

КАЧЕСТВЕННАЯ ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ — ОСНОВА ОБЪЕКТИВНОСТИ КОНТРОЛЬНО- ИЗМЕРИТЕЛЬНЫХ МАТЕРИАЛОВ

Оксана Викторовна Гурбунова,

учитель английского языка МБОУ СОШ № 28, г. Мытищи, inshakovaox@mail.ru

Ольга Анатольевна Иванова,

учитель английского языка ГБОУ «Школа № 1095», г. Москва, olga1979iv@mail.ru

• оценка уровня знаний • тестирование • интерпретация результатов тестирования • метод графического анализа • тестовые технологии

В области образования одной из актуальных является проблема оценки уровня знаний (подготовленность) учащихся по различным предметам. Объективность такой оценки необходима прежде всего для принятия верного решения при проектировании деятельности учащихся по окончании средней общеобразовательной школы.

Тестирование — это стандартизированный метод оценки знаний, умений, навыков учащихся, который помогает выявить и сформировать индивидуальный темп обучения, пробелы в текущей итоговой подготовке. Но качество результатов будет максимально объективным только при условии правильной интерпретации результатов.

Необходимость разработки методов интерпретации вызвана стремлением выявить истинные различия в уровне знаний испытуемых, сравнить результаты между собой. Это стремление продиктовано достаточно серьезной причиной, связанной с тем, что «сырые» или тестовые баллы не дают информации о реальном уровне знаний. Интерпретация результата и его отнесение к категории плохих или хороших зависят от целого ряда факторов. Прежде всего от того, как распределились результаты остальных испытуемых, от трудности и от количества заданий в тесте. Методы многофакторного анализа результатов тестирования являются средством объектив-

ной экспертизы содержательных компонентов качества образования и личностного развития: знаний, умений, навыков, творческого опыта учебной деятельности, ценностно-смыслового отношения к обучению, ключевых компетенций и др.

Адекватность интерпретации достигается путём сопоставления результатов анализируемой выборки или отдельного испытуемого с определёнными нормами выполнения теста — множеством показателей, устанавливаемых эмпирически в процессе стандартизации теста. Соотнесение результата испытуемого с нормами выполнения позволяет установить, соответствует ли данный результат среднему или насколько он выше или ниже среднего результата выполнения теста.

Результаты тестирования — это очень важный момент в процессе обучения. Они представляют собой не просто итоговый балл учащихся за тот или иной тест, а позволяют подробно проанализировать процессы выполнения теста одним учащимся или группой учащихся. Подробный анализ результатов тестирования позволяет учителю увидеть основные типичные ошибки учащихся и ещё раз обратить на них внимание. Кроме того, у преподавателя появляется возможность отследить качество тестовых заданий.

Для всестороннего анализа результатов тестового контроля данные представляются

в таблицах, на графиках, гистограммах и диаграммах, наиболее полно отображающих количественные и качественные показатели исследуемых массивов участников контрольно-оценочного процесса.

Наиболее удобным и показательным для представления статистических данных является метод графического анализа. Парные и множественные сравнения объектов наблюдения одного и того же или разных уровней проводятся по средним первичным или тестовым баллам, по медианным значениям или процентам выполнения заданий. Между собой могут сравниваться статистические данные нескольких объектов: классы в школе, школы в районе, районы в регионе и т. д.

В настоящее время существуют два теоретических подхода к созданию тестов: классическая теория и современная теория IRT (Item Response Theory). Оба подхода базируются на последующей статистической обработке так называемого сырого балла (raw score), то есть балла, набранного в результате тестирования. Только после проведения многократных статистических обработок можно говорить о создании теста с устойчивыми параметрами качества (надёжностью и валидностью).

Для обработки данных, полученных на этапе тестирования, воспользуемся пакетом MS Office и электронными таблицами MS Excel.

После сбора эмпирических данных необходимо провести статистическую обработку – с помощью выше перечисленного пакета программ. Этап математико-статистической обработки разобьём на ряд шагов.

Итак, наши исходные результаты тестирования, проведённого в классе N.

Если за каждый правильный ответ на задание испытуемому давать 1 балл, а за неправильный ответ или пропуск задания — 0 баллов, то профиль ответов учащихся будет иметь вид последовательности из единиц и нулей. Поскольку каждая единица или нуль появляются в результате взаимодействия испытуемого с заданием, то наиболее адекватной формой представления наблюдаемых результатов выполнения теста служит матрица, т.е. прямоугольная таблица, сводящая воедино профили ответов учащихся и профили заданий теста (столбцы из оценок всех учащихся по каждому заданию теста).

Шаг 1. Формирование матрицы тестовых результатов

Номера испытуемых, i	Номера заданий j										Индивидуальные баллы (множество X_i)
	1	2	3	4	5	6	7	8	9	10	
1	1	1	1	1	1	1	0	0	0	0	6
2	1	1	0	0	0	0	0	0	0	0	2
3	0	0	0	0	0	0	0	1	0	0	1
4	1	1	0	1	1	1	1	1	1	1	9
5	1	0	1	0	1	1	0	0	0	0	4
6	1	1	1	0	0	0	0	1	0	0	4
7	1	1	1	1	0	1	0	0	0	0	5
8	1	1	1	1	0	0	0	0	0	0	4
9	1	1	1	1	1	1	1	1	1	0	9
10	1	1	1	1	1	0	1	0	0	0	6
11	0	0	0	0	0	0	0	0	0	0	11
12	1	1	1	1	1	1	1	1	1	1	12
Число правильных ответов (множество Y_j)											

Справа в вертикальном столбце содержатся индивидуальные баллы учащихся X_i , которые получаются суммированием единиц

по горизонтали в каждом профиле ответов. Сложение единиц в столбцах по профилям ответов на n заданий теста ($n=10$) позволя-

ет получить числа Y_j , соответствующие количеству правильных ответов на каждое задание. С помощью матрицы можно выполнить ряд расчётов, интерпретация результатов которых позволяет сделать важные выводы относительно качества заданий теста и получить достаточно точные оценки параметра испытуемых в том случае, если тест соответствует определённым критериям качества.

Результаты ответов учеников на задания тестов оцениваются в дихотомической шкале: за каждый правильный ответ учащийся получает один балл, а за неправильный ответ или за пропуск задания — нуль баллов.

Пользуясь электронными таблицами MS Excel, создаём таблицу.

Таблица 1.1

Тема 4.1		A	B	C	D	E	F	G	H	I	J	K
		Номера заданий j										
Номера испытуемых i		1	2	3	4	5	6	7	8	9	10	
1												
2	i											
3	1	1	1	1	1	1	1	0	0	0	0	
4	2	1	1	0	0	0	0	0	0	0	0	
5	3	0	0	0	0	0	0	0	1	0	0	
6	4	1	1	0	1	1	1	1	1	1	1	
7	5	1	0	1	0	1	1	0	0	0	0	
8	6	1	1	1	0	0	0	0	1	0	0	
9	7	1	1	1	1	0	1	0	0	0	0	
10	8	1	1	1	1	0	0	0	0	0	0	
11	9	1	1	1	1	1	1	1	1	1	0	
12	10	1	1	1	1	1	0	1	0	0	0	
13	11	0	0	0	0	0	0	0	0	0	0	
14	12	1	1	1	1	1	1	1	1	1	1	

Шаг 2. Преобразование матрицы тестовых результатов

На втором шаге из матрицы тестовых результатов устраняются строки и столбцы, состоящие только из нулей или только из единиц. В приведённом выше примере таких столбцов нет, а строк только две. Одна из них, нулевая строка, соответствует ответам одиннадцатого испытуемого, который не смог выполнить правильно ни одного задания в тесте.

В этом случае вывод довольно однозначен: *тест непригоден для оценки знаний такого ученика. Для выявления его уровня знаний тест необходимо облегчить, добавив несколько более легких заданий, которые, скорее всего, выполнит правильно большинство остальных испытуемых группы.*

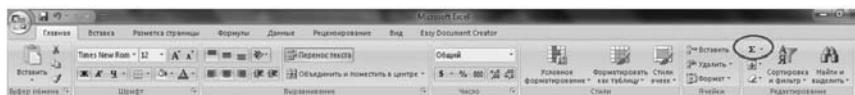
Столь же непригоден, но уже по другой причине, тест для оценки знаний двенадцатого ученика, который выполнил правильно все

без исключения задания теста. Причина непригодности теста заключается в его излишней лёгкости, не позволяющий выявить истинный уровень подготовки двенадцатого ученика. Возможно, двенадцатый ученик знает много чего другого и в состоянии выполнить по контролируемым разделам содержания гораздо более трудные задания, которые просто не были включены в тест.

Таким образом, на данном шаге необходимо удалить из матрицы данных 11 и 12 строки.

Шаг 3. Подсчёт индивидуальных баллов испытуемых и количество правильных ответов на каждое задание теста

Индивидуальный балл испытуемого получается суммированием всех единиц, полученных им за правильное выполнение задания теста. В Excel для суммирования данных по строке можно воспользоваться кнопкой «Автосумма» на панели инструментов Стандартная.



Для удобства полученные индивидуальные баллы (X_i) приводятся в последнем столбце матрицы результатов.
 Число правильных ответов на задания теста (Y_i) также получается суммированием единиц, но уже расположенным по столбцам.

Таблица 1.2

Тема 4.1												
	A	B	C	D	E	F	G	H	I	J	K	L
16	Номера заданий j											Индивидуальные баллы (число X _i)
17	Номера испытуемых i	1	2	3	4	5	6	7	8	9	10	
18	i	1	2	3	4	5	6	7	8	9	10	
19	1	1	1	1	1	1	1	0	0	0	0	6
20	2	1	1	0	0	0	0	0	0	0	0	2
21	3	0	0	0	0	0	0	0	1	0	0	1
22	4	1	1	0	1	1	1	1	1	1	1	9
23	5	1	0	1	0	1	1	0	0	0	0	4
24	6	1	1	1	0	0	0	0	1	0	0	4
25	7	1	1	1	1	0	1	0	0	0	0	5
26	8	1	1	1	1	0	0	0	0	0	0	4
27	9	1	1	1	1	1	1	1	1	1	0	9
28	10	1	1	1	1	1	0	1	0	0	0	6
29	Число правильных ответов (число Y _i)	9	8	7	6	5	5	3	4	2	1	50

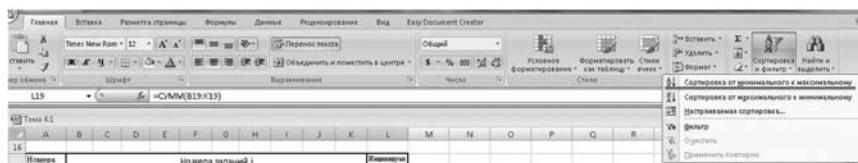
Шаг 4. Упорядочение матрицы результатов

Для анализа обычно используется упорядоченная матрица, в которой не только задания ранжированы по нарастанию трудности (см. табл. 1.2), но и баллы испытуемых расположены по убыванию или нарастанию сверху вниз (табл. 1.3).
 Значения индивидуальных баллов необходимо отсортировать по возрастанию, для этого в MS Excel:

- Выделим блок ячеек, содержащих номера испытуемых, матрицу результатов и индивидуальные баллы. Начинать выделение необходимо со столбца X (индивидуальные баллы);
- На панели инструментов *Стандартная* нажимаем на кнопку *Сортировка* по возрастанию. Матрица результатов примет вид, изображённый в табл. 1.3

Таблица 1.3

Тема 4.1												
	A	B	C	D	E	F	G	H	I	J	K	L
16	Номера заданий j											Индивидуальные баллы (число X _i)
17	Номера испытуемых i	1	2	3	4	5	6	7	8	9	10	
18	i	1	2	3	4	5	6	7	8	9	10	
19	3	0	0	0	0	0	0	0	0	1	0	1
20	2	1	1	0	0	0	0	0	0	0	0	2
21	5	1	0	1	0	1	1	0	0	0	0	4
22	6	1	1	1	0	0	0	0	1	0	0	4
23	8	1	1	1	1	0	0	0	0	0	0	4
24	7	1	1	1	1	0	1	0	0	0	0	5
25	1	1	1	1	1	1	1	0	0	0	0	6
26	10	1	1	1	1	1	0	1	0	0	0	6
27	4	1	1	0	1	1	1	1	1	1	1	9
28	9	1	1	1	1	1	1	1	1	1	0	9
29	Число правильных ответов (число Y _i)	9	8	7	6	5	5	3	4	2	1	50



Шаг 5. Графическое представление данных

По данным матрицы обычно проводят графическую интерпретацию распределений для трудности заданий и индивидуальных баллов испытуемых, которые представляют в виде полигона, гистограммы или сглаженной кривой.

Для построения кривых упорядочим результаты тестирования и подсчитаем частоту получения баллов.

Для этого формируем несгруппированный ряд:

1. Мышкой выделяем колонку A и нажимаем клавиши «Ctrl+C».
2. Щёлкаем мышкой на свободном поле и нажимаем клавиши «Ctrl+V».
3. Так же поступаем с колонкой X. Внизу столбца появляется значок «чемоданчика», нажимаем на него и выбираем «Только значения».

Номера испытуемых	Индивидуальные баллы (множество X_i)
i	
1	9
2	4
3	5
4	19
5	11
6	11
7	13
8	12
9	23
10	18

раскрываем чемоданчик и выбираем "ТОЛЬКО ЗНАЧЕНИЯ"

У нас получился несгруппированный ряд.

N	O	P	Q
	Номера испытуемых	Индивидуальные баллы (множество X_i)	
	i		
	1	6	
	2	2	
	3	1	
	4	9	
	5	4	
	6	4	
	7	5	
	8	4	
	9	9	
	10	6	

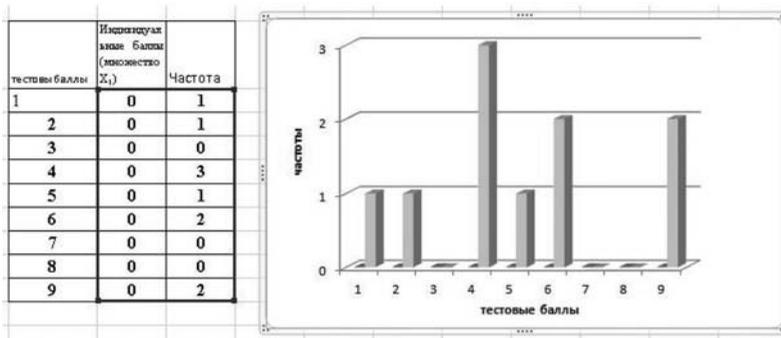
4. Используя фильтр, ранжируем колонку «Индивидуальные баллы» по возрастанию. У нас получился сгруппированный ряд.

Номера испытуемых	Индивидуальные баллы (множество X_i)
i	
3	1
2	2
5	4
6	4
8	4
7	5
1	6
10	6
4	9
9	9

5. Теперь подсчитываем частотное распределение баллов.

Индивидуальные баллы (множество X_i)	Частота
1	1
2	1
3	0
4	3
5	1
6	2
7	0
8	0
9	1

6. По ряду частотного распределения можно получить графическое представление результатов тестирования в виде гистограммы — последовательности столбцов, каждый из которых опирается на единственный (разрядный) интервал, а высота его пропорциональна частоте наблюдаемых баллов. Чтобы построить такого рода гистограмму, в столбце «Индивидуальные баллы» выставляем значения 0. Это необходимо для того, чтобы на горизонтальной оси появились значения тестовых баллов. Для подписи осей используем «Макет», «Подписи», «Название осей». На рисунке синим (тёмным) прямоугольником отмечена область построения.



Данные результаты позволяют выявить учащихся, наиболее восприимчивых к получению новых знаний, а также «проблемных» учащихся, не имеющих серьезной мотивации к обучению по данной дисциплине и требующих административного воздействия, и степень стратификации учащихся на отдельные группы, определяемые способностями и мотивацией.

«Плотность распределения баллов» позволяет судить о характере распределения результатов для данной выборки тестируемых (групп тестируемых).

Чтобы посчитать процент правильного выполнения заданий, необходимо количество правильных ответов умножить на сто и разделить на общее количество вопросов.

Чтобы в Excel провести расчёт, необходимо выполнить следующие операции:

1. Щёлкнуть на ячейку, в которой хотим увидеть результат.
2. В окне функций впечатываем следующую формулу:

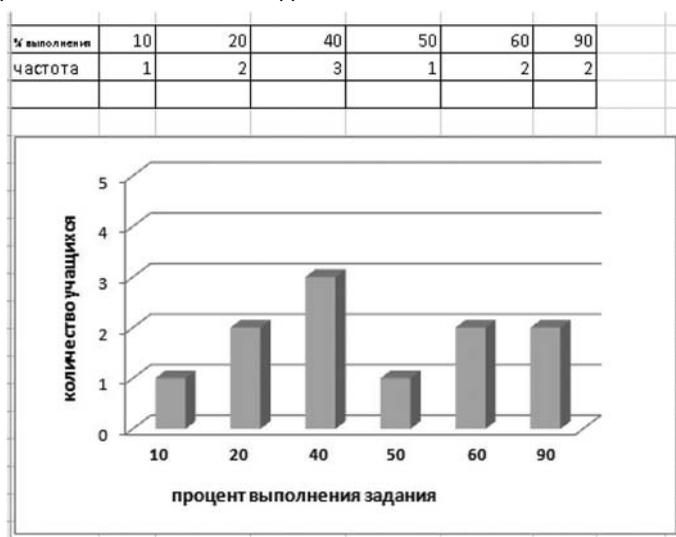
$$= \text{адрес строки} \cdot \text{на кол-во правильных ответов} / 100$$

В нашем случае формула будет выглядеть следующим образом:

Номера испытуемых, i	Индивидуальные баллы (множество X_i)	Частота	% выполнения
72	3	1	$=B72*100/10$
73	2	1	20
74	5	3	40
75	6	3	40
76	8	3	40
77	7	1	50
78	1	2	60
79	10	2	60
80	4	9	90
81	9	9	90

3. Нажимаем Enter.

Пользуясь информацией из таблицы: «% выполнения» и «частота», строим диаграмму «Плотность процентного выполнения заданий».



Каждый столбик на гистограмме показывает количество тестируемых, результаты которых лежат в процентном интервале. По гистограмме определяется характер распределения результатов для данной группы тестируемых и могут быть выделены подгруппы тестируемых с различным уровнем подготовки.

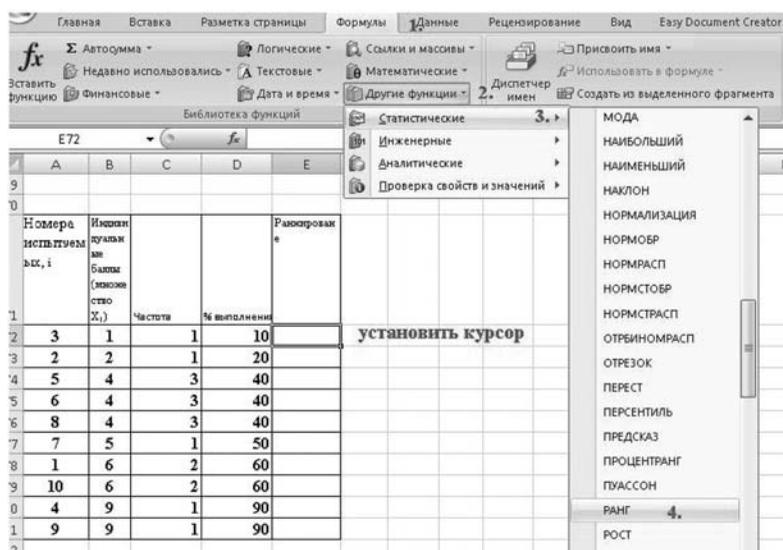
В профессионально разработанных нормативно-ориентированных тестах типичным является результат, когда приблизительно 70% учеников выполняют правильно от 30 до 70% заданий теста, а наиболее часто встречается результат в 50%.

«Ранжирование результатов тестирования»

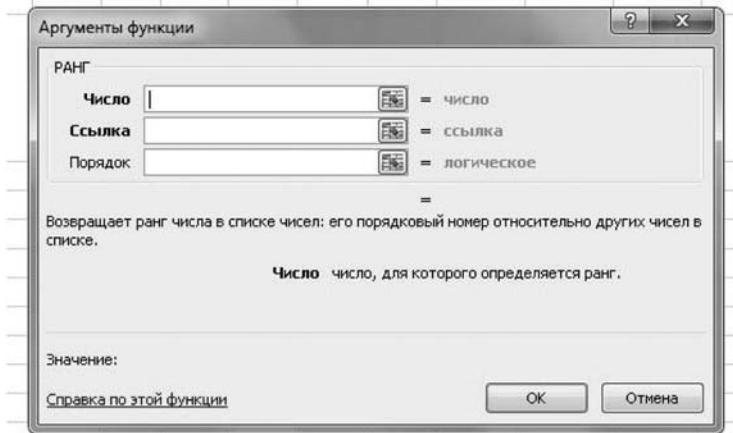
Этот метод анализа результатов тестирования может помочь при составлении рейтинга правильности выполнения тестовых заданий.

Чтобы в Excel выполнить данную операцию, необходимо:

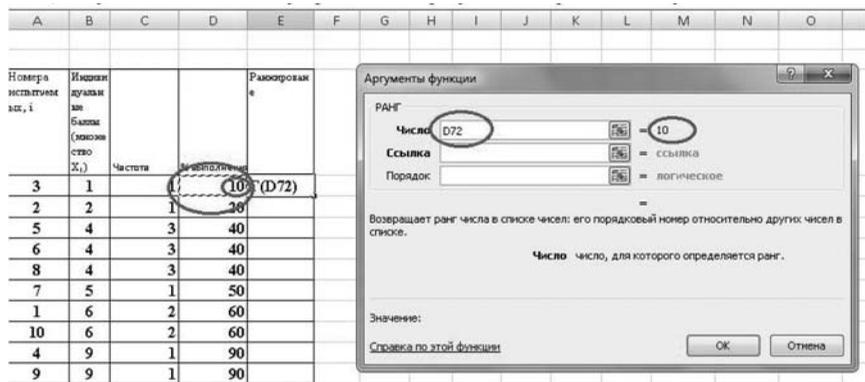
1. Установить курсор в ячейку, в которую хотите получить результат.
2. На панели инструментов выбрать вкладку «Формулы», «Другие функции», «Статистические», «Ранг».



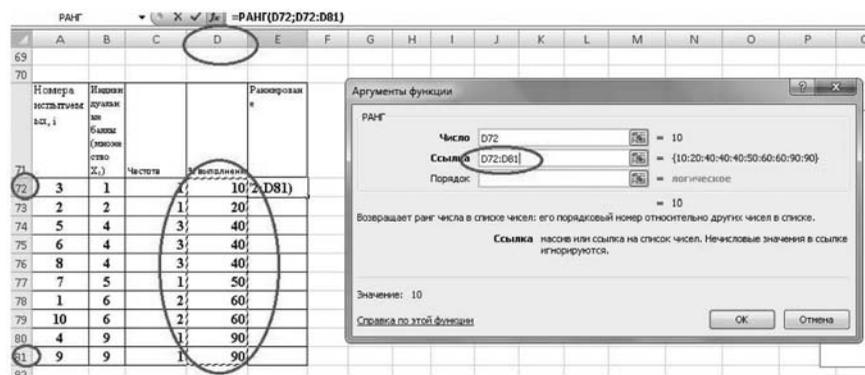
3. Откроется диалоговое окно.



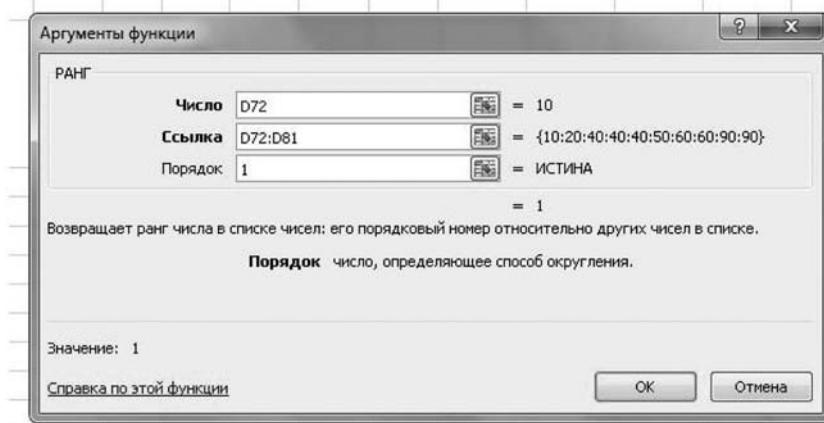
4. Щёлкнуть мышкой в ячейку процентного результата первого испытуемого.



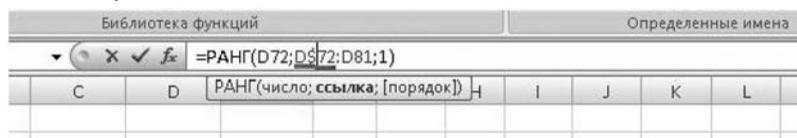
5. Перевести курсор мышки в позицию «Ссылка» в диалоговом окне функции. Курсором выделить диапазон ранжирования, т.е. все полученные результаты.



6. Перевести курсор в позицию «Порядок» в диалоговом окне функции. Впечатать цифру «1», если хотим, чтобы ранжирование было организовано в порядке возрастания, либо «0» – если в порядке убывания, нажать «OK» и «Enter».



7. Далее в окне «Функция» редактируем получившуюся функцию. Она должна иметь следующее содержание: =РАНГ(D72;D\$72:D81;1). Ко второму адресу ячейки в формуле мы добавили значок доллара \$.



8. Устанавливаем курсор в конце формулы и нажимаем «Enter».

9. Устанавливаем курсор в ячейку получившегося результата (в строке формул вы увидите формулу ранжирования) и нажимаем клавиши «Ctrl+C».

	А	В	С	Д	Е	Г	Н	И
69								
70								
	Номера испытываемых, i	Индикаторные баллы (возможство X _i)	Частота	% выполнения	Ранжирование			
71								
72	3	1	1	10	1			"Ctrl+C"
73	2	2	1	20				
74	5	4	3	40				
75	6	4	3	40				
76	8	4	3	40				
77	7	5	1	50				
78	1	6	2	60				
79	10	6	1	60				

10. Переводим курсор в следующую ячейку и нажимаем клавиши «Ctrl+C».

Номера испытуемых, i	Идентификационные баллы (множество X_i)	Частота	% выполнения	Ранжирование
3	1	1	10	1
2	2	1	20	
5	4	3	40	

11. Эту операцию проделываем до конца списка.

Номера испытуемых, i	Идентификационные баллы (множество X_i)	Частота	% выполнения	Ранжирование
3	1	1	10	1
2	2	1	20	2
5	4	3	40	3
6	4	3	40	3
8	4	3	40	3
7	5	1	50	7
1	6	2	60	9
10	6	2	60	9
4	9	1	90	11
9	9	1	90	11

По списку ранжирования можно выделить различные группы учащихся по их степени усвоение материала.

«**Коэффициент решаемости**» призван оценить уровень подготовки тестируемых по всем вопросам в выбранном тесте.

1. Для этого возвращаемся к первоначальной таблице.

Тема 4.1

Номера испытуемых, i	Номера заданий j										Индивидуальные баллы (множество X_i)
	1	2	3	4	5	6	7	8	9	10	
1	1	1	1	1	1	1	0	0	0	0	6
2	1	1	0	0	0	0	0	0	0	0	2
3	0	0	0	0	0	0	0	1	0	0	1
4	1	1	0	1	1	1	1	1	1	1	9
5	1	0	1	0	1	1	0	0	0	0	4
6	1	1	1	0	0	0	0	1	0	0	4
7	1	1	1	1	0	1	0	0	0	0	5
8	1	1	1	1	0	0	0	0	0	0	4
9	1	1	1	1	1	1	1	1	1	0	9
10	1	1	1	1	1	0	1	0	0	0	6
Число правильных ответов (множество Y_j)	9	8	7	6	5	5	3	4	2	1	50

2. Устанавливаем курсор в ячейку, в которой хотим получить процент учащихся, правильно выполнивших задание.

9	1	1	1	1	1	1	1	1	1	0	9
10	1	1	1	1	1	0	1	0	0	0	6
11	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	10
Число правильных ответов (количество Y.)	10	9	8	7	6	6	4	5	3	2	60
% выполнения заданий											

3. В строку формы вносим формулу определения процента правильно выполнивших задание.

Для этого ставим знак «=», потом адрес ячейки, содержащей количество учащихся, правильно выполнивших задание, умножаем (ставим знак «*») на количество всех учащихся, выполнявших задание и делим (ставим знак «/») на 100.

РАНГ X ✓ f_x =A\15*100/10

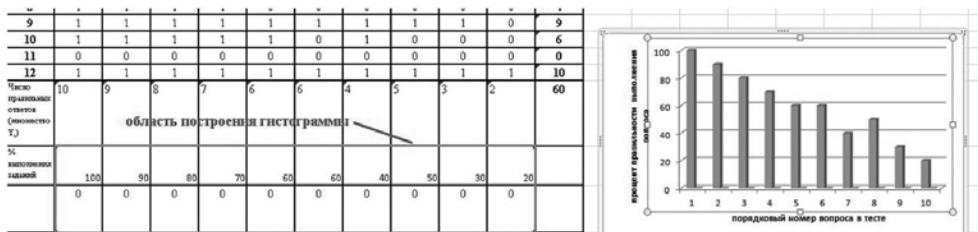
	AT	AU	AV	AW	AX	AY	AZ	BA		
4		2	1	1	0	0	0	0		
5		3	0	0	0	0	0	0		
6		4	1	1	0	1	1	1		
7		5	1	0	1	0	1	1		
8		6	1	1	1	0	0	0		
9		7	1	1	1	1	0	1		
10		8	1	1	1	1	0	0		
11		9	1	1	1	1	1	1		
12		10	1	1	1	1	1	0		
13		11	0	0	0	0	0	0		
14		12	1	1	1	1	1	1		
15		Число правильных ответов (количество Y.)	10	9	8	7	6	6	4	
16		% выполнения заданий	*100/10							

адрес ячейки

4. Пользуясь клавишами «Ctrl+C» и «Ctrl+V», получаем результаты по всем вопросам.

9	1	1	1	1	1	1	1	1	1	0	9
10	1	1	1	1	1	0	1	0	0	0	6
11	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	10
Число правильных ответов (количество Y.)	10	9	8	7	6	6	4	5	3	2	60
% выполнения заданий	100	90	80	70	60	60	40	50	30	20	

5. Пользуясь мастером создания диаграмм, создаём гистограмму решаемости вопросов, использовавшихся в процессе тестирования.



Анализируя получившуюся гистограмму, можно оценить степень освоения дидактических единиц и выявить разделы, освоенные на недостаточном уровне, а также правильность составления теста по принципу от простого к сложному.

Значения коэффициентов решаемости заданий рассчитываются как отношение числа испытуемых, решивших задание, к общему

числу прошедших тестирование. При анализе результатов педагогических измерений по карте коэффициентов решаемости можно придерживаться следующей классификации уровней трудности заданий: лёгкие задания – коэффициент решаемости от 0,7 до 1,0, задания средней трудности – коэффициент решаемости от 0,4 до 0,7 и задания повышенной трудности – коэффициент решаемости менее 0,4.



Тестовые технологии — эффективный инструмент для диагностики и проведения мониторинговых исследований. Но сами по себе тесты не войдут в образовательную

практику учебных заведений, они могут быть востребованы, только тогда, когда учитель сам станет разработчиком новых программно-аналитических материалов. □