАГ 1. Составление матрицы результатов

1.1. В первый столбец таблицы записываются ФИО обучающихся (или порядковый номер обучающегося в классном журнале). Столбцы таблицы, начиная со второго – соответствуют номерам тестовых заданий государственной итоговой аттестации.

Строки таблицы соответствуют результатам выполнения тестовых заданий обучающимися (участниками апробации).

1.2. Для оценки результатов выполнения тестовых заданий каждым обучающимся используется дихотомическая шкала: «1» выставляется за правильный ответ или правильное решение тестового задания и «0» за неправильный ответ и неправильное решение тестового задания или частично верный ответ.

Таким образом, на пересечении строк и столбцов находится число, соответствующее результату выполнения (правильно/неправильно) конкретного тестового задания конкретным обучающимся.

1.3. В последнем столбце таблицы записывается количество заданий, которое правильно выполнил каждый обучающийся. Это количество правильно выполненных заданий называется индивидуальный балл обучающегося ( $X_i$ ).

Пример 1.1.

**Пример матрицы результатов № 1**

ФИО обучающегося (i)	Номер задания (j)										Индивидуальный балл ( $X_i$ )
	1	2	3	4	5	6	7	8	9	10	
Антонов	1	1	1	1	1	1	1	0	0	0	7
Волк	1	1	0	0	0	0	0	0	0	0	2
Громова	0	0	0	0	0	0	0	1	0	0	1
Зверева	1	1	0	1	1	1	1	1	1	1	9
Иванко	0	1	1	1	1	0	0	0	0	0	4
Кузьмин	1	1	1	0	0	0	0	0	0	0	3
Кудин	1	1	1	0	0	1	0	0	0	0	4
Петров	1	1	1	0	0	1	0	0	0	0	4
Притула	1	1	1	1	1	1	0	1	0	1	8
Чурикова	1	1	1	0	1	1	1	0	0	0	6
<b>Суммарное значение правильных ответов (<math>R_j</math>)</b>	<b>8</b>	<b>9</b>	<b>7</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>3</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b><math>\Sigma=48</math></b>

1.4. В последней строке таблицы подсчитывается и записывается сколько

обучающихся правильно ответили на конкретное тестовое задание. Этот показатель является суммарным значением правильных ответов на конкретное тестовое задание ( $R_j$ ).

### ШАГ 2. Выбраковка тестовых заданий

Выбраковку тестовых заданий проводят следующим образом:

1) удаляют тестовые задания, которые не смог выполнить никто из обучающихся, трудные задания (удаляют из матрицы результатов столбцы, состоящие полностью из 0);

2) удаляют тестовые задания, которые выполнили все обучающиеся, легкие задания (удаляют из матрицы результатов столбцы, состоящие полностью из 1);

И первые и вторые задания не обладают дифференцирующей способностью и не позволяют разграничивать сильных и слабых обучающихся, однако дают представление о сложности заданий для обучающихся.

Тестовые задания, которые выполнили все обучающиеся можно отнести к базовому уровню, те задания, с которыми на справился никто следует заменить на другие задания.

### ШАГ 3. Определение трудности тестовых заданий и их места в контрольном комплексе заданий

Для примерной оценки сложности тестовых заданий необходимо элементы матрицы результатов проранжировать от большего значения к меньшему по суммарному значению правильных ответов на каждое тестовое задание (величине  $R_j$ )

Ниже приведен пример матрицы результатов № 2, полученной после упорядочения элементов по величине суммарного значения правильных ответов на конкретное задание.

#### Пример 1.2.

**Пример матрицы результатов № 2**  
(упорядочение элементов матрицы результатов по количеству правильных ответов на тестовые задания)

ФИО обучающегося (i)	Номер задания (j)										Индивидуальный балл ( $X_i$ )
	2	1	3	6	5	4	7	8	10	9	
Зверева	1	1	0	1	1	1	1	1	1	1	9
Пригула	1	1	1	1	1	1	0	1	1	0	8
Антонов	1	1	1	1	1	1	1	0	0	0	7
Чурикова	1	1	1	1	1	0	1	0	0	0	6
Иванко	1	0	1	0	1	1	0	0	0	0	4
Кудин	1	1	1	1	0	0	0	0	0	0	4
Петров	1	1	1	1	0	0	0	0	0	0	4
Кузьмин	1	1	1	0	0	0	0	0	0	0	3
Волк	1	1	0	0	0	0	0	0	0	0	2
Громова	0	0	0	0	0	0	0	1	0	0	1
Суммарное значение правильных ответов ( $R_j$ )	9	8	7	6	5	4	3	3	2	1	$\Sigma=48$
Трудность заданий ( $P_j$ )	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,3	0,2	0,1	

Мерой трудности заданий будем считать долю правильных ответов, которая определяется путем деления числа правильных ответов по каждому заданию ( $R_j$ ) на общее число обучающихся, участвующих в пробном тестировании ( $N$ ).

$$P_j = R_j/N$$

Вычислим трудность выполнения каждого задания и впишем ее в строку, которую добавим ниже.

Коэффициент трудности задания изменяется в диапазоне от 0 до 1 и может быть выражен в долях или %. В таблице 1 представлены значения показателя трудности тестовых заданий и степень трудности тестовых заданий.

Таблица 1

Степень трудности тестовых заданий

Значение показателя трудности тестовых заданий (в долях)	Степень трудности тестовых заданий
0 – 0,19	очень сложные тестовые задания
0,2 – 0,49	сложные тестовые задания
0,5 – 0,69	тестовые задания средней сложности
0,7 – 0,89	легкие тестовые задания
0,9 – 1	очень легкие тестовые задания

Будем исходить из того, чем больше правильных ответов дано на задание, тем оно легче для данной группы обучающихся.

Самыми легкими является заданиями были задание № 2 ( $P_j = 0,9$ ), т.е. с данным заданием справились 90% обучающихся и задание № 1 ( $P_j = 0,8$ ), т.е. с данным заданием справились 80% обучающихся.

Оптимальным значением степени трудности задания является показатель  $P_j = 0,5 \pm 0,1$ .

Таким образом, задания № 2, 1, 3, 6, 5, 4 можно отнести к базовому уровню трудности, задания № 7, 8 – к повышенному уровню трудности и задания № 9, 10 – к высокому уровню трудности.

Статистическая трудность позволяет определить место тестового задания в контрольном комплексе заданий. В начале ККЗ, помещают тестовые задания базового уровня трудности, затем задания повышенного уровня и в конце ККЗ помещают задания высокого уровня трудности.

#### **ШАГ 4. Определение дифференцирующей способности тестовых заданий (дискриминативности)**

Тестовые задания должны обладать способностью различать обучающихся с различным уровнем подготовки: сильные обучающиеся должны справляться с заданиями повышенного и высокого уровня трудности, менее подготовленные – с заданиями базового уровня.

Если же складывается ситуация, когда сильные обучающиеся не отвечают на задание правильно, а слабые – отвечают, то в этом случае задание не только не дифференцирует обучающихся, но и вносит дезинформацию в их оценивание. Про такие задания говорят, что они имеют отрицательную дискриминативность.

Показатель дифференцирующей способности тестовых заданий (показатель дискриминативности), рассчитывается по формуле:

$$D_n = P_n^1 - P_n^2$$

где,

$P_n^1$  – доля обучающихся, правильно выполнивших конкретное задание (n) в подгруппе из сильных обучающихся. В матрице результатов отбирают 25% – 27% обучающихся с самыми высокими результатами.

$P_n^2$  – доля обучающихся, правильно выполнивших конкретное задание (n) в подгруппе из слабых обучающихся. В матрице результатов отбирают 25% – 27% обучающихся с самыми низкими результатами.

В нашем случае сильные обучающиеся: Зверева, Притула, Антонов и слабые обучающиеся: Кузьмин, Волк, Громова.

Пример 1.3.

**Пример матрицы результатов № 3**  
(определение дифференцирующей способности тестовых задания)

ФИО обучающегося (i)	Номер задания (j)										Индивидуальный балл (X <sub>i</sub> )
	2	1	3	6	5	4	7	8	10	9	
Сильные обучающиеся											
Зверева	1	1	0	1	1	1	1	1	1	1	<b>9</b>
Притула	1	1	1	1	1	1	0	1	1	0	<b>8</b>
Антонов	1	1	1	1	1	1	1	0	0	0	<b>7</b>
Кол-во правильных ответов	3	3	2	3	3	3	2	2	2	1	
$P_n^1$	1,00	1,00	0,67	1,00	1,00	1,00	0,67	0,67	0,67	0,33	
Слабые обучающиеся											
Кузьмин	1	1	1	0	0	0	0	0	0	0	<b>3</b>
Волк	1	1	0	0	0	0	0	0	0	0	<b>2</b>
Громова	0	0	0	0	0	0	0	1	0	0	<b>1</b>
Кол-во правильных ответов	2	2	1	0	0	0	0	1	0	0	
$P_n^2$	0,67	0,67	0,33	0,00	0,00	0,00	0,00	0,33	0,00	0,00	
Показатель дискриминативности ( $D_n$ )	0,33	0,33	0,33	1,00	1,00	1,00	0,67	0,33	0,67	0,33	

Далее находим показатели  $P_n^1$  и  $P_n^2$  для каждого тестового задания. Для этого количество правильных ответов на каждое задание делим на количество отвечавших на него. И из первого показателя вычитаем второй.

Показатель  $D_n$  варьирует в пределах:  $-1 \leq D_n \leq 1$

Если задание правильно выполнили больше сильных обучающихся, чем слабых, то показатель дискриминативности  $D_n > 0$ . Если задание правильно выполнили больше слабых обучающихся, чем сильных, то показатель дискриминативности  $D_n < 0$ . Если задание правильно выполнили одинаковое количество сильных и слабых обучающихся, то  $D_n = 0$ , а значит, данное тестовое задание не позволяет дифференцировать обучающихся по уровню их подготовки.

В таблице 2 приведены значения показателей дискриминативности тестовых заданий и их общепринятая интерпретация.

Значения показателя дискриминативности тестовых заданий

Показатели дискриминативности тестовых заданий	Значение показателя дискриминативности тестовых заданий
0 – 0,20	задание должно быть исключено из теста или полностью переделано
0,21 – 0,30	задание нуждается в пересмотре
0,31 – 0,40	требуется небольшая коррекция задания
0,41 – 0,60	задание функционирует удовлетворительно
0,61 – 1,0	задание с отличной дискриминативностью

На основании матрицы результатов № 3 высокой степенью дифференциации (отличным показателем дискриминативности) обладают тестовые задания № 4, 5, 6, 7, 10, нуждаются в небольшой коррекции тестовые задания № 1, 2, 3, 8, 9. Тестовых заданий, которые нуждаются в пересмотре в данном контрольном комплексе заданий нет.