

УДК 004.9

АНАЛИЗ НАДЕЖНОСТИ ТЕСТА ДЛЯ ОЦЕНКИ УЧЕБНЫХ КОМПЕТЕНЦИЙ СТУДЕНТОВ

И.П. Овсянникова, О.Б. Ганак

Могилевский государственный университет продовольствия, г. Могилев, Республика Беларусь

Одним из направлений совершенствования процесса обучения в вузах является разработка оперативной системы контроля учебных компетенций студентов, позволяющей объективно оценивать знания студентов, выявляя имеющиеся пробелы и определяя способы их ликвидации. Наиболее популярной формой промежуточного и итогового контроля является компьютерное тестирование, позволяющее дать объективную оценку знаний студентов по изучаемой дисциплине. Компьютерное тестирование не заменяет традиционных методов контроля знаний студентов, а дополняет их, позволяет существенно сократить затраты времени при обработке результатов, формирует базы данных об успеваемости, позволяет получить результаты сразу после завершения тестирования

Инструментом для измерения по шкале достижений студента является правильно сконструированный качественный тест, который соответствует предмету обучения, его задачам и служит развитию системного подхода к изучению учебной дисциплины. Одной из важнейших качественных характеристик теста является его надежность. Надежность – это показатель точности педагогического измерения и устойчивости результатов тестирования к воздействию посторонних или случайных факторов. Тест считается надежным, если он дает одни и те же (или очень близкие) показатели для каждого испытуемого при повторном тестировании. При этом необходимо, чтобы сами испытуемые не изменили свой уровень подготовки перед вторым тестированием, а их мотивация к получению наилучших результатов осталась прежней. Оценка надежности тестов проводится различными методами, которые по способу осуществления можно условно разделить на две группы [1]. Первая группа методов базируется на двукратном тестировании, проводимом с помощью одного и того же теста либо с помощью двух параллельных форм теста. Вторая группа предполагает однократное тестирование при оценке надежности теста.

Для оценки надежности тестов авторы использовали метод расщепления. Данный метод применялся для оценки качества измерительных материалов при проведении итогового контроля знаний студентов 1 курса по дисциплине «Информатика». Оценка надежности строится на подсчете корреляции между двумя наборами результатов выполнения одного и того же теста или двух его параллельных форм. Чем выше корреляция, тем надежнее тест. Хорошим коэффициентом надежности теста считается тот, когда показатель колеблется в пределах $0,8 < r < 1$.

Метод расщепления теста удобен в практическом применении, так как ограничивается однократным тестированием. Он основан на допущении параллельности двух половин теста и предполагает деление результатов тестирования на две части: данные по нечетным заданиям теста (X) и по четным (Y). Коэффициент надежности вычисляется как коэффициент корреляции Пирсона по формуле:

$$r = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}}{\sqrt{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}} \sqrt{\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}}} \quad (1)$$

где x_i – индивидуальный балл i -го студента при ответе на тест по нечетным заданиям;
 y_i – индивидуальный балл i -го студента при ответе на тест по четным заданиям;
 n – число студентов.

Поскольку подсчет надежности ведется по расщепленному тесту, который в 2 раза короче, то оценка надежности корректируется по формуле Спирмена – Брауна:

$$r = \frac{2r_{pac}}{1 + r_{pac}} \quad (2)$$

В случае, когда выполнение задания оценивается дихотомически (1 балл – правильно; 0 баллов – неправильно), можно использовать способ оценки надежности с применением формулы Кьюдера – Ричардсона Коэффициент надежности находится по следующей формуле:

$$r = \frac{m}{m-1} \left(1 - \frac{\sum_{j=1}^m p_j q_j}{S_x^2} \right) \quad (3)$$

где m – число заданий в тесте;

p_j – доля правильных ответов на j -е задание теста;

q_j – доля неверных ответов на j -е задание, ($q_j=1-p_j$);

S_x^2 – дисперсия индивидуальных баллов, рассчитываемая из соотношения

$$S_x^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1} \quad (4)$$

где n – число студентов;

x_i – индивидуальный балл i -го студента.

Для итогового контроля знаний 18 студентов был разработан тест, содержащий 20 тестовых заданий. Математико-статистическая обработка результатов тестирования осуществлялась с помощью пакета MS EXCEL, применялась дихотомическая шкала оценивания. По результатам тестирования была сформирована матрица размерностью 18×20 , состоящая из 0 (неверный ответ) и 1 (верный ответ). Далее для расчета надежности теста потребовались следующие действия:

- 1) подсчитать индивидуальные баллы испытуемых (сумма баллов по строкам) X_i ;
- 2) найти количество правильных ответов на каждое задание (сумма баллов по столбцам) R_j ;
- 3) подсчитать количество неправильных ответов на каждое задание W_j ;
- 4) долю правильных ответов p_j и долю неправильных ответов на каждое задание q_j ;
- 5) рассчитать коэффициент надежности по формуле 3. Фрагмент матрицы представлен в таблице 1.

Таблица 1 – Результаты тестирования

Порядковый номер студента	Ответы на вопросы								X _i
	1	2	3	4	5	6	7	8	
1	1	1	1	0	1	1	1	1	7
2	0	1	1	1	1	1	1	1	7
3	1	1	0	1	1	0	0	1	5
4	0	1	1	1	1	1	0	0	5
5	1	0	1	1	0	0	0	1	4
6	1	1	0	0	0	0	0	0	2
R _j	4	5	4	4	4	3	2	4	
W _j	2	1	2	2	2	3	4	2	
p _j	0.500	0.625	0.500	0.500	0.500	0.375	0.250	0.500	
q _j	0.250	0.125	0.250	0.250	0.250	0.375	0.500	0.250	

Далее надежность теста рассчитывалась по формулам Спирмена – Брауна. Бинарная матрица результатов тестирования была разбита на две части: первую половину матрицы составили ответы по нечетным заданиям, вторую – по четным. Были подсчитаны индивидуальные баллы студентов по нечетным заданиям x_i и индивидуальные баллы по четным заданиям y_i в соответствующих строках. Подсчитан коэффициент корреляции Пирсона для двух половин теста по формуле 1, и полученное значение коэффициента надежности теста скорректировано по формуле Спирмена – Брауна.

Результаты расчета коэффициента надежности теста с помощью разных методов представлены в таблице 2.

Таблица 2 – Коэффициенты надежности теста

Используемые формулы	Значение
По формуле Кьюдера-Ричардсона	0,8353
По формуле Спирмена-Брауна	0,8148

Нижним пределом значения коэффициента надежности принято считать показатель 0,7. Если значение коэффициента надежности ниже, то надежность теста считается неудовлетворительной, так как возникает большая погрешность измерений.

Анализ результатов свидетельствует о том, что надежность разработанного теста является допустимой. Для увеличения коэффициента надежности некоторые тестовые задания необходимо заменить.

Качество диагностических материалов является одним из ключевых моментов контроля знаний. Во избежание неадекватной оценки знаний необходимо использовать только статистически обоснованные тестовые материалы, обладающие достаточным уровнем надежности.

Список литературы

1. Чельшкова М.Б. Теория и практика конструирования педагогических тестов: учебное пособие. М.: Логос, 2002- 432с.