

# Разработка контрольных измерительных материалов для диагностики структуры знаний обучаемых

Т.А. Снегирёва

**КАЧЕСТВО ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ЗАВИСИТ ОТ ОБЪЕКТИВНОЙ ОЦЕНКИ УЧЕБНЫХ ДОСТИЖЕНИЙ УЧАЩИХСЯ, СТУДЕНТОВ. ПОЭТОМУ ПРИ РАЗРАБОТКЕ КОНТРОЛЬНЫХ ИЗМЕРИТЕЛЬНЫХ МАТЕРИАЛОВ НЕОБХОДИМО ИСПОЛЬЗОВАТЬ СОВРЕМЕННЫЕ ПОДХОДЫ — ТЕХНОЛОГИИ, УЧИТЫВАЮЩИЕ ТРАДИЦИИ РОССИЙСКОГО ОБРАЗОВАНИЯ, ОСОБЕННОСТИ СОДЕРЖАНИЯ УЧЕБНЫХ ПРЕДМЕТОВ, ТРЕБОВАНИЯ ГОСУДАРСТВЕННЫХ ОБРАЗОВАТЕЛЬНЫХ СТАНДАРТОВ, ТЕХНОЛОГИИ, ОПИРАЮЩИЕСЯ НА СОВРЕМЕННЫЕ ТЕОРИИ ПЕДАГОГИЧЕСКИХ**

На Международной конференции «Оценка образовательных достижений в рамках национальных экзаменов», организованной Федеральным институтом педагогических измерений Министерства образования и науки РФ, которая состоялась в 2004 г. в Москве, были сформулированы общие подходы к разработке контрольных измерительных материалов для ЕГЭ.

При создании контрольных измерительных материалов (КИМ) для ЕГЭ реализуются современные подходы к объективной оценке подготовки учащихся с учётом традиций российской школы, учитываются особенности содержания учебных предметов, а также требования к созданию надёжного и объективного измерителя учебных достижений.

Контрольные измерительные материалы создаются как *тесты учебных достижений* в современном их понимании. Обеспечение качества КИМ-ов по каждому учебному предмету осуществляется при реализации следующих принципов<sup>1</sup>:

- соответствие структуры и содержания контрольных измерительных материалов основным целям ЕГЭ — определение уровня подготовки выпускников общеобразовательной средней (полной) школы с целью итоговой аттестации и отбора для поступления в вузы;

- ориентация на повышение качества образования и стимулирование развития общеобразовательной школы;

- адекватность содержания контрольных измерительных материалов основному минимуму содержания основного (общего) и среднего (полного) образования и программам для поступающих в вузы;

- представительность содержания контрольных измерительных материалов по отношению к основному минимуму основного (общего) и среднего (полного) образования и программам для поступающих в вузы;

- конструктивное описание содержания контрольных измерительных материалов, облегчающее составление и отбор соответствующих проверочных заданий;

- сочетание объективной и стандартизированной субъективной форм проверки;

- адекватность используемой формы задания (с выбором ответа, с кратким ответом, с развёрнутым ответом) проверяемым знаниям и умениям;

- недопустимость использования заданий, которые могут дис-

<sup>1</sup> Реализация эксперимента по введению единого государственного экзамена в Российской Федерации: Аналитический докл. междунар. конф. «Оценка образовательных достижений в рамках национальных экзаменов» (13–15 декабря 2004 года). М., 2004.

криминировать испытуемых по какому-либо основанию (гендерные различия, этнические различия, политические взгляды и др.);

- необходимость экспериментальной проверки контрольных измерительных материалов с целью определения валидности и надёжности заданий и работы в целом, а также системы оценивания отдельных заданий и работы в целом;

- учёт требований технологичности массовых процедур.

Объективность оценок учащихся за выполнение экзаменационной работы по основным учебным предметам обеспечивается, прежде всего, соответствием её содержания обязательному минимуму содержания основного общего и среднего (полного) общего образования.

Возможность получения объективных оценок учащихся обеспечивается также стандартизацией всего инструментария ЕГЭ — разработкой единых условий проведения экзамена, а также единых процедур проверки выполнения экзаменационных работ<sup>2</sup>.

Учитывая общие подходы и принципы при разработке контрольных измерительных материалов, реализация которых повышает их качество, ниже предлагается технология разработки педагогических тестовых материалов (ПТМ) для диагностики структуры знаний обучаемых.

Технология разработки ПТМ состоит из нескольких этапов. Приведём описание ряда ПТМ на примере медицинской и биологической физики.

### Подготовительный этап

#### 1. Определение целей тестирования

Необходимо чётко сформулировать, для каких целей создаются контрольно-измерительные материалы, в форме тестовых заданий (тесты, определяющие начальный уровень обученности; тесты входного, текущего, рубежного или

итогового контроля знаний; тесты для диагностики структуры знаний и др.).

2. *Определение контингента обучаемых* (учащиеся школ или профессиональных училищ, студенты, абитуриенты и т.д.).

3. *Определение ресурсных возможностей разработчика тестов*

На данном этапе в основном определяется форма организации процедуры тестирования (компьютерное или на бумажных носителях) и способы анализа результатов тестирования.

4. *Определение вида ПТМ* (гомогенный или гетерогенный тест)<sup>3</sup>.

5. *Определение подхода к разработке ПТМ* (нормативно-ориентированные, критериально-ориентированные или смешанные тесты)<sup>4</sup>.

В нашем случае *критериально-ориентированный тест* оценивает уровень подготовленности каждого тестируемого относительно требований учебной программы или её части. Индивидуальный результат при таком тестировании соотносится с заранее запланированным результатом (критерием).

*Нормативно-ориентированный тест* оценивает уровень подготовленности каждого тестируемого относительно других тестируемых (относительно тестовой нормы). Тестовая норма — это среднее статистическое значение тестового балла, определяемое для данного теста. Отсюда название теста: нормативно-ориентированный. Тестовая норма определяется предварительно по результатам апробации теста на представительной выборке учащихся или после проведения массового тестирования, если представительную выборку заранее установить невозможно<sup>5</sup>.

Разрабатываемые нами тесты для диагностики структуры знаний обучающихся — это смешанные тесты. «Нормой», определяющей структуру теста, является требование ГОС по соответствующей учебной дисциплине. «Критерием» — уровень формирования структуры знаний в рамках выбранной модели обучения.

### Отбор содержания учебного материала, по которому составляются ПТМ

#### 1. Разработка обобщённой таксономической модели структуры диагностируемых знаний

Обобщённая таксономическая модель структуры диагностируемых знаний представлена в виде таблицы, в которой сопряжены виды знаний, классы дескрипторов учебного предмета и уро-

<sup>2</sup> Реализация эксперимента по введению единого государственного экзамена в Российской Федерации: Аналитический докл. междунар. конф. «Оценка образовательных достижений в рамках национальных экзаменов» (13–15 декабря 2004 года). М., 2004.

<sup>3</sup> Аванесов В.С. Композиция тестовых заданий. М.: Центр тестирования МО РФ, 2002.

<sup>4</sup> Чельщикова М.Б. Теория и практика конструирования педагогических тестов: Учеб. пособ. М.: Исслед. центр проблем качества подготовки специалистов. 2001.

<sup>5</sup> Чулков П.В., Михалёва Т.Г. Составление тестовых заданий по математике: Учеб. пособие. М.: Центр тестирования МО РФ. 2003

вень их усвоения. Это основа для конструирования учебного тезауруса.

**2. Построение информационно-семантической структуры (ИСС) учебной дисциплины (разработка учебного тезауруса дисциплины)**

Содержательную область ПТМ для диагностики структуры знаний определяет учебный тезаурус медицинской и биологической физики<sup>6</sup>, разработанный на основе метода групповых экспертных оценок<sup>7</sup>.

**3. Разработка спецификации теста**

Основой для конструирования теста является спецификация, представленная таблицей — матрицей, в которой в одном столбце расположены виды диагностируемых знаний и соответствующие классы дескрипторов учебного тезауруса медицинской и биологической физики, а в другом — уровни сложности тестовых заданий ТЗ.

В качестве примера в табл. 1 приведена спецификация теста для диагностики структуры зна-

ний студентов по теме «Кинематика колебаний». Структура теста представлена в табл. 2.

**4. Анализ и экспертиза теста**

Оценка качества тестов ( $\eta$ ) производится по формуле (1)

$$\eta = C_1S + C_2R + C_3G + C_4L + C_5K;$$

в которой учтены 5 критериев качества теста: его содержательность ( $S$ ), репрезентативность структуры ( $R$ ) и ГОС ( $G$ ), пригодность для диагностики уровней обученности ( $L$ ) и видов знаний ( $K$ )<sup>8</sup>;  $C_1 - C_5$  — соответствующие коэффициенты «важности» критериев 1–5; их сумма должна быть равна единице (условие нормировки).

<sup>6</sup> Снегирёва Т.А. Диагностика структуры знаний обучаемых на основе тезаурусного и квалиметрического подходов: Дис.... канд. пед. наук. Ижевск, 2001.

<sup>7</sup> Черепанов В.С. Экспертные оценки в педагогических исследованиях. М.: Педагогика, 1989.

<sup>8</sup> Шихова О.Ф. О сертификации качества педагогических тестовых материалов // Образование и наука: Вестник Уральского отделения РАН. № 4. Екатеринбург, 2004. С. 50–55.

Таблица 1

Спецификация теста «Кинематика колебаний»

№	Виды диагностируемых знаний	Классы дескрипторов	Число ТЗ по уровням сложности			
			I	II	III	Σ
	<b>Фактуальные</b>	<b>Понятия</b>				
1	Колебания: гармонические колебания, периодические колебания, непериодические колебания		4			4
2	Основные характеристики колебания (смещение, амплитуда, период, частота, циклическая частота, фаза, начальная фаза)		16			16
3	Спектр (линейчатый, сплошной)		4			4
4	Амплитудно-частотная характеристика (АЧХ)			2		2
	<b>Классификационные</b>	<b>Классификации</b>				
1	Классификация колебаний по природе		2			2
2	Классификация колебаний по форме		2			2
	<b>Системные</b>	<b>Формулы</b>				
1	Формула гармонического колебания		4	12		16
2	Формулы (периода, частоты, циклической частоты, фазы)		8	8	4	20
	<b>Системные</b>	<b>Законы</b>				
1	Сложение гармонических колебаний с одинаковыми частотами			4	16	20
2	Сложение гармонических колебаний с частотами, кратными основной частоте			4		4
3	Теорема Фурье			2		2
	<b>Алгоритмические</b>	<b>Графические объекты</b>				
1	Графики: гармонического, периодического, непериодического колебаний				4	4
2	Определение по графику амплитуды, периода, начальной фазы			8	8	16
3	Векторная диаграмма			4	4	8
	Итого:		40	44	36	120

Структура теста «Кинематика колебаний»

№ п/п	Виды диагностируемых знаний	Классы дескрипторов	Общее число дескрипторов в ИСС	Количество дескрипторов в тесте	Уровень сложности ТЗ		
					I	II	III
1	Фактуальные	Понятия	21	20	24	2	
2	Классификационные	Классификации	2	2	4		
3	Системные	Формулы	6	6	12	20	4
4	Системные	Законы	3	3		10	16
5	Алгоритмические	Графические объекты	10	6		12	16
	Общее количество		42	37	40	44	36

Значения критериев 1–5 и их весовые коэффициенты определяются методом групповых экспертных оценок<sup>9</sup>. Например, для теста «Кинематика колебаний», оценка его качества составила

$$\eta = 0,3 \cdot 0,88 + 0,1 \cdot 1,89 + 0,2 \cdot 0,43 + 0,1 \cdot 0,67 + 0,3 \cdot 0,57 = 0,77.$$

#### 5. Составление тестовых заданий в соответствии со спецификацией, компоновка теста

Разработанные нами гомогенные тесты для диагностики структуры знаний являются тестами с заданиями закрытой формы, расположенными в порядке возрастающей трудности. Тест состоит из 4 вариантов, в каждом из которых 30 заданий. При создании отдельных вариантов тестов выдерживается их параллельность по структуре, содержанию, уровню сложности; используются также единые подходы к оценке отдельных заданий.

При составлении теста разработаны тестовые задания трёх уровней сложности. Тестовые задания I-го и II-го уровней сложности составляют инвариантную часть гомогенного педагогического теста и определяют основное ядро диагностируемых знаний студентов;

ТЗ III-го уровня сложности представляют собой вариативную часть теста, которая позволяет дифференцировать индивидуальные различия студентов, выяснить степень сформированности их аналитического мышления. Ниже приведены их характеристики<sup>10</sup>.

#### ТЗ I-го уровня сложности (ученический

уровень) предполагают узнавание формул, законов; общее представление о свойствах, явлениях (их констатацию); знание условных обозначений, единиц измерения, табличных данных, констант;

ТЗ II-го уровня сложности (алгоритмический уровень) предполагают воспроизведение формул и законов по памяти, типовые расчёты по формулам; качественную и количественную оценку свойств, явлений;

ТЗ III-го уровня сложности (эвристический уровень) предусматривают применение знания свойств, явлений, законов для решения практических задач; решение задач, требующих умения аналитически мыслить, самостоятельно разрабатывать алгоритм решения; задач по трансляции и трансформации знаний.

#### 6. Анализ и экспертиза тестовых заданий

Композиционное построение теста принимается после согласования с экспертами. Экспертиза гомогенного педагогического теста заключалась в том, что эксперты должны были определить число ТЗ по уровням сложности и оценить предложенные ТЗ (I, II или III-го уровней сложности) для обеспечения содержательной валидности теста.

Исследователи выделяют множество различных видов валидности (содержательная, соответствия, прогностическая и т.д.), но, по сути, важнейшими для педагогического тестирования являются два вида: по содержанию и по критерию<sup>11</sup>.

Содержательная валидность определяет «репрезентативность содержания дидактического теста выборке измеряемой области навыков и знаний учебного предмета»<sup>12</sup>. При определении содержательной валидности необходимо установить соответствие заданий теста целям диагностики, в качестве которых могут выступать: требования ГОС, требования к профессиональ-

<sup>9</sup> Черепанов В.С. Экспертные оценки в педагогических исследованиях. М.: Педагогика, 1989.

<sup>10</sup> Беспалько В.П. Слагаемые педагогической технологии. М.: Педагогика, 1989.

<sup>11</sup> Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М.: Прометей, 2000.

<sup>12</sup> Казаринов А.С., Култышева А.Ю., Мирошниченко А.А. Технология адаптивной валидности тестовых заданий: Учеб. пособие. Глазов: Изд-во ГГПИ, 1999. С. 21.

ной подготовке, установленные уровни обученности и т.п.

Критериальная валидность (или валидность соответствия) показывает соответствие между результатами тестирования и другими формами контроля, но с использованием в них тех же заданий, что и в тесте.

Количественной мерой валидности служит обычно коэффициент корреляции Пирсона или коэффициент связи, данные для расчёта которых получаются путём оценки группой опытных независимых экспертов предъявляемых им ПТМ по тому или иному критерию<sup>13</sup>.

### Апробация и рецензирование пробного теста

Как показывает анализ исследований, апробацию разработанного теста целесообразно осуществить на репрезентативной выборке, что обеспечивает более достоверные результаты<sup>14</sup>.

### Статистическая обработка результатов апробации теста

Статистическая обработка результатов тестирования проводится на основании методологии педагогической квалиметрии и математического аппарата теории ИРТ<sup>15</sup>.

Уровень усвоения знаний обучаемых рассчитывается по формуле

$$K = n/N, \quad (2)$$

где  $N$  — число ТЗ I, II или III –го уровней сложности;  $n$  — число правильных ответов испытуемого на ТЗ соответствующего уровня сложности. Оценка уровня усвоения знаний на соответствующем уровне сложности ТЗ проведена на основании вычисленного коэффициента усвоения (по формуле 2) и 12-балльной шкалы оценки знаний по В.П. Беспалько<sup>16</sup>.

Уровень структуры знаний определялся по формуле

$$\gamma = Xi/B, \quad (3)$$

где  $Xi$  — индивидуальный балл испытуемого;  $B$  — максимально возможное количество бал-

лов, которые может получить испытуемый при выполнении всего теста.

### Коррекция тестовых заданий (исключение неудовлетворительных и отбор удовлетворительных в конечный вариант теста)

Необходимо отметить, что практически все этапы технологии разработки контрольно-измерительных материалов осуществляются на основе метода групповых экспертных оценок<sup>17</sup>.

ПТМ, разработанные на основе предложенной технологии, прошли неоднократную апробацию на выборке из 140 студентов первого курса лечебного, педиатрического и стоматологического факультетов, обучающихся на кафедре физики Ижевской государственной медицинской академии.

Объективность результатов диагностики структуры знаний студентов обеспечивается, прежде всего, соответствием содержания ПТМ содержанию рабочей программы курса, составленной на основе требований ГОС, а также стандартизацией условий проведения компьютерного тестирования. □

<sup>13</sup> Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М.: Прометей, 2000.

<sup>14</sup> Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М., 2000; Чельшкова М.Б. Теория и практика конструирования педагогических тестов: Учеб. пособ. М., 2001; Чулков П.В., Михалёва Т.Г. Составление тестовых заданий по математике: Учеб. пособие. М.: Центр тестирования МО РФ, 2003.

<sup>15</sup> Аванесов В.С. Композиция тестовых заданий. М.: Центр тестирования МО РФ, 2002; Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М., 2000; Снегирёва Т.А. Диагностика структуры знаний обучаемых на основе тезаурусного и квалиметрического подходов: Дис.... канд. пед. наук. Ижевск, 2001; Чельшкова М.Б. Теория и практика конструирования педагогических тестов: Учеб. пособ. М., 2001.

<sup>16</sup> Беспалько В.П. Слагаемые педагогической технологии. М.: Педагогика, 1989.

<sup>17</sup> Черепанов В.С. Экспертные оценки в педагогических исследованиях. М.: Педагогика, 1989.