

ПРОЕКТИРОВАНИЕ И РЕАЛИЗАЦИЯ КОМПЬЮТЕРНОГО КОНТРОЛЯ УРОВНЯ ЗНАНИЙ СТУДЕНТОВ-МЕДИКОВ ПО АНГЛИЙСКОМУ ЯЗЫКУ

Наталья Чернышева
Петрозаводский государственный университет
formed@psu.karelia.ru

Опубликовано в ж. Педагогические Измерения №1 2011 г.

Аннотация

В статье рассматриваются основные принципы создания тестов в применении к проектированию и практической реализации системы компьютерного контроля знаний студентов-медиков по английскому языку. Приводятся особенности построения тестовых заданий, а также рассматривается научное обоснование теста в рамках классической теории тестов. Описывается разработанная методика определения статистических характеристик теста. Дается краткое описание тестов, входящих в комплекс: исходный, текущие, рубежные и итоговый.

Ключевые слова: тест, тестовое задание, статистические характеристики.

С чего начинается разработка теста

Очевидно, любой разработке предшествует желание внести изменения в обычный процесс обучения, которые помогут сделать его более эффективным, а также интересным для студентов. Одним из направлений совершенствования учебного процесса является осуществление четкого контроля за качеством усвоения материала. Традиционные методы контроля при обучении иностранным языкам, такие, как устный опрос и письменные контрольные работы имеют существенные недостатки. Прежде всего, это субъективность оценки уровня знаний. Кроме того, их использование связано со значительными затратами труда и времени преподавателя не только на подготовку материалов для контроля, но и на проверку и анализ работ студентов. Получить необходимую информацию позволяет контроль на основе использования программных средств ИКТ,

предусматривающий компьютерную обработку данных тестирования и представления результатов обработки.

Следует заметить, что иногда тестом ошибочно называют любую проверочную работу, будь то словарный диктант или контрольная работа по грамматике. Здесь имеется в виду следующий вид контроля: «тест - это научно-обоснованный метод, представляющий систему заданий специфической формы, возрастающей трудности, определенного содержания, позволяющий качественно оценить структуру знаний и эффективно измерить их уровень»¹. [1] Но это то, что должно получиться в идеале.

Практическое использование тестов связано с решением многих серьезных проблем. Прежде всего, это подготовка материала для теста. Т.к. стандартных наборов тестов для большинства дисциплин не существует, возникает необходимость самим преподавателям разрабатывать тесты. К тому же, практически невозможно найти тест, который по содержанию и по форме соответствовал бы всем требованиям преподавателя.

Разработка теста начинается с постановки цели тестирования и составления плана теста. Это помогает сосредоточиться на тех основных грамматических и лексических вопросах, знание которых ставится задачей проверить с помощью данного теста. Программа обучения по предмету должна быть оптимально отражена в тесте. В неязыковом вузе обучение иностранному языку приобретает профессиональную направленность. В нашем случае это медицина. Следовательно, перед преподавателями стоит задача: подготовить будущих докторов к работе с медицинской литературой на иностранном языке и сформировать умение использовать устные формы общения в рамках специальности. Для овладения этими видами деятельности необходимо усвоение языкового материала,

¹ Аванесов В.С. Основы педагогических измерений// Педагогические измерения, №1, 2004, с.15-21.

характерного для медицинской научной литературы: модели предложений, грамматические формы, лексические единицы, и т.д. Таким образом, основной задачей является отбор знаний, умений и навыков по иностранному языку в качестве объектов контроля с помощью тестов.

Постановка цели тестирования зависит от того, будет ли данное тестирование осуществлять текущий контроль знаний студентов, рубежный контроль по отдельным темам или разделам, итоговый контроль по всей программе учебной дисциплины. Этот список следует дополнить очень важным видом контроля – исходным, который позволяет быстро и объективно оценить готовность студентов к восприятию программы обучения на данном этапе и, что не менее важно, позволяет оценить в процессе обучения эффективность работы студента, группы студентов и самого преподавателя. В Табл. 1 представлен пример плана теста.

Таблица 1. Пример плана теста (1 курс, 1 семестр).

Содержание предмета	Количество заданий
Грамматика	19
включая: Indefinite Tense	3
Continuous Tense	2
Согласование времен	1
Причастия	2
Степени сравнения прилагательных	2
Модальные глаголы	1
Страдательный залог	3
Артикль	1
Союзы	1
Местоимения	3
Лексика	7
В с е г о:	26

Тест по грамматике не должен содержать сложного лексического материала. При этом грамматика должна быть представлена в том объеме, в каком она изучалась в данном курсе. К примеру, в тест по временам английских глаголов для неязыковых специальностей не должен включаться задания, требующие знаний всех тонкостей использования глагольных форм.

Оптимальное отображение содержания теста предполагает также выбор соответствующей тестовой формы. Выбор формы тестового задания – полностью прерогатива преподавателя. Самой популярной формой, как среди российских, так и зарубежных разработчиков теста, является задание с выбором правильного ответа из числа предложенных (закрытой формы). Например, задание на выбор правильной глагольной формы.

Нажимайте на клавишу с номером правильного ответа:

1. THE PROBLEM .

- 1) is studied
- 2) is studying
- 3) studied
- 4) are studied

AT THE LABORATORY.

Или задание на использование предлога.

2. DO YOU USUALLY GO OUT

- 1) on
- 2) in
- 3) at
- 4) for

FRIDAY EVENING?

Однако и здесь можно разработать более интересные варианты, предложить тестируемому выбрать правильный перевод предложения. При этом в любом неязыковом вузе обучение иностранному языку обязательно приобретает профессиональную направленность.

3. ПРАВИЛЬНЫЙ ПЕРЕВОД ПРЕДЛОЖЕНИЯ «ПО-ВИДИМОМУ, ИЗМЕНЕНИЯ ПОГОДЫ ВЛИЯЮТ НА СОСТОЯНИЕ НАШЕГО ЗДОРОВЬЯ

- 1) Weather changes appear to influence the state of our health.
- 2) Weather changes appear as the influence on the state of our health.
- 3) Weather changes and influences the state of of our health.
- 4) Weather changes are influenced by the state of our health.

В данном случае для того, чтобы выбрать правильный перевод, надо проанализировать все четыре предложения.

Усовершенствованным вариантом этой формы заданий являются задания с выбором нескольких правильных ответов. В этом случае вариантов ответов может быть предложено больше. Например, задание:

Выберите все правильные ответы.

4. WHAT BONES DOES THE SKELETON OF THE TRUNK CONTAIN?

- 1) the skull
- 2) the spinal column
- 3) the ribs
- 4) the breastbone
- 5) the foot
- 6) the vertebra
- 7) the extremity.

Можно включать задания на составление пар глаголов, имеющих похожее написание, но противоположное значение, таких, как to improve (улучшать) - to impair (ухудшать, нарушать). Следует подчеркнуть, что предпочтительным является контроль лексики в контексте, без использования русского языка. Дело в том, что беспереводные способы семантизации (определения значения) слова развивают догадку, увеличивают практику в языке, создают опоры для запоминания, усиливают ассоциативные связи. Знание учащимися значений каждого слова группировки позволяет выполнить тестовое задание, например:

5. НЕ ЯВЛЯЕТСЯ СИНОНИМОМ ДЛЯ ТРЕХ ОСТАЛЬНЫХ СЛОВО

- 1) important
- 2) subsequent
- 3) significant
- 4) considerable

В тех случаях, когда варианты для выбора соответствия состоят не из одного, а из группы слов, более удобно и более понятно для тестируемых использование заданий в фасетном варианте, когда тот же набор вариантов для выбора предлагается в нескольких заданиях, в каждом случае для иной ситуации. Например,

Выберите все правильные ответы.

6. WHAT TAKES PLACE WHEN ONE BREATHES IN?

- 1) the diaphragm contracts and goes down
- 2) the diaphragm becomes relaxed
- 3) the volume of the chest increases and lungs extend
- 4) the volume of the chest decreases and the lungs contract
- 5) the pressure in the lungs becomes less
- 6) the pressure in the lungs becomes higher

7. WHAT TAKES PLACE WHEN ONE BREATHES OUT?

- 1) the diaphragm contracts and goes down.
- 2) the diaphragm becomes relaxed.
- 3) the volume of the chest increases and lungs extend.
- 4) the volume of the chest decreases and the lungs contract.
- 5) the pressure in the lungs becomes less.
- 6) the pressure in the lungs becomes higher.

Что касается количества тестовых заданий, то здесь приходится искать компромисс между желанием проверить усвоение всего пройденного материала подробно и разумной длиной теста. Максимальное число заданий 60–80. Однако, обычно тесты, рассчитанные на 30 – 45 минут, включают 25 – 30 заданий.

Пока в комплексе тестов не нашлось места заданиям открытой формы; возможно, и для них придет время. Следует отметить, что тест не является застывшей структурой, изменения заданий по мере необходимости могут вноситься. Кроме того, используемый нами сервер не

дает возможности использовать в качестве заданий аудио материалы, поэтому нет возможности проверить навыки аудирования.

Итак, предположим, что цель определена, материал подобран, форма заданий выбрана. Однако, то, что получилось в результате, правильнее определить как набор заданий в тестовой форме или «предтест». К сожалению, часто считается, что на этом разработка теста заканчивается. Показательны в этом смысле результаты исследования, проведенного методом анкетирования преподавателей профессионального образования по уровню владения компьютерными программами, позволяющими создавать тестовые задания. Так, 80% преподавателей не используют программы статистической обработки данных и только 10% имеют опыт создания тестовых заданий для компьютерного тестирования². Думается, что это не исключение, а скорее правило. Устранение данной проблемы авторы совершенно справедливо видят в организации курсов по разработке электронных образовательных средств, в том числе и тестовых программ, что, в частности, и активно практикуется в ПетрГУ.

Научное обоснование теста

В отличие от традиционных средств контроля тесты должны быть научно обоснованы, что предполагает оценку соответствия характеристик теста – надежности и дифференцирующей силы заданий - определенным критериям. Только качественные тестовые материалы, обладающие достаточным уровнем надежности, обеспечивают эффективное тестирование. Для оценки качества теста проводится его апробирование на репрезентативной выборке и обработка результатов тестирования.

Существует два варианта организации тестирования: безмашинный, т.е. бумажный и машинный, т.е. компьютерный. Несомненными

² Материалы конференции «Тестирование в сфере образования: проблемы и перспективы развития», Красноярск, 2010, с. 160.

преимуществами компьютерного тестирования являются такие, как снижение до минимума затрат труда и времени на проверку работ, повышает оперативность проверки, а также вызывает большой интерес у обучающихся.

Однако, для проведения компьютерного тестирования, помимо соответствующей технической оснащённости, необходимо иметь, по крайней мере, три компонента: структуру электронной базы данных тестовых заданий, программное обеспечение для внесения тестовых заданий в электронную базу данных, программное обеспечение для проведения предварительной обработки ответов. Помимо всего, требуется наличие определенной компьютерной грамотности как у обучающихся, так и у преподавателей, участвующих в проведении тестирования.

Следует отметить, что время примитивных самодельных программ уже ушло, и на рынке программного обеспечения предлагается немало вариантов для проведения автоматизированного контроля знаний. В качестве основного инструмента для реализации контроля знаний в форме компьютерного тестирования в нашем университете используется информационная среда WebCT (Web Course Tools). WebCT обладает рядом преимуществ, таких, как понятный как для разработчиков, так и для пользователей интерфейс; предлагает большой набор различных сервисов, как для проведения дистанционного обучения, так и для мониторинга учебного процесса.

Система контроля знаний состоит из следующих компонентов:

- 1) программа, позволяющая проходить тест студентам;
- 2) программа, выполняющаяся на сервере и собирающая результаты тестов в базу данных;
- 3) программа для конструирования новых тестов и изменения существующих;

4) программа для просмотра результатов тестов и формирования статистики.

Подготовленный контрольный тест помещается на сервер дистанционных ресурсов Web СТ, и может использоваться на занятиях преподавателями английского языка. Преподаватель может выбрать режимы тестирования, в частности ограничение времени тестирования, и способ отображения результатов тестирования для студентов: только итоговый процент выполнения или весь протокол тестирования. Использование режима «случайный порядок ответов на задания» означает, что даже при последовательном тестировании порядок следования ответов будет другим.

В разделе «Представление» можно просмотреть результаты всей группы тестируемых, отдельных групп и каждого студента в режиме «Общая статистика». Режим «Статистика по вопросам» позволяет оценить выполнение отдельных заданий теста всеми студентами и отдельной группой, так что преподаватель может сразу оценить насколько трудным является то или иное задание для данного контингента студентов (Табл. 2).

Таблица 2. Итоги выполнения теста одной из групп в режиме «Статистика по заданиям».

№№ Заданий	Количество	Правильный:			Классификация	Среднее значение	Стандартное отклонение
		Вся группа	Лучшие 25%	Худшие 25%			
1-1	10	80	100	33	0.59	80.0%	42.2
1-2	10	90	100	66	0.26	90.0%	31.6
1-4	10	50	100	33	0.61	50.0%	52.7
1-5	10	50	66	33	0.25	50.0%	52.7
1-10	10	60	66	33	0.47	60.0%	51.6
1-6	10	80	100	100	0.03	80.0%	42.2
1-14	10	40	66	33	0.20	40.0%	51.6
1-12	10	50	66	0	0.61	50.0%	52.7
1-8	10	40	66	33	0.49	40.0%	51.6

1-9	10	100	100	100	0.00	100.0%	0.0
1-11	10	50	100	0	0.53	50.0%	52.7
1-15	10	60	100	33	0.38	60.0%	51.6
1-16	10	30	33	0	0.34	30.0%	48.3
1-17	10	90	100	66	0.26	90.0%	31.6

В таблице1 в разделе в колонке «Количество правильных ответов в % (вся группа)» представлена уже рассчитанная трудность задания, в колонке «Классификация» даны коэффициенты дифференцирующей силы заданий, определенные по методу «крайних» групп. Как показали проведенные нами расчеты, значения индексов дифференцирующей силы заданий с достаточной точностью совпадают с коэффициентами, определенными как точно-бисериальные коэффициенты.

В современной литературе достаточно подробно изложены основные этапы расчета статистических характеристик теста в рамках классической теории тестов. Однако далеко не всегда у преподавателя имеется достаточно времени для освоения математического аппарата, лежащего в основе этих расчетов, в особенности у преподавателей гуманитарного цикла. В результате, не смотря на широкое распространение тестового контроля знаний студентов, вопрос об оценке качества тестовых материалов остается нерешенным в полной мере. В литературе или указывается, с помощью какой программы выполняется статистическая обработка, или приводится весь математический аппарат с большим количеством формул.

Поэтому одним из направлений данной работы было создание такой методики, которая позволила бы любому преподавателю, вне зависимости от его математической подготовки и уровня владения компьютером, проверить качество разработанного им теста. Самой часто упоминаемой и используемой в отечественных статьях программой является приложение MS Excel из пакета офисных программ компании Microsoft MS Office.

Причины столь широкого распространения этого программного обеспечения - это наличие русскоязычной версии, тесная интеграция с MS Word и Power Point. Кроме того, MS Excel - это электронные таблицы с достаточно мощными математическими возможностями со встроенными статистическими функциями. Все это дает возможность использовать Excel для расчета статистических характеристик тестов.

Программа состоит из двух этапов, которые подробно описаны:

импорт результатов теста из программной среды WebCT в таблицы Excel;

выполнение расчетов с помощью программы Excel.

На стадии импорта выполняется преобразование таблицы итогов тестирования в дихотомическую матрицу. Для этого в таблице общих результатов заменяются буквенные ответы в соответствии с таблицей общих итогов (Табл.3) на 1 или 0. В таблице общих итогов курсивом выделены правильные ответы.

Таблица 3. Таблица общих итогов.

Заголовок	Количество	правильных ответов в %:			Классификация	Частота					
		Вся группа	Лучшие 25%	Худшие 25%		-	a	b	c	d	
1-1	140	83	97	60	0.51	0	11	7	5	12	6
1-10	140	72	97	46	0.50	0	3	102	19	16	
1-11	140	35	60	8	0.42	0	50	15	66	9	
1-2	140	65	88	40	0.45	0	14	92	14	20	
1-4	140	50	76	24	0.41	0	71	21	16	32	
1-5	140	49	69	35	0.31	0	69	57	6	8	
1-6	140	88	97	84	0.26	0	3	124	3	10	
1-14	140	35	58	20	0.31	1	49	50	24	16	
1-12	140	60	93	40	0.40	1	85	24	8	22	
1-8	140	59	76	60	0.14	1	6	46	83	4	
1-9	140	74	93	53	0.41	0	12	104	13	11	
1-15	140	67	90	40	0.48	0	94	9	9	28	
1-16	140	59	76	35	0.42	0	83	9	32	16	

1-17	140	66	90	33	0.51	0	93	10	22	15
1-18	140	63	86	44	0.45	0	18	11	89	22

На второй стадии выполняются расчеты общего балла, набранного студентами в тесте, среднего значения тестовых баллов (функция СРЗНАЧ), отклонений от среднего, квадратов отклонений от среднего и их сумма, суммы баллов по отдельным вопросам теста, дисперсия (функция ДИСПР), а также вариация тестовых баллов и надежность теста. В методике дано пошаговое описание всех процедур вычислений. Надежность определяется по методу Kuder-Richardson³. Следует заметить, что использование данной методики помогает и тем, кто хорошо знаком с математическим аппаратом оценки качества теста, но эти работы выполняет время от времени.

В литературе даются различные значения допустимых значений коэффициента дифференцирующей способности задания, т.е. значений, определяющих, является ли задание тестовым, от 0,15 до 0,3. Однако, всеми принимается, что при корреляции свыше 0,4 вероятность включения задания в тест заметно возрастает. Низкий коэффициент корреляции помогает выявить сомнительные задания, которые вряд ли могут считаться тестовыми, эти задания устраняются из проектируемого теста или изменяются. К примеру, в нашем случае (Табл.2) задание 1-8 обладает слабой дифференцирующей способностью, поэтому оно вряд ли будет тестовым.

Кроме того, по частоте выбора того или иного ответа, выбранного в задании, можно судить о том, как «работает» тот или иной дистрактор. Дистрактор, который не был выбран ни одним из тестируемых, следует изменить или, по возможности, удалить.

³ Цит.: В.С. Аванесов, Проблемы качества педагогических измерений, 2008, <http://testolog.narod.ru/EdMeasmt2.html>

Что касается надежности тестовых результатов, то мнения по поводу минимально приемлемого значения расходятся. Это 0,7 или 0,6 (диапазон 0 – 1). Как правило, надежность качественных тестовых результатов колеблется в пределах (0,8 - 0,95). Следует отметить, что сама по себе надежность не является самоцелью, но достаточно высокая надежность теста определяет точность измерений и устойчивость результатов теста к действию случайных факторов.

Комплекс тестов

Необходимость систематического контроля усвоения материала не вызывает сомнений. Как показали исследования российских и зарубежных ученых, компьютерное тестирование - это эффективный метод контроля знаний студентов, который помогает сделать процедуру контроля более привлекательной для обучающихся и в то же время обеспечивает объективность оценки знаний обучающихся. Кроме того, считается, что у человека, адаптированного к работе с заданиями в тестовой форме, результаты компьютерного тестирования на 11% выше, чем у того, кто сталкивается с ними впервые⁴.

Наиболее ярко все преимущества использования компьютерного тестирования проявляются на начальном этапе обучения, т.к. определение исходного уровня владения иностранным языком с помощью теста на первом занятии дает возможность преподавателю получить объективное представление о знаниях каждого учащегося и группы в целом и реально сопоставить достижения студентов с исходными результатами.

Начиная с 2004 года, исходное тестирование по английскому языку проводится на медицинском факультете ПетрГУ. В связи с отсутствием готовых тестов, соответствующих нашим требованиям, был составлен

⁴ Анастази А. Психологическое тестирование, - М., 1982.

авторский тест. Грамматический аспект был выбран, как базовый при оценке владения иностранным языком⁵. Кроме того, уровень подготовки студентов различается, поэтому все задания основаны на несложной лексике.

При планировании теста руководствовались теми вопросами, которые являются определяющими на переходном этапе от школы к вузу и входят в вводно-коррективный курс, например: использование артиклей, предлогов места и времени, неопределенного времени глаголов, модальные глаголы и т.д. Тест состоял из заданий с выбором одного правильного ответа из числа предложенных. Разработанный тест прошел экспериментальную апробацию с целью проверки его качества с помощью статистических методов, на основании полученных результатов тест корректировался.

Результаты тестирования позволили выявить те задания, который слабо изучены в школьном курсе и на которые следует обратить особое внимание в процессе дальнейшего обучения, а также общий уровень знаний. Окончательный вариант теста, внедренного в учебный процесс, состоит из 25 заданий и полностью соответствует классическим требованиям, предъявляемым к тестам. Надежность теста – 84%, дифференцирующая сила заданий – 0,21 – 0,62. Спецификация теста приведена в Приложении 1.

На этапе апробирования и внедрения исходного теста отношение преподавателей к этому новому методу было неоднозначно. Сначала было непонимание цели данного вида работы, недоверие к точности оценки знаний студентов, определенной этим методом, отношение к тесту, как к виду работы, который просто разнообразит урок.

⁵ Kitao S.R. and Kitao K. Testing Grammar/ TESL-EJ II (6, June):1-3, 1996.

Переход на компьютерное тестирование вызвал большой интерес и преподавателей, и студентов, однако породил новые проблемы: неготовность преподавателей к работе с группой студентов в компьютерном классе, различный уровень компьютерной грамотности студентов. После разработки подробной инструкции и для преподавателей, и для студентов, проходящих тестирование, все проблемы такого рода были разрешены. Показателем востребованности исходного тестирования может служить количество студентов, которые его проходили в течение ряда лет. Эти цифры представлены в таблице 4.

Таблица 4. Количество студентов, проходивших исходное тестирование на медицинском факультете.

Год	2004	2005	2006	2007	2008	2009	2010
Студенты	49	34	42	140	148	160	183

В практике обучения иностранному языку требует совершенствования текущий контроль, так как он позволяет не только улучшить обратную связь, контролировать ход усвоения студентами учебного материала, но и производить коррекцию процесса обучения. Для текущего контроля знаний были разработаны тесты по таким темам, как «Времена английских глаголов», «Предлоги» и т.д. Одним из достоинств этих тестов является то, что они не включают профессиональную лексику, поэтому могут использоваться для контроля знаний по данным темам в широком диапазоне специальностей

Тест по теме «Предлоги» включает вопросы с выбором одного правильного ответа из нескольких. Поскольку раздел посвящен сопоставлению трех основных предлогов, используемых как предлоги места, так и времени in, on и at, то в разных ситуациях были предложены

именно эти три варианта. Во вторую часть включены задания с выбором одного правильного ответа из четыре возможных.

Наиболее сложным является подготовка надежного и обоснованного контрольного материала для рубежного и итогового тестирования, который мог бы обеспечить объективную оценку уровня знаний студентов, а также эффективную обратную связь в управлении процессом обучения. Тесты для рубежного и итогового контроля для студентов медицинского факультета включают задания трех видов: лексические, грамматические и задания на понимание текста. Целью тестирования в общем смысле является проверка готовности студентов к работе с неадаптированной медицинской литературой на английском языке. В более узком смысле - это проверка готовности студента к экзамену, по результатам выполнения итогового теста студент допускается до сдачи экзамена, который проводится по традиционной методике. Основные характеристики тестов представлены в Табл. 5.

Таблица 5. Основные характеристики тестов комплекса.

Тест	Кол-во заданий	Форма заданий			Стат. характеристики теста		
		1 ответ	несколько ответов	соотв.	Трудность заданий	Дифф. Способность	Надежность
1. Предлоги	25	25	нет	нет	0,43-0,95	0,15-0,52	82%
2. Времена глаголов	24	21	1	2	0,49-0,86	0,15 -0,56	87%
3. Рубежный 1 курс, 1 семестр	26 (19 -гр, 7 - л)	26	нет	нет	0,37-0,74	0,20-0,52	76%
4.	28	21	4	3	0,34-0,95	0,25 - 0,54	77%

Рубежный 1 курс 2 семестр	(12 -гр, 13 - л, 3 - т)							
5. Итоговый 2 курс 2 семестр	25 (10 -гр, 13 - л, 2 - т)	17	1	7	0,33-0,85	0,20-0,65	84%	

В графе «Количество заданий» указано, сколько грамматических (гр.), лексических (л.) заданий и заданий по тексту (т) включено в тест. Трудность заданий определена как отношение числа правильных ответов на задания к общему числу ответов, т.е. тестируемых.

Тестирование проводится в компьютерном классе в присутствии преподавателя. Продолжительность тестирования – 45 минут. После завершения тестирования в режиме Просмотр результатов студент получает свой результат: количество набранных баллов и процентное соотношение правильных ответов к общему числу заданий.

Особое место занимает оценивание студентов по результатам тестирования, т.е. перевод процентного соотношения в оценку по 5-балльной шкале. И тут, судя только по материалам недавней конференции по тестированию⁶, имеются существенные разногласия. Зачетный интервал предлагается определить как 20 – 100%, 40 – 100%, 50 – 100% и т.д. Однако общеизвестно, что об усвоении знаний можно говорить, если коэффициент усвоения имеет значение $K > 0,7$ (или 70%). Исходя из этого, оценивались результаты рубежного и итогового тестирования: менее 70% - «2», 70 -80% - «3», 80-90% - «4», 90-100% - «5». Студенты, которые не справились с тестом,

⁶ Материалы конференции «Тестирование в сфере образования: проблемы и перспективы развития», Красноярск, 2010, с. 160.

получают персональный план повторения пройденного материала и тестируются повторно. Таким образом, появляется возможность индивидуализации подхода к обучаемым.

Заключение

Разработка качественных тестов требует больших затрат времени и труда, которые компенсируются в значительной мере при их применении в учебном процессе. Для подготовки тестов наряду с владением предметом от преподавателя требуются знания основных принципов составления тестов, статистических методов проверки их качества и соответствующего уровня владения компьютерными программами.

Проверка качества тестов по результатам его апробирования является неотъемлемой частью разработки теста. Разработанная методика вычисления параметров теста в таблицах Excel позволяет облегчить эту задачу.

Наряду с такими общепризнанными преимуществами компьютерного тестирования, как технологичность, сокращение затрат времени и труда преподавателей, стандартизация процедуры контроля знаний в пределах факультета позволяет объективно оценить эффективность работы отдельных студентов, студенческих групп, а также эффективность работы преподавателей. В продолжение работы планируется разработка рубежного теста (1 семестр 2 курса), а также текущих тестов по отдельным темам.

Приложение 1.

Тест «Определение исходного уровня владения английским языком.

Общие данные

1. Предмет – Английский язык.
2. Разработчик – Чернышева Н.Д.
3. Период разработки – 2003- 2007 г.г.
4. Апробация теста – 2004, 2005, 2006, 2007г.г.

Спецификация

1. Цель создания теста	Входная аттестация
2. Специальности, для которых планируется использование теста	Медицинский факультет: спец. лечебное дело, педиатрия, фармация
3. Перечень исходных документов, использованных при разработке теста	УМКД по английскому языку, составлена в соответствии с ГОС ВПО специальности 060101(лечебное дело) и 060103 (педиатрия)
4. Вид теста	Гомогенный
5. Подход к разработке теста	Нормативно-ориентированный
6. Число заданий	25
7. Количество и процентное содержание заданий каждой формы	100% с выбором ответов из числа предложенных
8. Время выполнения теста	45 минут

Статистические характеристики заданий и теста

1. Трудность заданий (% правильных ответов)	40 – 80%
2. Статистические данные о дистракторах	Все дистракторы ненулевые
3. Дифференцирующий уровень заданий (точечно-бисериальный коэффициент корреляции)	90% - выше 0,4
4. Результаты проверки правдоподобности гипотезы о нормальном распределении эмпирических результатов (для тестов с нормативно-ориентированной интерпретацией результатов)	Распределение результатов теста соответствует нормальному, $X_{ср.} = 15,46 = 3S_x$ ($S_x = 5,25$); асимметрия - 0,2
5. Надежность тестовых результатов	84%
6. Валидность оценок	Содержательная, тест соответствует программе вводно-коррективного курса
7. Анализ корреляционных связей заданий	$K > 0$
8. Трудность теста	14.41 (ср. значение баллов – 15.46)