

Бородина А. И., канд. экон. наук, доцент, Корытко О. В., ассистент, БГЭУ, г. Минск

СИСТЕМА ТЕСТОВ КАК ИЗМЕРИТЕЛЬНЫЙ ИНСТРУМЕНТ И УСЛОВИЕ ПОВЫШЕНИЯ УРОВНЯ ЗНАНИЙ СТУДЕНТОВ

1. Роль системы тестов как измерительного инструмента в области образования на современном этапе

Тесты как измерительный инструмент используются в большинстве стран мира. Их разработка и использование основано на мощной теории и подтверждено многочисленными эмпирическими исследованиями. Тестология как теория и практика тестирования существует более 120 лет, и за это время накоплен громадный опыт использования тестов в различных сферах человеческой деятельности, включая образование.

Увеличению масштабов тестирования в вузах и школах способствует повышение информатизации всех уровней образования, разработка государственных образовательных стандартов, переход к массовому высшему образованию. Компьютерное тестирование реализуется при проведении текущих контрольных срезов знаний учащихся и студентов, итогового контроля и диагностики знаний, содержит в себе значительный обучающий потенциал. На Западе сложилась практика: чем выше уровень развития тестового контроля, тем выше рейтинг вуза.

Преимущество тестов как измерительного инструмента знаний в том, что они оказываются значительно более качественным и объективным способом оценивания. Объективность тестирования достигается путем стандартизации процедуры его проведения (на всех этапах проведения тестирования невозможно внести субъективную составляющую в оценку) и путем стандартизации и проверки качества заданий и тестов целиком. Тесты являются более емким инструментом оценки знаний, так как ориентированы на определение уровня усвоения ключевых понятий, навыков, знаний, а не на констатацию наличия у учащихся определенной совокупности усвоенных знаний. Процедура тестирования обеспечивает равные условия для всех ис-

пытуемых, имеет единые критерии оценки и эффективна с экономической точки зрения.

Научной теории создания тестов и проблемам практического использования в образовательной среде посвящено немало работ отечественных и зарубежных исследователей – В.С. Аванесова, Т.Д. Макаровой, А.Н. Майорова, Н. Гронлунда, П. Клайна, В.К. Гайды и Захарова В.П., а также инструкции, нормативные документы, стандарты CITO, NEAB.

2. Теоретические основы создания педагогических тестов

Согласно определению В.С. Аванесова, педагогический тест – это система заданий определенного содержания, возрастающей трудности, специфической формы, позволяющая качественно и эффективно измерить уровень и оценить структуру подготовленности учащихся.

Тест, как система, обладает составом, целостностью и структурой. Тест состоит из заданий, правил их применения, оценок за выполнение каждого задания, рекомендаций по интерпретации тестовых результатов. Целостность означает взаимосвязь заданий, их принадлежность общему измеряемому фактору. Каждое задание теста выполняет отведенную ему роль и потому ни одно из них не может быть изъято из теста без потери качества измерения. Структура теста образует способ связи заданий между собой, когда каждое задание связано с другим через общее содержание и общую часть вариации тестовых результатов. Проявлению системного качества тестов способствует и единая дисциплинарная общность заданий, реализующая идею подготовленности испытуемых по определенной учебной дисциплине (гомогенный тест).

В качестве другого системообразующего фактора выступает время. Каждый тест имеет оптимальное время тестирования, уменьшение или превышение которого снижает качественные показатели теста.

Специфическая форма тестовых заданий означает, что задания теста представляют собой не вопросы и не задачи, а задания, сформулированные в форме высказываний, истинных или ложных. Определенное содержание тестовых заданий означает использование только такого контрольного материала, который соответствует содержанию учебного курса. Содержание тестовых заданий существует в одной из четырех основных форм: 1) задания с выбором одного

или нескольких правильных ответов из числа предложенных; 2) задания открытой формы; 3) задания на установление соответствия; 4) задания на установление правильной последовательности.

Важнейшим тестообразующим фактором является трудность заданий. В тесте задания должны различаться по уровню трудности. Трудность заданий может определяться двояко: 1) умозрительно, на основе предполагаемого числа и характера умственных операций, необходимых для успешного выполнения заданий; 2) после эмпирического апробирования заданий, с подсчетом доли неправильных ответов.

В отличие от других педагогических контрольных материалов тестовые задания проходят процесс специальной апробации. Ключевым здесь является момент эмпирической проверки тестовых свойств заданий на основе матрицы тестовых результатов. Матрица тестовых результатов представляет результаты всех испытуемых по всем проверяемым заданиям и позволяет определить общий тестовый балл каждого испытуемого, долю неправильных ответов каждого испытуемого по всем заданиям, долю неправильных ответов всеми испытуемыми по каждому конкретному заданию – трудность тестового задания, вариацию тестовых баллов испытуемых в каждом задании, вариацию тестовых баллов испытуемых по всем заданиям и дифференцирующую способность тестовых заданий.

Для заданий с дихотомической оценкой мера вариации тестовых баллов испытуемых в каждом задании определяется как произведение долей правильных и неправильных ответов испытуемых на это задание. Если на конкретное задание правильно или неправильно отвечают все испытуемые, то такое задание становится для данной группы испытуемых не тестовым, так как между тестируемыми нет вариации. Нулевая вариация означает практическую необходимость переработки или удаления задания из проектируемого теста. Мерой вариации баллов испытуемых по всем заданиям является стандартное отклонение тестовых баллов от среднего арифметического тестового балла по группе.

Задание в тестовой форме нельзя назвать тестовым, если оно не коррелирует с суммой баллов по всему тесту. Коррелируемость задания с общим тестовым баллом представляет собой более точную и технологичную меру дифференцирующей способности задания.

Дифференцирующая способность задания (дискриминативность задания) отражает меру соответствия между успешностью решения одной задачи и всего теста по выборке испытуемых и является показателем пригодности задачи для теста. Дискриминативность задания рассчитывается на основе коэффициента корреляции Пирсона и его модификации (Point-biserial коэффициента корреляции).

Анализ тестовых свойств заданий на основе показателей вариации баллов, дискриминативности заданий позволяет подготовить первый вариант теста.

Требования к тесту как к измерительному инструменту содержат требования к расчету показателей качества теста. Определение качества теста сводится к определению меры надежности и вопросов валидности тестовых результатов.

Надежность характеризует точность теста как измерительного инструмента, его «помехоустойчивость». На практике используются 3 основных метода оценки надежности теста: 1) повторное тестирование (ретестирование) – определение помехоустойчивости; 2) расщепление группы – определение помехоустойчивости; 3) расщепление теста – определение внутренней согласованности. Как правило, надежность определяется как коэффициент корреляции между двумя рядами оценок тестируемых, полученных в разных условиях. Так как тестовый балл испытуемого содержит в себе 2 компоненты – истинную и ошибочную, проведя повторное тестирование и получив второй ряд оценок, можно определить насколько тесно связаны между собой оба ряда оценок. Истинные оценки двух рядов измерений остаются постоянными, а ошибочные – разными. Чем больше истинная компонента, тем выше корреляция между этими двумя рядами оценок, а чем больше ошибочная компонента, тем корреляция меньше. Для определения надежности гомогенных тестов учебных достижений коэффициент корреляции теоретически более обоснованно считать по формуле Kuder-Richardson-8.

Понятие «валидность» означает меру пригодности тестовых результатов для определенной цели. Валидность особенно важна для тестов, ориентированных на критерий, поскольку определение надежности для этих тестов затруднено. Валидность зависит от качества заданий, их числа, от степени полноты и глубины охвата содержания учебной дисциплины в заданиях теста, от баланса и распределе-

ния заданий по трудности, от интерпретации тестовых результатов, от подбора выборочной совокупности испытуемых и другого. Часть показателей валидности может быть определена статистическими методами, однако судить только по ним о валидности теста некорректно. Валидность получается из экспертных оценок и специфична для каждого конкретного теста.

3. Трехэтапная система тестов как условие повышения уровня и оптимизации структуры знаний студентов

Для повышения уровня и оптимизации структуры знаний студентов экономического вуза по курсу «Основы информатики и вычислительной техники» с минимальными затратами времени и учетом индивидуальных способностей студентов авторами создана и внедрена в педагогический процесс система тестов на трех уровнях сложности по разделу «Арифметические и логические основы ЭВМ».

Дифференциация тестов по степени сложности дает возможность:

- самостоятельного выбора студентом уровня контроля и оценки знаний по изучаемому разделу дисциплины;
- продвижения от низшего уровня сложности к высшему с получением соответствующего количества баллов на каждом уровне сложности;
- глубокого изучения материала, развития логики и творческого подхода к решению нетипичных задач (третий уровень сложности теста).

Технология создания и применения трехуровневой системы тестирования знаний студентов состоит из следующих этапов:

1. Создание системы тестов по трем уровням сложности в соответствии с научными основами теории тестов.

2. Рецензирование экспертами тестовых заданий и корректировка системы тестов.

3. Апробация системы тестов с помощью процедуры компьютерного тестирования на широкой выборке испытуемых с целью:

- определения оптимального времени тестирования на каждом уровне сложности;
- эмпирической проверки тестовых свойств заданий на основе показателей трудности, вариации баллов, дифференцирующей способности заданий и последующей корректировки содержания теста для получения окончательного варианта;

- определения качества теста на основании показателей надежности тестовых результатов.

4. Проведение педагогического эксперимента по выявлению эффективности применения трехэтапной системы тестов с формированием контрольной и экспериментальной группы студентов.

5. Анализ результатов эксперимента.

6. Разработка рекомендаций по проектированию и применению на практике трехэтапной системы тестов.

Создание и применение на практике трехэтапной системы тестирования знаний студентов по курсу ОИиВТ позволяет:

- осуществить текущий промежуточный контроль знаний студентов с дифференциацией по степени сложности и минимальными затратами времени;

- осуществить самоконтроль знаний студентов по каждому модулю дисциплины с дифференциацией по степени сложности;

- получить значительный обучающий эффект в процессе процедуры тестирования;

- активизировать познавательные интересы студентов и повысить мотивацию к изучению дисциплины.

Литература:

1. Аванесов В.С. Композиция тестовых заданий. М.: Адепт, 1998.

2. Аванесов В.С. Научные проблемы тестового контроля знаний. М.: 1994.

3. Гронлунд Н. Тесты достижений в конструировании. Лондон: Прентис-Холл, 1982.

4. Клайн П. Введение в психометрическое программирование: Справочное руководство по конструированию тестов. Киев, 1994.

5. Майоров А.Н. Теория и практика создания тестов для системы образования. М.: 2001.

6. Макарова Т.Д. Итоговое тестирование. Дидактика-2000. Педагогическое общество России. 1999.