

# Педагогические задания в тестовой форме

Вадим Сергеевич Аванесов, профессор, доктор педагогических наук, [testolog@mail.ru](mailto:testolog@mail.ru)

## Проблемная ситуация

В предыдущей<sup>1</sup> и настоящей статье сделана попытка построить теорию педагогических заданий, объединяющую педагогические, метрические, статистические и математические методы исследования в этой области на основе педагогики, а также исследование вопросов качества заданий, безотносительно какой-либо одной учебной дисциплины.

Актуальность такой работы объясняется тем, что задания начинают приобретать ведущую роль в научной организации компьютеризованного процесса образования и самообразования, контроля и самоконтроля, особенно в связи с существенным увеличением доли самостоятельной работы учащихся и студентов, а также в организации дистанционного обучения. Такая ситуация вызвала заметное позитивное изменение в отношении педагогической общественности к заданиям в тестовой форме, чего нельзя пока сказать о тестах.

К настоящему времени теория педагогических заданий ещё не оформилась как научно-педагогическая теория. Одна из причин — её междисциплинарность. С одной стороны, это должна быть теория, основной предмет которой — формирование принципов разработки педагогических заданий. Однако в самой педагогической науке до сих пор всё ещё нет направления, которое занималось бы этим предметом. С другой стороны, задания много лет исследуются в западных теориях педагогических и психологических тестов. Основным недостаток таких теорий — слабая связь статистических и математических методов с педагогикой и недостаточное внимание к педагогическим методам исследования.

Педагогическая наука пронизана идеей ведущей роли учителя в учебном процессе.

В то же время современные компьютер-

ные образовательные технологии, где ведущую роль играют учебные тексты и задания в тестовой форме, позволяют получать образование без повседневного участия преподавателя, дистанционно или самостоятельно, что возможно лишь при наличии качественных методических материалов и заданий. Разумеется, это не следует толковать как недооценку важной роли учителя в общей постановке процесса образования и самообразования.

Вопрос в другом. В наше время заметно изменились некоторые функции учителя и учебных материалов. Заметно возрастает роль учебных заданий и текстов для самостоятельной работы, что объективно позволяет уменьшить трудозатраты учителя, повысить качество усвоения материала за счёт активизации мыслительной деятельности учащихся и уменьшить трудоёмкость выполняемой учащимися учебной работы. Это вопрос эффективности организации учебной работы, который решается при очень важном условии — организации уровня образования на основе шкалированных по трудности заданий в тестовой форме. Можно высказать предположение, что с течением времени будет усилена роль заданий в тестовой форме для организации уровня образования и самообразования, а также контроля и самоконтроля в компьютеризованной форме.

В последние шесть лет интерес к разработке педагогических заданий существенно возрос в связи с проводившимся в России экспериментом по введению Единого экзамена (далее ЕГЭ). Этот «эксперимент» сопровождался созданием так называемых контрольных измерительных материалов (далее КИМов), на основе систем своеобразных взглядов, которые можно назвать КИМологией и ЕГЭведением.

В истории науки и литературы различное понимание явлений и процессов приносило,

<sup>1</sup> См. статью автора в ШТ, 2007. № 3.

как правило, положительные результаты, если в сфере управления и социуме проявляются способности решать возникающие противоречия. К нашему времени обнаружилось противоречия в вопросах истолкования предмета, составных и структурных элементах, в методах обоснования качества заданий и в существенных лексических различиях в области педагогических измерений. Нынешняя ситуация оставляет мало надежд на культурное решение упомянутых противоречий. А это означает, что к этим вопросам придётся возвращаться ещё не раз. Правительство надеется ввести ЕГЭ в полном масштабе к 2009 году, но уже не как эксперимент.

### Научная проблема

Рассмотрим проблему построения такой теории педагогических заданий, которая могла бы представлять логически непротиворечивую систему научного знания, теорию, дающую целостный взгляд на существенные свойства заданий, что открывает дорогу к существенному улучшению образовательного процесса, потому что в сферу применимости такой теории входят задания, применяемые как в учебном процессе, так и при контроле уровня подготовленности учащихся и студентов, выпускников школ, вузов и всех других образовательных заведений.

В теории должны найти своё место научные определения, аксиомы, постулаты, гипотезы, принципы, формы, а также вопросы определения содержания заданий, и, кроме того, методы обоснования качества заданий и тестов, представляющих системы таких заданий. Предполагается, что такая теория создаст основу для разработки качественных методов педагогических измерений, в которых качество проверяется по известным критериям надёжности и валидности тестовых результатов.

### Цели и задачи теории

*Цель* создания теории педагогических заданий — формирование логически непротиворечивой системы научного знания, дающей целостный взгляд на существенные свойства заданий, применяемые в учебном процессе. Очевидно, что в осмысленном педагогическом процессе содержание и мера трудности предлагаемых заданий зависят от уровня подготовленности испытуемых,

хотя форма какое-то время может быть одной и той же. Необходимость достижения сформулированной здесь цели вытекает из современной практики расширяющегося и массивного применения заданий в тестовой форме в автоматизированных системах обучения и контроля знаний.

*Задачи* теории педагогических заданий:

- определение исходных и основных понятий данной теории;
- формулирование системы гипотез, аксиом и ведущих постулатов теории, необходимых для создания качественных заданий;
- разработка общих принципов формулирования заданий;
- исследование формы и содержания заданий;
- определение логических требований к разработке заданий в тестовой форме;
- исследование возможностей математического моделирования процесса ответов испытуемых на задания педагогического теста и прогнозирования вероятности правильного ответа в зависимости от уровня трудности задания и уровня подготовленности испытуемых.

### Исходные определения

К теории педагогических заданий применимы требования обоснованной, логически непротиворечивой системы научного знания, дающей целостный взгляд на существенные свойства педагогических заданий, независимо от конкретной учебной дисциплины и от уровня обучения. В дополнение к ранее сформулированным определениям добавляются три других:

- *предмет* теории педагогических заданий составляет система понятий и их определений, характеристики заданий, формы заданий, общие для всех учебных дисциплин, а также инвариантные принципы отбора содержания педагогического контроля для всех учебных дисциплин;
- *объект* настоящего исследования — процесс разработки, исследования и практического применения педагогических заданий для целей педагогического измерения и улучшения практики учебного процесса;
- *педагогические измерения* можно определить как процесс представления интересующего свойства испытуемых и других объектов педагогической науки и практики в виде

числовой величины. В самом общем виде величиной можно назвать всё то, что может быть больше или меньше, что может быть присуще объекту в большей или меньшей степени; числовая величина — такая, которая может быть выражена числом<sup>2</sup>. Таким образом, измерение есть установление числового соотношения между проявлениями интересующего свойства у изучаемых объектов. Измерение интересующего свойства испытуемых производится в предположе-

нии, что у каждого из них есть интересующее свойство в каком-то количестве. Если выясняется, что у испытуемого нет данного свойства, то это даёт основание для исключения данного испытуемого из предполагаемой выборки лиц, обладающих данным свойством<sup>3</sup>.

Обобщая всё ранее написанное, можно сказать, что по первой задаче получены следующие результаты:

- даны новые определения шести исходным<sup>4</sup> и десяткам основных понятий двух теорий — педагогических заданий и общей педагогической теории измерений. Среди понятий — «педагогическое задание»<sup>5</sup>, «задание в тестовой форме», «тестовое задание», «композиция заданий в тестовой форме», «педагогический тест» и многие другие. Предполагается, что при благоприятных условиях эти понятия помогут создать языковую основу сущностного понимания основ педагогических измерений у специалистов, владеющих русским языком. Нынешнее засилье калькированных переводов с английского языка, а также введение понятий ЕГЭведения, может отодвинуть нормальное развитие научных проблем педагогических измерений.

Актуальность такого подхода вытекает из нынешнего неупорядоченного состояния лексики педагогических измерений, где в последние годы предпринята массивная попытка внедрить в теорию педагогических измерений неадекватную лексику так называемых «ЕГЭ-КИМов»<sup>6</sup>.

Трудность проведения данной работы заключается в том, что вместо развития педагогических измерений мы столкнулись с пропагандистскими попытками намеренного изменения смысла вещей путём изменения их наименования и внедрения в сознание социума удобных мифов и метафор о ЕГЭ как об универсальном средстве решения множества социальных и образовательных проблем.

Когда Конфуция спросили, как он стал бы управлять государством, мудрец ответил: я бы начал с возвращения словам их смысла.

В основу многих предложенных определений положена диалектическая идея о понятии как форме мысли, отображающей предметы и явления в их наиболее общих и существенных признаках, а также как итоге познания предмета, характеризующем су-

<sup>2</sup> Петров Ю.А., Никифоров А.Л. Логика и методология научного познания. М.: Изд-во Моск. ун-та, 1982. С. 64.

<sup>3</sup> На странице 42–43 статьи В.С. Аванесова в журнале «Педагогические измерения», 2006, № 2 оказалась незамеченной автоматическая компьютерная конвертация греческой буквы ( $\theta$  тета) в латинскую букву  $q$ . Это произошло в конце стр. 42, после слов « вслед за классиком американской тестологии F.M. Lord... Там надо поменять латинскую букву  $q$  на греческую  $\theta$ , где  $\theta$ , означает значение шкалированного тестового балла испытуемого  $i$ . И далее, по смыслу, до пункта 7, стр. 43. То же явление имело и в одной прежней публикации автора в журнале «Педагогическая диагностика». Автор и редакция приносят читателям свои извинения.

<sup>4</sup> Определение исходных понятий теории педагогических измерений. ПИ, 2005. №2.

<sup>5</sup> Аванесов В.С. Основы педагогической теории измерений. ПИ, 2004. № 1.

<sup>6</sup> Например, шесть лет в стране проводился так называемый эксперимент по введению единого госэкзамена. Естественно поставить вопрос: — что такое ЕГЭ? В распространенном в Госдуме проекте Федерального Закона «О внесении изменений в Закон Российской Федерации «Об образовании» и Федеральный закон «О высшем и послевузовском профессиональном образовании» в части введения единого государственного экзамена» (где определения обязательны), написано следующее: «Единый государственный экзамен представляет собой форму независимой оценки уровня учебных достижений обучающихся с использованием заданий стандартизированной формы (контрольные измерительные материалы), выполнение которых позволяет установить уровень учебных достижений обучающихся по освоению федерального компонента государственного образовательного стандарта среднего (полного) общего образования». Из чего законодатель должен понять, что ЕГЭ — это форма независимой (от кого?) оценки с использованием формы неопределённых материалов!

В появившемся, наконец, в открытой печати, так называемом «аналитическом докладе» вместо определения ЕГЭ на стр. 9 написано: «В основе ЕГЭ лежат две основные идеи: идея повышения объективности оценки и идея совмещения в одной процедуре школьного выпускного экзамена и вступительного экзамена». См.: «Национальные экзамены в системе оценки качества образования: Материалы и тезисы докладов Международной конференции. 12–14 декабря 2005 г. М.: Уникум—Центр, 2006. Читатель может вывести, что в правительственном эксперименте совмещены идея повышения объективности оценки и идея совмещения двух экзаменов. Если отставить в сторону смысловую бессодержательность используемых при этом слов, то опять получается, что ЕГЭ — это форма формы и совмещение совмещения!

Обращает на себя внимание и лукавое название издания — «Национальные экзамены», вместо привычного «Единого государственного экзамена», что можно истолковать как начало осмысления негодности названия и сути самого ЕГЭ.

ущественные признаки предмета, известные в настоящее время. Существенные признаки — это такая группа признаков предмета, каждый из которых, отдельно взятый, необходим, а все, вместе взятые, достаточны, чтобы с их помощью можно было отличить данный предмет от всех остальных.

### Объективные свойства педагогических заданий

В предыдущей статье все свойства педагогических заданий были разделены на три класса: объективные, субъективные и интерсубъективные<sup>7</sup>.

*Объективные* свойства заданий включают содержание, форму и технологичность. Рассмотрим ключевые вопросы формы тестовых заданий.

Форма определяется как способ связи организации, упорядочения и существования содержания в общей композиции тестовых заданий<sup>8</sup>. Сложность проблемы состоит в том, что относительно формы существует противоречие между теоретическим и практическим мышлением. Большинству практиков тестового процесса форма заданий кажется знакомой и вполне понятной, а потому они не видят здесь никаких проблем. Соответственно, практики не считают нужным что-то менять в форме заданий с выбором одного правильного ответа — в форме, сложившейся сто лет назад в Европе, и, как выясняется<sup>9</sup>, сотни лет назад в Средней Азии<sup>10</sup>.

Есть основания утверждать, что практики не хотят учиться формотворчеству и методике разработки новых тестовых заданий. Из-за такой позиции практика тестирования в России начала деградировать. Она полна примеров неадекватных форм тестирования. Самый типичный случай — массированное использование заданий с выбором одного правильного ответа из трёх-пяти ответов для проверки чуть ли не всех видов знаний<sup>11</sup>, что даёт нам пример массового заблуждения относительно измерительных возможностей данного варианта заданий с выбором одного правильного ответа.

Если практики видят в форме привычное средство создания теста или, скорее, набора заданий, похожего на тест, то теоретик обращает внимание на достоинства и недостатки существующих форм, на поиск дополнительных возможностей автоматизации тестового процесса, а также на поиск

таких форм, которые позволяют проверить знания полнее, глубже, точнее, быстрее и объективнее. Вспомним Г. Спенсера: хорошее содержание всегда должно быть в хорошей форме! Вот почему исследование формы становится, быть может, самой важной частью теории.

Теоретическое мышление находит отношение практиков к форме ограниченным и вредным, тормозящим развитие тестовой культуры. От правильности формы зависит понимание содержания заданий, технологичность тестового процесса, точность и эффективность измерений. Малейшее нарушение требования формы часто делает задание непонятным. Например, при нечётких инструкциях испытуемые не знают, как им отвечать на задания. Нарушение требований формы также является одним из самых распространённых источников погрешностей педагогического измерения. Можно выразиться ещё более определённо: пока преподаватели не овладеют формами, нет надежд на появление качественных тестов.

Обобщая полученные результаты по вопросам формы, можно сказать, что исследованы четыре формы тестовых заданий, предложена система принципов их композиции<sup>12</sup>, разработаны новые варианты и системы заданий в тестовой форме<sup>13</sup>.

### Угадывание правильных ответов

Результаты тестирования посредством самых распространённых заданий с выбором одного правильного ответа из 3–5 предлагаемых на выбор, подвержены известным искажениям — тем большим, чем меньше число ответов в каждом задании. В таких заданиях, наряду с одним правильным ответом, есть несколько неправильных, но правдоподобных ответов. Последние называются дистракторами (от англ. to distract — отвлекать); число дис-

<sup>7</sup> **Аванесов В.С.** Основы теории педагогических заданий. ПИ, 2004. № 2. С. 26–62.

<sup>8</sup> **Аванесов В.С.** Форма тестовых заданий. М.: Центр тестирования, 2005. С. 8.

<sup>9</sup> **Джуринский А.Н.** История зарубежной педагогики: Учеб. пособие для студентов вузов. М., 1998. С. 8–9.

<sup>10</sup> **Калдыбаев С.К.** Педагогический аспект становления и развития тестов. ПИ, 2006. № 2. С. 91–92.

<sup>11</sup> О видах знаний см.: **Аванесов В.С.** Знание как предмет педагогического измерения. ПИ, 2005. № 3. С. 3–31.

<sup>12</sup> **Аванесов В.С.** Композиция тестовых заданий. М.: Центр тестирования, 2003.

<sup>13</sup> **Аванесов В.С.** Форма тестовых заданий. М.: Центр тестирования, 2005; **Аванесов В.С.** Применение тестовых форм в Rasch Measurement. ПИ, 2005. № 4. С. 3–20.

тракторов обычно колеблется от одного до четырёх.

В заданиях с выбором одного правильного ответа из трёх вероятность угадывания равна  $1/3$ , что приводит к тому, что одна треть всех заданий может быть решена не за счёт знания учебного материала, а за счёт ответов испытуемым наугад. В тесте, состоящем, скажем, из тридцати заданий, таких «правильных» ответов может оказаться около десяти, за это педагоги могут выставить привычные три балла. Но это ошибочная практика. Не случайно студентам и школьникам нравятся задания с тремя ответами, где всегда есть реальная возможность угадывать. Именно таких заданий с тремя и четырьмя ответами особенно много в пособиях по русскому и иностранным языкам, а также по русскому языку как иностранному. Если ввести коррекцию на угадывание, то всё равно остаётся проблема неточности измерений и неэффективности используемой шкалы.

Задания с выбором одного правильного ответа имеют два существенных дефекта: 1) достаточно высокую вероятность угадывания правильного ответа и 2) возможность правильного ответа на задания без выполнения самих заданий, посредством поочередной подстановки предлагаемых к заданию ответов. Угадывание правильных ответов в теории педагогических измерений исследовалось неоднократно; оно рассматривается как источник погрешностей измерения — тем большее, чем больше доля угадываемых правильных ответов. Для коррекции тестовых баллов испытуемых, полученных вследствие возможного угадывания, используется формула

$$X_{ci} = X_i - \frac{W_i}{k - 1}$$

где  $X_{ci}$  — скорректированный на догадку тестовый балл испытуемого. Отсюда и смысл индекса:  $c$  от англ. corrected, символ  $i$  обозначает номер испытуемого.

$X_i$  — тестовый балл испытуемого  $i$ , без коррекции;

$W_i$  — число ошибочных ответов у того же испытуемого.

Эта формула используется при предположении, что испытуемый не знает правильного ответа ни на одно задание, и по всему тесту пытается отвечать наугад. В ней наиболее

вероятное число ответов, которое можно угадать, ничего не зная, вычитается из набираемой суммы баллов.

Если взять, для примера, тест, состоящий из 30 заданий с четырьмя ответами, то в случае 20 правильных и 10 неправильных ответов получим:

$$X_{ci} = 20 - \frac{10}{4 - 1} = 16,6,$$

или, округлённо, 17 баллов. Из структуры этой формулы видно, что с увеличением количества правильных ответов число вычитаемых на догадку баллов в заданиях с четырьмя ответами заметно уменьшается. Из чего видно, что хорошо подготовленных испытуемых коррекция баллов на догадку не должна беспокоить, в то время как у плохо подготовленных испытуемых коррекция отнимает много баллов. Например, в случае, когда есть всего десять правильных и двадцать неправильных ответов, скорректированный тестовый балл оказывается равным  $X_{ci} = 10 - 20 / 4 - 1 = 3,333$ , или округлённо, трём баллам, что означает существенное понижение исходных тестовых баллов испытуемых.

Немногим лучше дела обстоят в заданиях с выбором одного правильного ответа из пяти ответов. Такие задания широко применяются во всех российских и зарубежных центрах тестирования. При выборе из пяти ответов правильные ответы могут быть угаданы примерно в пятой части от общего числа заданий. В итоге испытуемые получают баллы, которые они не заслужили. Это и есть одна из самых распространённых форм искажения тестовых результатов за счёт усталой и несовершенной формы заданий.

Второй дефект зримо проявляется в заданиях по математике с выбором одного правильного ответа, где большинство испытуемых пробуют подставлять предлагаемые ответы способом перебора, и, таким образом, находят правильный ответ.

Задания с выбором одного из 4–5 ответов применяются на едином экзамене. При этом коррекция результатов на догадку не проводится. Тем самым тестовые баллы содержат явные источники погрешностей измерения. При этом есть много дистракторов, которые вообще не выбирают испытуемые, вследствие чего повышается вероятность угадывания правильного ответа и снижается точ-

ность измерений. Для противодействия этому отрицательному явлению обязательно необходимо проводить так называемый «дистракторный анализ», суть которого кратко можно свести к подсчёту процентов выбора каждого ответа во всех заданиях. Хотелось бы подчеркнуть, что без дистракторного анализа тестов не бывает.

Специалисты по ЕГЭ приняли решение использовать задания открытой формы, где угадывание практически невозможно. Ответы по таким заданиям в ЕГЭ оцениваются в два раза выше, поскольку задания открытой формы труднее для испытуемых, чем задания с выбором одного правильного из 4–5 ответов. Надо ли убеждать авторов таких решений в простых истинах: трудность заданий зависит, главным образом, не от формы, а от содержания заданий! Таким образом, принятая в ЕГЭ система оценивания порождает ещё один источник погрешностей измерения.

Ошибки оценивания в этих частях столь велики, что по некоторым дисциплинам результаты испытуемых не коррелируют положительно с результатами испытуемых по первой части «А», а иногда связаны и отрицательно. Вопрос: можно ли складывать полученные по частям баллы в общую оценку за ЕГЭ, и если это делается, то каковы научные основания? Здесь скрывается третий источник ошибок измерения, вызванный *неверным толкованием* тестовых форм и их возможностей, ошибками сложения баллов, не обладающих свойством аддитивности.

Задания открытой формы из-за их нетехнологичности для массовых исследований не годятся. Ведь не случайно зарубежные тестовые службы в массовых обследованиях используют только задания с выбором, используя далее автоматизированную считку результатов с бланков посредством оптических сканеров очень высокой производительности, от 5 до 50 тысяч бланков в час. Хороший пример успешного применения такой техники даёт Национальное тестирование Республики Казахстан.

В России для проведения ЕГЭ применяется массовый ручной труд, сканирование и прочие атрибуты неэффективных технологий.

### Уменьшение вероятности угадывания правильного ответа

Задания, в которых довольно высока вероятность угадывания правильных ответов,

для качественного педагогического тестирования не пригодны, слишком велики при этом ошибки измерения. Кроме того, при догадке сильно теряют моральная и воспитательная стороны тестового метода. Принцип развития в них заменяется принципом угадывания. В качестве конструктивного выхода из тупика, в котором заблудилась практика тестирования, автор предлагает переходить к использованию заданий с выбором нескольких правильных ответов из примерно 10–12 предлагаемых. В таких заданиях вероятность угадать комбинацию правильных ответов оказывается менее одной тысячной<sup>14</sup>.

Вот почему, во-первых, надо скорее менять формы тестовых заданий. Вместо заданий с выбором одного правильного из 3–5 ответов надо переходить (где это оправданно) к заданиям с выбором нескольких правильных ответов. Посмотрим на некоторые примеры:

#### 1. IRREGULAR VERBS

- |             |            |
|-------------|------------|
| 1) start    | 9) begin   |
| 2) swim     | 10) jump   |
| 3) speak    | 11) wear   |
| 4) look     | 12) print  |
| 5) run      | 13) obtain |
| 6) read     | 14) finish |
| 7) think    | 15) draw   |
| 8) complete |            |

#### 2. «О» ПИШЕТСЯ В СЛОВСОЧЕТАНИЯХ

- 1) упл...тнить сроки
- 2) обн...жить пороки
- 3) р...птать на судьбу
- 4) прил...скать собаку
- 5) отк...заться от услуг
- 6) опр...вдать поступок

<sup>14</sup> **Аванесов В.С.** Форма тестовых заданий. М.: Центр тестирования, 2005; Аванесов В.С. Проблема формы тестовых заданий: В сб. Тестовые формы контроля по русскому языку как иностранному (РКИ). II Всероссийская научно-практическая конференция. Доклады и сообщения. Центр международного образования МГУ им. М.В. Ломоносова, 12–13 апр. 2005 г. / Под ред. Н.П. Андрушиной, О.А. Усковой. М., 2005. С. 20–30; Применение тестовых форм в Rasch Measurement // Педагогические Измерения, 2005. № 4. С. 3–20; **Аванесов В.С., Брусенина Н.Д., Барер Г.М.** Новые формы научной организации самостоятельной работы студентов. <http://testolog.narod.ru> 2005 г.

- 7) угр...жать расправой
- 8) в...пиющий произвол
- 9) раск...лить сковороду
- 10) усл...жнить обстановку

К таким заданиям даётся инструкция: «Вашему вниманию предлагаются задания, в которых могут быть один, два, три и больше правильных ответов. Обведите кружком номера всех правильных ответов!»

Преимущество таких заданий — в технологичности, что позволяет, при правильной организации тестового процесса, исключить элементы субъективизма в оценивании. Кроме того, испытуемые не знают точное число правильных ответов, что делает такие задания труднее. Помимо повышения трудности и технологичности, задания этого варианта позволяют проверить знания полнее, глубже и точнее. В таких заданиях испытуемые могут получить от нуля до трёх баллов, что повышает вариацию данных и, как следствие, точность измерения.

Во-вторых, помимо заданий с выбором нескольких правильных ответов в педагогических измерениях можно использовать сдвоенные задания. При условии, что они рассматриваются как одно общее задание, имеющее один номер и один общий балл. Иначе возможен эффект цепных заданий, недопустимый в тесте, из-за нарушения требования аксиомы локальной независимости, рассматриваемой ниже.

### 3. ЧТО ВЫ ЧИТАЕТЕ?.. — СПРОСИЛ (А)

- |            |               |
|------------|---------------|
| 1) Офелия  | 5) Лаэрт      |
| 2) Полоний | 6) Горацио    |
| 3) Клавдий | 7) Гертруда   |
| 4) Гамлет  | 8) Фортинбрас |

«СЛОВА, СЛОВА, СЛОВА» — последовал ответ

- |            |                |
|------------|----------------|
| 1) Офелии  | 5) Лаэрта      |
| 2) Полония | 6) Гамлета     |
| 3) Клавдия | 7) Гертруды    |
| 4) Горацио | 8) Фортинбраса |

За правильный ответ в каждом из сдвоенных заданий испытуемый получает по одному баллу. Таким образом, за такое сдвоенное задание

теста оценки могут варьироваться от нуля до двух баллов. Эта оценка благотворно отражается на повышении различающей способности заданий и на точности измерения.

В-третьих, для создания параллельных вариантов теста полезно применять так называемые фасеты. Все элементы одного фасета считаются одинаково трудными по содержанию. Здесь они представлены в фигурных скобках. Каждому испытуемому компьютерная программа подбирает один вариант из фасета:

}	Краткость есть душа ума
	Порвалась дней связующая нить
	Весь мир — тюрьма!
	Какого обаянья ум погиб!
	Всех слушай, но беседуй редко с кем
	Я пал (а), чтоб встать!

### СКАЗАЛ (А)

- |               |                |
|---------------|----------------|
| 1) Гамлет     | 6) Лаэрт       |
| 2) Офелия     | 7) Актёр       |
| 3) Полоний    | 8) Горацио     |
| 4) Клавдий    | 9) Гертруда    |
| 5) Фортинбрас | 10) Розенкранц |

### Субъективные и интересубъективные свойства заданий

К субъективным свойствам относятся *привлекательность и трудоёмкость заданий*<sup>15</sup>. Привлекательность выражается мерой интереса учащихся к заданиям, трудоёмкость — мерой трудозатрат. Одна из причин образовательного кризиса заключается в недостатке привлекательных заданий и в отсутствии точного учёта трудоёмкости заданий, предлагаемых учащимся и студентам.

*Интерсубъективные свойства заданий* возникают как следствие измерения и усреднения субъективных свойств, что позволяет перевести субъективные свойства в группу интересубъективных или иначе, объективированных свойств заданий. Одна из двух важнейших целей педагогических измерений — получение объективированных оценок интересующих свойств заданий на линейной шкале. Вторая цель — получение на той же шкале оценок уровня подготовленности учащихся (испытуемых). Использованное здесь

<sup>15</sup> Аванесов В.С. Основы теории педагогических заданий. ПИ, 2006. № 2. С. 43–44.

разом, за такое сдвоенное задание

словосочетание «интерсубъективные оценки» означает, что это оценки уже не субъективные, но ещё и не объективные. Этот промежуточный класс оценок полезно ввести для философского осмысления проблемы соотношения объективного и субъективного в процессе педагогических измерений.

Субъективизм — это мировоззренческая позиция, игнорирующая объективный подход к действительности, отрицающая объективные законы природы и общества. Интерсубъективизм, как позиция, позволяет отразить усреднённые оценки и точки зрения, создающие предпосылки для отказа от крайних точек зрения и приближения к некоторому объективно функционирующему типу мышления относительно интересующих объектов. Решающую роль при этом играет обоснованный выбор множеств заданий, испытуемых, экспертов и методов. В тестовой практике чаще других применяются методы стратифицированного отбора заданий в тест, случайного отбора испытуемых из тщательно сформированной генеральной совокупности и выбора методов, дающих наименьшее смещение выборочных статистик от параметров генеральной совокупности<sup>16</sup>.

### Гипотезы

Каждая научная теория имеет гипотезы, подлежащие эмпирической проверке. Гипотезой называют научно обоснованное предположение о сущности, связях и причинах интересующих явлений. В теориях педагогических измерений и заданий могут использоваться гипотезы об адекватности формы и содержания заданий целям и задачам образовательной деятельности. Наиболее часто применяемые гипотезы — о наличии тестовых свойств у заданий, или, иначе, о пригодности заданий для включения в тест.

При исследовании качества педагогических заданий используются ряд гипотез.

Первой можно назвать гипотезу адекватности содержания заданий целям и задачам измерения. В процессе проверки содержательной валидности результатов тестирования содержание некоторых заданий оказывается неадекватным цели измерения или не соответствующим измеряемому свойству (качеству). Особенно много нареканий вызывают задания по литературе и математике. Например, литературоведчески невежественным считается вопрос:

«Укажите, кто является выразителем авторского мировоззрения в поэме Н.А. Некрасова «Кому на Руси жить хорошо?»»

- 1) странники
- 2) Матрёна Тимофеевна
- 3) Григорий Добросклонов
- 4) Ермил Гирич»

Все перечисленные ответы к этому вопросу могут быть признаны правильными, поскольку и странники, и упомянутые герои, и поэма в целом выражают авторское мировоззрение<sup>17</sup>. К мнению известного учителя-словесника хотелось бы добавить, что подобные вопросы не имеют позитивного отношения не только к литературе, но и к форме тестовых заданий, не говоря уже о методике педагогических измерений. Недемократический и вредный запрет на апелляцию, существовавший во время «введения эксперимента» относительно содержания подобных вопросов оказал очень плохую услугу организаторам ЕГЭ, лишив их возможности общественной критики совершаемых ошибок. Например, в США и Казахстане подобного запрета нет, а потому возникающие там критические замечания используются для удаления неудачных заданий из банка Национального Тестирования.

Вторая гипотеза — адекватности используемых форм контроля знаний целям измерения, а иногда и задачам образовательной деятельности в целом. Практика полна примерами неадекватных форм тестирования. Самый типичный случай — массированное использование заданий с выбором одного правильного ответа из трёх-пяти ответов для проверки чуть ли не всех видов знаний<sup>18</sup>, что указывает на состояние массового заблуждения относительно измерительных возможностей данного варианта заданий с выбором одного правильного ответа. Особенно губительно влияние такого заблуждения на математическое образование. Вместо выполнения заданий и определения правильного ответа многие испытуемые во время тестирования заняты выбраковкой заведомо неподходящих ответов и угадыванием подходящего ответа. Гипотеза адекватности формы

<sup>16</sup> Аванесов В.С. Основы теории педагогических заданий. ПИ, 2006. № 2.

<sup>17</sup> Айзерман Л. Зачем я сегодня иду на урок литературы. М.: Изд-во Захаров, 2005. С. 21.

<sup>18</sup> О видах знаний см.: Аванесов В.С. Знание как предмет педагогического измерения. ПИ, 2005. № 3. С. 3–31.

выбора одного правильного из четырёх-пяти ответов ошибочна из-за высокого уровня вероятности угадывания. Не случайно в трёхпараметрической модели измерения появился параметр вероятного угадывания  $c_j$ , который в нашей литературе неточно называется «параметром угадывания». F.M. Lord называет его параметром «псевдоугадывания».

Третья гипотеза, обычно выдвигаемая при проверке тестовых свойств педагогических заданий, — это гипотеза связи ответов испытуемых на отдельное задание (вектор-столбец  $X_j$ ) с суммой баллов испытуемых по всему тесту (вектор-столбец  $Y_j$ ). Подтверждение этой гипотезы корреляционными методами обычно даёт надежду на включение задания в тест. При отсутствии достоверной связи ниже 0,300 задание лишается возможности быть включённым в тест. Пример статистической проверки такой гипотезы читатель найдёт в предыдущей статье<sup>19</sup>.

Четвёртая гипотеза — соответствие уровня трудности заданий уровню подготовленности испытуемых. Эта гипотеза проверяется специальными математико-статистическими пакетами Winsteps и RUMM при сравнении двух гистограмм: распределения тестовых результатов испытуемых и уровня трудности заданий (Person — Item Location Distribution), а также на основе статистического вывода относительно так называемых Item Fit и Person Fit.

Пятая гипотеза — об одномерности получаемой шкалы. Одномерность означает, что все задания разрабатываемого теста измеряют в основном одно и тоже интересующее свойство. Эта гипотеза имеет не абсолютный, а статистический характер. На английском языке она часто называется assumption of unidimensionality. На русском языке это можно назвать предположением (гипотезой) одномерности заданий теста для измерения интересующего свойства испытуемых. Эта гипотеза практически никогда и нигде в полной мере не подтверждается. Тем не менее, она позволяет использовать теорему умножения вероятностей правильных ответов на задани-

мости нет. Каждое задание теста должно быть индикатором наличия того свойства, которое призван измерять тест. Задания, обладающие такими свойствами, образуют гомогенный тест.

## Аксиомы

К семи ранее сформулированным аксиомам<sup>20</sup> педагогической теории измерений необходимо добавить четыре другие.

Первая — это аксиома так называемой локальной независимости, которую можно перевести на русский язык следующим образом: для испытуемых одинакового уровня подготовленности вероятность правильного ответа на одно задание не должна зависеть от вероятности правильного ответа на любое другое задание теста. Казалось бы, эта аксиома противоречит интуитивно понимаемой связи между вероятностью правильного ответа испытуемых на трудные и лёгкие задания: правильные ответы на трудные задания как будто гарантируют большую вероятность правильного ответа на лёгкие задания, на чём и основана кумулятивная модель измерений Л. Гутмана.

Вторая аксиома касается вопроса гомогенности заданий теста. Утверждается принадлежность содержания всех заданий одной, и только одной интересующей дисциплине, иначе свойство гомогенности можно назвать предметной чистотой заданий теста. Эта аксиома имеет не абсолютный характер, а скорее похожа на исходное предположение, открывающее дорогу к созданию теста из имеющихся заданий. Каждое задание теста должно, по идее, быть индикатором наличия того свойства, которое призван измерять тест. Задания, обладающие такими свойствами, образуют тест. Результаты полученного при этом теста обладают свойством валидности по содержанию.

Аксиома гомогенности (одномерности) практически никогда и нигде в полной мере не подтверждается. Тем не менее, она позволяет использовать теорему умножения вероятностей правильных ответов на задания теста. Из-за инструментальной ценности этой полезной теоремы альтернативы аксиоме одномерности нет.

Третья аксиома сформулирована Г. Рашем. Она утверждает: правильность (или наоборот, неправильность) ответов испытуемых

<sup>19</sup> Аванесов В.С. Основы теории педагогических заданий. ПИ, 2006. № 2. С. 49–52.

<sup>20</sup> Аванесов В.С. Проблема качества педагогических измерений. ПИ, 2004. № 2. С. 3–27.

ных ответов на задания теста. Из-за инструментальной ценности этой полезной теоремы альтернативы гипотезе одно-

на задания теста зависит от разности между уровнем подготовленности испытуемого под номером  $i$  и уровнем трудности заданий под номером  $j$ . Формально эта аксиома записывается так:  $P = f(\theta_i - \beta_j)$ . С ростом уровня подготовленности испытуемых вероятность правильного ответа на задание фиксированного уровня трудности возрастает, что вполне согласуется со здравым смыслом. Этому же автору удалось поставить в соответствие здравому смыслу и этой гипотезе функцию

$$P_j(\theta) = \{x_{ij} = 1 \mid \beta_j\} = \frac{\exp(\theta_i - \beta_j)}{1 + \exp(\theta_i - \beta_j)} \quad (1)$$

где  $P_j(\theta)$  — это вероятность правильного ответа на задание под номером  $j$ , вместо  $x$  стоит разность  $\theta_i - \beta_j$ .

$x_{ij} = 1$ , если ответ испытуемого  $i$  на  $j$ -ое задание правильный;

$\theta$  — уровень подготовленности (знаний), латентная переменная. Поэтому индекс  $i$  при символе опущен.

$\beta_j$  — уровень трудности  $j$ -го задания теста, измеряемой на латентном континууме трудности заданий.

$\exp$  — константа  $e$ , иррациональное число, равное (округлённо) 2,71828.

Четвёртая аксиома касается качества измерения. Для того, чтобы точность измерения была высокой, уровень трудности заданий должен соответствовать уровню подготовленности испытуемых. Симметрично можно говорить об адекватности уровня подготовленности испытуемых уровню трудности заданий<sup>21</sup>. В случае правильности данной аксиомы принимается, что математическая функция

$P = f(\theta_i - \beta_j)$  верно отображает связь между уровнем подготовленности испытуемых, трудностью заданий и вероятностью правильных ответов. При этом предполагается, что значения  $\theta_i$  и  $\beta_j$  истинны, безошибочны, отражают соответствующие уровни на латентной шкале, а потому называются параметрами:  $\theta_i$  называется параметром уровня подготовленности испытуемого, а  $\beta_j$  — параметром трудности задания.

### Постулаты

Кроме аксиом, для случаев создания гомогенного теста в математической теории педагогических измерений (IRT) принимаются два основных постулата: 1) При условии

правильной организации процесса тестирования, минимизации угадывания правильных ответов и предотвращения возможностей списывания вероятность правильного ответа на задание  $j$  зависит от уровня подготовленности испытуемого.

2) Зависимость вероятности правильного ответа на задание  $j$  от уровня подготовленности испытуемого можно выразить функцией (1).

Если функция (1) пригодна для описания эмпирических данных (что проверяется статистически), то возникают некоторые привлекательные свойства параметров: это, в частности, их сравнительная устойчивость в различных выборках испытуемых и заданий, а также относительная независимость одних параметров от других.

Задания, в которых довольно высока вероятность угадывания правильных ответов, для качественного педагогического тестирования не пригодны, слишком велики при этом ошибки измерения. Кроме того, при систематическом угадывании ответов сильно теряют моральная и воспитательная стороны тестового процесса. Так как принцип формирования знаний заменяется принципом угадывания правильных ответов, подлинное обучение заменяется изучением способов «сдачи» учебных предметов. В качестве конструктивного выхода из тупика, в котором заблудилась нынешняя практика тестирования и проведения ЕГЭ, автор предлагает переходить к использованию заданий с выбором нескольких правильных ответов из примерно 10–12 ответов. В таких заданиях вероятность угадать комбинацию правильных ответов оказывается менее одной тысячной<sup>22</sup>.

<sup>21</sup> Ахиллесова пята ЕГЭ — это безнадёжное нарушение четвёртой аксиомы ввиду ошибочной метрической конструкции госэкзамена, порождающей парадокс: чем больше так называемых КИМ отвечает требованию ЕГЭ, тем ниже точность измерений.

<sup>22</sup> **Аванесов В.С.** Форма тестовых заданий. М.: Центр тестирования, 2005; **Аванесов В.С.** Проблема формы тестовых заданий. В сб. Тестовые формы контроля по русскому языку как иностранному (РКИ). II Всероссийская научно-практическая конференция. Доклады и сообщения. Центр международного образования МГУ им. М.В. Ломоносова, 12–13 апр. 2005 г. / Под ред. Н.П. Андрюшиной, О.А. Усковой. М., 2005. С. 20–30; **Аванесов В.С.** Применение тестовых форм в Rasch Measurement // Педагогические измерения, 2005. № 4. С. 3–20; **Аванесов В.С., Брусенина Н.Д., Барер Г.М.** Новые формы научной организации самостоятельной работы студентов. <http://testolog.narod.ru>, 2005 г.

### Применение заданий в тестовой форме в учебном процессе

В современном образовательном процессе нет проблемы более важной и одновременно более сложной, чем организация самостоятельной работы студентов. Важность этой проблемы связана с новой ролью самостоятельной работы: она постепенно превращается в ведущую форму организации учебного процесса. В результате самообразовательной деятельности студентов идёт процесс приобретения, структурирования и закрепления знаний. Сейчас роль самостоятельной работы настолько возросла, что её приходится специально планировать, создавать для неё специальные формы и методы, выделять время, помещения и технические ресурсы.

Сложность проблемы заключается в необходимости оптимизации сочетания времени на лекционные занятия и на выполнение самостоятельной работы по различным дисциплинам. Сейчас это редко превышает соотношение 1:1, в то время как в Европейских странах Болонского процесса и в США отмечается устойчивая тенденция снижения общего времени на чтение лекций и повышения времени самостоятельной работы студентов (в примерном соотношении 1:3). Именно такое, трёхкратное превышение времени на самостоятельную работу студентов по сравнению с лекционной формой занятий считается наиболее эффективным для улучшения качества подготовки специалистов.

Для достижения отмеченного оптимума предстоит проделать невероятно большую работу по созданию подходящих форм и методов организации индивидуальной работы студентов. Предстоит научиться планировать самостоятельную работу, получать информацию о реальных затратах времени студентов на выполнение домашних работ по различным учебным дисциплинам, искать способы ликвидации нехватки аудиторного фонда, компьютерной техники, а также разрабатывать на кафедрах специальные задания нового поколения — такие задания, которые были бы интересны по содержанию и одновременно позволяли бы студентам работать самостоятельно. Не менее трудно обеспечить компьютерную поддержку автоматизированной оценки результатов самостоятельной работы студентов.

Полезно подчеркнуть, что повышение роли самостоятельной работы не снижает ценность лекционных форм работы со студентами. Интересно прочитанная лекция выполняет ориентирующую, организующую, мотивирующую, систематизирующую, познавательную и другие важные функции. В этом смысле ни одна другая форма не может успешно соперничать с лекцией, а потому роль интересных лекций никогда не снизится.

Правда, однако, такова, что не все вузовские лекторы и лекции интересны студентам. Ряд учебных материалов, которые необходимо запомнить, бесполезно давать в форме лекций, потому что такие материалы могут быть усвоены либо в процессе самостоятельной работы, либо при работе в группах или делении студентов на пары. Именно при работе в парах у студентов появляется возможность вербализации изучаемых понятий и формирования профессиональной речи. Вот почему появились такие учебные материалы, которые специально создаются для изучения в процессе самостоятельной работы: это материал, подлежащий заучиванию, апробации, вербализации, структурированию, закреплению и совершенствованию в процессе индивидуальной работы.

*Закономерности.* Можно сформулировать пять объективных педагогических закономерностей организации самостоятельной работы студентов:

- по мере продвижения студентов от курса к курсу значение и объёмы самостоятельной работы растут;
- творческая составляющая в самостоятельной работе студентов имеет тенденцию к увеличению на старших курсах;
- в процессе организации самостоятельной работы студентов возрастает потребность в тьюторской роли педагогов. Тьютор — это педагог, всегда готовый помочь каждому студенту его группы преодолеть индивидуальные затруднения при изучении учебных дисциплин;
- компьютерная поддержка организации самостоятельной работы студентов становится абсолютно необходимой — как для оперативной выдачи учебных материалов, так и для автоматизированного учёта учебных достижений студентов;
- возникла необходимость специального повышения педагогической квалификации

профессорско-преподавательского состава в вопросах научной организации самостоятельной работы студентов.

К новым формам научной организации самостоятельной работы студентов можно отнести применение систем так называемых цепных, текстовых и тематических заданий в тестовой форме, а также ситуаций с интегративным содержанием и с вариативными заданиями в тестовой форме. Перечисленные формы рассматриваются как феномен, требующий научно-педагогического исследования и специальной организации. По сути, можно говорить о новом направлении в создании технологичных форм самостоятельной работы студентов.

Помимо заданий с выбором нескольких правильных ответов в самостоятельной работе студентов желательно использовать новые, так называемые сдвоенные задания. Посмотрим примеры:

### 1. МЕТОД ГЛУБОКОГО ФТОРИРОВАНИЯ ПРЕДЛОЖЕН

- 1) П. Дауге,      3) А. Лимбергом,
- 2) П. Леусом,    4) А. Кнаппвостом.

### ЭТОТ МЕТОД ПРИМЕНЯЕТСЯ ПРИ ЛЕЧЕНИИ

- 1) кариеса,
- 2) пульпита,
- 3) гиперестезии,
- 4) эрозии эмали,
- 6) клиновидного дефекта,
- 5) периодонтита.

### 2. ДЛЯ УМЕНЬШЕНИЯ КРОВОТОЧИВОСТИ ДЁСЕН РЕКОМЕНДУЮТСЯ ПАСТЫ

- 1) детские,
- 2) отбеливающие,
- 3) гигиенические,
- 4) противокариозные,
- 5) противовоспалительные,
- 6) снижающие чувствительность зубов;

### СОДЕРЖАЩИЕ

- 1) карбамид,
- 2) нитрат калия,

- 3) хлоргексидин,
- 4) хлорид стронция,
- 5) фторид натрия,
- 6) солевые добавки,
- 7) монофторфосфат,
- 8) экстракты лекарственных растений.

За правильный ответ в каждом из сдвоенных заданий испытуемый получает по одному баллу. Таким образом, оценки за такие задания могут варьироваться от двух до нуля. Это обстоятельство благотворно отражается и на повышении точности измерения в связи с тем, что растёт дисперсия баллов в каждом отдельном задании, а вместе с тем и дисперсия тестовых баллов по тесту в целом.

□