

УДК 371.69:004.3

## ОЦЕНКА ЗНАЧИМОСТИ РАЗЛИЧНЫХ ВИДОВ ТЕСТОВЫХ ЗАДАНИЙ МЕТОДОМ ЭКСПЕРТНОГО РАНЖИРОВАНИЯ

Вардомацкая Е.Ю.

*Витебский государственный технологический университет, г. Витебск, Беларусь, elena101.vard@gmail.com*

**Аннотация.** Рассмотрены подходы к формированию заданий для текущего тестового контроля знаний. Выявлены типы вопросов, наиболее значимые с точки зрения студентов и с точки зрения преподавателей при компьютерном тестировании.

**Ключевые слова.** Компьютерное тестирование, экспертное ранжирование, коэффициент конкордации Кендалла, коэффициент вариации, коэффициенты значимости.

Учебными программами большинства дисциплин информационного направления для студентов факультета экономики и бизнес-управления Витебского государственного технологического университета в качестве одной из форм текущего контроля знаний предусмотрено независимое тестирование. Для получения достоверных результатов оценки знаний при использовании такой формы контроля первостепенное значение имеет качество предлагаемых студентам тестовых заданий: соответствие учебной программе, корректность, точность и компактность формулировок, логичность изложения, однозначность понимания вопроса, уровень сложности, значимость задания, предполагаемое время на подготовку ответа, оптимальное количество вариантов ответа. Процесс подготовки качественных тестовых заданий достаточно трудоемкий и ответственный, поэтому, как правило, выполняется группой специалистов – профессионалов в своей области, с обязательной оценкой интеллектуального индекса трудности заданий и учетом особенностей восприятия задания студентами. При анализе результатов интеллектуальный индекс сложности теста позволяет преподавателю отследить процесс мыслительной деятельности студента, сделать соответствующие выводы и при необходимости внести коррективы в состав тестовых заданий.

Использование компьютерных технологий тестирования автоматизирует процесс тестового контроля и расширяет возможности его совершенствования и адаптации для разных групп тестируемых.

В Витебском государственном технологическом университете в качестве программной среды для контроля знаний студентов всех форм обучения используется система дистанционного обучения (СДО) Moodle. Эта среда позволяет пользователю создавать, хранить и редактировать практически неограниченные базы вопросов разных категорий а, значит, компоновать тесты разного уровня сложности, корректируя их по мере необходимости. Так, например, студентам первого курса дневной формы обучения факультета экономики и бизнес-управления при проведении текущего контроля знаний по дисциплине «Информационные технологии» предлагаются тесты, интеллектуальный уровень сложности которых представлен шкалой вида [1]:

25 % заданий решаются посредством сравнения и выявления соответствия объектов или операций, что соответствует низкому интеллектуальному индексу;

25 % заданий решаются посредством мыслительной операции обобщения (средний интеллектуальный индекс);

25 % заданий на систематизацию и классификацию (высокий интеллектуальный индекс);

25 % заданий творческого характера (высший интеллектуальный индекс).

При этом за правильный ответ на вопрос из соответствующей группы выставляется разный балл.

Такое распределение вопросов по группам сложности обусловлено тем, что, содержание учебной дисциплины «Информационные технологии» во многом опирается на знания школьной программы и цель ее изучения - систематизировать, расширить и углубить уже имеющиеся знания по цифровой грамотности и информационным технологиям.

По мнению преподавателей-практиков и ученых-специалистов в области педагогики и психологии [1], для комфортного восприятия тестового задания большое значение имеет не только содержание вопроса, но и предлагаемый тип ответа на него. Используя разные типы ответов, простое задание можно «сделать» более сложным для ответа, и наоборот, сложное задание – упростить. Правильно подобранный вариант ответа на вопрос или задание теста позволяет более точно оценить глубину знаний студентов, минимизировать случаи угадывания ответа или автоматического заучивания, что способствует формированию прочных фундаментальных и остаточных знаний.

Цель нашего исследования – проанализировать и сравнить значимость возможных типов вопросов теста с точки зрения преподавателей и с точки зрения студентов.

Метод исследования – экспертное ранжирование [2].

Инструментарий исследования – интегрированная система Statistica и табличный процессор MS Excel.

Задача исследования – определить весомость вопроса каждого вида в тесте для текущего контроля знаний в общей совокупности вопросов теста, что и позволит провести их ранжирование. Предпочтительными, значимыми для экспертов, следует признать те вопросы, коэффициент значимости которых выше.

СДО Moodle предоставляет возможность формировать тесты разного уровня интеллектуальной сложности, используя и комбинируя тестовые вопросы и задания различных типов, предполагающие различ-



ные варианты ответов. Причем, тип ответа не всегда характеризует уровень сложности вопроса. Задача преподавателя составить вопрос таким образом, чтобы максимально уравновесить уровень сложности задания и комфортность восприятия студентами формата ответа.

Для проведения исследования были сформированы две группы экспертов.

Первая группа экспертов в составе восьми человек – преподаватели кафедры математики и информационных технологий, имеющие опыт использования различных форм тестового контроля знаний студентов по дисциплинам естественно-математического профиля.

Вторая группа экспертов в составе тридцати человек – студенты первого курса дневной формы обучения специальности «Маркетинг», текущий контроль знаний которых выполняется в форме компьютерного тестирования.

Для формирования информационного обеспечения данного исследования проводился анонимный опрос экспертов, что обеспечивает получение наиболее полных и достоверных сведений.

Результаты ранжирования показателей (типов вопросов) оформлены в виде таблиц Excel, где в столбцах обозначены факторы (виды вопросов), в строках – эксперты. При этом оценка качества экспертов, выполненная по отклонению от средней оценки экспертной группы для каждого эксперта составила величину  $> 8.5$ , а это значит, что мнению каждого из них можно доверять.

Показателем согласованности мнений экспертов является коэффициент конкордации Кендалла. Значения этих коэффициентов для каждой группы экспертов и значимость каждого из них, проверенная по критерию согласия Пирсона  $\chi^2$ , приведены в таблице 1.

Таблица 1 – Коэффициенты конкордации

Эксперты	Коэффициент конкордации	$\chi^2_{\text{расч}}$	$\chi^2_{\text{табл}}$
Преподаватели	0,81	33,34	14,07
Студенты	0,87	158,11	43,77

В обоих случаях значения коэффициентов конкордации выше 0,8, табличное значение критерия  $\chi^2$  меньше расчетного. Это свидетельствует о высокой степени согласованности мнений экспертов относительно значимости видов вопросов тестовых заданий [2].

Степень согласованности мнений экспертов по значимости каждого типа вопроса определяется коэффициентом вариации  $S$ , а значимость каждого во-

проса в тесте – значением коэффициента весомости вопроса  $k$ .

Результаты расчетов этих показателей для каждой группы экспертов приведены в таблице 2, где цифрами 1...7 обозначены типы вопросов.

Таблица 2 – Критерии значимости вопросов.

Тип вопроса	1	2	3	4	5	6	7
<b>Студенты</b>							
Коэфф. вариации $S$	22,7	26,1	9,8	25,1	15,4	11,9	6,3
Коэфф. весомости $k$	<b>0,30</b>	<b>0,17</b>	0,10	0,15	0,09	0,12	0,08
<b>Преподаватели</b>							
Коэфф. вариации $S$	27,3	12,0	9,1	11,8	31,5	37,4	37,4
Коэфф. весомости $k$	<b>0,15</b>	0,05	0,13	0,07	<b>0,14</b>	<b>0,23</b>	<b>0,23</b>

Коэффициенты вариации, полученные по результатам анкетирования студентов, характеризуют степень согласованности мнений экспертов преимущественно выше средней ( $S \leq 25\%$ ). При этом предпочтительными вариантами вопросов являются вопросы (см. таблица 2), предполагающие ответ в форме множественный выбор (тип вопроса 1), верно/неверно (тип вопроса 2) и вставка пропущенного слова (тип вопроса 4). Для этих вопросов  $k \geq 1/n \geq 0.143$ . По мнению преподавателей наиболее значимыми для оценки знаний студентов являются вопросы, подразумевающие числовой (тип 6), вычисляемый (тип 7) или короткий (тип 5) ответ, а также ответ в форме множественного выбора (тип 1).

Полученные данные являются показательными и позволяют корректировать стратегию и тактику проверки степени усвоения материала. При составлении заданий для текущего компьютерного контроля знаний необходимо учитывать как требования учебных программ по формированию необходимого уровня теоретических и практических знаний будущих специалистов, так и особенности восприятия студентами различных типов заданий, связанные со спецификой изучения учебных дисциплин.

### Литература

1. Красильникова, В.А. Подготовка заданий для компьютерного тестирования: Методические рекомендации / В.А. Красильникова. – Оренбург, ИПК ГОУ ОГУ, 2004. 31 с.
2. Похабов, В.И. Экономико-математическое методы и модели: практикум / В.И. Похабов Минск.: БНТУ, 2003, –130с.

## ASSESSMENT OF THE SIGNIFICANCE OF DIFFERENT TYPES OF TEST TASKS USING THE EXPERT RANKING METHOD

E.U. Vardomatskaya

Vitebsk State Technological University, Vitebsk, Belarus, elena101.vard@gmail.com

**Abstract.** Approaches to the formation of tasks for current test control of knowledge are considered. The types of questions that are most significant from the point of view of students and teachers during computer testing have been identified.

**Keywords.** Computer testing, expert ranking, Kendall concordance coefficient, coefficient of variation, significance coefficients.