

# МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ ИЗМЕРЕНИЯ ОБУЧЕННОСТИ

**Валентин Идиатулин,**  
ФГОУ ВПО Ижевская ГСХА;  
root @ isa.nivad.ru

Рассмотрены методологические основания измерения структуры обученности. Указано, что пропозициональная репрезентация учебного знания в полной мере соответствует форме тестовых заданий, что способствует обучению и контролю. Отмечено, что точность, надёжность и валидность тестовых результатов в наибольшей мере обеспечивают пригодность заданий в тестовой форме для дидактических исследований и управления обучением.

Показано, что для измерения обученности не характерны предположения о случайном компоненте тестовых баллов и их нормальные распределения. Параметры распределений тестовых баллов направленно меняются в процессе обучения. Оценивание обученности требует анализа как её индивидуальной структуры, так и содержания оцениваемых видов учебной деятельности.

***Ключевые слова:** обученность, методология, педагогические измерения.*

Цель настоящей работы — исследование методологических оснований измеримости свойств педагогических объектов.

Задачи — выделение свойств, доступных измерению, выявление их измеримых параметров, а также определение накладываемых на методы измерения требований для обоснования их выбора.

Предмет исследования — измерение обученности как результата обучения, выраженного в действиях обучаемого, т.е. совокупности знаний, умений и навыков по данной учебной дисциплине, и ее структурных составляющих — градаций объёмов знаний и уровней владения ими.

## **Введение**

Педагогические измерения стали частью методологии педагогики, обобщающей принципы и способы добывания знаний о педагогической действительности, методы их преобразования и интерпретации.

Наиболее технологичным средством измерений обученности служат тесты, способные выполнять все дидактические функции педагогического контроля и позволяющие конструировать системную технологию организации учебного процесса. Обучающая функция весьма полно реализуется при регулярном модульно-тематическом тестировании в процессе самостоятельной работы над заданиями теста при поиске правильного решения, она активизируется при последующем их обсуждении с участием преподавателя. Обучающая функция обеспечивается деятельностью студента, направленной на усвоение и приобретение знаний, совершенствование навыков и умений при выполнении тестовых заданий, а также последующем обсуждении ошибочных и правильных решений. Закрепляющая и обобщающая функции реализуются при предъявлении тестовых заданий в их логической последовательности в оригинальной форме представления учебного материала.

Развивающая функция реализуется при выполнении заданий, ориентированных на зону ближайшего развития обучаемых, а также заданий эвристического уровня, имеющих учебно-проблемный характер. Активизация обучения прояв-

ляется в соревновательном характере тестирования, в действенной мотивации, возникающей в условиях педагогического контроля. Она может быть усилена элементами дидактической игры в малой группе.

Систематически контролируемая учебная деятельность формирует ответственность, создаёт определённую мотивацию и выполняет воспитывающую функцию, особенно эффективную при соответствующем учебным целям педагогическом стимулировании. Воспитание ответственности обеспечивается регулярностью тестирования, объективностью его результатов, их гласностью. Постоянство и индивидуальность контроля не оставляют безразличным ни одного обучаемого, они становятся действенным психологическим фактором его совершенствования в учебной дисциплине.

При регулярном текущем тестировании выявляются не только достижения, но и конкретные пробелы в знаниях и умениях отдельных обучаемых, а также общие закономерности усвоения. По данным тестирования может быть проведена индивидуальная диагностика степени обученности каждого студента и построена дифференциальная карта структуры его обученности. Диагностирующая функция

наиболее полно реализуется при обработке совокупности индивидуальных тестовых баллов, что даёт возможность построения карты уровней обученности как для академической группы, так и для всего потока студентов. Анализ структуры обученности позволяет преподавателю оперативно проводить коррекцию учебного процесса, осуществлять эффективное замкнутое управление им в реальном времени, обеспечивая обратными связями организующую функцию тестирования. Тестовая диагностика является надёжным средством измерения обученности, изучения её динамики, исследования эффективности дидактических процессов или циклов.

Педагогические тесты стали важным фактором научно-педагогических исследований, объективной и надёжной мерой дидактического эксперимента, придающей учебному процессу исследовательский характер. Одна из важных особенностей педагогического измерения — активное воздействие на объект, меняющее состояние последнего, напоминает ситуацию в неклассической физике и наводит на предположения о стохастическом характере объекта измерения. Вероятностные модели получили весьма широкое распространение в педагогике.

## Измерение в метрологии и педагогике

Историческое развитие любой самой точной науки начиналось с качественных понятий и их субъективных оценок, затем устанавливались сравнительные степени и границы качеств, причём далеко не всегда степень качества могла быть выражена числом. До сих пор в ходу ранговая шкала Мооса твёрдости минералов, шкала Бофорта для оценки силы ветра, сейсмическая шкала Рихтера и др. Традиционная порядковая шкала академических оценок основана на опыте и интуиции преподавателя, арифметические операции с ними некорректны и имеют весьма условное значение. Аналогичные оценки в США просто обозначаются буквами латинского алфавита, поэтому и не делаются попытки определить их средние значения. Следует заметить, что отечественные матрикулы также содержат лишь качественные оценки, и их нумерация есть лишь дань сложившейся устойчивой практике, которая отражает общую потребность в оперировании числом для анализа результатов деятельности обучаемых, обучающихся, целых коллективов и образовательных учреждений.

Допустимые статистические операции ограничивают выбор типа шкалы. Полностью они выполнимы только в про-

порциональной шкале отношений, имеющей естественный нуль. Метрической является и шкала интервалов, позволяющая вычислять среднее значение (математическое ожидание признака) и стандартное отклонение от него. Для ранговой шкалы адекватна медиана — срединное значение упорядоченного признака, для которого вероятности больших и меньших значений одинаковы. Эквидистантность (равноудалённость) рангов практически не реализуема и только претендует на точность шкалы. Для номинальной шкалы справедлива мода — наиболее частое значение признака.

Над значениями рангов нельзя производить обычные арифметические операции, только с введением единицы измерения можно сводить отношения между величинами к отношениям между числами. Принципиальное значение имеет то, что точность измерений в ранговой шкале даже падает с ростом их числа, тогда как в интервальной и пропорциональной растёт. Точность наиболее высока в отсутствие случайных ошибок, отсутствие же систематических ошибок определяет правильность измерения. Постоянство статистических результатов характеризует их воспроизводимость и устойчивость при повторных измерениях, определяет этим их надёжность.

Она характеризует степень точности измерения, которая не может быть выше его стабильности, что и выявляет корреляция повторных измерений. Коэффициент надёжности мало информативен, во многих отношениях предпочтительнее стандартная ошибка с известной доверительной вероятностью или другая оценка погрешности. Измерение надёжно, если его повторение даёт тот же результат в пределах погрешности.

С методологической точки зрения фундаментальной характеристикой измерения является валидность получаемых результатов. Она отражает степень соответствия показателя свойству объекта измерения, т.е. измеряет ли он то, что должен измерять. Валидность не абсолютна, она зависит от вида критерия. В оценке обученности различают валидность содержания, определяемую по степени охвата учебного материала, функциональную валидность по степени владения этим материалом; валидность соответствия результатов, полученных разными методами, прогностическую валидность соответствия с последующими измерениями, конструктивную валидность модели объекта измерения и другие.

Ни валидность, ни надёжность измерения не достижимы, если не обеспечивается его объективность. В педагоги-

ке наблюдаемы не латентные процессы, а внешние признаки их проявления. Измерению доступны не всякие, а только конвергентные результаты, которые могут быть представлены в однозначно лучшем или наиболее правильном виде. Дивергентные достижения могут приводить, в принципе, к разным, но одинаково ценным результатам, диверсификация которых не имеет объективной меры. Креативная сфера неизмерима даже методами экспертной квалиметрии, когнитивная простирается от знания фактов до аргументированных оценок.

Для измерения, т.е. количественной оценки какого-либо свойства объекта существует проблема представления, то есть доказательства того, что для эмпирической системы с отношениями можно построить гомоморфную числовую систему<sup>1</sup>. Определение измерения как процедуры сравнения величины с некоторой однородной её единицей применимо только к экстенсивным величинам, которые отвечают принципу аддитивности, иначе используется существующая связь между измеряемой интенсивной и какой-либо экстенсивной величиной. В естествознании измеряемые свойства носят название величин — таковы, например, размеры, массы или заряды объектов; неизмеряемые вели-

чинами не являются, это может быть цвет объекта, его материал, форма и т.п.

Для педагогики характерны косвенные измерения, выводимые из ответов на задания, но не сводимые к ним. Возможность формализации педагогических объектов является необходимым условием их измеримости. Спорным кажется утверждение, что любое качество (свойство) личности может быть выражено в единицах той или иной меры. Обычно считаются неизмеряемыми престиж, репутация, надёжность, ответственность, добросовестность, настроение. Они не имеют выражения в каких-либо естественных единицах. Должна быть объективная уверенность, что степень проявления качества окажется упорядоченной в рамках определённой шкалы. При этом шкалы разных индивидуумов не всегда совпадают и не могут быть наложены одна на другую. Возникает необходимость в некоторой эталонной шкале, не зависящей от процедуры и субъектов измерения.

С энциклопедической точки зрения мерой знания является количество его элементов. Элементами научного знания могут быть идеи и формы их представления или их автономные фрагменты. В дидактике часто используется понятие учебных элементов, как пере-

1

*Михеев В.И.,  
Потапова А.Ю.*

Методы измерений в педагогических исследованиях // Профессиональное образование. 2002. №12.

несённых из науки в учебные программы объектов изучения. Их число и количество логических связей между ними заведомо ограничены объёмом учебного материала дисциплины и порогом насыщения у учащегося. Учебные элементы характеризуются степенью научной абстракции, уровнями и рядом качеств усвоения. Задачей контент-анализа ставится выделение некоторого набора достаточно постоянных и однородных элементов знания. Их перечень можно было бы составить, приписывая каждому элементу ранговый индекс значимости, определяемый, например, частотой встречаемости в литературе или употребления в учебной и научной деятельности. Такое упорядочение элементов знаний делает их, хотя бы в принципе, доступными для наблюдения. Однако одномерное усвоение знаний ограничено возможностями обучаемых и совершенно бессодержательно, поскольку даже если возможно вообще выделение каких-либо элементарных единиц знания, то они существуют лишь в конкретной связи с другими. Наиболее важной характеристикой знания является совокупность содержательных ассоциаций и комбинаций элементов разных порядков, а также глубина мыслительной деятельности по их усвоению и применению.

## Многомерность структуры обученности

Представляется недостаточно обоснованной тенденция одномерного шкалирования оценки знаний и умений, несмотря на выявляемую многомерность их структуры. Для измерения обученности была бы идеальной пропорциональная шкала, однако, ее достижение бывает затруднено. Различные свойства, характеристики, параметры знания несводимы друг к другу. Многомерность проявляется в анализе общей и индивидуальной культуры как совокупность параметров значимости, абстрактности, специфичности и связности элементов знания. Уровни усвоения учебных элементов также ориентированы на различные виды деятельности, каждый из которых объективно оценивается независимо от других. Введение весовых характеристик нередко произвольно и неоднозначно, оно часто вызвано необходимостью ранжирования обучаемых, что характерно для всех рейтинговых систем.

Элементы знания существуют в их конкретной связи с другими. Их количество экстенсивно характеризует широту эрудиции, а число и сила связей — глубину и продуктивность ума. Когнитивная сфера многомерна во многих отношениях: по классам знаний, по

языкам их выражения, по факторам общих способностей, по способам мышления, по семантическому дифференциалу, по формам представления информации, по её характеристикам, по структуре ценностей. В структуре обученности традиционно выделялась триада: знания, умения, навыки; эпистемология также всегда различала декларативное и процедурное знание. Обученность как результат обучения проявляется в деятельности обучаемого, которая может быть репродуктивной и продуктивной. Каждый вид разделяется на уровни, разные авторы детализируют их по-разному, но общие черты совпадают: репродуктивными уровнями являются фактуальный (декларативный) и операционный (процедурный), продуктивными — эвристический и творческий.

Эта таксономия когнитивных целей достаточно полна для диагностики структуры обученности с оговоркой, что творческий уровень в меньшей степени является результатом обучения, он имеет не деятельностную, а личностную природу, практически плохо измерим из-за отсутствия эталона объективно нового, его элементы могут включаться в эвристический уровень. Таким образом, в структуре обученности можно выделять уровень знания фактов, терминов, понятий, опреде-

лений, формул, формулировок, т. е. всего того, что можно выучить; уровень умений выполнять операции по освоенному правилу, алгоритму, образцу..., т.е. того, чему можно научиться; уровень эвристического анализа нестандартных ситуаций, для чего необходимо ещё и развитие определённых способностей, интуиции и логики. Структура обученности с необходимостью должна находить отражение в форме тестовых заданий, предназначенных для её измерения.

Богатый опыт тестирования, накопленный в психологии и социологии, и перенесённый в педагогику в качестве средства измерения обученности, равно как и экспертные методы, не всегда в нужной мере учитывают особенности дидактического объекта. Тестовый балл данного студента не является случайной величиной, на чём основываются многие теории тестирования, он не может быть охарактеризован ни устойчивой частотой, ни аксиоматической вероятностью. Неопределённой величиной является результат голосования эксперта, поскольку вероятность его априорно неизвестна, иначе эксперта можно было бы заменить, например, бросанием игральной кости. Обработка и интерпретация эмпирических результатов требуют известной осторожности.

Абсолютизация математических критериев часто при-

водит к игнорированию содержательных особенностей конкретных наук. Теория вероятностей применима только к статистически устойчивым событиям, частоты любых подпоследовательностей которых совпадают, либо к событиям, которые могут быть охарактеризованы аксиоматической вероятностью. Ни тестовый балл данного студента, ни оценка эксперта такими свойствами не обладают, их вероятности априорно не известны, они статистически неустойчивы, не наблюдаемы многократно в однородных условиях, поэтому являются величинами не случайными, а вообще говоря, неопределёнными, такого рода величины всегда сосуществуют с детерминированными и случайными. Задача математической статистики, заключающаяся в нахождении приближённых вероятностей каких-либо событий по эмпирическим данным, тоже разрешима только в условиях статистической устойчивости. Математический аппарат теории измерений основан на непосредственных независимых, как правило, равноточных измерениях одной величины. Экспертная оценка во многих случаях является единственно возможной, хотя и не работает на творческом и акмеологическом уровнях, а также в инновационных областях.

Обученность может быть соотнесена с результатами дру-

гих, которые определяют социальную норму, с прежними результатами обучаемого, что характеризует индивидуальную норму, либо с поставленными целями, которые представляют критериальную норму. Последняя наиболее отвечает целям и задачам систем образования и признается потребителями образовательных услуг. Тесты с критериальной интерпретацией результатов не предполагают высокой различающей способности, они не имеют целью ранжирование обучаемых по какому-либо нормативному признаку, а являются тестами учебных достижений каждого испытуемого<sup>2</sup>.

Статистически определяемые характеристики — средний тестовый балл, дисперсия и среднеквадратичное отклонение, асимметрия и эксцесс — не дают оценок надёжности и точности отдельного измерения, критерий выполнения теста может быть весьма жёстким. Критериальный тестовый балл должен быть достаточно далёк от математического ожидания числа правильных ответов при их случайном выборе и находиться за пределами двух-трёх стандартных отклонений. Чтобы избежать элементов угадывания ответов, можно требовать краткого обоснования выбора наиболее правильного из них, т.е. формулировку, расчёт, объяснение и т.п. При этом тестовое задание превращается в

## 2

*Идиатулин В. С.*

Опыт конструирования и эмпирической проверки тестов уровней обученности по физике // Вопросы тестирования в образовании. 2003. №5. С. 71–88.

*Идиатулин В. С.* Тесты уровней обученности как средство диагностики учебных достижений // Вопросы тестирования в образовании. 2002. №4. С. 18–30.

истинное или ложное суждение с вариантом доказательства этого. Такая форма сохраняет все преимущества закрытого типа, включая известные оперативность и объективность проверки, но также предоставляет дидактически мотивированную возможность коммуникативного взаимодействия с обучаемым.

### **Измеримые параметры обученности**

Классическая теория тестов в большей степени относится к тестам, ориентированным на соотношение с социальной нормой; их широкому внедрению способствовали работы В.С. Аванесова<sup>3</sup>. Потребность в определении степени достижения учебных целей, отвечающей некоторому установленному критерию, вызвала к жизни распространение тестов учебных достижений, валидность результатов которых с самого начала находится на переднем плане<sup>4</sup>. При выборе таких тестов в качестве средства диагностики обученности К. Ингенкамп особо отмечает их учебное единство, использование как разговорного, так и формального языков с выбором ответов. Конструирование тестовых заданий должно в наиболее полной мере соответствовать учебной цели, определение этого возможно экспертным путём,

хотя оценка сложности заданий экспертами достоверна только при выполнении редко реализуемых условий. Для тестов уровней обученности формулы классической теории не работают, распределение ответов не может предполагаться нормальным, определение его вида и параметров становится самостоятельной задачей эмпирического исследования. Однако часто математические модели опираются на предположение о нормальном распределении эмпирических данных, хотя исследования, как автора настоящей работы, так и многих других, показывают, что нормальное распределение параметров педагогических объектов наблюдается обычно в тех случаях, когда предпринимаются специальные меры по их нормализации.

Основная дидактическая цель использования тестов с критериально-ориентированной интерпретацией тестовых результатов заключается в оптимизации учебного процесса по выбранным параметрам, при этом достаточно бывает критериев, установленных преподавателем. В более широких масштабах потребуются соглашения или экспертная оценка критериев. Дидактические цели могут быть сформулированы практически весьма точно, для контроля их достижимости как раз и нужна диагностика, которая даёт информацию не только

**3**

*Аванесов В.С.*  
Научные основы тестового контроля знаний. М.: Исследовательский центр, 1995. 136 с.

*Аванесов В.С.*  
Композиция тестовых заданий. Учебная книга. М.: Центр тестирования, 2002. 240 с.

**4**

*Ингенкамп К.*  
Педагогическая диагностика /Пер. с нем. М.: Педагогика, 1991. 240 с.

о конечных результатах, но и о приводящих к ним конкретных дидактических процессах, в том числе и ситуативно обусловленных, что позволит сразу выделять педагогические находки в обеспечении условий эффективного обучения, использовать полученную информацию в педагогическом поиске<sup>5</sup>.

К преимуществам тестовых измерений справедливо относят высокую объективность контроля, его независимость от взаимоотношений преподавателя и студента. Точность измерений обеспечивает большую обоснованность и дифференцированность оценки, а скорость обработки данных позволяет работать с большим контингентом обучаемых. Тестовому измерению в большей мере доступна проверка степени усвоения просто организованного материала, чем глубинного понимания, поэтому многие считают, что тесты оценивают не развитие творческих способностей, а лишь умение оперировать знаниями, тем не менее тестирование даёт возможность выявить талантливых и способных, делая это, конечно же, особым справедливым и весьма объективным способом.

По форме тест представляет собой систему достаточно кратких заданий, объединённых общей логикой построения. Область содержания теста охватывает по возможности на-

иболее полно тот объём знаний, умений и навыков, которыми должен овладеть обучаемый в учебном процессе. При преобразовании в тестовые задания учебные цели необходимо конкретизируются и операционализируются. Учебная дисциплина состоит обычно из разделов, в которых выделены темы, состоящие из ограниченного числа учебных элементов, которые сохраняют самостоятельный смысл, а их дальнейшее расчленение ведёт к утрате включённой в них учебной и научной информации. Смысловые учебные элементы должны быть представлены в виде некоторых утверждений, отражающих положительное содержание учебной дисциплины. Логическая структура теста как раз и использует эту форму их представления, которая в зависимости от выбора испытуемого превращается в истинное или ложное высказывание, кодируемое единицей или нулём. Интеррогативная форма задания исключает эту возможность, так как вопросы не могут быть истинными или ложными, а ответы на них, как правило, многовариантны, неопределённые, многословны и, в конечном счете, нетехнологичны. В вопросе важную роль играют интонация, логические ударения, расстановку и акценты на которых непросто выразить однозначно.

Знания как совокупность усвоенных элементов опыта человека и общества, проверенные общественной практикой, воспроизводятся обучаемым в речи и деятельности. Наименьшая значимая единица знания, выделяемая в отдельное утверждение, элемент семантической памяти, связывающий идеи и понятия, в когнитивной психологии называется пропозицией<sup>6</sup>. Современная психолингвистика<sup>7</sup> подтверждает пропозициональный характер доречевых когнитивных референтов, близость к ним по структуре фраз и предложений облегчает оперирование последними. Любая осмысленная информация, в конце концов, перекодируется в вербальную форму, которая является основной в обучении. Понимание также основано на пропозициях, упорядоченных в когнитивных структурах — понятно уже выраженное знакомым образом. По этой причине представляется важным на этапе закрепления переводить учебный материал в систему пропозиций, тематически и логически связанных друг с другом.

Трактовка обученности как результата обучения, выраженного в действиях обучаемого, снимает многие проблемы, хотя и не все и не всегда. Поэтому столь стандартизованные требования предъявляются не только к форме задания, но и к

самой процедуре тестирования. Спецификация теста устанавливает соответствие области содержания теста и тестовых заданий предъявляемым требованиям и предметной области учебной дисциплины. Они содержат общее описание того, что измеряет тест, включает в себя инструкцию по проведению тестирования, характерные формы и содержание заданий, правила представления ответов на задания, критерии их оценки и способ оценивания.

Таким образом, надо решить, какие знания, умения, представления, навыки должны быть и могут быть измерены. Для целей измерения в структуре обученности по каждой дисциплине возможно выделение минимального объема мировоззренческих знаний, относящихся к данному предмету, которые должны оставаться в памяти любого обучаемого, к ним примыкают базовые знания, которые необходимы для усвоения учебного материала самой изучаемой дисциплины. Следующая категория включает остальные программные знания, которые нужны в первую очередь для обеспечения готовности к усвоению дисциплин учебного плана и, в определённой мере, к профессиональной деятельности. Эти три градации могут быть дополнены блоком специальных сверхпрограммных знаний, которые либо рас-

6

*Солсо Р.*  
Когнитивная психология. М.: Тривола, 1996. 600 с.

7

*Леонтьев А.А.*  
Основы психолингвистики. М.: Смысл, 1997. 287 с.

считаны на индивидуальные потребности обучаемых<sup>8</sup>, либо отражают требования специальности, вуза, региона.

Достаточно надёжно измерим фактуальный уровень владения знаниями, т.е. умение узнавать, идентифицировать, воспроизводить основные факты, сведения, термины, определения, формулировки, формулы, принципы, законы, положения изучаемой дисциплины, т.е. всё то, что можно узнать, усвоить, выучить и потом воспроизвести по памяти или выделить среди подобного. Форма тестовых заданий с выбором правильного ответа на этом уровне наиболее уместна, т.к. оценка выбора может быть сделана вполне однозначно, а открытая форма ответа бывает трудно верифицируема. Воспроизведение и узнавание как мнемонические процессы не имеют существенных различий, и нет особых оснований полагать, что задания с выбором ответа измеряют нечто иное, чем задания со свободным ответом<sup>9</sup>.

Операционно-алгоритмический уровень владения изученным материалом измерим столь же уверенно и надёжно. Он предполагает выполнение действий и операций по освоенному образцу, правилу, рецепту, алгоритму, т.е. умение делать всё то, чему можно научиться. Подразумевается, что действия и операции настолько

однозначно выполнимы, что для оценки их правильности достаточно сравнения полученного результата с эталоном ответа, причём валидность задания обеспечивается исключением операций, которые не являются результатом обучения по данной дисциплине. На этом уровне одинаково применимы задания как открытой, так и закрытой формы, однако последняя уместней в дидактическом и методологическом отношениях. Первое обеспечивается разработкой дистракторов, мотивированных типичными ошибками обучаемых, для последующего их анализа и исправления; открытая форма в этом случае не даёт информации. Методологически существенно то, что закрытая форма имеет естественную меру вероятности случайного выбора, с которой возможно сравнение эмпирических частот.

Умение проводить эвристический анализ и строить процедуры операций без известного однозначно алгоритма, но с установленным эталоном решения, также доступно тестовому измерению. Объективно тесты могут измерять те же признаки, что и традиционные формы проверки, и не могут то, что не могут они. Эвристический анализ опирается во многом на догадку, развитую интуицию, способность к трансформации усвоенного на ещё не из-

8

*Родионов Б.У.,  
Татур А.О.*  
Стандарты и тесты в образовании. М.: МИФИ (ТУ), 1995. 48 с.

9

*Ингенкамп К.*  
Педагогическая диагностика /Пер. с нем. М.: Педагогика, 1991. 240 с.

веданные обучаемым области. Этот уровень усвоения включает всё то, что нельзя просто выучить, чему нельзя и научиться, не обладая определёнными задатками, которые обуславливают развитие способностей, позволяющих успешно овладевать такой деятельностью; от творческого уровня он отличается, пожалуй, лишь тем, что результат деятельности только субъективно неизвестен, хотя для обучаемого он является открытием, пусть даже небольшого масштаба. Закрытая форма заданий эвристического уровня чётко выделяет эталон правильного ответа, но налагает особые требования на искусство составления дистракторов. Они должны включать не неправильные ответы, а, например, ответы, которые соответствуют другим условиям, что и потребует их эвристического анализа.

Метрологический принцип сопоставимости результатов измерения однородных величин и независимости их от объекта и субъекта измерений требует обеспечения единства точности, методики и средства измерения. По этой причине нежелательно сочетание в одном тесте разных форм тестовых заданий, а тем более нетестовых и псевдотестовых. Они ориентированы на разные виды умственной деятельности, меры которых трудно сопоставимы.

## Тестовые задания как инновационный фактор обучения

Потребность в объективной информации об учебных достижениях вызвала резкий рост масштабов использования тестов. Тестовая форма способна проверить усвоение не только фрагментов знаний, которые входят в содержание экзаменационного билета, но и, в определённой мере, весь подлежащий проверке учебный материал на требуемых уровнях владения им. Как элемент педагогической практики педагогический тест ещё не завершил стадию полного признания педагогической общественностью, иногда причина этому кроется в тех случаях, когда его использование ни методологически, ни психологически, ни методически несостоятельно. Так, зачастую не воспринимается тестирование по литературе, поскольку считается, что знать кое-какие факты и разбираться в литературе — не одно и то же, тестам не донести слово писателя, не приохотить к чтению, а скорее заставить забыть про художественный вкус.

Именно так воспринимается тестирование Л. Айзерманом в статье «Преданная литература», которая опубликована в журнале «Континент» в 2000 году (№ 104, с. 314–335). Неприятие вызывает как раз распространение тестов на облас-

ти, недоступные им, либо их неудачное составление и применение. Тесты уровней обученности должны измерять качество усвоения учебного материала данной дисциплины: объём знаний и уровни владения ими, т.е. то, что охватывается понятием «обученность». Результаты валидны, если дадут низкий балл недоучившемуся и высокий — обученному; они надёжны, если дают один и тот же балл одинаково знающим. Способствует этому унификация тестов по форме, избегание ухищрений, которые больше выявляют интеллект и способности, нежели знания и умения.

Тесты уровней обученности не предназначены для измерения трудоспособности, профпригодности, интеллекта, прилежания или личностных характеристик. Измерение последних в массовом порядке во многих странах запрещено, оно недопустимо и опасно без учёта психологического состояния испытуемых. Объективная мера обученности тоже не всегда воспринимается психологически комфортно, поскольку не оставляет места на скидки из-за предположения возможности необъективного оценивания. Думается, что каждый вправе знать об уровне своих знаний, но это необязательно должно становиться достоянием других. Этика и право также не препятствуют определению

когнитивных возможностей и обученности студентов.

Вопрос о том, измеряют ли тесты интеллектуальные способности, давно стоял перед психологами, причём довольно надёжно установлено, что интеллектуальные тесты не фиксируют каких-либо процессов, а только сравнивают их результаты. Задания, репрезентативные для одних социальных или этнических групп, не всегда были пригодны для других. Даже сама степень знакомства с процедурой тестирования может породить различия в тестовых баллах, не связанные ни со способностями, ни с уровнем обученности. Решение задач, выполнение тестовых заданий всегда подключают весь арсенал когнитивных процессов, таких как восприятие, память, распознавание, мышление; причём практически невозможно таким путём выделить отдельную способность и приходится анализировать их в совокупности. Определений интеллекта так много, что это свойство индивида едва ли является измеримым. Даже такие психологические способности, как память, невозможно рассматривать как отдельные свойства, которые можно измерять<sup>10</sup>, поскольку в познавательной деятельности имеем дело не со свойствами, а с процессами, зависящими от конкретной ситуации, мотивации, установки и других факторов.

Тестирование является довольно сложной технологической процедурой, разработка, проверка и применение тестов требуют высокой квалификации и опыта работы. Возможности тестов заведомо ограничены степенью формализуемости учебного материала. Как и любой другой метод контроля, тест проверяет в большей степени конечный результат, а не способы его получения. Ограничение времени тестирования и возможность свободно распределять его между отдельными заданиями аналогичны подчинению правилам шахматных соревнований, когда глубокое погружение в размышления над каким-то одним заданием (ходом) приводит к цейтноту, и только накопленный опыт позволяет вырабатывать правильную стратегию в распределении ресурса времени.

Но такие же проблемы встают при выполнении любой контрольной работы. Тест не оценивает навыки связного письменного изложения — для этого нужны другие средства, зато его логическая форма способствует их формированию. Тест не оценивает коммуникативные способности, однако его результат оказывается всегда мотивированным поводом для проведения собеседования со студентом. Критика как тестов, так и любых других методов измерений лишь указывает на

правильный путь — совместить и адекватно использовать достоинства разных способов и со знанием дела минимизировать их недостатки.

К подобному выводу приходит и М. Панин<sup>41</sup>, однако он видит выход в накоплении условных баллов за каждый вид учебной деятельности, включая посещение лекций, практических и лабораторных занятий, чтение текстов, аннотирование и конспектирование, контрольные работы, т.е. в соединении меры количества с оценкой качества. Гораздо продуктивнее разделить эти два понятия и оценивать только качество выполненных работ, которые предусмотрены учебным планом, причём привычной педагогической оценкой. Их совокупность становится intersубъективной, когда разные виды работ проводят и оценивают разные преподаватели. Она становится объективизированной, когда сопровождается тестовым мониторингом тех уровней усвоения, доступных измерению, и сопоставлением результатов.

Результат тестирования должен давать возможность объективной диагностики качества усвоения и оказывать стимулирующее влияние на действия субъектов обучения, что достижимо при опоре на индивидуальные диагностические данные. Тестовый балл является объективной мерой обученнос-

11

Панин М.  
Морфология рейтинга //  
Высшее образование в  
России. 1998.  
№1. С. 90–94.

ти, но оценка этой меры остаётся субъективной. Оценивание уровня обученности по В.П. Беспалько<sup>12</sup> связывается с коэффициентом усвоения определённого уровня с жёсткой их иерархией и автономным оцениванием. Б.У. Родионов рекомендует оценивать удовлетворительно усвоение базового объема на фактуальном и операционном уровнях, а отлично — то, что превышает либо программный объём, либо эвристический уровень усвоения, поскольку знания даются легче, чем умения.

Исследования автора показывают, что априорное установление соответствия тестового балла какой-либо академической оценке практически невозможно без изучения структуры обученности, её динамики, соотношения иерархии её уровней с состоянием контингента обучаемых. Для полноты изучения необходима технология обучения, основанная на когнитивной квалиметрии обученности, исследовательский характер которой позволяет разрешать эти вопросы.

Проблема угадывания всегда служила основой критики заданий с выбором одного правильного ответа. Из наших результатов<sup>13</sup> следует, что угадывание отсутствует у студентов: в области вероятного угадывания (для пяти вариантов ответов она лежит вблизи 20%

степени усвоения) эмпирические распределения имеют минимум частоты. Если бы угадывание имело место для всех, то здесь был бы максимум, если бы для некоторой части, то локальный максимум, который при большом размахе и СКО распределения мог бы перекрываться последним, — как раз это отмечалось при входном контроле слушателей курсов довузовской подготовки.

Без анализа эмпирических распределений обученности не следует априорно вводить коррекцию на угадывание, там, где оно отсутствует. Упорный труд над заданиями, поиск правильного ответа — всё это требует напряжения ума, работы мысли и следует отделять их от бездумного гадания. Смешивание этих двух процессов как раз и служит почвой для сторонних критиков, которые не давали себе труда вникать в существо тестирования как умственной деятельности, происходящей на высоком для тестируемого уровне.

Психология и педагогика не связывают повышение качества обучения с регулярным контролем, он лишь заставляет работать систематически; более того, считается, что быстрая оценка не способствует обучению, так как в условиях неопределённости обучаемый продолжает работать над ситуацией. Это, может быть,

## 12

*Беспалько В.П.,  
Татур Ю.Г.*

Системно-методическое обеспечение учебно-воспитательного процесса подготовки специалистов: Учебно-методическое пособие. М.: Высшая школа, 1989. 144 с.

## 13

*Идиатулин В.С.*

Квалиметрия уровней обученности // Стандарты и качество. 1999. №1. С. 80–82.;  
*Идиатулин В.С.* Когнитивная технология обучения студентов. Ижевск: Шеп, 2002. 180 с.

верно в отношении контроля, сопровождающегося выставлением оценки. Диагностика обученности отделена от её оценивания, она указывает на конкретные пробелы, над устранением которых надо работать. Неопределённость ситуации поиска решения может сохраняться достаточно долго, а указание на то, что уже найденный ответ неверен, лишь подталкивает к продолжению работы, а не завершает её, как это делает оценка. Диагностика не будет отвлекать от учебной деятельности, если тестовые задания прямо включены в её содержание. Учебный процесс подчиняется объективным законам, и теория даёт решение многих дидактических задач, однако многофакторный характер обучения, зависимость его от конкретной ситуации и компенсационные возможности дидактических систем всякий раз заставляют искать пути разрешения проблем, в том числе и не решаемых общей теорией.

## **Заключение**

Широкое распространение тестовых форм контроля знаний и внедрение тестовых технологий в педагогическую практику требуют упорядочения сложившейся в тестологии и педагогическом тестировании терминологии. Измерение обученности и выявление действующих на

учебный процесс факторов отражают объективную потребность системы образования и превращают его в деятельность, увлекающую педагога-исследователя. Вместе с тем сложность педагогических объектов вызывает необходимость методологического анализа оснований измерительных процедур и интерпретации результатов измерения. Многомерность когнитивной сферы заставляет выделять в структуре обученности адекватные целям обучения несводимые друг к другу измеримые параметры.

Результаты измерения не могут априорно считаться содержащими случайный компонент, его наличие необходимо проявляется в распределениях критериальных баллов степени обученности. Пропозициональная репрезентация фактуального учебного материала должна использоваться в утвердительных выражениях тестовых заданий для вербализации когнитивных референтов латентного знания. Форма заданий с выбором правильных ответов далеко не исчерпала своих возможностей хотя бы потому, что исследованиями в когнитивной психологии установлено, что объём хранения полученной информации в три-четыре раза превышает объём её воспроизведения, по которому обычно судят о её усвоении<sup>14</sup>.