

УДК 37:001.12/.18

С.П. Мирошниченко**ПРОЦЕССЫ ТЕСТИРОВАНИЯ И РАЗВИТИЯ В ОБРАЗОВАНИИ**

При проведении и обработке результатов Российского государственного педагогического тестирования используются модели теории Item Response Theory (IRT), которая в переводе на русский язык получила название «Теория моделирования и параметризации педагогических тестов» (ТМППТ или ТППТ). Основоположающим в этой теории является предположение о том, что учебные достижения учащегося в определенной области знаний можно оценивать с помощью объективного скрытого (латентного) параметра, называемого уровнем подготовленности. В свою очередь, каждое задание, решаемое учащимся, также характеризуется скрытым объективным параметром, называемым уровнем трудности задания. Скрытые параметры учащихся и заданий не поддаются непосредственному измерению, однако их значения могут быть оценены по результатам решения теста большим количеством учащихся с использованием методов статистического анализа. Для вычисления статистических оценок скрытых параметров заданий и учащихся необходимо задаться математической моделью зависимости вероятности решения конкретного задания конкретным участником (вероятности успеха) от уровней подготовленности участника и трудности заданий (функцией успеха). В качестве функций успеха чаще всего используют однопараметрическую модель Раша и двухпараметрическую модель Бирнбаума. Однопараметрическая модель Раша предполагает, что вероятность успеха однозначно определяется одним параметром $x = \theta - \delta$ – разностью между уровнями подготовленности учащегося θ и трудности задания δ . Параметры θ и δ измеряются на интервальной логистической шкале с областью определения $(-\infty; +\infty)$, которая в последствии преобразуется в 100-балльную шкалу тестовых баллов. Модель Раша является упрощенной моделью и ее основным достоинством является простота ее применения.

Более сложной моделью является двухпараметрическая модель Бирнбаума, в которую по сравнению с моделью Раша добавлен еще один параметр – коэффициент дискриминации задания d , характеризующий способность задания измерять уровень подготовленности испытуемого, т.е. другими словами, параметр качества задания. При использовании модели Бирнбаума первичные баллы уже не являются достаточными статистиками, т.е. испытуемые, решившие одинаковое число заданий, получают разные тестовые баллы. Тестовый балл определяется не только количеством правильно решенных заданий, но и тем, какие именно задания решил испытуемый. При этом достаточными статистиками при определении тестовых баллов являются суммы коэффициентов дискриминации правильно решенных заданий. Таким образом, тестовый балл испытуемого определяется первичным баллом, взвешенным по качеству решенных заданий, определяемому их коэффициентами дискриминации. Хорошие задания вносят существенный вклад в тестовый балл, а плохие, некорректные и плохо сформулированные задания влияют на него слабо. Конечно, это шаг вперед по сравнению с моделью Раша, однако трудность решенных заданий в рамках модели Бирнбаума не влияет на тестовый балл.

С нашей точки зрения, модели Раша и Бирнбаума имеют следующие недостатки:

- ◆ трудность решенных тестовых заданий не влияет на тестовый балл;

- ◆ не учтены такие важные факторы, как внимательность испытуемых и неравномерность их подготовленности по различным разделам курса.

Эти недостатки могут быть устранены путем разработки новых (или модернизации существующих) моделей функции успеха, учитывающих кроме уровня подготовленности испытуемых и другие их индивидуальные качества. В частности, параметр d двухпараметрической модели Бирнбаума может быть интерпретирован не как параметр задания, а как параметр участника тестирования. При этом параметр d характеризует однородность подготовки учащегося. Величина d (коэффициент дискриминации учащегося), измеряет способность учащегося отличать трудные задания от легких. Если считать уровень подготовленности учащегося при решении конкретных заданий случайной величиной, то величины θ и d^{-1} могут интерпретироваться как математическое ожидание и среднеквадратичное отклонение этой случайной величины. Поскольку число независимых факторов, влияющих на результат выполнения каждого задания, достаточно велико, можно предположить, что выполняются условия центральной предельной теоремы, вследствие чего уровень подготовленности учащегося имеет нормальное распределение вероятностей. Функция успеха описывается интегральной функцией нормального распределения, хорошей аппроксимацией которой является модель Бирнбаума.

Если наряду с коэффициентом дискриминации учащегося в модели учитывать коэффициенты дискриминации заданий, то коэффициент d в модели Бирнбаума приобретает смысл обобщенного коэффициента, являющегося функцией коэффициента дискриминации учащегося d_y и коэффициента дискриминации задания d_s . При этом разность между уровнем подготовленности учащегося и уровнем трудности задания будет случайной величиной, закон распределения которого является композицией законов распределения уровней подготовленности учащегося и трудности задания. Дисперсия этого распределения является суммой дисперсий названных исходных распределений. С учетом этого, обобщенный параметр d может быть определен из выражения $d = (d_y^{-2} + d_s^{-2})^{-1/2}$. Введение дополнительного параметра, характеризующего учащегося, позволяет учитывать при выставлении баллов не только количество, но и трудность правильно выполненных заданий. В частности, из двух учащихся, набравших одинаковые высокие первичные баллы, больший тестовый балл получит учащийся, правильно решивший большее количество трудных заданий. В то же время, среди учащихся, набравших невысокие первичные баллы, наибольший тестовый балл получит учащийся, решивший большее количество простых заданий. Логика такого определения баллов заключается в том, что учащиеся, набравшие высокие первичные баллы, не решили простые задания, скорее всего, случайно. В то же время, среди учащихся, набравших низкие первичные баллы, решение трудных заданий чаще всего будет случайным, а не решение достаточно простых – закономерным.

Оценка уровня подготовленности, используемая в настоящее время, измеряет уровень подготовленности на уровне вероятности успеха $p = 0,5$. Однако, если в дальнейшей профессиональной деятельности учащийся будет правильно решать только половину поставленных задач, то это явно не может быть признано удовлетворительным. В связи с этим вызывает интерес измерение подготовленности на других, более высоких уровнях вероятности успеха, например, на уровне $p = 0,9$ – это будет оценка твердых, устойчивых, хорошо воспроизводимых знаний, позволяющих выполнять задания практически безошибочно, другими словами – оценка умений. Для приведения оценок знаний и умений к единой шкале будем считать, что оценки знаний учащегося, имеющего единичный коэффициент дискримина-

ции, совпадают на всех уровнях. Тогда, оценка знаний в логитах θ_p на уровне вероятности p будет определяться из выражения

$$\theta_p = \theta + \left(1 - \frac{1}{d}\right) \ln \frac{p}{1-p}.$$

В частности, на уровне вероятности $p = 0,9$, учащийся, имеющий коэффициент дискриминации, $d = 0,7$ при уровне знаний $\theta = 2$ логита будет иметь оценку умений, равную 1,058 логит, а учащийся с тем же уровнем знаний и коэффициентом дискриминации 1,5, будет иметь оценку уровня умений 2,732 логит.

Измерение структуры подготовленности по индивидуальной характеристической функции участника тестирования. В 2001 году результаты Всероссийского государственного централизованного тестирования были впервые обработаны с использованием модели Бирнбаума. Считалось, что индивидуальные характеристические функции всех участников имеют одинаковую форму и отличаются друг от друга лишь сдвигом вдоль оси логит. Однако, форма индивидуальных характеристических функций, в зависимости от структуры подготовки, также может быть различной. Информация, необходимая для построения эмпирической индивидуальной характеристической функции участника, содержится в матрице ответов участника на задания с известной трудностью и дифференцирующей способностью. Исследуя форму индивидуальной характеристической функции участника. М.Б. Челышкова [1] предлагает в зависимости от вероятности успеха разбить диапазон трудностей заданий, решаемых участником, на ряд областей. В работе делается вывод о том, что существование широкой зоны ближайшего развития у учащегося связано с наличием большого числа пробелов в структуре знаний и свидетельствует о его низкой обучаемости. Также отмечается, что ширина зоны ближайшего развития (a , следовательно, и величина коэффициента d) не определяет однозначно качество обучения и в зависимости от подготовленности может быть рассмотрено и как позитивное явление, свидетельствующее о высоких потенциальных возможностях обучаемого, и как негативное, говорящее о пробелах в знаниях и низкой динамике обучения. Отсюда можно сделать вывод, что модель Бирнбаума и вычисленный на ее основании коэффициент d не могут использоваться для однозначной числовой оценки качества подготовленности участника.

На наш взгляд, противоречивость выводов о роли зоны ближайшего развития в процессе обучения объясняется следующими причинами. Измерение подготовленности на уровне вероятности успеха $p = 0,5$, принятое в ТМППТ, является психологически непривычным по сравнению с классическим измерением по пятибалльной или десятибалльной шкалам, где оценка «отлично» соответствует знаниям на уровне вероятности успеха $p = 0,9 - 0,95$. В качестве примера традиционного оценивания знаний по четырёхбалльной шкале рассмотрим стандарт Таганрогского государственного радиотехнического университета [2], в котором регламентируется выставление оценок по результатам итогового рейтингового контроля знаний: 85% – 100% – отлично; 70% – 84% – хорошо; 55% – 64% – удовлетворительно; 0% – 54% – неудовлетворительно. Оценке "отлично" соответствуют в среднем знания на уровне $p = 0,925$; а значение $p = 0,5$, используемое для измерения в ТМППТ, соответствует оценке "неудовлетворительно". Второй причиной вышеназванных противоречий является несовершенство модели Бирнбаума. Вывод о негативном влиянии широкой зоны ближайшего развития является следствием симметричности графика характеристической функции участника в модели Бирнбаума относительно точки перегиба. На самом деле ширина различных зон может не зависеть (или слабо зависеть) друг от друга, при этом форма индивидуальной харак-

теристической функции участника может быть более сложной, в частности, - асимметричной и для ее описания требуются более сложная модель.

Приведенный анализ показывает, что для получения возможности измерения структуры подготовленности (структуры знаний) участников тестирования должны быть решены следующие задачи:

- ◆ Разработка методики измерения структуры подготовленности на основе информации, содержащейся в индивидуальной характеристической функции участника.
- ◆ Построение математической модели, объединяющей характеристические функции заданий и участников и позволяющей описывать характеристические функции участников различной, в том числе и асимметричной формы.
- ◆ Разработка методики построения и сглаживания индивидуальных характеристических функций участников тестирования.

Данные задачи успешно решены в работе [3].

Функция структуры знаний. Индивидуальная характеристическая функция участника (ИХФУ) тестирования несёт в себе информацию о его структуре знаний. Оцениваемая в ТМППТ подготовленность участника θ представляет собой оценку математического ожидания подготовленности, ничего не говорящую о структуре знаний. Информация о структуре знаний участника может быть получена в результате измерения значений индивидуальной характеристической функции на различных уровнях вероятности успеха. В работе [4] ИХФУ разбивается на 5 интервалов (навыки, умения, недавно полученные знания, получаемые знания, перспективные знания). Использование характеристических функций участников тестирования, различных по форме, требует построение новых моделей. В результате рассмотрения механизма формирования характеристических функций может быть предложена трехпараметрическая модель тестирования, в которую наряду с дифференцирующей способностью задания входит параметр, характеризующий дифференцирующую способность участника тестирования и определяющий крутизну наклона характеристической кривой. В отличие от однопараметрической и двухпараметрической моделей, предложенная трехпараметрическая модель дает возможность одновременного учета различных дифференцирующих способностей разных заданий и участников тестирования. Использование трехпараметрической модели позволяет оценивать не только уровень подготовленности участника тестирования, но и структуру его знаний.

Проведенные исследования показывают возможность разработки технологии второй оценки при тестировании. Характер (вид) зависимости ИХФУ не только однозначно оценивает структуру знаний, но и показывает потенциальные возможности будущего студента. Кроме того, рассматривая ИХФУ как кривую развития и проводя тестирование в течение определенного времени (через месяц, семестр и т.д.), можно построить кривую развития студента (школьника), т.е. контролировать процесс развития учащегося. Данная технология представляет интерес не только для педагогики, но и психологии.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Чельшкова М.Б.* Адаптивное тестирование в образовании (теория, методология, технология). – М.: ИЦКПС, 2001. – 165 с.
2. Основные положения по организации и проведению рейтинг контроля в системе РИТМ для студентов дневной формы обучения. Стандарт Таганрогского государственного радиотехнического университета. – Таганрог, 1994.

3. Глушенко А.А., Иванцов В.В., Клово А.Г., Радомская М.В. Измерение и оценка качества образования. – М.: Изд-во «Международная Педагогическая Академия», 2003. – 237 с.
4. Клопченко В.С., Мирошниченко С.П., Таран В.А. Методология и теория процессов развития (часть II). – М.: Изд-во «МПА-ПРЕСС», 2005. – 194 с.

УДК 681.3.066:378.1

В.Н. Василевская

ИСПОЛЬЗОВАНИЕ МУЛЬТИМЕДИА КАК ОДНОГО ИЗ СРЕДСТВ ИНТЕНСИФИКАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ПРИ ИЗУЧЕНИИ ИНОСТРАННОГО ЯЗЫКА

Компьютеризация образовательного процесса получила распространение и применение во всем мире и успешно практикуется во всех ведущих ВУЗах. Все большее внимание уделяется изучению и использованию новейших компьютерных разработок для изучения различных учебных дисциплин, в том числе и иностранного языка. Современное компьютерное обеспечение расширяет образовательные возможности, как студентов, так и преподавателей. Но, в то же время, поток информации настолько огромен, что необходимо умело в нем лавировать и извлекать необходимые сведения, потратив на поиск минимум времени и усилий.

Из года в год поток информационных материалов растет, что ведет к увеличению учебной нагрузки студентов. С другой же стороны, количество часов, выделяемых на аудиторные занятия, заметно уменьшается. Эти факторы определяют повышенное внимание к требованиям, предъявляемым к самостоятельной работе студентов.

Актуальная в последнее время индивидуализация обучения направлена на то, чтобы процесс обучения стал личностно-ориентированным, то есть соответствовал интересам и потребностям будущих специалистов. Но, в разрез с этим, уменьшение аудиторной нагрузки снижает возможность реализации субъект-субъектного взаимодействия. Эти трудности вполне преодолимы с помощью персонального компьютера и мультимедийных технологий.

Самостоятельная работа является одной из важнейших форм обучения в ВУЗе – обязательной составляющей учебной, научной и исследовательской деятельности студентов. Ее эффективность в значительной мере определяет качество профессиональной подготовки будущих специалистов. К сожалению, практика показывает, что у студентов самостоятельная работа вызывает вполне определенные сложности, они не умеют рационально распределять и планировать свое время, выделять главное, самостоятельно принимать решения и делать выводы. В такой ситуации преподаватель должен научить студентов работать самостоятельно, направлять их, а не жестко регламентировать их деятельность, заинтересовать обучаемых какой-либо проблемой таким образом, чтобы они почувствовали личную заинтересованность в ее самостоятельном решении.

Анализ работ, посвященных исследованию вопросов организации самостоятельной работы (Л.Г. Вяткин, М.Г. Гарунов, П.И. Пидкасистый, С.Ф.Шатилов и др.), показал, что в методической литературе нет единого определения термину «самостоятельная работа». Все исследователи трактуют это понятие по-своему, но все сходятся в едином мнении, что самостоятельная работа представляет собой