

ПЕДАГОГИЧЕСКИЕ ТЕСТЫ КАК ИНСТРУМЕНТ ИЗМЕРЕНИЯ ОЦЕНКИ ЗНАНИЙ СТУДЕНТОВ И КАЧЕСТВА ОБУЧЕНИЯ

Коломийцев Ю.Н.¹

Самарский государственный технический университет
443100, г. Самара, ул. Молодогвардейская, 244
E-mail: kolomytsev@bk.ru

Рассмотрены вопросы практического применения централизованных измерительных педагогических тестов для контроля остаточных знаний студентов. Дан анализ причин, снижающих достоверность и надёжность результатов тестирования. Названы факторы, учёт которых необходим при разработке и проведении тестирования в ходе процедуры аккредитации вуза.

Ключевые слова: педагогические тесты, внешний контроль, цели, педагогические измерения, погрешности, дисциплина, курс обучения.

По мнению исследователей, инновационное инженерное образование невозможно без разработки современной системы внешнего контроля за освоением образовательных программ [1]. Главным инструментом такого контроля в настоящее время является процедура тестирования «остаточных» знаний студентов, которая проводится по конкретной дисциплине через некоторый промежуток времени после окончания курса обучения и сдачи экзамена (зачёта). Контроль выполняется посредством тестов, составленных для данной дисциплины. По результатам тестирования судят как о степени усвоения учебного материала, так и об объёме знаний студентов по данной дисциплине, которые они обнаруживают спустя время. Вместе с тем, как указано в «Методических рекомендациях по разработке педагогических тестов контроля качества обучения студентов» (приложение к СТП 12 100-02), «...основные цели тестирования рассматриваются в связи с задачами внутривузовского управления качеством подготовки специалистов, а также с задачами итоговой аттестации выпускников, государственной аттестации и аккредитации вуза» [2]. Таким образом, тестирование выполняет функции, выходящие далеко за рамки контроля «остаточных» знаний студентов.

Фактически тестирование применяется и как некоторый «арбитражный инструмент» для оценки качества работы преподавателей и педагогического уровня кафедр. Однако такая оценка не может быть достоверной без анализа качества и пригодности самих тестов для проверки конкретного вуза, а также согласования количества и структуры часов, отведённых на дисциплину учебным планом, с содержанием теста.

Каким общим требованиям должен отвечать тест при столь серьёзных задачах, возлагаемых на него?

Тесты не разрабатываются в аттестуемых вузах, а поставляются в них централизованно. Поэтому, формально соответствуя образовательному стандарту специальности, они не учитывают особенностей учебных планов, согласованных с профилирующими кафедрами конкретных вузов. Именно эти особенности отражают интересы производства и работодателей и способствуют успешному трудоустройству будущих инженеров. В соответствии с упомянутыми «Методическими рекомендациями» [2] педагогические тесты должны разрабатываться как измерительные инструменты для объективной оценки как объёма знаний студентов, так и уровня их подготовленности. Тесты должны давать органам управления образованием достоверную информацию об общих результатах обучения, что достигается посредством таких измерительных инструментов и процедур, с помощью которых можно надёжно и точно (с наперед заданной доверительной вероятностью в границах доверительных интервалов) определять объём знаний студентов, а также уровень их подготовленности.

Специалисты указывают на трудности теоретического характера при создании тестов: «...классические результаты теории измерений плохо применимы к измерениям в тестировании» [3], «...в современной литературе нет ясности относительно теоретической основы педагогических измерений» [4]. Эту точку зрения выражают как отечественные (В. Аванесов, Ю.М. Нейман, В.А. Хлебников и др.), так и зарубежные учёные (классические работы Лорда, а также распространённая на Западе математическая теория измерения латентных качеств личности –

¹ Коломийцев Юрий Николаевич (к.т.н., доцент), каф. теоретической и общей электротехники.

Latent Trait Theory (LTT)). Степень достоверности тестового контроля обусловлена наличием объективно существующих погрешностей в данных на всех этапах тестирования. В [5] выделяются основные источники погрешностей педагогических измерений посредством тестов:

1. Недоступность для прямого измерения как уровня подготовленности тестируемого, так и уровня трудности задания. Эти две величины тесно связаны между собой, тогда как законы математической статистики основываются на гипотезах независимости оцениваемых величин.

2. Использование при оценке латентных параметров (с использованием модели Раша) гипотезы нормального распределения баллов, что справедливо только в случае применимости закона больших чисел к обрабатываемым данным. Это не может быть соблюдено в процедуре тестирования.

3. Необходимость использования больших выборок при тестировании ($N \gg 100$) для получения достоверных результатов. Реальные выборки имеют гораздо меньший объем, а их увеличение невозможно из-за роста трудоемкости теста.

Таким образом, даже с точки зрения теории проблема создания идеальных тестов в настоящее время не решена. Тем не менее, доступность, удобство и простота компьютерной процедуры тестирования по любой дисциплине делают тестовый метод оценки знаний студентов доминирующим.

Проанализировав опыт проведения тестовых проверок и оставив в стороне дискуссионные вопросы современных теорий педагогических измерений, сформулируем ряд методических рекомендаций и требований к тестовым заданиям с целью повышения достоверности и надёжности результатов тестирования.

Уровень сложности заданий по каждому разделу курса:

1. Должен учитывать, является тестируемая дисциплина специальной или общей для проверяемой специальности.

2. Должен учитывать особенности конкретного курса, которые, как правило, согласовываются с профилирующими кафедрами. Например, для электротехнологических и материаловедческих специальностей не требуется углубленное изучение электропривода и магнитных цепей, поэтому сложность вопросов по этим разделам курса должна быть минимальной.

3. Должен определяться глубиной изучения материала, т.е. не может устанавливаться без учёта количества часов, отведённых на изучение данного раздела для студентов конкретной специальности в конкретном вузе.

4. Должен учитывать структуру часов, отведённых на изучение дисциплины, в частности наличие или отсутствие практических и лабораторных занятий, контрольных и курсовых работ по данной дисциплине. Если в данном вузе таковые не предусмотрены, задача теста – выявить уровень овладения теорией, поэтому тест не должен содержать заданий на решение задач. Если таковые предусмотрены, в тест включаются задачи, решение которых не требует применения вычислительных средств и не перегружено вычислительными операциями.

5. С учётом времени, отведенного на тест (обычно это один академический час), и количества вопросов теста (обычно от 18 до 30) количество задач в тесте не должно превышать число разделов (6-10 разделов в курсе, по одной на раздел). Это требование логически обусловлено п. 6.

6. Согласно «Методическим рекомендациям» [2] на ответ на один вопрос тестового задания у студента должно уходить не более 2-5 минут.

7. Время проведения тестовых проверок должно согласовываться с учебным планом конкретного вуза. Принято назначать тестовые проверки после окончания курса обучения по данной дисциплине. При этом из-за разницы в учебных планах специальностей по одной и той же дисциплине одновременно проходят тестирование студенты, сдавшие экзамен и полгода, и год тому назад. Очевидно, результаты будут разные, хотя критерии оценки одинаковые. Всё это не учитывается централизованными тестами. Представляется правильным проводить тестирование не после окончания курса по данной дисциплине, как это сейчас принято, а после сдачи экзамена по ней. В некоторых случаях эти сроки разнесены более чем на два семестра, что ставит под сомнение результаты тестирования.

Из проблем, которые обнаруживаются при централизованном тестировании и снижают его достоверность, следует отметить:

1. Опечатки (ошибки) в материалах тестирования. Практика показывает, что их процент достаточно высок, и задания оцениваются проверяющими как неверно решенные. Следует подчеркнуть, что таких ошибок не возникает при разработке и изготовлении материалов в самом проверяемом вузе.

2. Очевидно, что любой тест должен соответствовать образовательному стандарту по данной дисциплине. Это условие необходимое, но не достаточное. Известно, что в зависимости от ресурсов конкретного вуза даже в рамках одного образовательного стандарта для одной и той же дисциплины, но для разных специальностей лекционный и другой материал дифференцируют, что повышает качество обучения. Это отражается в учебно-методическом комплексе, который составляется для каждой дисциплины и специальности и, как правило, согласовывается с профилирующими кафедрами. В некоторых вузах при недостатке и экономии ресурсов лекционные потоки объединяют и дифференциация учебного материала по разным специальностям становится невозможной, т.к. студентам разных специальностей читается один и тот же материал. Обе эти особенности никак не учитываются при централизованном тестировании, что, безусловно, снижает его достоверность.

Заключение.

Важнейшими преимуществами тестовой системы контроля знаний являются:

1. Стандартизация процедуры проведения проверки показателей качества заданий и тестов в целом.

2. По сравнению с экзаменами тестирование ставит всех студентов в одинаковые условия как в процессе контроля, так и в процессе оценки, исключая субъективизм преподавателя. По данным английской аттестационной ассоциации NEAB, тестирование позволяет снизить количество апелляций более чем в три раза, сделать процедуру оценивания одинаковой для всех учащихся вне зависимости от места проживания и типа образовательного учреждения.

3. Широта охвата дисциплины: тесты включают в себя задания по всем темам дисциплины, в то время как в экзаменационный билет обычно включают вопросы по 2-5 темам. Это позволяет выявить знания учащегося по всему курсу, исключив элемент случайности, который имеет место при вытаскивании билета. При помощи тестирования любой проверяющий, даже не специалист в данной области, может установить уровень знаний студента по предмету в целом и по отдельным его разделам.

Однако необходимо отметить и недостатки:

1. Разработка качественного тестового инструментария – длительный, трудоемкий и дорогостоящий процесс, что усугубляется некорректностью повторного применения теста из-за «эффекта привыкания». Иначе говоря, каждый тест должен быть разового применения, т.е. для тестирования студентов одной специальности по одной дисциплине в разных группах требуется целый набор тестов равной сложности. Кроме того, необходимо обеспечить конфиденциальность тестовых заданий.

2. Результаты тестирования, хотя и включают в себя информацию о пробелах в знаниях по конкретным разделам, не позволяют судить о причинах пробелов; это легко выясняется преподавателем при экзаменационной форме контроля.

3. Широта охвата тем при тестировании и жёсткое ограничение времени на каждый ответ (при продолжительности тестирования в один академический час и при 18-24 вопросах – менее трёх минут на один ответ), в отличие от экзамена, не позволяет студенту провести сколько-нибудь глубокий анализ вопроса теста, что искажает результаты и вносит в тестирование элемент случайности (угадывания).

Из сказанного следуют три простых вывода:

1. Для повышения информативности тестов, их надёжности и достоверности обязателен учёт особенностей учебных программ проверяемого вуза. Практика показывает, что централизованные тесты по большинству дисциплин эти особенности не учитывают.

2. Наиболее целесообразно разрабатывать тесты с привлечением специалистов проверяемых вузов.

3. В своём настоящем виде содержание централизованных педагогических тестов и процедура проведения тестирования не отвечают в полной мере функциональному назначению тестов как «арбитражного инструмента» для оценки остаточных знаний студентов и для аттестации и аккредитации проверяемых вузов.

СПИСОК ЛИТЕРАТУРЫ

1. Бутенков С.А., Сальников В.А., Бутенков Д.С. Методика и средства индивидуального тестирования в вузе. – Таганрог: Изд-во ТРТУ, 2005. – 21 с.
2. Методические рекомендации по разработке педагогических тестов контроля качества обучения студентов (приложение к СТП 12 100-02). – М.: Центр тестирования МО РФ, 2002. – 152 с.

3. Нейман Ю.М., Хлебников В.А. Педагогическое тестирование как измерение. – Ч. 1. – М.: Центр тестирования МО РФ, 2002. – 68 с.
4. Аванесов В.С. Композиция тестовых заданий. – М.: Центр тестирования, 2002. – 52 с.
5. Хлебников В.А. Вводная статья // Вопросы тестирования в образовании. - 2002. - № 2.

Поступила в редакцию – 21/I/2012,
в окончательном варианте – 21/II/2012.

UDC 371.263

PEDAGOGICAL MEASURING TESTS AS AN INSTRUMENT OF STUDENTS' KNOWLEDGE ESTIMATION AND EDUCATION QUALITY CONTROL

Yu. N. Kolomytsev

Samara State Technical University
244 Molodogvardeiskaya st., Samara, 443100
E-mail: kolomytsev@bk.ru

The article describes the points of the practical application of centralized pedagogical measuring tests for the residual control of the students' knowledge. The article also gives the analysis of the factors which decrease the trustworthiness and reliability of the test results. The article mentions the factors which should be considered by the elaboration and carrying out of the tests during the procedure of the high school accreditation.

Key words: *pedagogical tests, external control, purposes, pedagogical measurings, discipline, errors, course.*

Original article submitted – 21/I/2012,
revision submitted – 21/II/2012.

Yuri N Kolomytsev. Candidate of technical science, associate professor of the chair of theoretical and common electrotechnology of the Samara State Technical University.