

УДК 004.054:371.277.2

**Захарова О.А.,
Черкесова Л.В.,
Акишин Б.А.,
Манаенкова О.Н.,
Никишина Т.Г.**

МЕТОДОЛОГИЧЕСКИЕ И ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СОЗДАНИЯ ПЕДАГОГИЧЕСКИХ ТЕСТОВ ПО ИНФОРМАТИКЕ (ЧАСТЬ 1)

Ключевые слова: педагогический тест, тестирование, оценка знаний, дистанционное обучение, статистическая обработка, тестовый балл, конструирование теста, валидность теста, рейтинговая шкала оценки.

Введение

В конце XX в. возникло явное противоречие между потребностями развивающегося информационного общества, а также особенностями личности каждого обучающегося студента и сложившейся системой образования. Поиск решения этого противоречия привел лучших педагогов к идеям гуманистической педагогики, которая выступает антиподом традиционной, авторитарной педагогики. Основные ценности гуманистической педагогики были положены в основу новых концептуальных подходов к современному образованию, таких как личностно ориентированный подход, конструктивизм и др.

В развитие новых подходов внесли свой вклад многие ученые, среди которых следует выделить имена Дж. Дьюи, К. Роджерса, А. Маслоу, Дж. Брунера, Г. Олпорта, Л.С. Выготского, А.Н. Леонтьева, Ш.А. Амонашвили и др.

Еще ранее идея гуманистического обучения разрабатывалась в трудах И.Г. Песталоцци, который полагал, что главная задача педагога в отношении учеников – помочь им раскрыться, развить собственные идеи. Схожие идеи высказывал и М.В. Ломоносов, утверждавший, что образование есть не сумма знаний, а правильный образ мышления, и Я.А. Коменский, призывавший к реализации элементарного умственного образования, предусматривающего свободную ориентацию в числах, форме и языке, а также дополнительного образования соответственно врожденным способностям.

Определяющее влияние на формирование принципиально новых способов обучения, основанных на принципах гуманистической педагогики, оказал процесс глобальной информатизации и повсеместного внедрения информационных и коммуникационных технологий. В начале XXI в.

развитие Интернет-обучения превратилось в фактор глобального значения, открывающий новые перспективы в совершенствовании мировой образовательной системы.

Сегодняшний уровень развития компьютерных технологий закладывает реальный фундамент для создания информационно-образовательной среды поддержки профессионального обучения, одним из главных компонентов которой является система мониторинга и оценки знаний обучающихся [1].

Европейский и американский опыт использования дистанционных технологий показывает, что наиболее перспективной моделью для высшего и дополнительного образования является модель, основанная на интеграции традиционных и дистанционных форм обучения, а также реализации на их основе распределенных и сетевых информационных технологий [2].

Донской государственный технический университет (ДГТУ) на протяжении ряда лет интенсивно изучает и внедряет опыт открытых университетов в области использования информационно-коммуникационных технологий в образовании. Технологическая поддержка профессионального обучения в ДГТУ базируется на использовании как собственных зарегистрированных программных разработок, так и стандартных открытых инструментальных средств, позволяющих осуществлять мониторинг и контроль качества обучения [3].

Педагогическая концепция системы мониторинга и оценки знаний обучающихся студентов, реализуемая в ДГТУ на основе информационно-образовательной виртуальной среды, является результатом обобщенного анализа законодательства Российской Федерации, работ ведущих теоретиков и практиков, занимающихся данной проблемой в области теоретико-

методологических основ профессионального инженерно-технического образования (А.А. Вербицкий, М.А. Данилов, В.И. Загвязинский, В.В. Краевский, Н.В. Кузьмина, Я.Е. Львович, М.Г. Минин, Н.Ф. Талызина и др.), теории и практики информатизации образования (Я.А. Ваграменко, И.Е. Вострокнутов, Л.Х. Зайнутдинова, О.А. Козлов, Т.А. Лавина, М.П. Лапчик, Л.П. Мартиросян, И.П. Мухаметзянова, Н.И. Пак, И.В. Роберт и др.), автоматизации управления образовательным процессом (С.Г. Данилюк, А.А. Павлов, Ю.А. Романенко, В.И. Сердюков и др.), личного опыта и взгляда многих авторов на методологические основы, цели, задачи, закономерности и принципы, содержание, технологию формирования и реализации системы подготовки специалистов на основе информационно-образовательной среды, а также на подходы к управлению качеством подготовки специалистов технического направления и основные пути ее совершенствования.

Под информационным взаимодействием, реализуемым в процессе повышения квалификации, будем понимать «взаимодействие между пользователями, основанное на процессе передачи-приема информации, представленной в любом виде при реализации обратной связи, развитых средств ведения интерактивного диалога, при обеспечении автоматизации процессов сбора, обработки, продуцирования, архивирования, транслирования информации» [4].

Информационное взаимодействие требует определенной технологической реализации; в современном исполнении оно осуществляется средствами информационных и коммуникационных технологий. Для определения показателя качества высшего профессионального образования необходима

объективная оценка достижений обучаемого студента. Такая оценка успехов в учебе осуществляется, как правило, стандартизированными процедурами, при реализации которых все обучаемые находятся в равных условиях [5].

Принято считать, что педагогический тест, в отличие, например, от обычной контрольной работы, можно рассматривать как своеобразный измерительный инструмент, имеющий определенную разрешающую способность и точность.

Нельзя забывать, что такой объект измерения, как знания студентов, чрезвычайно специфичен, и поэтому результаты тестирования существенно зависят от возможностей преподавателей разумно формализовать этот объект в рамках применения контекстного [6] и личностно ориентированного [7] подходов.

Традиционно используется следующий подход к оценке знаний испытуемых: итоговый балл вычисляется как алгебраическая сумма оценок за выполнение каждого задания теста. Достоинством подобных первичных баллов является простота и наглядность, поэтому они часто используются на практике в качестве результирующей меры подготовленности участников тестирования.

Пионерской работой в области педагогического тестирования является работа датского математика Г. Раша. Эта ставшая классической работа дала толчок интенсивному развитию теоретической базы педагогического тестирования, ныне широко применяемого в образовании, и в частности в ЕГЭ [8].

Результатом развития этой теории за рубежом (в Европе и Америке) стала современная теория тестирования *Item Response Theory (IRT)*, основной целью которой является разработка математической модели процесса тестирования, параметрами чего служат различные характеристики участников тестирования и самого педагогического теста как

комплекса тестовых заданий. На русский язык термин IRT можно перевести как «Теория моделирования и параметризации создания педагогических тестов и самого процесса тестирования» [9].

В процессе педагогического тестирования большая роль отводится правильному составлению тестовых заданий. Они должны удовлетворять требованиям «информационной наполненности» как способности информации предоставлять возможность ее нового толкования и интерпретации через период времени. Подобное учебное задание – это средство интеллектуального развития, образования и обучения, способствующее активизации учебного процесса, повышению качества знаний студентов, а также возрастанию эффективности труда преподавателей [10].

Большое значение имеет процесс конструирования тестового задания, создания его формы [11]. Повсеместно считается, что существует множество подобных форм. Однако известный отечественный ученый В.С. Аванесов выделяет всего четыре их вида. Остальные формы, по его мнению, представляют собой варианты этих четырех базовых форм: либо одной из них, либо комбинаций двух или нескольких. Такое смешение форм тестовых заданий нежелательно, поскольку это затрудняет понимание и оценку результатов тестирования [12].

Тестовые задания закрытой формы с выбором одного или нескольких правильных ответов

Для таких заданий необходимо сформулировать в утвердительной форме предложение с неизвестным компонентом. Подстановка правильного ответа на место такого компонента превращает утверждение в истинное высказывание, а неправильного ответа – в ложное, что покажет незнание студентом конкрет-

ного учебного материала. При составлении подобных заданий преподавателю следует руководствоваться следующими принципами [13; 14].

- **Принцип противоречия**, использующий логическую схему типа «А и не А».

Пример. Если увеличить разрешение экрана видеомонитора от режима 800 × 600 пикселей до режима 1024 × 768 пикселей, то скорость работы компьютерных программ (здесь и далее жирным шрифтом выделены правильные ответы):

1. Уменьшится.
2. Не уменьшится.

- **Принцип противоположности**, использующий правила трехзначной логики.

Пример. При увеличении масштаба изображения в 6–8 раз в растровом графическом редакторе (например, MS Paint) качество изображения:

1. Ухудшается.
2. Не изменяется.
3. Улучшается.

- **Принцип исключения** из списка ответов лишней записи.

Пример. Файловая система операционной системы Windows не позволяет:

1. Задавать способы организации файлов.
2. Реализовывать доступ к содержимому файла.
3. Организовывать саму файловую структуру.
4. Осуществлять шифрование информации в файлах.
5. Манипулировать (копировать, переименовывать, удалять) файлами.

- **Принцип однородности**, при котором ответы относятся к одному роду, виду.

Пример. Основные функции операционной системы ПК заключаются:

1. В упрощении доступа к внутренней и внешней памяти ПК.

2. В распределении ресурсов внутренних и внешних устройств ПК.

- 3. **В управлении временем процессора и доступом к оперативной памяти ПК.**

4. В упрощении доступа пользователя к ресурсам персонального компьютера и распределении их между конкурирующими процессами и задачами.

5. В выполнении профилактики аппаратуры и обеспечении надежности и эффективности функционирования вычислительной системы ПК.

- **Принцип кумуляции**, при котором содержание следующего варианта ответа вбирает в себя содержание предыдущего и добавляется еще один компонент.

Пример. Внутреннюю память персональной ЭВМ составляют:

1. Регистры.
2. Регистры и оперативное запоминающее устройство (ОЗУ).

3. **Регистры, ОЗУ и постоянное запоминающее устройство (ПЗУ).**

4. Регистры, ОЗУ, ПЗУ и накопитель на жестких магнитных дисках (НЖМД).

5. Регистры, ОЗУ, ПЗУ, НЖМД и накопитель на оптических дисках (НОД).

- **Принцип сочетания** в ответах по два, по три или более слов в каждом ответе.

Пример. К векторным графическим редакторам можно отнести:

1. Adobe Photoshop и 3D-Studio Max.
2. **3D-Studio Max и AutoCAD.**
3. AutoCAD и MS Paint.
4. MS Paint и Corel Draw.
5. Corel Draw и Adobe Photoshop.

- **Принцип градуирования**, при котором используются оценочные понятия или количественные значения конкретных величин.

Пример. Длина имени файла в операционной системе MS Windows достигает:

1. 15 символов.
2. 31 символа.

3. 63 символов.
4. 127 символов.
5. **255 символов.**
6. 511 символов.

- **Принцип удвоения противопоставления**, при котором применяется некое противоречие и используется сочетание нескольких ответов, что позволяет предъявить повышенные требования к уровню и структуре знаний студентов.

Пример. У матричного принтера по сравнению со струйным принтером:

1. Выше быстродействие и выше качество печати.
2. Выше быстродействие и ниже качество печати.
3. Ниже быстродействие и выше качество печати.
4. **Ниже быстродействие и ниже качество печати.**
5. Такое же быстродействие и такое же качество печати.

- **Принцип фасеточности** содержания задания (фасет – это шаблонная, или трафаретная форма записи нескольких вариантов одного и того же задания).

Пример. Размер файла в растровом графическом редакторе MS Paint (Adobe Photoshop и др.) максимален (минимален и др.) при использовании типа файла:

1. **24-разрядный рисунок.**
 2. 256-цветный рисунок.
 3. 16-цветный рисунок.
 4. Монохромный рисунок.
 5. Рисунок в формате JPEG.
 6. Рисунок в формате TIFF.
- **Принцип импликации** (конструкция «если, то...»).
- Пример.* Если при сканировании документа увеличить разрешение (количество точек на дюйм) с 300 до 400 dpi, то размер полученного файла:
1. Уменьшится.
 2. Не изменится.
 3. **Увеличится.**

- **Принцип сопоставления**, при котором формулируются задания с вариантами ответов, правильными в различной мере, что проверяют сопоставительные знания.

Пример. С увеличением разрешающей способности экрана видеомонитора ПК до SVGA размеры каждого пиксела («зерна») составляют:

1. Менее 0,55 мм.
 2. Менее 0,45 мм.
 3. Менее 0,35 мм.
 4. **Менее 0,25 мм.**
 5. Менее 0,15 мм.
 6. Менее 0,05 мм.
- **Принцип сочетания принципов.** Например, в одном тестовом задании может присутствовать сочетание принципов однородности и фасеточности.

Пример. Операционная система реального времени (разделения времени и др.) предназначена для выполнения следующих функций управления ЭВМ:

1. **Управление внешними по отношению к ЭВМ технологическими процессами.**
2. Обеспечение удобства работы группы пользователей.
3. Наиболее эффективное использование ресурсов ЭВМ.
4. Диалоговый доступ нескольких пользователей через терминалы.
5. Выделение каждому пользователю ресурсов ЭВМ «по очереди».

- **Принцип сочетания различных качественных понятий.**

Пример. При вводе текста, иллюстрации или графического изображения через сканер, например с помощью программы Adobe Photoshop, увеличение разрешения сканирования (числа точек на дюйм) может привести к:

1. **Повышению качества изображения и увеличению размера файла.**
2. Ухудшению качества изображения и уменьшению размера файла.

3. Уменьшению размера файла и улучшению качества изображения.

4. Увеличению размера файла и ухудшению качества изображения.

5. К тому, что качество изображения и размер файла не изменятся.

Тестовые задания открытой формы

В тестовых заданиях открытой формы испытуемый студент должен вписать свой ответ в специально отведенное для этого место. Подобные задания открытой формы формулируются в виде утверждений, которые превращаются в истинное высказывание, если ответ правильный, и становятся ложными высказываниями, если ответ неправильный. При составлении подобных заданий используются следующие принципы [15; 16].

- **Принцип композиции**, при котором нужно вписывать несколько ответов.

Пример. Диапазон значений целого типа данных word (shortint, integer, longint и др.) в языке программирования Turbo Pascal 7.0 составляет значения _____, причем такой тип данных занимает в памяти компьютера _____ байт.

Правильные ответы: 0–65535; 2.

- **Принцип параллельности**, который достигается с помощью рассмотренного выше принципа вариативности (фасеточности, шаблонности).

Пример. Численный метод Эйлера (Рунге–Кутта, Ньютона и др.) служит для решения _____.

Правильный ответ: дифференциальных уравнений с помощью интегрирования.

- **Принцип обратимости**, при котором для некоторого высказывания нужно сформулировать обратное утверждение и доказать его истинность.

Пример. Имеется прямое утверждение: «Процесс преобразования программы, написанной на алгоритмическом языке программирования

высокого уровня, в машинные коды называется _____».

Правильный ответ: трансляцией.

Тогда обратное утверждение будет таковым: «Трансляцией называется процесс преобразования компьютерной программы, написанной на алгоритмическом языке программирования высокого уровня, в _____».

Правильный ответ: в машинные коды.

- **Принцип краткости для тестовых заданий открытой формы** (по принципу «чем меньше слов (в формулировании задания), тем меньше недоразумений»).

Пример. Значение переменной n после выполнения фрагмента алгоритма:

$n := 0;$

нц для k от 1 до 1000

если $\text{mod}(k, 30) = 1$ или $\text{mod}(k, 5) = 2$
то $n := n + 1$

все

кц

равно _____.

Правильный ответ: 234.

- **Принцип импликации для тестовых заданий открытой формы** («если, то»).

Пример 1. Если при вводе текста документа в текстовом процессоре MS Word нужно набрать математическую формулу, то при использовании опции меню «Вставка/Объект» нужно выбрать строку _____.

Правильный ответ: редактор формул Microsoft Equation 3.0 (Math Type).

Пример 2. Числа $x = 220_{(3)}$, $y = 130_{(4)}$ и $z = 22_{(15)}$ заданы в троичной, четверичной и пятнадцатиричной системах счисления соответственно. Сумма тех из них, которые в десятичной системе счисления кратны трем, равна _____.

Правильный ответ: $x = 220_{(3)} = 24_{(10)}$; $y = 130_{(4)} = 28_{(10)}$; $z = 22_{(15)} = 32_{(10)}$. Кратно трем здесь только число 24, поэтому сумма кратных чисел равна 24.

- **Принцип неотрицательности**, согласно которому в формулировках заданий не рекомендуется использовать отрицательные частицы, так как это нарушало бы логическое правило: «определение не должно быть отрицательным».

Пример. В процессе выполнения компьютерная программа и обрабатываемые ею данные должны храниться в _____.

Правильный ответ: в оперативной памяти персонального компьютера.

Тестовые задания на установление соответствия элементов одного множества элементам другого множества

Этот вид тестовых заданий позволяют проверить так называемые ассоциативные знания, существующие в каждой учебной дисциплине. Принципы составления подобных тестовых заданий были рассмотрены в работе [17]:

- **Принцип композиции.**
- **Принцип вариативности** (фасеточности, шаблонности).

Пример. Заданы два множества. Установить соответствие:

Назначение программы	Программа
1. Текстовые процессоры	а) Microsoft Internet Explorer
2. Электронные таблицы	б) Microsoft Outlook
3. Системы управления базами данных	в) Writer
4. Редакторы презентаций	г) Microsoft Excel
5. Почтовые программы	д) Corel Draw
6. Программы-переводчики	е) Microsoft Access
7. Программы-браузеры	ж) Microsoft Paint
8. Графические редакторы	з) Opera
9. Языки программирования	и) Microsoft Word
	к) Adobe Photoshop
	л) Super Calc
	м) Prompt
	н) Microsoft Power Point
	о) Netscape Navigator
	п) Stylus
	р) Google Chrome
	с) Impression
	т) Paradox
	у) Mail.ru
	ф) Delphi
	х) C++

Правильные ответы: 1 – в, и; 2 – г, л; 3 – е, т; 4 – ж, с; 5 – б, у; 6 – м, п; 7 – а, з, о, р; 8 – д, ж, к; 9 – ф, х [18].

Тестовые задания на установление правильной последовательности действий – вычислений, операций, команд, шагов алгоритмов и т.д.

Цель введения таких тестовых заданий – формирование алгоритмического мышления, алгоритмических знаний, умений и навыков [19; 20].

Пример. Чтобы скопировать фрагмент документа из одного окна текстового процессора MS Word в другое окно, нужно выполнить следующие действия:

- а) открыть другое окно MS Word;
- б) вставить из буфера содержащийся там фрагмент;
- в) в первом окне выделить нужный фрагмент документа;
- г) установить курсор в нужное место второго окна с документом;
- д) скопировать его в буфер.

Правильный ответ: в, д, а, г, б.

Статистическая обработка тестовых заданий

Актуальным вопросом остается следующая статистическая обработка результатов педагогического тестирования. Для того чтобы проанализировать существующие математические модели оценки результатов педагогического тестирования, имеет смысл провести сравнение их достоинств и недостатков.

Известно три основных подхода к разработке педагогических тестов [21]:

1. Критериально-ориентированные тесты, подразделяющиеся на два основных типа: квалификационные и содержательные – и предназначенные для установления уровня квалификации и знаний испытуемыми содержания изучаемой дисциплины.

2. Нормативно-ориентированные тесты, предназначенные для сравнения уровня подготовленности испытуемых между собой.

3. Смешанный тип, сочетающий оба вышеуказанных типа педагогических тестов.

В практике педагогического тестирования чаще всего встречается критериально-ориентированный подход к разработке тестовых заданий, формирующих тест. При этом используется специальная моделирующая методика, которая позволяет при конструировании теста использовать научно обоснованный подход к отбору тестовых заданий и столь же продуманный и психологически обоснованный подход к их расположению в тесте, состоящем из 40–50 тестовых заданий, когда на обдумывание ответов студенту выделяется не более 1 минуты, что вполне достаточно при наличии твердых знаний по изучаемой дисциплине, по которой сдается тест [22].

Исторически к созданию педагогического теста выделяются два основных подхода: как к совокупности некоторого множества тестовых заданий; а также как к интерпретации результатов его выполнения испытуемыми (студентами).

В первом подходе, получившем развитие в рамках классической теории педагогического тестирования, уровень знаний испытуемых оценивается с помощью их индивидуальных баллов. Недостатками такого подхода считаются отсутствие дифференциальной ошибки измерения и зависимость оценок от трудности тестовых заданий. С помощью такого теста нельзя объективно оценить значение параметров, характеризующих уровень знаний испытуемых и трудности тестовых заданий, а также выразить значение этих параметров в индивидуальной рейтинговой шкале конкретного студента.

Второй подход заключается в использовании новой теории: в отличие от классической теории, современная теория педагогического тестирования IRT является частью более общей теории латентно-структурного анализа, в которой распределение переменных предполагается непрерывным на оси латентной переменной. Вместе с тем IRT не всегда дает явные преимущества в процессе моделирования теста. Исследователи отмечают, что при использовании однопараметрической и двухпараметрической моделей тестирования преимущества не всегда очевидны и незначительны [23].

Однако при статистической оценке тестовых заданий с помощью трехпараметрической модели ситуация существенно изменяется в пользу IRT, поэтому именно ее рекомендуют использовать при конструировании теста. Отметим, что применение IRT не всегда предпочтительно в силу сложности математического аппарата и наличия ряда ужесточенных требований к качеству выборки, по которой собирается статистика для конструирования теста. Кроме того, в IRT используются не прямые методы оценки параметров заданий, поскольку вначале проводится подгонка статистических данных под требование моделей измерения, а лишь затем получаются оценки параметров тестовых заданий.

В рамках IRT конструирование теста осуществляется согласно заданной предварительно информационной функции теста. В процессе моделирования можно выделить последовательность шагов, позволяющую применить информационные функции IRT при конструировании теста [24]:

- задать целевую функцию конструирования педагогического теста и выбрать подход к его разработке (либо нормативно-ориентированный, ли-

бо критериально-ориентированный, либо смешанный, гибридный тип, как говорилось выше);

- получить первоначальные представления о положении на оси измеряемой переменной данных об оцениваемой выборке испытуемых (студентов);
- выбрать планируемую точность педагогических измерений;
- задать форму целевой информационной функции теста (на выделенном интервале шкалы логитов – логарифмических оценок параметров тестирования);
- выбрать планируемый вид распределения оценок трудности тестовых заданий;
- отобрать тестовые задания, информационные функции которых наиболее приемлемо заполняют область под целевой информационной функцией;
- добавлять задания, вычисляя каждый раз количество информации в различных точках оси измеряемой переменной для каждого из вновь образованных тестов;
- продолжать отбор тестовых заданий из банка заданий, пока очередная информационная функция формируемого педагогического теста не обеспечит аппроксимацию (в приемлемой степени) заданной целевой информационной функции;
- выполнение перечисленных шагов предполагает завершение предварительной работы по созданию банка шкалированных тестовых заданий и спецификации, обеспечивающей содержательную валидность (т.е. пригодность) теста.

В результате такого целенаправленного отбора из банка тестовых заданий определенного содержания, нужной трудности и дифференцирующей способности можно сформировать такой

педагогический тест, который имеет высокую содержательную валидность для измерения уровня знаний испытуемых студентов и способен обеспечить планируемую точность педагогических измерений на выбранном интервале оси измеряемой переменной.

При критериально-ориентированном подходе для повышения информативности теста очень важно определить количество тестовых заданий, необходимых для оценки порогового балла для отсева испытуемых студентов, не прошедших критериальный отбор, означающий успешное выполнение теста.

В качестве критерия используется сумма тестовых баллов испытуемых студентов, полученная при выполнении пробного варианта педагогического теста.

В процессе моделирования теста нужно собирать статистические данные и вычислять дисперсию индивидуальных баллов испытуемых студентов после каждого очередного добавления тестовых заданий и удлинения теста. Как только дисперсия перестанет возрастать, длину теста можно зафиксировать как оптимальную, при условии удовлетворительного заполнения области под целевой информационной функцией и высокой содержательной валидности теста [25].

Вычислив значения информационных функций для нескольких (хотя бы двух) педагогических тестов, можно вычислить функцию сравнительной эффективности. Получение значения функции сравнительной эффективности позволит оценить педагогический эффект, например, при удалении из теста заданий повышенной сложности, а также при замене одних тестовых заданий на другие либо замене тестовых заданий средней трудности на задания более сложные или более легкие. Таким образом, рассмотренная нами функция сравнительной эффективно-

сти позволяет моделировать педагогический тест без сбора дополнительной эмпирической информации.

В отличие от других контролируемых педагогических материалов, тесты всегда проходят специальный процесс предварительной апробации.

Создание теста начинается с общего педагогического замысла, однако затем разработчикам требуется перейти от обычных заданий в тестовой форме к тестовым заданиям как тестовым единицам, отвечающим заранее заданным требованиям общей суммарной трудности, технологичности, вариативности баллов и коррелируемости тестовых заданий с выбранным критерием.

Для проверки тестовых свойств заданий используется матрица результатов тестирования в виде нулей и единиц, приведенная к верхнетреугольному виду путем упорядочения в порядке убывания. Сложение элементов матрицы по строкам и по столбцам здесь имеет особый педагогический смысл: сложение по строке (Y_j) дает тестовый балл испытуемого, а по столбцу (R_j) дает количество правильных ответов испытуемых на конкретное задание теста. Очевидно, что это количество зависит от числа испытуемых N , поэтому удобнее использовать нормированный статистический показатель (P_j), который можно применять в группах с различным числом испытуемых студентов [27]:

$$P_j = R_j / N.$$

В классической теории тестирования этот показатель выступал в качестве показателя трудности, однако из-за его смысловой неточности, заключающейся в том, что увеличение нормированного значения указывает не на возрастание трудности, а, наоборот, на возрастание легкости, в настоящее время предпочтение отдается противоположной статистике q_j ,

определяемой как отношение числа неправильных W_j ответов к числу испытуемых студентов N :

$$q_j = W_j / N.$$

Имеет смысл вычислить среднеарифметический тестовый балл M для того, чтобы определить стандартное отклонение S_y результата тестирования каждого испытуемого студента от среднего значения по данной выборке. Эта величина является удобной мерой вариации тестовых баллов. Так как дисперсия тестовых баллов S_y^2 результатов тестирования является стандартным показателем вариации, то ее вычисление можно производить по формуле:

$$S_y^2 = \frac{SS_y}{N-1},$$

где величина SS_y – сумма квадратов отклонений значений баллов результатов тестирования от среднего арифметического, определяемого выражением:

$$SS_y = \sum Y^2 - (\sum Y)^2 / N.$$

Следующим этапом проверки валидности теста и тестовых свойств заданий является вычисление классического коэффициента корреляции Пирсона:

$$r_{xy} = SP_{xy} / \sqrt{SS_x \times SS_y},$$

где SP_{xy} – сумма парных произведений; SS_x – сумма квадратов отклонений баллов испытуемых от среднего арифметического балла в интересующем задании; SS_y – сумма квадратов отклонений тестовых баллов испытуемых студентов от среднего арифметического балла по всему тесту.

Если значение корреляции стремится к единице ($k \rightarrow 1$), то тест хорош, т.е. обладает валидностью – пригодностью и способностью оценки знаний.

Нулевое значение коэффициента корреляции ($k = 0$) свидетельствует

об отсутствии у заданий системных свойств, присущих тесту. В этом случае тест практически бесполезен, является неустойчивым и требует доработки.

Если же значение корреляции стремится к минус единице ($k \rightarrow -1$), то тест плох, непригоден, валидностью не обладает и его использовать нельзя. В этом случае разработчикам теста следует его полностью переработать [28; 29].

Полученные коэффициенты корреляции нужно свести в корреляционную матрицу, и после ее анализа можно приступить к созданию теста, который не будет содержать заданий, имеющих отрицательный коэффициент корреляции.

В современных технологиях адаптивного обучения и контроля используется логарифмическая оценка уровня знаний, определяемая как $\ln \frac{p_i}{q_i}$. Также вводится логарифмическая оценка трудности задания $\ln \frac{q_j}{p_j}$. Таким образом, существуют понятия логита уровня знаний, равного $\ln \frac{p_i}{q_i}$, и логита трудности задания, равного $\ln \frac{q_j}{p_j}$.

Логарифмические оценки привели к попытке сравнения их посредством вычитания. Эффективность такого сравнения оказала огромное влияние на развитие педагогической теории и практики. С помощью ЭВМ можно сопоставить значения логита трудности задания и логита уровня знаний и на этой основе подобрать очередное задание в системах адаптивного обучения и контроля знаний.

Используя математический аппарат IRT, можно конструировать тесты, обладающие высокими показателями надежности и валидности оценки знаний.

После опроса группы студентов, проведенного по тестовым материа-

лам, подготовленным по рекомендациям одной из двух теорий тестирования (классической теории или теории IRT), перед преподавателем встает задача сопоставления результатов тестирования с принятыми в высшей школе: двадцатипятибалльной (рейтинговой) шкалой, применяемой в вузах во время промежуточных аттестаций в середине семестра, а также пятибалльной (экзаменационной) и дихотомической (зачет – незачет) шкалами оценивания знаний испытуемых студентов. При этом возникает следующая ситуация.

Качественный тест образует шкалу, которая представляет собой числовую систему, выражающую отношения между различными испытуемыми студентами. Полученные исходные тестовые баллы следует сгруппировать и построить гистограмму, отражающую зависимость баллов, набранных испытуемыми студентами, от частоты их появления, для каждого тестового задания.

На основании полученных результатов и расчетов каждый преподаватель может самостоятельно сделать вывод об усвоении учебного материала каждым испытуемым студентом и осуществить переход к принятым шкалам оценки.

В зависимости от уровня педагогического контроля можно использовать различные подходы к выставлению итоговых баллов по шкалам оценки знаний, используемым в современной российской высшей школе [30].

Выводы

В практической педагогической деятельности при составлении тестовых заданий, контролирующих знания студентов по любой дисциплине, по нашему мнению, следует руководствоваться методологическими и дидактическими принципами разработки (конструирования) теста, изложенными выше.

На кафедре «Математики и информатики» Донского государственного технического университета широко используется автоматизированная система итогового и промежуточного тестирования для контроля знаний студентов по различным математическим и информационным курсам [31]. Разработан обширный банк тестовых заданий как для бланкового (бумажного) тестирования студентов в лекционных аудиториях, так и для компьютерного тестирования, проводящегося в компьютерных классах. В базе тестовых заданий этой системы широко использованы все формы тестовых заданий, что позволяет провести тщательный анализ уровня знаний студентов.

Педагогические тестовые материалы прошли апробацию на различных выборках испытуемых студентов и показали соответствие требованиям, предъявляемым к качеству тестовых подсистем программ дистанционного обучения [32] и дистанционного образования в целом [33].

(Продолжение в следующем номере)

Библиография

1. *Keeves, J.P.* (Ed.), 1988. Educational Research, Methodology and Measurement: An International Handbook. Oxford: Pergamon Press.
2. *Dede, C.*, 1996. The Evolution of Distant Education: Emerging Technologies and Distributed Learning. The American Journal of Distance Education, 10 (2): 4–36.
3. *Захарова О.А.* Виртуальная образовательная среда в профессиональной подготовке и системе повышения квалификации. Ростов н/Д: Изд. центр ДГТУ, 2011.
4. *Роберт И.В.* О понятийном аппарате информатизации образования // Информатика и образование. 2003. № 1, 2; 2004. № 5.
5. *Захарова О.А., Месхи Б.Ч.* Стратегия развития инженерного образования: опыт ДГТУ // Дистанционные технологии в инженерной педагогике. 2-е изд., доп. Ростов н/Д, 2011.
6. Контекстный подход в психологии, педагогике и менеджменте: межвузовский сб. науч. трудов / науч. ред. А.А. Вербицкий. М.: МГГУ им. М.А. Шолохова, 2011.
7. *Зимняя И.А.* Педагогическая психология. М., 2002.
8. *Rasch, G.*, 1980. Probabilistic Model for Some Intelligence And Attainment Tests. Chicago: Univ. of Chicago Press.
9. *Crocer, L. and J. Algina*, 1986. Introduction to Classical and Modern Test Theory. N.Y.: HBJCP.
10. *Daft, R.L. and R.H. Lengel*, 1986. A proposed integration among organizational information requirements, media richness, and structural design. Management Science, 32: 554–571.
11. *Gronlund, N.E.*, 1998. How to Construct Achievement Test. Englewood Cliffs, NJ: Prentice Hall.
12. *Аванесов В.С.* Композиция тестовых заданий. М.: Адепт, 1998.
13. *Аванесов В.С.* Указ. соч.
14. *Черкесова Л.В., Зотова Т.П.* Методологические основы создания педагогических тестов на основе математического моделирования // Известия вузов. Электромеханика. Спец. юбил. вып. Информационные технологии и управление. Новочеркасск: ЮРГТУ, 2001. С. 178–189.
15. *Аванесов В.С.* Указ. соч.
16. *Черкесова Л.В., Зотова Т.П.* Указ. соч.
17. *Аванесов В.С.* Указ. соч.
18. *Черкесова Л.В., Зотова Т.П.* Указ. соч.
19. *Аванесов В.С.* Указ. соч.
20. *Черкесова Л.В., Зотова Т.П.* Указ. соч.
21. *Gronlund, N.E.* Op. cit.
22. *Gronlund, N.E.* Op. cit.
23. *Челышкова М.Б.* Разработка педагогических тестов на основе современных математических моделей. М.: Исследовательский центр проблем качества подготовки специалистов, 1995.
24. *Челышкова М.Б.* Теория и практика конструирования педагогических тестов. М.: Исследовательский центр проблем качества подготовки специалистов, 2000.
25. *Челышкова М.Б.* Теория и практика конструирования...
26. *Челышкова М.Б.* Теория и практика конструирования...
27. *Челышкова М.Б.* Теория и практика конструирования...
28. *Челышкова М.Б.* Разработка педагогических тестов...
29. *Челышкова М.Б.* Теория и практика конструирования...
30. *Сорокопуд Ю.В.* Повышение эффективности отечественной системы подготовки преподавателей высшей школы: пути решения. М.: Изд-во Военного университета, 2010.
31. *Захарова О.А.* Корпоративное обучение в системе повышения квалификации // Современные педагогические и информационные технологии в системе образования // И.В. Докучаева [и др.]. Новосибирск: Изд-во Центра развития научного сотрудничества, 2013. С. 200–245.
32. *Комаринский С.М.* Основные результаты экспертного оценивания качества тестовых под-

систем программ дистанционного обучения // Известия Южного федерального университета. Педагогические науки. 2014. № 4. С. 117–124.

33. Самойлов А.Н., Быкасова Л.В. Дистанционное образование: культура, среда, социализация // Известия Южного федерального университета. Педагогические науки. 2014. № 2. С. 37–42.

Bibliography

1. Keeves, J.P. (Ed.), 1988. Educational Research, Methodology and Measurement: An International Handbook. Oxford: Pergamon Press.
2. Dede, C., 1996. The Evolution of Distant Education: Emerging Technologies and Distributed Learning. The American Journal of Distance Education, 10 (2): 4–36.
3. Zakharova, O.A., 2011. Virtual academic environment in vocational training and system of professional skill improvement. Rostov-on-Don: published by Don State Technical University. (rus)
4. Robert, I.V., 2003. About the conceptual framework of informatization in education. Computer science and education, 1, 2. (rus)
5. Zakharova, O.A. and B.Ch. Meskhi, 2011. Strategy of engineering education development: experience Don State Technical University. Remote technologies in engineering pedagogics. 2nd ed., corrected. Rostov-on-Don. (rus)
6. Verbitsky, A.A. (Ed.), 2011. Contextual approach in psychology, pedagogy and management: interuniversity collection of papers. Moscow: published by Moscow State University of M.A. Sholokhov. (rus)
7. Zimnyaya, I.A., 2002. Pedagogical psychology. Moscow. (rus)
8. Rasch, G., 1980. Probabilistic Model for Some Intelligence and Attainment Tests. Chicago: Univ. of Chicago Press.
9. Crocer, L. and J. Algina, 1986. Introduction to Classical and Modern Test Theory. N.Y.: HBJCP.
10. Daft, R.L. and R.H. Lengel, 1986. A proposed integration among organizational information requirements, media richness, and structural design. Management Science, 32: 554–571.
11. Gronlund, N.E., 1998. How to Construct Achievement Test. Englewood Cliffs, NJ: Prentice Hall.
12. Avanesov, V.S., 1998. Composition of test tasks. Moscow: published by Adept. (rus)
13. Avanesov, V.S. Op. cit.
14. Cherkesova, L.V. and T.P. Zotova, 2001. Methodological background of creating pedagogical tests on the basis of mathematical modeling. News of higher schools. Electro mechanics. Special edition. Information technologies and management. Novocherkassk: published by Novocherkassk Polytechnic Institute: 178–189. (rus)
15. Avanesov, V.S. Op. cit.
16. Cherkesova, L.V. and T.P. Zotova. Op. cit.
17. Avanesov, V.S. Op. cit.
18. Cherkesova, L.V. and T.P. Zotova. Op. cit.
19. Avanesov, V.S. Op. cit.
20. Cherkesova, L.V. and T.P. Zotova. Op. cit.
21. Gronlund, N.E. Op. cit.
22. Gronlund, N.E. Op. cit.
23. Chelyshkova, M.B., 1995. Development of pedagogical tests on the basis of modern mathematical models. Moscow: published by Research Center on Problems of Experts Preparation Quality. (rus)
24. Chelyshkova, M.B., 2000. Theory and practice of creating pedagogical tests. Moscow: published by Research Center on Problems of Experts Preparation Quality. (rus)
25. Chelyshkova, M.B. Theory and practice of creating...
26. Chelyshkova, M.B. Theory and practice of creating...
27. Chelyshkova, M.B. Theory and practice of creating...
28. Chelyshkova, M.B. Development of pedagogical tests...
29. Chelyshkova, M.B. Theory and practice of creating...
30. Sorokopud, Y.V., 2010. Increasing efficiency of the national system of preparation of teachers of higher school: solutions. Moscow: published by Publishing house of Military University. (rus)
31. Zakharova, O.A., 2013. Corporate training in the system of professional skills improvement. In: Dokuchayeva, I.V. et al. Modern pedagogical and information technologies in education system. Novosibirsk: Publishing house of the Center of Development of Scientific Cooperation: 200–245. (rus)
32. Komarinsky, S.M., 2014. The main results of expert assessment of qualities of test subsystems of distant learning programs. News-Bulletin of Southern Federal University. Pedagogical Sciences, 4: 117–124. (rus)
33. Samoylov, A.N. and L.V. Bykasova, 2014. Distant education: culture, environment, socialization. News-Bulletin of Southern Federal University. Pedagogical Sciences, 2: 37–42. (rus)