

Особенности разработки педагогических тестов как компоненты информационных технологий

Константин Кречетников

Дальневосточный государственный университет, г. Владивосток

krechet@tovmi.dvgu.ru

Опубликовано в ж. «Педагогические Измерения» № 4, 2007

Аннотация

Рассматриваются особенности разработки и применения педагогических тестов в составе автоматизированных компьютерных тестовых программных оболочек и компьютерных обучающих программ. Особый акцент сделан на специальную подготовку учебного материала перед разработкой тестов.

Ключевые слова: альтернативы, гипертекст, дескрипторы, ссылки, связи, тезаурус, юниты

Педагогический контроль является неотъемлемой компонентой образовательного процесса, имеет значимость в любой технологии обучения, преследует различные цели и выполняет различные функции, определяющие формы контроля. С проблемой контроля всегда было связано множество вопросов, на некоторые из которых все еще не дано определенного ответа. Существующая система педагогического контроля не лишена многих недостатков, главными из которых являются: субъективизм, отсутствие регулярности контроля и чётких критериев оценки¹. Одним из путей преодоления указанных недостатков, автоматизации и объективизации контроля и самоконтроля является использование педагогических тестов.

С внедрением тестов в педагогике появилась возможность использовать статистические подходы к оценке эффективности образовательной среды, позволяющие объективно судить о качестве её проектирования. Кроме того, тестовый контроль способствует совершенствованию методов, способов, методик и приемов обучения, позволяя оценивать их результативность на основе объективных критериев и превращая тем самым педагогическое проектирование в допускающую измерение точную науку.

Тестовый контроль обеспечивает, по сравнению с традиционными формами контроля, целый ряд *преимуществ*, таких как: большая объективность и, как следствие, большее стимулирующее воздействие на познавательную деятельность обучающегося;

1. Аванесов В.С. Композиция тестовых заданий. – М.: Центр Тестирования, 2002. – 240 с.

эксперименты показывают, что даже на значительных выборках отличия субъективной и объективной оценок составляют до 20 % ²;

- широкая область использования: от самоконтроля до оценивания эффективности функционирования образовательной среды в целом;

- возможность использования в составе информационных образовательных технологий;

- обеспечение процесса интеграции дисциплин (междисциплинарные тесты).

Основным недостатком тестирования считается лишение обучающихся практики устной речи и участия в дискуссиях. Поэтому при использовании тестов необходимо практиковать и другие формы контроля и взаимодействия с обучающимися (семинары, теле-, видео- и обычные конференции, диспуты, обсуждения, деловые игры), обеспечивающие компенсацию указанного недостатка.

По-особому тесты востребованы в образовательной среде, в которой широко используются *информационные технологии*. В этом случае обучение должно начинаться с входного тестового контроля, сопровождаться самоконтролем и заканчиваться итоговым тестированием.

Особенности использования тестовых заданий различной формы

Содержание педагогического теста может быть представлено в четырёх основных формах³:

- *задания открытой формы* (со свободными ответами, – когда испытуемому необходимо самостоятельно дописать слово, словосочетание, знак, формулу); в заданиях открытой формы испытуемому не надо ничего выбирать, а требуется только дополнить само задание правильным ответом; эта форма контроля полностью исключает догадку, за которую критикуются задания с выбором;

- *задания с выбором одного или нескольких правильных ответов*; одной из модификаций данной формы тестовых заданий являются задания с выбором одного, наиболее правильного ответа (если все предложенные варианты ответов являются правильными, но каждый из них соответствует истине в разной степени);

- *задания на установление соответствия* – где требуется сопоставить между собой два списка, причём сопоставление должно производиться поэлементно;

2. Кибакин М.В., Лапшов В.А., Чмыхова Е.В. Мониторинг успешности изучения курса физики старшеклассниками: проблема оптимизации объема усваиваемых знаний // Труды СГУ. – М., 1999. – Вып. 10. – С. 94 – 101.

3. Аванесов В.С. Композиция тестовых заданий. – М.: Центр Тестирования, 2002. – 240 с.

– задания на установление правильной последовательности. Они применяются, когда требуется установить порядок расчётов, операций, алгоритм действий, хронологию событий и т.д.

Для проведения анализа ответа обучающегося в заданиях *открытой формы* необходимо⁴:

1) выделить систему понятий курса. Понятия, объединенные в словарь, помогут, во-первых, определить соответствие содержания учебного материала поставленным автором целям обучения, а, во-вторых, служат точками входа в семантическую сеть курса.

2) структурировать понятия, организовать каждое в виде дескрипторной словарной статьи, содержащей определение понятия. Словарная статья содержит главный дескриптор, синонимы, ближайшие вышестоящие (родовые) дескрипторы, ближайшие нижестоящие (видовые) дескрипторы, ассоциативные дескрипторы.

3) представить каждую словарную статью в виде графа с расширенными связями между понятиями и ссылками на процедуры обработки понятия.

Предлагаемая технология позволяет проектировщику автоматизированной обучающей системы более свободно формулировать задания, а для анализа ответов указывать наборы словарных статей, содержащих контролируемые понятия. При этом обучающийся получает возможность вводить ответ в произвольной форме, а компьютер анализирует не только отдельные предложения ответа, но и весь дискурс в целом. Такой семантический анализ осуществляется по следующему алгоритму:

1) выбор словарной статьи из базы знаний, сопровождающей курс, для анализа ответа на конкретный вопрос;

2) поиск вхождения в ответ дескрипторов словарных статей либо фрагментов словарных статей с указанием связей между дескрипторами;

3) при полном соответствии ответа (или его части) словарной статье обучающемуся выдается положительное сообщение, после чего анализ завершается;

4) при частичном соответствии или несоответствии анализ продолжается. Просматривается цепочка словарных статей, связанных с контролируемым понятием; при этом учитывается количество шагов-отступлений в цепочке от требуемых словарных статей; обучающийся может дать более общий ответ, ввести более широкие понятия, а также, наоборот, детализировать его, углубить; в первом случае оценка ответа понижается, во втором – повышается (хотя могут быть и обратные варианты).

4. Апатова Н.В. Информационные технологии в школьном образовании. – М.: РАН, 1994. – 227 с.

Задания с выбором одного или нескольких правильных ответов из числа предложенных имеют хорошую перспективу для проверки глубины, прочности, точности и полноты освоения понятий, кроме того, они очень подходят для проверки классификационных знаний⁵. Успешность использования заданий данной формы во многом определяется удачностью подбора дистракторов (отвлекающих вариантов ответа). Оценка тестовых заданий по трём критериям - временные затраты, количество попыток, вероятность угадывания - позволила сделать вывод, что оптимальное число вариантов ответов в задании – 4 или 5⁶. Следует отметить, что ситуации, в которых необходимо сделать выбор из нескольких вариантов, часто вызывают у обучающихся чувство неуверенности. Французский исследователь П. Де Бек предполагает, что "на начальных этапах выбора наиболее важным фактором является недостаток информации, а на поздних этапах основными источниками сомнений становятся проблемы, связанные с субъективной ценностью альтернатив"⁷. Он выделяет три аспекта "субъективной ценности": отсутствие ясности, конфликтный характер альтернатив и равную привлекательность альтернатив.

Задания на установление соответствия и правильной последовательности рекомендуются для проверки соответственно ассоциативных и алгоритмических знаний, очень важных в век информационных технологий и повышенных требований к интеллекту человека. Исследования американских учёных показали, что решение аналогий, построенных на основе учебного материала, позволяет делать выводы не только о наличии или отсутствии у обучающихся тех или иных понятий, но и о структуре знаний обучающихся, сформированных у них стратегических когнитивных процессах⁸. Кроме того, задания данного типа позволяют проверить освоение сразу нескольких ветвей учебного материала, что обеспечивает их повышенную проверочную емкость.

Подходы к разработке педагогических тестов

Разработка педагогического теста – многоплановый процесс, основанный на достижениях современной теории педагогических измерений. Элементы содержания заданий должны быть не только значимыми, но и достаточными для контроля. Но репре-

5. Аванесов В.С. Форма тестовых заданий. М. Центр тестирования, 2006. – 155стр.

6. Кречетников К.Г. Проектирование креативной образовательной среды на основе информационных технологий в вузе. – Москва: Изд-во Госкоорцентр, 2002. – 296 с.

7. De Boeck P. Two Factors of Preference Certainty in the Process of Choosing an Educational Program // European Review of Applied Psychology. Paris: Editions du centre de psychologie app Uquee. – 1999. – Vol. 49. – № 1. – P. 23 – 29.

8. Alexander P.A., Murphy P.K., Kulikowich J.M. What responses to Domain-Specific Analogy Problems Reveal About Emerging Competence: A New Perspective on an Old Acquaintance // Educational Psychology. – 1998. – Vol. 90. – № 3. – P. 397 – 406.

зентативность теста не предполагает обязательного использования всех значимых элементов содержания, т. к. многие из них могут быть связаны между собой в общей структуре учебной дисциплины, включены один в другой, иерархически соподчинены⁹.

Один из подходов, облегчающий процесс разработки тестовых заданий, заключается в *построении структурированной модели знаний в виде ориентированного семантического графа*¹⁰. В вершинах графа располагаются модули знаний и понятия со своими свойствами, а их взаимосвязи определяются отношениями. Существенным преимуществом подобного представления модели знаний в виде семантического графа является системность знаний предметной области и наглядность её структуры. Построение модели знаний является аналогичным разработке алгоритма с помощью блок-схемы. Возможны технологии "сверху вниз" и "снизу вверх". При наличии семантического графа эксперту несложно формировать тестовые задания по каждому модулю знаний или понятию.

Семантическая сеть для отображения модели предметной области может быть наглядно представлена в форме гипертекста¹¹. *Гипертекстовое представление* позволяет подробно структурировать цели и содержание образования. Дуги, связывающие вершины (дескрипторы понятий гипертекста), несут значительную смысловую нагрузку, поэтому построение хорошей модели предмета в виде гипертекста требует разработки специальной методики. Речь идет о создании и анализе сетей, в узлах которых – высказывания на естественном языке. Устанавливая связь между узлами, составитель гипертекста утверждает тем самым истинность составного высказывания, полученного с помощью связки. Структура изложения содержания, заключенного в гипертексте, должна быть максимально приближена к структуре познавательной деятельности обучающегося.

Для составления научно-обоснованного, сбалансированного и объективного педагогического теста рационально также использовать "*учебный таксономический тезаурус*"¹². Данный тезаурус должен разрабатываться на основе метода групповых экспертных оценок с использованием математического аппарата педагогической квалиметрии. Он сопрягается с таксономической моделью уровней обученности и классификатором знаний и способно-

9. Зиновкина М.М. Формирование творческого технического мышления и инженерных умений студентов технических вузов: Дис. ... д-ра пед. наук. – М., 1989. – 326 с.

10. Пак Н.И., Симонова А.Л. Компьютерная диагностика знаний в системах дистанционного образования // Дистанционное образование. – 2000. – № 2. – С. 17 – 21.

11. Сметанин Ю.М. Концептуальное моделирование компонентов учебного процесса // Развитие системы тестирования в Удмурдской Республике: Тез. докл. регион. научно-практ. конф. 1 февраля 2001 г. – Ижевск: ИжГТУ, 2001. – С. 43 – 49.

12. Кречетников К.Г. Проектирование креативной образовательной среды на основе информационных технологий в вузе. – Москва: Изд-во Госкоорцентр, 2002. – 296 с.

стей. Учет таксономии учебных целей позволяет конструировать задания различной когнитивной сложности с заранее заданными параметрами. Из составленного учебного тезауруса отбираются те дескрипторы учебного предмета, которые способны составить содержательную область педагогического теста. Для этого предлагается составить таблицу-матрицу, в которой на одной оси расположены содержательные структуры, а на другой – уровни сложности тестовых заданий, что является основой для конструирования теста.

О.М. Карпенко и В.Ю. Переверзев предлагают следующую последовательность составления тестов, по каждому разделу дисциплины¹³:

- 1) определение целей и задач тестирования, отбор материалов для теста;
- 2) создание плана теста и спецификаций тестовых заданий;
- 3) составление тестовых заданий и их экспертный анализ;
- 4) проведение пробного тестирования и анализ его результатов; расчёт уровня трудности заданий, дифференцирующей способности, эффективности дистракторов;
- 5) выбор стандартов оценивания (критериального балла);
- 6) оценка параметров качества теста (подробнее рассмотрены ниже).

Описанная методика проверена в Современном гуманитарном университете и его многочисленных филиалах. В соответствии с модульным принципом, каждая дисциплина этого университета разбивается на крупные тематические модули – *юниты*. В свою очередь, юниты состоят из дидактических единиц (тем). Дальнейшая структуризация основана на применении сравнительно-тезаурусного подхода, позволяющего представить тему списком понятий и фактов.

Понятия раскрываются через определения, в которых выражены их семантические связи (*линки*) с другими уже известными понятиями. Среднее число линков в определении понятия равно 4 – 5, или 400 – 500 линков на юнит. Создание такого количества тестовых заданий по юниту (из расчета: один линк – одно задание) нецелесообразно. Поэтому в качестве контрольной единицы тестирования введён модульный объект, являющийся дидактическим объединением знаний и умений, согласно их смысловому содержанию. При этом важно, чтобы совокупность контрольных единиц достаточно полно покрывала всю тематическую область юнита.

При разработке тестов рекомендуется использовать классификатор знаний, разработанный В.С. Аванесовым¹⁴, который содержит 17 видов знаний. При объединении некоторых видов знаний получается классификатор, включающий 9 видов: фактуальные,

13. Карпенко О.М., Переверзев В.Ю. Технология тестового контроля успеваемости студентов в СГУ // Труды СГУ. Серия "Психология и социология образования". – М., 1999. – Вып. 10. – С. 132 – 136.

14. Аванесов В.С. Композиция тестовых заданий. – М.: Центр Тестирования, 2002. – 240 с.

сравнительные, ассоциативные, причинные, классификационные, алгоритмические, системные, абстрактные и методологические знания.

Для диагностики указанных видов знаний целесообразно использовать различные формы тестовых заданий. Например, для проверки фактуальных знаний целесообразно использовать задания на выбор одного или нескольких правильных ответов; ассоциативных и классификационных – задания на установление соответствия и правильной последовательности и т. д. С целью диагностики как можно большего количества видов знаний предлагается использовать внутри одного теста разнообразные формы тестовых заданий.

В целом же составление тестовых заданий – это большое искусство, которое требует напряжённого творческого труда, поиска, полёта мысли, глубокого знания своего предмета и психологии обучающихся.

Особенности компьютерных тестов

Технологическое преимущество заданий тестовой формы проявляется в их соответствии требованиям автоматизации рутинных компонентов обучения и контроля знаний. Это позволяет быстро регистрировать ответы и объективно их оценивать по заранее разработанным правилам. Задания в тестовой форме легко вводятся в компьютер, компактно отображаются на экране монитора, хорошо различаются по форме и смыслу. Анализ матрицы ответов обучающихся обеспечивает возможность оценить качество изложения материала, скорректировать содержание и методику образования. Кроме того, используя профиль ответов каждого обучающегося, можно индивидуализировать образование.

Для организации тестирования важное значение имеет выбор места, времени проведения тестирования, обеспечение комфортных условий для испытуемых (невысокий уровень шума, хорошее освещение, оборудованное рабочее место).

В максимальной степени требования к организации тестирования могут быть соблюдены *при использовании компьютера*. Педагогическое тестирование в настоящее время "становится одной из самых актуальных информационных технологий образования"¹⁵. В сочетании с персональными ЭВМ и программно-педагогическими средствами тесты помогают перейти к адаптивному обучению и контролю знаний – высокоэффективным, но и редко применяемым у нас формам организации образовательного процесса.

Тесты применяются в компьютерных обучающих программах: в блоках входного, промежуточного и итогового контроля знаний. Блоки тестового контроля выдают исход-

15. Матушанский Г.У. Проектирование педагогических тестов для контроля знаний // Информатика и образование. – 2000. – № 6. – С. 7 – 11.

ную информацию для адаптивного блока компьютерной обучающей программы, управляющего ходом образовательного процесса. Возможность автоматизации тестирования предоставляет большие возможности для дистанционного контроля и самоконтроля обучающихся.

Использование тестовых заданий в автоматизированных контрольно-обучающих программах позволяет обучающимся самостоятельно обнаруживать пробелы в структуре своих знаний. Это говорит о значительном обучающем потенциале тестовых заданий, использование которого является одним из эффективных направлений практической реализации принципа единства и взаимосвязи обучения и контроля.

Использование компьютерного тестирования придаёт обучению *мощный мотивационный импульс*. Как показывают результаты анкетирования, 64,6% обучающихся считают компьютерный тестовый контроль более объективным в сравнении с устным опросом (18,5%) или письменной контрольной работой (16,9%)¹⁶.

Помимо вышесказанного, *сравнивая бланковые и компьютерные тесты, необходимо отметить следующие преимущества последних:*

- уменьшение затрат на тиражирование материалов;
- обеспечение комфортных условий для работы с тестом в удобное для тестируемого время;
- сокращение числа посторонних факторов, негативно влияющих на достоверность результатов тестирования;
- повышение эффективности тестирования за счет уменьшения продолжительности теста и использования новых типов заданий, обладающих большей проверочной ёмкостью;
- случайная перестановка заданий в тесте или ответов в задании делают невозможным "механическое" копирование номеров правильных ответов;
- доступность результатов тестирования сразу после окончания теста;
- возможность автоматизированного составления матриц ответов и быстрой статистической обработки результатов тестирования на компьютере.

Компьютерные адаптивные тесты предоставляют дополнительные преимущества:

- 1) подстройку под индивидуальные возможности тестируемого;
- 2) повышение точности оценки уровня подготовки тестируемых благодаря использованию банка заданий разного уровня трудности;

16. Кречетников К.Г. Проектирование креативной образовательной среды на основе информационных технологий в вузе. – Москва: Изд-во Госкоорцентр, 2002. – 296 с.

3) упрощение внесения изменений в банк новых тестовых заданий, которые будут автоматически учтены адаптивным алгоритмом.

Однако при компьютерном тестировании *возникает ряд дополнительных проблем*, таких как¹⁷:

- влияние на результаты тестирования навыков работы на компьютере (перехода от бумаги и ручки к монитору, клавиатуре и мыши);
- изменение процедуры и тактики выполнения теста (например, адаптивный алгоритм не позволяет вернуться к пропущенным заданиям);
- необходимость обучения технического персонала и оснащения компьютерных классов необходимым аппаратным и программным обеспечением;
- влияние продолжительной работы за компьютером на здоровье;
- неготовность и нежелание тестируемых выполнять тест на компьютере (возможность выбора бумажного варианта);
- недостаточные возможности ряда персональных компьютеров по представлению информации и ограниченные пропускные возможности телекоммуникационных сетей;
- высокая стоимость организации и проведения тестирования.

В отличие от других педагогических контрольных материалов тестовые задания проходят процесс специальной апробации с обязательной эмпирической проверкой на типичной выборочной совокупности испытуемых и применением статистических методов обработки данных. Для проверки тестовых свойств заданий используется матрица результатов тестирования, представляющая собой результаты испытуемых по всем проверяемым заданиям. Проводится ряд расчетов, позволяющих определить качество теста. Надежность и валидность тестов могут быть значительно повышены, если их качественные и содержательные характеристики будут связаны со статистическими данными, полученными при обработке больших массивов результатов тестирования испытуемых.

Экспертиза педагогических тестовых материалов предусматривает оценку качества их содержания, правильности форм заданий и статистических характеристик. И компьютерные тесты намного удобнее с точки зрения экспертизы, так как могут сразу после тестирования выдавать результаты в матричном виде для статистической обработки.

Таким образом, тестирование является одной из наиболее технологичных форм проведения автоматизированного контроля (самоконтроля) с управляемыми параметрами качества. Применение тестов в составе информационных образовательных технологий обеспечивает возможность индивидуализации обучения, предоставляет каждому

17. Нардюжев В.И., Нардюжев И.В. Модели и алгоритмы информационно-вычислительной системы компьютерного тестирования. – М.: Прометей, 2000. – 148 с.

обучающемуся объективное и надёжное средство самоконтроля, оценивающее успешность его продвижения по оси саморазвития.