

## МЕСТО ТЕСТИРОВАНИЯ В СИСТЕМЕ МЕТОДОВ КОНТРОЛЯ И ОЦЕНКИ ЗНАНИЙ

**Светлана Сиренко**

Белорусский государственный университет, г. Минск  
ssn27@mail.ru

### Введение

Как известно, обучение — процесс двухсторонний. Посредством него изменяется, а точнее развивается, не только учащийся, но и учитель. И одними из самых мощных стимулов к пересмотру подходов и собственных позиций в образовательном процессе являются контроль и оценка результатов обучения.

Для преподавателя результаты проверки дают возможность своевременно выявить и устранить трудности учащихся и, что немаловажно, недочёты в собственной педагогической деятельности. Не секрет, что для многих преподавателей именно оценка знаний и умений обучающегося является наиболее трудной задачей, особенно, если приходится ставить низкую отметку. И конечно, не всегда низкая оценка мотивирует учащегося к работе над собой. В случае отсутствия заранее оговоренных и принятых обеими сторонами критериев, низкая отметка вызывает негодование обучающихся, порождает упреки педагогу в его необъективности. Тем не менее, все участники процесса обучения реагируют на результаты оценивания более гибко и адекватно, чем на призывы или указания работать лучше.

Можно ли снизить стресс обучающихся при проведении контроля их знаний, повысить точность и объективность оценки? Как педагоги относятся к тестированию как методу оценки знаний и умений? С какими трудностями и стереотипами в понимании роли тестирования и составлении тестовых материалов сталкивались слушатели курсов переподготовки по специальности «Педагогические измерения» при изучении дисциплины «Разработка и использование педагогических тестов» (Республиканский институт высшей школы, Минск). Какие наиболее распространённые ошибки при проектировании теста у них встречались? Ответим на эти вопросы, а также поделимся опытом проведения занятий.

## Взаимосвязь тестирования с другими методами контроля и оценки знаний

По мнению ряда исследователей (В.В. Давыдов, В.В. Сериков, В.А. Сластенин, А.В. Хуторской и др.), процесс обучения тем эффективнее, чем более вовлечён учащийся в самостоятельную деятельность по разрешению учебных задач. При этом могут применяться развивающие технологии, основанные на активных, рефлексивно-деятельностных формах и методах обучения, проектно-исследовательские методы, проблемно-модульное обучение и т.п. Тогда при условии хорошо подобранного содержания учебного материала и технологий обучения создаются возможности для развития у учащихся и студентов более прочных знаний, умений самостоятельно принимать решения, способности к переносу знаний и умений в новые ситуации. Однако проблемы объективной диагностики этих важных качеств и способностей, а также разработки соответствующего инструментария являются актуальными и не полностью решёнными в настоящее время.

Методы контроля и оценки знаний обучающихся, как правило, соответствуют целям, содержанию и методам обучения

(действительно, нельзя проверить творческие умения через решение репродуктивных задач). В педагогической практике наиболее часто применяются так называемые традиционные методы оценки знаний и умений: устные и письменные опросы, решение задач, контрольные и самостоятельные работы, экзамены, практикумы и др.

К новым методам, наиболее соответствующим развивающим технологиям обучения, можно отнести защиту исследовательских проектов, участие в коллективном обсуждении проблем, разрешении проблемной ситуации и т.п. Их особенностью является то, что оценивается не только конечный продукт, но и деятельность учащегося по его достижению. Несмотря на большое число методов контроля, нередко встаёт вопрос об *объективности* выставяемой оценки (особенно при устном ответе), *широте* охвата проверяемых знаний и умений (например, на экзамене проверяются знания по одной-двум темам), *экономичности* метода оценки и его *эффективности*.

Тестирование существенно отличается от всех аналогичных методов оценки, используемых в педагогической практике, поскольку предъявляет чётко разработанные, единые требования к процессу, проверочным материалам и методам их обработки и интерпретации результатов.

Методология

Безусловно, тесты имеют свои достоинства, недостатки и ограничения применения, о которых уже много говорилось в ряде публикаций журнала «Педагогические измерения». Они не заменяют, а *дополняют* другие методы диагностики, а также могут им предшествовать. Так, например, эффективное применение активных форм и методов обучения возможно лишь при наличии у учащихся определённого уровня знаний. Выявлять этот необходимый уровень первоначальной подготовки обучающихся удобно именно с помощью тестов.

Также следует заметить, что тесты являются неотъемлемой частью ряда педагогических технологий, например, технологии модульного обучения или технологии полного усвоения знаний. Именно тест с заданиями возрастающей трудности позволяет выявить не только уровень подготовки, но и структуру знаний учащихся, а точнее её отклонение от «идеальной» структуры (т.е. указать на наличие «дыр» в знаниях).

Если рассмотреть уровни освоения обучающимися деятельности, воспользовавшись одной из наиболее популярных классификаций<sup>1</sup>, то эффективность использования тестовых заданий для их диагностики снижается по мере приближения к эвристическому (применение в незнакомой ситуации)

и творческому (создание объективно новой информации) уровню. Для выявления эвристических и творческих умений обучающихся более эффективно использовать проблемные задания и нетиповые (возможно, учебные или научно-исследовательские) задачи, задания или проекты. Однако на уровне воспроизведения деятельности или знаний по памяти и на уровне понимания (т.е. применения знаний в знакомой ситуации) тесты вполне уместны и эффективны. Более наглядно это положение представлено на рисунке, на котором соотношены уровни освоения деятельности и возможности использования тестов и других видов контрольных заданий. Выбор формы тестового задания определяется многими факторами, о которых подробно указано, например, в работах В.С. Аванесова<sup>2</sup>. Не претендуя на исчерпывающий ответ, на рисунке нами выявлены формы тестовых заданий, которые соответствуют проверяемым знаниям и умениям.

Следует отметить, активизация обучающей функции тестирования помогает снизить стресс перед другими, нетестовыми, формами контроля. Действительно, обучающийся своевременно может уточнить и повторить наиболее важные аспекты учебного материала, устранить пробелы в знаниях, по-

## 1

Классификация по  
В.П. Беспалько.

*Беспалько В.П.*

Педагогика и прогрессивные технологии обучения: Учеб. для студентов пед. вузов / В.П. Беспалько. М., 1995. 336 с.

## 2

*Аванесов В.С.*

Композиция тестовых заданий. М.:

Центр тестирования, 2005.



лучить объективную оценку, проанализировать собственный прирост знаний и, при желании, сравнить их со знаниями других обучающихся. Требования теста, условия его проведения и интерпретация результатов одинаковы для всех испытуемых, субъективная составляющая оценивания сведена к минимуму. Именно это позволяет студенту или ученику соотносить свой успех с приложенными усилиями, а не с субъективным мнением преподавателя. Таким образом, обоснованное использование тестов поможет учащемуся более чётко ориентироваться в учебном материале, иметь объективное представление о собственных успехах, своевременно ликвидировать недочёты в подготовке, а также снизит стресс получить несправедливую оценку.

### Актуальные проблемы использования и проектирования тестов

В 2007/2008 уч. г. на базе РИВШ была организована переподготовка специалистов по специальности «Педагогические измерения». Одной из важных составляющих подготовки являлась дисциплина «Разработка и использование педагогических тестов», рассчитанная на 58 аудиторных часов (авторы соответствующей учебной программы доценты С.Н. Сиренко и В.Д. Скаковский). Опыт работы со слушателями этой категории позволил проанализировать отношение учителей и преподавателей вузов и вузов Беларуси к тестам, трудности, возникающие при использовании и проектировании ими тестов.

Следует отметить, что отношение к тестам и тестированию среди слушателей курсов неоднозначное. Негативные реакции на применение тестов в педагогической практике связано, прежде всего, с неполным пониманием педагогами целей тестирования и принципов интерпретации тестовых баллов. Так, например, тяжело воспринимаются учителями и учащимися низкие тестовые баллы, полученные на централизованном тестировании, когда итоговая оценка по предмету была высока. В описанном случае учитель слышит в свой адрес нарекания со стороны руководства и родителей, хотя при выставлении оценки пользовался необходимыми критериями.

В данной ситуации педагогу рекомендуется помнить о том, что итоговая оценка на экзамене выставляется на основе соответствия знаний учащегося определённым содержательным критериям, а централизованное тестирование направлено на построение рейтинга поступающих, т.е. выбора из числа абитуриентов, чей уровень знаний удовлетворяет критериям лучших. Именно из-за различия целей этих оценочных мероприятий отличается и подходы к выбору содержания проверочных материалов, их сложности, продолжительности проведения, а также к интерпретации баллов.

При составлении тестовых заданий часто встречались задания, при составлении которых слушатели курсов преследовали несколько целей. Например, от ученика требовалось правильно узнать страну по её описанию, представленному на иностранном языке, а также без ошибок записать её название. Здесь делается попытка одним заданием проверить и то как учащийся умеет понимать иностранную речь, как он усвоил отличительные признаки той или иной страны, а также умение правильно писать. При правильном ответе на такое тестовое задание учитель, безусловно, получит достаточно полную информацию о знаниях ученика. А как оценивать и интерпретировать неправильный или частично правильный ответ? Такое задание следует разбить на несколько, выполнение которых укажет, что конкретно знает или не знает ученик.

Немало споров вызвал вопрос о том, сколькими баллами следует оценивать частично правильный ответ или правильный ответ на трудное задание, с которым мало кто из учащихся справляется. Исходя из своего предыдущего педагогического опыта, слушатели стремились оценить частично правильный ответ, используя доли баллов. Так, если заданием предусмотрено 4 правильных варианта для выбора, а респондент ука-

зывает только 2, то предлагалось, что за такой ответ учащийся получит 0,5 балла.

Использование дробных чисел для оценки частично правильного ответа на задание теста ведёт к тому, что учащийся может иметь отрывочные знания по всем вопросам теста и получить достаточно высокий тестовый балл, который может соответствовать удовлетворительной отметке в журнале. Можно с уверенностью сказать, что это именно незнание материала и угадывание ответов. Описанный выше подход к обработке результатов тестирования ведёт к получению недостоверной информации о знаниях обучающихся, снижению объективности оценки, создаёт прецеденты для опротестования результатов. Введение дробей затрудняет подсчёт тестовых баллов и не даёт существенного прироста качества при интерпретации результатов.

Оценка более трудного задания большим количеством баллов (например, двумя, тремя или более) также таит в себе опасность искажения реальной картины. На результаты тестирования сильного ученика, ответившего на задание, указанный весовой коэффициент не повлияет, однако сильно исказит результаты слабого ученика. Действительно, угадывание слабым учащимся ответа нивелирует его различия с более

подготовленными респондентами и невыполнение нескольких заданий, которые оцениваются меньшим числом баллов.

В данной ситуации, нужно помнить, что решение о присвоении весовых коэффициентов тестовому заданию должно идти не «на глазок», а на основе математических методов исследования качества тестов и обработки его результатов, при наличии репрезентативной выборки респондентов. Дихотомическая (0 баллов — неправильный ответ, 1балл — правильный) шкала оценки тестового задания является самой простой, но действенной. Не следует стремиться увеличить разброс тестовых баллов за счёт таких искусственных и неоправданных мер, лучше изменить некоторые параметры теста: сложность заданий, длину теста или время его выполнения.

Интересен оказался вопрос о возможности применения тестов и особенностях конструирования тестовых заданий для респондентов разного возраста. Например, обсуждались вопросы о возможностях тестирования учащихся, которые ещё не умеют читать (это могут быть дети начальных классов или учащиеся, недавно приступившие к изучению иностранного языка), организации процесса тестирования с такими детьми. Актуальными оказались вопросы невербальной поддержки

тестовых заданий (использование рисунков, схем, моделей), возможности использования интегративных тестов для различных возрастных категорий. Ответы на эти вопросы обсуждались на основе анализа специфики ведущего вида деятельности и психологических особенностей личности в данном возрастном периоде.

Самым трудоёмким оказался для слушателей курсов переподготовки раздел, связанный с проверкой качества тестовых заданий и качества теста в целом. Связано это с необходимостью наличия у них определённых математических знаний и умений из области математической статистики, владения компьютерными средствами. Большинство курсантов этими умениями обладают в недостаточном объёме.

В этой связи был сделан акцент не столько на нахождение сложных для подсчёта коэффициентов, а на более полное понимание и умение правильно интерпретировать наиболее простые, информативные из них, а также те, которыми обычный педагог сможет воспользоваться в своей педагогической практике. Так, например, не всегда следует удалять из теста задание, на которое ответили все респонденты. Здесь следует ещё раз задуматься о цели тестирования. В том случае, если по результатам тестирования необ-

ходимо было определить, что именно из содержания учебной дисциплины испытуемый знает, то такие задания являются информативными и уместными. Их выполнение всеми респондентами говорит о том, что этот фрагмент содержания освоен всеми и в полном объёме.

В качестве заключительного контрольного мероприятия слушатели представляли разработанный по нескольким темам тест с пакетом документации к нему. В пакет входили: спецификация теста, сам тест, инструкции для проверяющих и респондентов, результаты апробации, математические расчёты по проверке качества заданий и теста в целом. В процессе апробации созданного измерительного инструментария слушателями анализировались и корректировались неудачные задания, эмпирически устанавливалось оптимальное время выполнения теста.

Все занятия проводились с использованием электронных презентаций, что высвобождало время на обсуждение проблемных вопросов, облегчало представление многочисленных примеров и схем. Математическая обработка результатов тестирования и проверки качества теста проводилась с использованием электронных таблиц. Для диагностики освоения содержания учебного материала использовались раз-

личные методы, в том числе и тестовые задания, которые выполняли функцию не только контрольного диагностического, но и материала для работы слушателей в качестве экспертов по содержанию и форме тестовых заданий, совместно проводилась математическая обработка и интерпретация результатов. Представим некоторые из заданий.

### Задания в тестовой форме

*Инструкция:* Вашему вниманию предлагаются задания, в которых могут быть один, два, три и большее число правильных ответов. Обведите кружком номера всех правильных ответов:

#### 1. ТЕСТ КАК СИСТЕМА ВКЛЮЧАЕТ

- a) задания
- b) баллы за выполнение заданий
- c) учебные материалы
- d) список употребляемых терминов
- e) правила выполнения заданий тестируемыми
- f) измерительное оборудование
- g) рекомендации по интерпретации тестовых баллов
- h) субъектов тестирования

#### 2. ЕСЛИ ТЕСТ ПОНИМАЕТСЯ КАК СИСТЕМА, ТО

- a) количество заданий равно 30
- b) разрабатывается банк параллельных заданий теста

- c) все задания должны иметь одну форму
- d) каждое задание имеет известную эмпирическую меру трудности
- e) в него включаются задания из смежных областей знания

#### 3. ПАРАЛЛЕЛЬНЫЕ ЗАДАНИЯ

- a) проверяют знания по темам, изучаемым студентами параллельно по другим предметам
- b) охватывают одно содержание
- c) имеют примерно одинаковую меру трудности
- d) входят в один вариант теста

#### 4. ПРИ ПОМОЩИ ПЕДАГОГИЧЕСКОГО ТЕСТА МОЖНО

- a) сравнить достижения испытуемого со средними показателями выполнения теста
- b) выявить меру выраженности проверяемого качества в настоящий момент с показателями более раннего периода
- c) определить степень одарённости тестируемого
- d) сопоставить результаты испытуемых по разным тестам
- e) выявить личностные характеристики испытуемых

#### 5. ПЕДАГОГИЧЕСКОЕ ТЕСТИРОВАНИЕ НАПРАВЛЕНО НА РЕШЕНИЕ ЗАДАЧ

- a) выявление умений учиться
- b) отбор учащихся
- c) построение рейтинга студентов
- d) выявление пробелов в знаниях испытуемых



ПЕД	
	измерения

- е) заключение о психолого-педагогической готовности к деятельности  
 ф) аттестация учащихся

**6. К ПРЕИМУЩЕСТВАМ ТЕСТИРОВАНИЯ ПЕРЕД ДРУГИМИ МЕТОДАМИ КОНТРОЛЯ ОТНОСЯТСЯ**

- а) универсальность проверяемых знаний  
 б) объективность оценки знаний  
 с) отбор учащихся  
 д) возможность сравнения достижений студентов  
 е) положительный эмоциональный фон  
 ф) качество измерения  
 г) системность оценки знаний и умений

**7. ФОРМА ЗАДАНИЯ ЯВЛЯЕТСЯ ПРАВИЛЬНОЙ, ЕСЛИ ОНА**

- а) соответствует содержанию  
 б) имеет стандартный вид  
 с) содержит четыре и более вариантов ответов  
 д) легко воспринимается всеми респондентами  
 е) исключает возможность появления ошибочных ответов у хорошо подготовленных лиц

**8. ВЫБОР ФОРМЫ ТЕСТОВОГО ЗАДАНИЯ ЗАВИСИТ ОТ**

- а) учебного предмета  
 б) цели тестирования  
 с) содержания учебного материала  
 д) уровня проверяемых знаний

- е) сложившихся традиций тестирования

**9. ТЕСТОВОЕ ЗАДАНИЕ ЛЮБОЙ ФОРМЫ ДОЛЖНО**

- а) иметь эталон правильного решения  
 б) дублировать некоторые задания в тесте  
 с) быть кратко по форме  
 д) начинаться со слов «почему», «где», «когда»  
 е) быть ключом для выполнения другого задания  
 ф) представлять собой вопросительное предложение

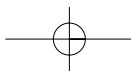
**10. УКАЖИТЕ ФОРМУ ТЕСТОВОГО ЗАДАНИЯ, КОТОРАЯ ОБЕСПЕЧИВАЕТ НАИБОЛЬШУЮ ДОСТОВЕРНОСТЬ ОЦЕНКИ ЗНАНИЙ СТУДЕНТОВ \_\_\_\_\_**

**11. К ПРИЁМАМ СОЗДАНИЯ ПРЕДМЕТНОЙ ЧАСТИ ТЕСТОВОГО ЗАДАНИЯ ОТНОСЯТСЯ**

- а) дублирование инвариантной части задания  
 б) фасетность элементов  
 с) взаимное исключение  
 д) обратимость  
 е) импликация

**12. ДЛИНА ТЕСТА ОПРЕДЕЛЯЕТСЯ КОЛИЧЕСТВОМ**

- а) страниц, который он занимает  
 б) временем, отведённым на его выполнение  
 с) заданий, которые он содержит



**13. Установите соответствие:**

Части тестового задания	Характеристика и назначение
1) Установочная	a) может отсутствовать
	b) содержит непосредственно вопрос или задачу
	c) содержит описание ситуации
	d) требует от студентов проявления активности
2) Предметная	e) предназначена для создания однотипных условий всех отвечающих
	f) включает вариативный и инвариантный компонент
Ответы:	1 . 2 .

Установить правильную последовательность:

**14. ЭТАПЫ ПРОЕКТИРОВАНИЯ ТЕСТА**

- a) выяснение ресурсных возможностей разработчиков
- b) определение целей и задач тестирования
- c) отбор тестового содержания

- d) подготовка пакета тестовых материалов
- e) предварительная апробация заданий
- f) составление заданий
- g) составление рабочего варианта спецификации теста
- h) уточнение содержания теста
- i) экспертиза теста

**ОТВЕТЫ:**

1 – a, b, e, g	4 – a, b, d	7 – a, d, e	10 – открытая	13 – 1: a c e 2: b d f
2 – b, d	5 – b, c, d, f	8 – b, c, d	11 – b, d, e	14 – b, a, g, c, f, e, h, i, d
3 – b, c	6 – b, d, f, g	9 – a, c	12 – c	

**Методология**

Методология

