

## **КОМПЬЮТЕРНОЕ ТЕСТИРОВАНИЕ В КОНТРОЛЕ ЗНАНИЙ ЭКОЛОГОВ-ПРИРОДОПОЛЬЗОВАТЕЛЕЙ**

В. П. Новиков

### **Введение**

Внедрение в учебный процесс тестовой формы контроля знаний в Югорском государственном университете проводилось в рамках общего направления перехода к контролю знаний в вузах России на уровень международных требований в соответствии с Болонским соглашением. Разработка тестовых заданий (ТЗ) начата в октябре 2005 года после организации в ЮГУ лаборатории инструментальных средств обучения и выхода распоряжения по университету «О разработке автоматизированных тестов» от 05.11.2004 г.

В качестве эксперимента компьютерное тестирование было введено в ЮГУ временным положением «О компьютерном тестировании знаний студентов в зимнюю сессию 2007/2008 гг.». Постоянное положение «О компьютерном тестировании знаний студентов» было утверждено 26 апреля 2010 года. Оба документа введены в действие в соответствии с решениями учебно-методического совета университета от 27.11.2006 г. и от 28.01.2007 г., а также в соответствии с приказом ректора от 02.11.2007 г. № 4/823. В положениях утверждалось (п. 1.2), что компьютерное тестирование является одной из форм промежуточной аттестации обучающихся (проведение экзамена и зачёта). В дальнейшем, в соответствии с требованиями ФГОС ВПО, создавались фонды оценочных средств для проведения не только промежуточной аттестации, но и текущего контроля успеваемости студентов. Эти фонды включали тесты и компьютерные тестирующие программы. Анкетирование студентов ЮГУ, проведенное в 2007 году, продемонстрировало положительное отношение большинства из них к тестированию, как альтернативе традиционной оценке знаний [12: 160].

В «Методических рекомендациях по проектированию оценочных средств ...» [8: 37], разработанных в рамках национального проекта «Формирование системы инновационного образования в МГУ им. М. В. Ломоносова», выделено три вида контроля качества знаний и компетенций у студентов вузов: устный опрос, письменные работы и контроль с помощью технических средств и информационных систем. Тест, как одна из форм контроля знаний, отнесён ко второму и третьему виду контроля. Письменные работы включают тесты в качестве простейшей формы контроля знаний. В письменной работе тест состоит из небольшого количества элементарных задач; может предоставлять возможность выбора из перечня ответов; занимает часть учебного занятия (10–30 минут) [8: 40]. В технических средствах и информационных системах и технологиях оценивания качества учебных достижений студентов электронные тесты являются эффективным средством контроля результатов образования на уровне знаний и понимания. Тесты здесь присутствуют в виде программ компьютерного тестирования: обучающих тестов и аттестующих тестов. Обучающие тесты предназначены для самоконтроля студента и определения траектории обучения. Аттестующие тесты могут использоваться как для проведения текущего контроля успеваемости в течение семестра, так и для проведения промежуточной и рубежной аттестации [8: 46–49]. В нашей работе тестирование представлено в виде обучающих и аттестующих тестов.

На кафедре экологии и природопользования Института природопользования ЮГУ тестирование применялось при подготовке специалистов экологов-природопользователей, бакалавров и магистров по направлению экология и природопользование для текущего контроля знаний в течение семестра и итогового контроля в период экзаменационной сессии.

В бланковой форме тесты разрабатывались и применялись по всем читаемым автором дисциплинам: экология (для технических специальностей), экологическое проектирование и экспертиза, экологические проблемы Югры, экологический туризм, особо охраняемые природные

территории, биологическое ресурсосведение, природопользование и экология Югры в условиях социально-экономической трансформации, международное сотрудничество в области охраны окружающей среды, устойчивое развитие человечества, образование для устойчивого развития.

В электронной форме компьютерное тестирование применялось по дисциплине: «Оценка воздействия намечаемой хозяйственной и иной деятельности на окружающую среду» (ОВОС). По ОВОС составлен сборник тестовых заданий [9, 10] (печатный и электронный вариант), использовавшийся в учебном процессе в качестве учебного пособия для подготовки к текущему и итоговому тестированию (как обучающее тестовое пособие). Информационной основой для разработки тестов по ОВОС принято учебное пособие для вузов «Экологическая экспертиза» коллектива авторов под редакцией профессора В. М. Питулько [16].

### **Разработка, структура и оценка тестов**

При разработке сборника тестовых заданий по ОВОС и написании данной статьи в качестве основных методических руководств использованы: работа В. И. Васильева и Т. Н. Тягуновой [2] по основам культуры адаптивного тестирования, учебное пособие М. Б. Челышковой [13], освещающее вопросы теории и практики конструирования педагогических тестов, а также методическое пособие к тестовым заданиям по информатике А. А. Кузнецова и др. [6]. При составлении и оформлении ТЗ принимались во внимание рекомендации, изложенные в сборниках тестов по биологии [1] и экологии [15].

В настоящее время в России идёт процесс становления системы тестирования в области образования, а тестовые технологии рассматриваются как одно из объективных средств контроля качества подготовки студентов [11: 5]. В целях осуществления единой политики Минобрнауки России в области конструирования профессиональных компьютерных тестов и получения объективных оценок уровня знаний, умений и навыков студентов образовательных организаций Центром образовательных коммуникаций и тестирования профессионального образования при Московском государственном университете печати (МГУП) разработаны методологические «Требования к программно-дидактическим тестовым материалам и технологиям компьютерного тестирования» [3]. Современные представления об особенностях методологии и технологии компьютерного тестирования, разработке и оценке качества банков ТЗ изложены в учебном пособии [11]. Обе работы [3, 11] предназначены преподавателям и сотрудникам образовательных учреждений, использующих компьютерное тестирование как для разработки тестовых материалов, так и для проверки уровня учебных достижений студентов.

В 2010 году автор повысил квалификацию в Центре тестирования профессионального образования при МГУП по программе «Современные технологии оценки качества в системе высшего образования» с получением соответствующего сертификата.

В сборнике тестовых заданий по ОВОС [9, 10] тесты тематические, т. е. тестовые задания составлены по каждой теме предмета. Один тест включает одну или две темы (в зависимости от объема). Всего в сборник вошло 18 тестов по 26 учебным темам. Сборник состоит из двух частей: 1-я часть – введение в ОВОС, 2-я часть – собственно ОВОС. Каждая часть объединяет по 9 тестов. Объем сборника или банк тестовых заданий (БТЗ) на принятой информационной основе составляет 1131 тестовое задание. Средний объем одного теста – около 63 тестовых заданий (от 37 до 80 ТЗ). Одно тестовое задание обычно содержит 5 вариантов ответа (85 % от количества ТЗ закрытой формы: одиночный и множественный выбор, т. е. один или несколько правильных ответов и неправильные ответы или дистракторы) [13: 115]. ТЗ с 6-тью вариантами ответа составляют 5 %, с 7-мью вариантами – 10 % заданий закрытой формы. Современными правилами к составлению ТЗ для заданий закрытой формы рекомендуется 4–6 вариантов ответов [11: 29]. Тестовые задания разрабатывались вручную, без использования программы «Конструктор тестов» по автоматизированному формированию заданий в тестовой форме.

В зависимости от важности вопроса (важности для изучения предмета) выделено три категории тестовых заданий: простые, средней сложности и сложные. Оценка их дифференцирована: за выполнение простого задания начисляется 1 балл, средней важности/сложности/трудности [2: 341] – 2 балла и сложного/важного/трудного – 3 балла. Понятия

«важность», «сложность» и «трудность» в контексте данной работы используются как синонимы. Задания повышенной важности в тестах обозначаются звездочками, соответственно 2 или 3. Простые задания звездочками не обозначаются.

Принято три формы тестовых заданий: закрытой формы – одиночный или множественный выбор; открытой формы – вставить пропущенное (задания на дополнение) [13: 181], установить соответствие. В зависимости от формы ТЗ принято три варианта их оценки:

1-й вариант – при оценке заданий «одиночного» или «множественного» выбора начисляется от 1-го до 3-х баллов за каждый правильный ответ (в зависимости от показателя важности вопроса);

2-й вариант – при оценке заданий «вставить пропущенное» число баллов (от 1-го до 3-х) засчитывается за пропущенное слово, так как допускается пропуск только одного ключевого слова;

3-й вариант – при оценке заданий «установить соответствие» засчитывается от 1-го до 3-х баллов за каждый правильный ответ.

Сочетание «формы» и «важности/сложности» в ТЗ может быть различным (табл. 1).

Таблица 1

**Оценка выполнения тестовых заданий в зависимости от их структуры (баллы)**

| Важность/<br>Сложность<br>заданий | Форма заданий      |                        |                         |                            |
|-----------------------------------|--------------------|------------------------|-------------------------|----------------------------|
|                                   | Одиночный<br>выбор | Множественный<br>выбор | Вставить<br>пропущенное | Установить<br>соответствие |
| Простые                           | 1                  | 1                      | 1                       | 1                          |
| Средней сложности                 | 2                  | 2                      | 2                       | 2                          |
| Сложные                           | 3                  | 3                      | 3                       | 3                          |

Как следует из таблицы 1, в результате комбинации «формы» и «важности/сложности» в ТЗ могут быть три варианта их оценки: 1-й вариант – за выполнение простого задания любой формы начисляется 1 балл; 2-й вариант – за выполнение задания средней важности/сложности любой формы начисляется 2 балла; 3-й вариант – за выполнение важного/сложного задания любой формы начисляется 3 балла. При оценке выполнения заданий формы «множественный выбор» и «установить соответствие», в которых имеется несколько вариантов правильных ответов, начисляется от 1-го до 3-х баллов за каждый из них.

Примеры форм и категорий тестовых заданий приведены из сборника [8].

Пример 1. Задание № 1.4 из теста № 1, темы 1: Введение в экологическую оценку [9: 6]. Простое задание закрытой формы, одиночный выбор: правильный ответ – 3; начисляется 1 балл. Название задания и варианты ответов приведены ниже.

**1.4. Указать истоки систем экологической оценки:**

1. Истоки систем экологической оценки связаны с Великой Октябрьской социалистической революцией 1917 г. и необходимостью переустройства народного хозяйства на социалистический лад.

2. Системы экологической оценки стали активно развиваться в XX веке в связи с открытием Колумбом Америки и необходимостью промышленного освоения нового континента.

3. Истоки систем экологической оценки восходят к земельным и горным законам Великобритании, Германии и США.

4. Экологическая оценка возникла вследствие закономерного развития экологии как всеобщей науки об окружающей среде.

5. Истоки систем экологической оценки связаны с усилением движения за охрану окружающей среды во второй половине XX столетия.

Пример 2. Тестовое задание № 1.1.\*\*\* из теста № 1, темы 1: Введение в экологическую оценку [9: 6]. Сложное/важное задание закрытой формы, множественный выбор: правильные ответы – 2 и 3; начисляется по три балла за каждый правильный ответ, всего 6 баллов. Название задания и варианты ответов приведены ниже.

**1.1.\*\*\* Основу российской системы экологической оценки составляет:**

1. Экологическое обоснование проектной документации.
2. Экологическая экспертиза.
3. ОВОС.
4. Экологическое нормирование и стандартизация.
5. Экологический аудит хозяйственной деятельности.

Пример 3. Тестовое задание № 1.3.\*\* из теста № 1, темы 1: Введение в экологическую оценку [9: 6]. Задание средней сложности/важности, открытая форма – вставить пропущенное слово; ключевое слово и его синонимы приведены под заданием; начисляется два балла за правильный ответ (синонимы равноценны ключевому слову).

**1.3.\*\* Вставить пропущенное:**

Экологическая оценка основана на простом принципе: легче выявить и \_\_\_\_\_ негативные для окружающей среды последствия деятельности на стадии её планирования, чем обнаружить и исправлять их на стадии осуществления этой деятельности.  
*предупредить; опередить; предварить;*

Пример 4. Тестовое задание № 6.27.\*\* из теста № 4, темы 6: Сфера применения процедуры ОВОС [9: 58]. Задание средней сложности/важности; форма – установить соответствие; начисляется по два балла за правильное решение каждого варианта задания, всего – 10 баллов; правильные соответствия: органы власти – 1, соответствует выполняемой функции – 2; 2 – соответствует 5; 3 – 4; 4 – 3; 5 – 1. Задание упрощено в сравнении с оригиналом.

6.27.\*\* Выбрать и установить правильные соответствия уровня компетенции государственных органов и выполняемых ими функций в сфере ОВОС в РФ.

| Органы власти   | Выполняемые функции   |
|---|---|
| 1. Федеральные органы власти  | 1. Установление условий, касающихся экологических аспектов экономической деятельности в трансграничном контексте.   |
| 2. Территориальные отделения федеральных органов власти                   | 2. Разработка стратегии ОВОС.   |
| 3. Органы власти субъектов РФ   | 3. Данные о характере деятельности, состоянии окружающей среды в районе реализации проекта и возможных воздействиях на окружающую среду.  |
| 4. Научно-исследовательские центры, вузы, неправительственные организации | 4. Осуществление экологического контроля за хозяйственной деятельностью; обладание специальными знаниями о характере деятельности и экологической обстановке в районе реализации проекта. |
| 5. Международные организации  | 5. Установление нормативов и стандартов, лицензирование и контроль хозяйственной и иной деятельности.   |

Представленность тестовых заданий различных форм в учебном пособии [9, 10] не одинакова. Структура тестов по количеству форм тестовых заданий приведена в таблице 2.

Таблица 2

**Формы тестовых заданий в учебном пособии по ОВОС [9, 10]**

| Формы                   | Задания    |      |
|-------------------------|------------|------|
|                         | Количество | %    |
| Одиночный выбор         | 362        | 32,0 |
| Множественный выбор     | 276        | 24,4 |
| Вставить пропущенное    | 485        | 42,9 |
| Установить соответствие | 8          | 0,7  |
| Всего:                  | 1131       | 100  |

Как следует из таблицы 2, наибольшее предпочтение при разработке ТЗ по ОВОС отдано закрытой форме: ТЗ одиночного и множественного выбора в сумме составляют более 56 % объема сборника. На втором месте по количеству представлены ТЗ открытой формы, в данном случае – «вставить пропущенное»: примерно 43 % банка ТЗ. Задания на установление соответствия в БТЗ единичны: составляют менее 1 % его объема.

Современные тенденции в разработке тестов связаны с отходом от моноформности и широким применением ТЗ на дополнение. Введение различных форм в тесте обусловлено многообразием проверяемых знаний и умений обучающихся [13: 208–209]. Различные формы ТЗ в тестологии выполняют функцию проверки различных компонентов обученности тестируемых. Считается, что ТЗ закрытой формы ориентированы на оценку знаний студентов, а ТЗ открытого типа – на оценку их интеллектуальных умений [6: 450]. С этой позиции составленный нами БТЗ по ОВОС нацелен, во-первых, на оценку знаний студентов по предмету и, во-вторых, на оценку их интеллектуальных умений.

В структуре сборника (печатный вариант) [9, 10] контрольные таблицы с ответами на ТЗ расположены в конце каждого теста. Синонимы всех ключевых слов в контрольные таблицы не включены. Здесь присутствует только одно основное ключевое слово. В тексте учебного пособия в структуре ТЗ «вставить пропущенное», как правило, присутствуют три варианта (синонимы) ключевых слов.

Шкала для оценки выполнения ТЗ составлена эмпирически, на основании собственного опыта, сведений из специальной методической литературы [1, 2] и консультаций с преподавателями кафедры экологии. Принято 4 уровня оценки выполнения теста: неудовлетворительно, удовлетворительно, хорошо, отлично. При оценке всего тематического теста неудовлетворительная оценка выставляется при количестве правильных ответов, не превышающих 49 % от всего их числа, удовлетворительная оценка – от 50 % до 74 % от числа правильных ответов, хорошая – от 75 % до 90 % , и отличная – при 91–100 % правильных ответов.

В зависимости от количества и важности ТЗ, входящих в один тест, общее число баллов, начисляемых за его выполнение, изменялось в пределах от 81 до 203, при среднем значении, равном 146 баллам. Границы балльных оценочных категорий рассчитывались для каждого теста отдельно посредством арифметической пропорции. Пример среднестатистической оценочной шкалы для всех 18-ти тестов сборника [9, 10] представлен в таблице 3.

Таблица 3

**Шкала для оценки результатов тестирования экологов-природопользователей по ОВОС**

| Оценки              |          |         |
|---------------------|----------|---------|
| Категории           | Проценты | Баллы   |
| Неудовлетворительно | 0–49     | 0–72    |
| Удовлетворительно   | 50–74    | 73–108  |
| Хорошо              | 75–90    | 109–131 |
| Отлично             | 91–100   | 132–146 |

Как следует из таблицы 3, представленная шкала имеет неравные оценочные интервалы, т. е. шкала неравномерная. Величины оценочных интервалов составляют 49, 25, 16 и 10 %, соответственно: 1 – неудовлетворительно, 2 – удовлетворительно, 3 – хорошо и 4 – отлично. Нелинейный характер оценочной шкалы выражает уравнение полинома 3-й степени при абсолютной величине критерия достоверности аппроксимации линией тренда (рис. 1).

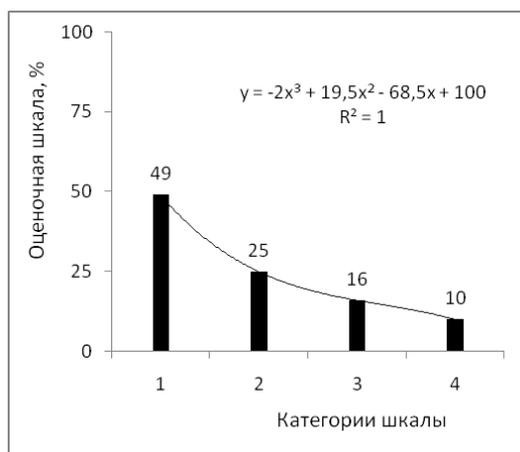


Рисунок 1 – Соотношение величин интервалов шкалы для оценки результатов тестирования экологов-природопользователей по ОВОС

В сборнике тестовых заданий по ОВОС [9, 10] оценочные шкалы с градациями баллов для конкретных тестов приведены в конце каждого теста, под контрольными таблицами. Это значительно облегчает пользование учебным пособием при самостоятельной подготовке студентов к тестированию.

### Процедура тестирования

Для компьютерного тестирования весь банк тестовых заданий по ОВОС разбит на шесть частей, по три тематических теста в каждой части. Каждая часть содержит от 163 до 208 тестовых заданий, в среднем – 189 шт. Тестирование по предмету проводилось шесть раз, отдельно по каждой части банка тестовых заданий. В ходе тестирования компьютерная программа перемешивает тестовые задания и случайным образом выбирает и представляет на монитор компьютера принятую международную норму: 60 тестовых заданий. В обозначении тестовых заданий на мониторе компьютера присутствует одна программная подсказка:  $\circ$  – один правильный ответ в данном тестовом задании,  $\square$  – несколько правильных ответов.

Результаты компьютерного тестирования показали, что одно тестирование по ОВОС одной группы студентов на принятой базе (60 тестовых заданий) занимает практически полную учебную пару (1,5–2 академических часа, включая перерыв), хотя наиболее способные студенты выполняют тестирование примерно за 40 минут. В нашем случае затраты времени на одно тестирование по ОВОС хорошо согласуются с предложениями тестологов по данному вопросу. Рекомендуемое время выполнения теста для студентов составляет 1,5 часа [2: 270]. Результаты компьютерного тестирования одной академической группы студентов в течение учебного года показаны на рисунке 2.

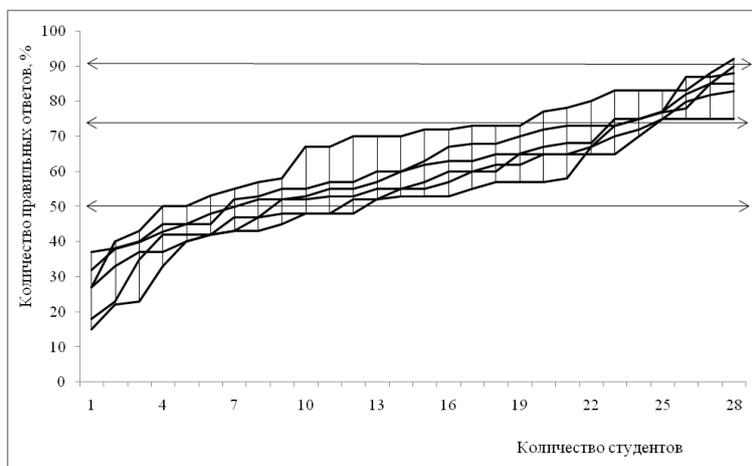


Рисунок 2 – Результаты компьютерного тестирования группы экологов-природопользователей по ОВОС

Тело графика на рисунке 2 содержит шесть рядов ранжированных ответов на задания теста. Горизонтальными линиями со стрелками показаны уровни оценочной шкалы: 50 % правильных ответов и выше – удовлетворительно; 75 % и выше – хорошо; 91 % и выше – отлично. По результатам 6-ти тестирований студенты данной группы получили 168 оценок: неудовлетворительно – 46, удовлетворительно – 89, хорошо – 32 и отлично – одна оценка.

### Полученные результаты

В течение семи учебных сезонов с 2008 по 2015 гг. протестировано 138 студентов из семи академических групп (табл. 4).

Таблица 4

#### Академические группы студентов, принимавшие участие в компьютерном тестировании по ОВОС в период с 2008 по 2015 гг.

| № группы         | 1         | 2         | 3         | 4         | 5         | 6         | 7         |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Шифр группы      | 2460      | 27706     | 27806     | 27906     | 27006     | 27116     | 27216     |
| Учебный год      | 2008–2009 | 2009–2010 | 2010–2011 | 2011–2012 | 2012–2013 | 2013–2014 | 2014–2015 |
| Кол-во студентов | 14        | 24        | 18        | 15        | 24        | 15        | 28        |

В течение учебного года проводилось по шесть компьютерных тестирований с каждой группой. Оценки выставлялись в соответствии с оценочной шкалой на основании количества набранных правильных ответов при каждом тестировании. Всего получено 828 оценок. В связи с присутствием у студентов стремления к улучшению первоначального результата посредством пересдачи тестов в течение семестра в обработку приняты результаты окончательного тестирования, т. е. итоговые оценки по каждому тестированию, вошедшие в зачёт или экзамен. Демо-тестирование с экологами-природопользователями не проводилось. Общие результаты тестирования приведены в таблице 5.

Таблица 5

#### Общие результаты компьютерного тестирования по предмету ОВОС для всех групп экологов-природопользователей за 2008–2015 учебные годы

| Оценки     | Категории оценок    |                   |        |         |       |
|------------|---------------------|-------------------|--------|---------|-------|
|            | Неудовлетворительно | Удовлетворительно | Хорошо | Отлично | Всего |
| Количество | 175                 | 336               | 221    | 96      | 828   |
| Процент    | 21,135              | 40,579            | 26,690 | 11,594  | 100   |

Из таблицы 5 следует, что количество неудовлетворительных оценок примерно соответствует 1/5 вероятности случайного выбора. Следовательно, неподготовленный студент, выбирая случайным образом ответы из 5-ти вариантов, не может получить положительную оценку. В соответствии с оценочной шкалой для получения удовлетворительной оценки необходимо набрать 50 % правильных ответов. При количестве вариантов ответов от 5 до 7 область случайного выбора составляет 14–20 %. В то же время область незначения или область отрицательного результата, т. е. неудовлетворительной оценки в соответствии с оценочной шкалой, составляет 49 %. Доля положительных оценок, т. е. качество подготовки по предмету на полученном материале, составила примерно 79 %.

Характер распределения тестовых оценок по результатам тестирования всех групп студентов за рассматриваемый период хорошо отображается полиномом четвёртой степени, при абсолютной величине уровня достоверности аппроксимации линией тренда (рис. 3).

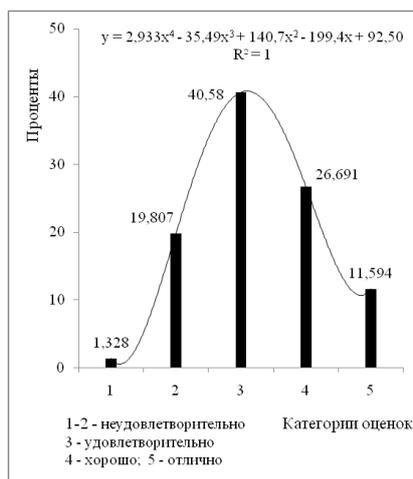


Рисунок 3 – Распределение тестовых оценок по ОВОС по категориям за 2008–2015 гг.

На рисунке 3 категория неудовлетворительных оценок разбита на две группы: «единицы» и «двойки». На «единицу» условно оценивались работы, в которых количество правильных ответов не превышало 37 баллов или 25 % правильных ответов по принятой оценочной шкале (табл. 3). Из всего массива полученных оценок (828 за 7 лет компьютерного тестирования) доля работ, выполненных на «единицу», составила немногим более одного процента (рис. 3). Вместе с работами, выполненными на «двойку», доля работ, выполненных неудовлетворительно, и составит потенциальную область случайного выбора или угадывания при тестировании неподготовленного студента.

При оценке результатов тестирования рекомендуется применять коррекцию баллов на догадку или угадывание ответа, так как считается, что без коррекции результаты отдельных слабых учащихся будут явно завышены [13: 123–137]. Однако если тест имеет достаточно большое количество заданий, то угадыванием можно пренебречь. Также для снижения эффекта угадывания разработчики тестов практикуют увеличение числа ответов в ТЗ [13: 124]. В заданиях с большим количеством дистракторов на первый план при выборе правильного ответа выходят знания, а не догадка. Поэтому чем больше дистракторов, тем труднее угадать правильный ответ [13: 137]. В нашей работе объём тестов при компьютерном тестировании можно считать оптимальным, так как он составляет международную норму: 60 ТЗ. Размер же одного тестового задания содержит предельно допустимое количество вариантов ответа: 5–7 в ТЗ закрытой формы. Поэтому коррекция полученных результатов тестирования на догадку в нашей работе не применялась.

Для сравнения результатов каждого тестирования между собой использованы осреднённые данные по выполнению каждого тестирования всеми группами студентов за весь рассматриваемый период. В таблице 6 приведены средние арифметические с утроенной ошибкой, рассчитанные для результатов каждого из 6-ти тестирований методами параметрической статистики по [5, 7].

Таблица 6

## Средний балл для каждого из 6-ти тестирований за период 2008–2015 гг.

| Номера тестирований | Количество оценок (n) | Сумма баллов | Средний балл $M \pm 3m$ |
|---------------------|-----------------------|--------------|-------------------------|
| 1                   | 138                   | 9010         | $65,289 \pm 4,279$      |
| 2                   | 138                   | 8741         | $63,341 \pm 4,491$      |
| 3                   | 138                   | 8765         | $63,514 \pm 5,124$      |
| 4                   | 138                   | 9776         | $70,841 \pm 4,186$      |
| 5                   | 138                   | 9348         | $67,739 \pm 5,165$      |
| 6                   | 138                   | 9481         | $68,703 \pm 5,047$      |

По каждому из шести тестирований получено 138 оценок, выраженных в количестве баллов правильных ответов (процент из 100 возможных). Средний балл получен как величина от деления суммы баллов, набранных всеми группами студентов по каждому тестированию за 7 лет, на количество студентов, принимавших участие в тестировании.

В настоящее время в теории тестов отрицается возможность применения количественных методов параметрической статистики, основанной на нормальном законе распределения признаков, для обработки результатов тестирования, являющихся объектами нечисловой природы [2: 73, 186, 298]. Утверждается необходимость перехода на методы непараметрической и интервальной статистики, поскольку нормальных распределений в реальности практически не существует [2: 297; Дубровский С.С. по [2: 297]. Однако признаётся, что при увеличении числа наблюдений в выборке распределение среднего значения приближается к нормальному [2: 506]. Поэтому при достаточно большом  $n$  (значительно  $> 50$ ) биномиальное распределение можно заменить нормальным [2: 499, 562]. Это допускает применение методов параметрической статистики для обработки результатов тестирования. В нашей работе выборки наблюдений являются большими: количество оценок по результатам тестирования составляет от 84 до 168 (табл. 6, 7). Поэтому применение методов параметрической статистики для их обработки вполне допустимо.

Достоверность средних арифметических (средних баллов) определена как отношение её величины к своей ошибке [5, 7]. Критерий достоверности средней арифметической рассчитан по формуле:

$$t_M = \frac{M}{m_M}.$$

Пример: ошибка средней арифметической для 1-го тестирования будет равна отношению среднего квадратичного отклонения ( $\sigma$ ) к корню квадратному из величины выборки:

$$m_M = \pm \sigma / \sqrt{n} = \pm 16,758 / \sqrt{138} = \pm 1,426.$$

Средняя арифметическая с утроенной ошибкой  $M \pm 3m = 65,289 \pm 4,279$  (табл. 6);

$$t_M = 65,289 / 4,279 = 15.$$

При критерии достоверности  $t_M=15$  вероятность достоверности для большой выборки ( $n=138$ ) составит  $>0,9999$  [5, с. 80–81], т. к. полученный критерий многократно превышает необходимый уровень достоверности при  $t \geq 3$  для вероятности, равной 0,997. Уровень значимости полученного результата составляет 0,0001, т. е. он может быть ошибочным только в одном случае из 10 тысяч.

По остальным 5-ти тестированиям критерии достоверности средних арифметических получены равными от 12 до 16, что свидетельствует о высоком уровне достоверности определения их средних арифметических.

Результаты 6-ти тестирований показаны на рисунке 4.

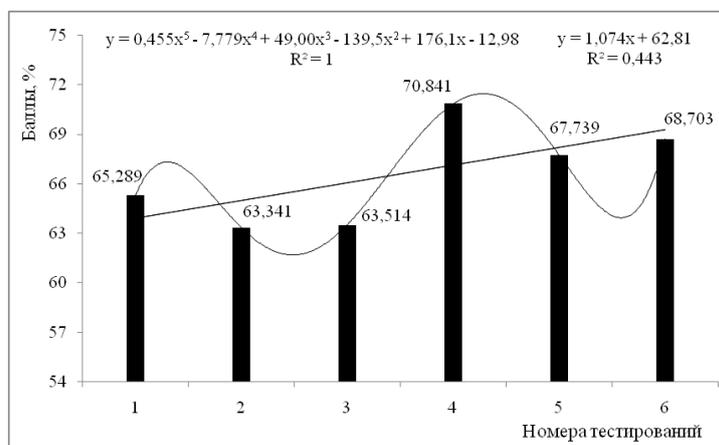


Рисунок 4 – Результаты 6-ти тестирований экологов-природопользователей по ОВОС за 2008–2015 гг.

Общий характер тестирования хорошо аппроксимируется линией полинома 5-й степени. Линейный же тренд свидетельствует о наличии тенденции увеличения успешности тестирования по второй половине изучения предмета. Обнаружена тенденция улучшения качества выполнения тестовых заданий от начала к концу семестра.

Данную закономерность можно объяснить наличием двух причин. Первая причина – тесты первой части сборника [9] содержательно труднее, чем тесты второй его части [10]. Особенно сложными являются тесты с № 4 по № 9: темы 6–14 в сборнике [9]. Сюда вошли вопросы, освещающие принципы проведения и национальную процедуру ОВОС, критериальную базу оценок воздействия и интегральные показатели техногенных воздействий, а также методы оценки интенсивности техногенных нагрузок на окружающую среду.

Повышенная трудность тестов, входящих в состав второго и третьего тестирований, хорошо иллюстрируется количеством «отличных» оценок, полученных студентами всех 7-ми групп в течение 2008–2015 гг. Всего по результатам шести сквозных тестирований получено 96 «отличных» оценок (табл. 5). Наименьшее количество «пятерок» получено по второму тестированию (рис. 5).

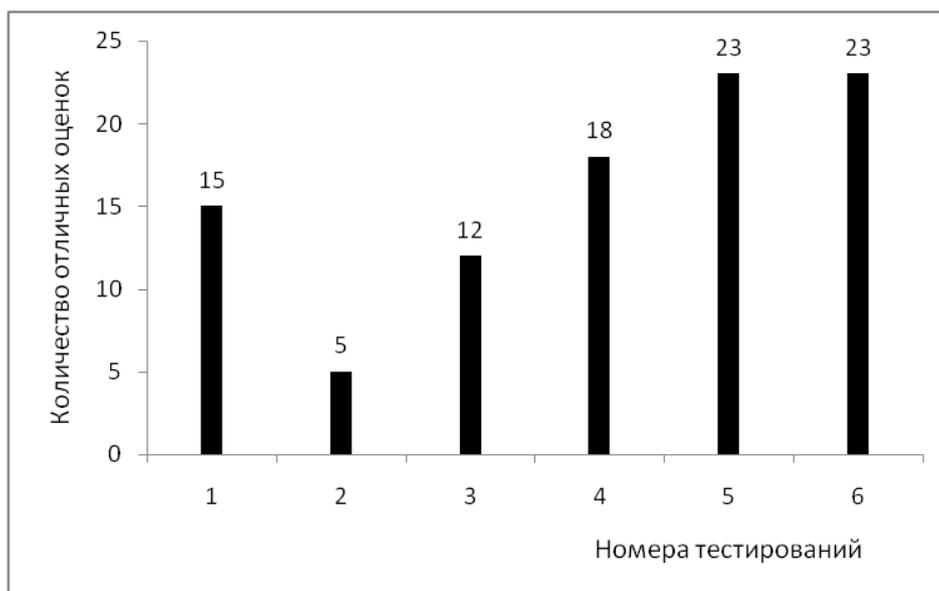


Рисунок 5 – Динамика «отличных» оценок по результатам 6-ти тестирований в 2008–2015 гг.

Следовательно, материалы второго и, в меньшей степени, третьего тестирований (рис. 5) по темам 6–14 сборника [9] являются наиболее сложными и трудными при изучении предмета ОВОС экологами-природопользователями.

Вторая причина роста успешности тестирования по содержанию второй части сборника тестов указывает на овладение студентами терминологией и понятийным аппаратом предмета в ходе его систематического изучения и регулярного текущего тестирования.

Для характеристики динамики прохождения тестирования в течение 2008–2015 гг. построено облако рассеяния по результатам 6-ти тестирований каждой группы студентов (рис. 6). Облако включает 828 точек (балльных оценок). Количество протестированных студентов, показанное на горизонтальной оси рисунка 6, привязано к каждой академической группе студентов. Поскольку размер студенческих групп различен (табл. 4), то их границы на горизонтальной оси графика, разделённой на равные интервалы по 23 единицы, не отражены.

Всего в компьютерном тестировании по ОВОС за 7 лет приняло участие 138 студентов. Данные на рис. 6 приведены в хронологическом порядке, что соответствует последовательности рассматриваемого периода от 2008-го к 2015 году. Расположение точек в облаке рассеяния отражает динамику успеваемости академических групп по ОВОС по годам от начала к концу периода.

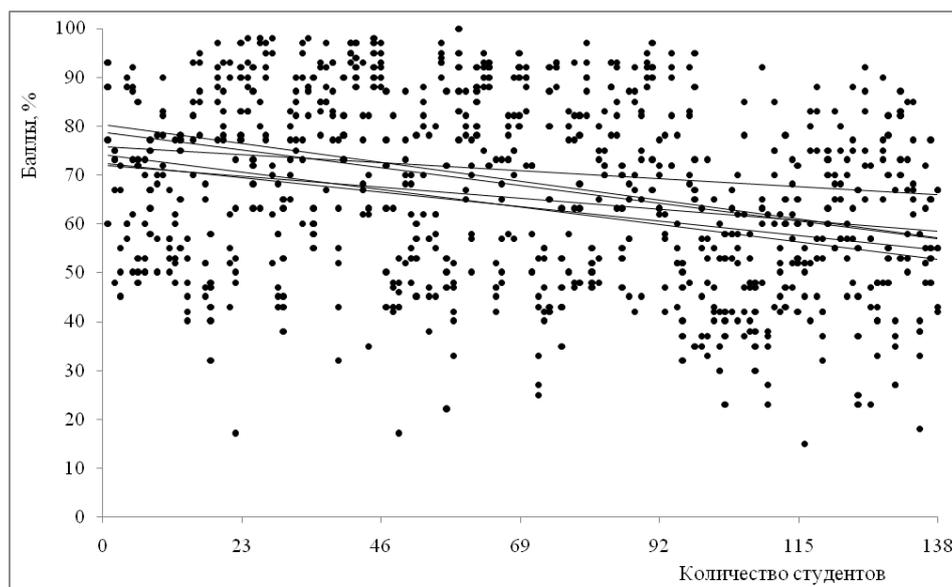


Рисунок 6 – Динамика тестирования экологов-природопользователей по ОВОС за 2008–2015 учебные годы

Направление линий тренда по каждому из 6-ти тестирований отражает состояние уровня успеваемости по предмету от начала к концу рассматриваемого периода. На основании интерпретации облака рассеяния результатов тестирования на рисунке 6 можно предположить, что уровень успеваемости по ОВОС у экологов-природопользователей закономерно снижался в течение 7 лет, на что указывает направление линий трендов. Для проверки данного предположения был определён средний балл по результатам тестирования для каждой группы студентов и проведено их сравнение. Результаты представлены в таблице 7.

Таблица 7

Средний балл для каждой из 7-ми групп студентов за период 2008–2015 гг.

| Номера групп | Количество оценок (n) | Сумма баллов | Средний балл $M \pm 3 m$ |
|--------------|-----------------------|--------------|--------------------------|
| 1            | 84                    | 5534         | $65,881 \pm 4,662$       |
| 2            | 144                   | 10866        | $75,458 \pm 4,356$       |
| 3            | 108                   | 7666         | $70,981 \pm 5,636$       |
| 4            | 90                    | 6520         | $72,444 \pm 5,586$       |
| 5            | 144                   | 9945         | $69,063 \pm 4,552$       |
| 6            | 90                    | 4678         | $51,978 \pm 5,133$       |
| 7            | 168                   | 9912         | $59,0 \pm 3,699$         |

Средний балл для отдельной группы получен как величина от деления суммы всех оценок, набранных студентами данной группы за 6 тестирований, на количество оценок. Пример расчёта среднего балла по группе № 1:  $\Sigma 5534$  балла /  $(14 \times 6) = 65,881$  балла.

Достоверность средних арифметических (средних) баллов для результатов тестирования каждой из 7-ми групп студентов определена аналогично оценке достоверности средних арифметических для 6-ти тестирований (см. выше). Критерии достоверности средних арифметических, представленных в таблице 7, получены равными от 10 до 17. Они многократно превышают принятый уровень достоверности ( $t \geq 3$ ), что указывает на высокую вероятность и уровень значимости полученных величин [5].

Изменение среднего балла по группам показано на рисунке 7. Результаты тестирования групп приведены в хронологическом порядке от 2008 к 2015 году (табл. 4).

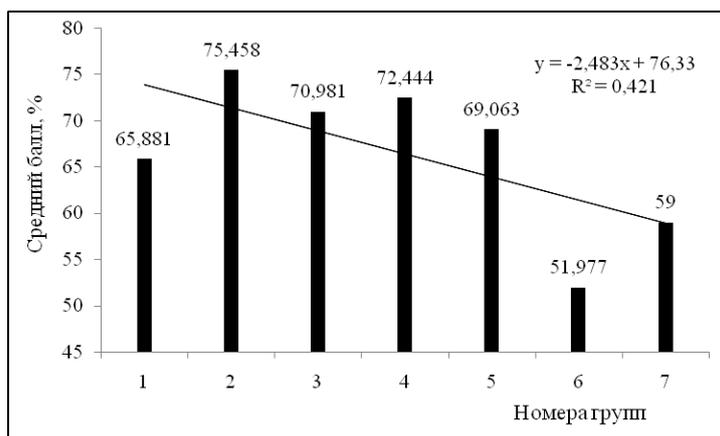


Рисунок 7 – Изменение среднего балла по результатам тестирования по каждой группе студентов за 2008–2015 гг.

При сравнении полученных результатов между группами подтверждается существенная разница в уровне средних оценок между группами 1–5 и 6–7, показанная на рисунке 6.

При сравнении выделенных групп по среднему баллу получены следующие результаты. Средний балл по группам № 1–5: 65,881; 75,458; 70,981; 72,444; 69,063 (табл. 7). Сумма – 353,827 баллов. Среднее =  $353,827 : 5 = 70,765$  балла. Средний балл по группам № 6–7: 51,977; 59,000. Сумма – 110,977 баллов. Среднее =  $110,977 : 2 = 55,489$  балла. Разница результатов:  $70,765 - 55,489 = 15,276$  баллов. Средний балл по группам 1–5 примерно на 15 % выше, чем по группам 6–7. Предположение о закономерном снижении уровня успеваемости по ОВОС у экологов-природопользователей в течение 7 лет не подтвердилось.

Причина обнаруженной разницы в результатах тестирования между группами, скорее всего, следующая: студенты групп 1–5 сдавали по предмету ОВОС экзамен, а студенты групп 6–7 – зачёт. В первые 5 учебных сезонов итоговым контролем по ОВОС был экзамен, в последние два года – зачёт. Вывод: при подготовке к экзамену студенты лучше готовятся и лучше сдают предмет, т. е. относятся к изучению предмета более ответственно. Тестологи [2: 378] подтверждают высокое значение мотивации тестируемых в ходе тестовых проверок: одно дело – самостоятельная проверка студентами уровня собственных знаний, например, в виде демо-тестирования, и совсем иное дело – тестирование в ходе официальной аттестации, с заведомо более высоким уровнем ответственности за конечный результат.

Одним из важнейших требований, предъявляемых к тестам, является их надёжность [6: 7]. Надёжность – это степень повторяемости или степень постоянства результатов тестирования [13: 64]. Оценка надёжности тестов по ОВОС проведена по [11: 64–68]. В данной работе предложено два способа оценки надёжности тестов, применяемых при повторном и при однократном тестировании. Оценка надёжности при повторном тестировании проводится между двумя последовательными тестами на одной и той же выборке испытуемых. В нашем случае при тестировании по ОВОС выборки испытуемых были разными, так как тестировались разные группы студентов различной величины в следующие друг за другом учебные годы. Мотивация при тестировании также различалась: подготовка к экзамену в первые пять лет и сдача зачёта в последние два года. Поэтому оценка надёжности тестов по ОВОС проведена по второму способу, применяемому при однократном тестировании.

Результаты всех тестирований за 7 лет приняты как дополняющие друг друга, т. е. представлены как однократное тестирование (одна выборка). Способ оценки надёжности при однократном тестировании заключается в расщеплении теста на две половины, на нечётные и чётные величины (по номеру местоположения в ряду данных). За 7 лет тестирования по ОВОС выборка участников составила 138 студентов. Величина расщеплённой выборки составила 69 студентов по каждому из шести ежегодных тестирований.

Вычисление коэффициента надёжности теста осуществляется на основе коэффициента корреляции Пирсона по формуле [11: 66]:

$$K = r_{xy} = \frac{\sum xy - \frac{\sum x \sum y}{N}}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{N}} \sqrt{\sum y^2 - \frac{(\sum y)^2}{N}}},$$

где  $K$  – коэффициент надёжности теста,

$r_{xy}$  – коэффициент корреляции Пирсона,

$x$  – величины нечётной половины данных,

$y$  – величины чётной половины данных,

$N$  – общее количество тестируемых студентов.

Полученный коэффициент ( $r_{xy}$ ) отражает надёжность теста вдвое меньшей длины ( $138 / 2 = 69$ ). Поэтому проведена коррекция полученных коэффициентов по формуле Спирмана-Брауна [11: 68]:

$$K = \frac{2r}{1+r},$$

где  $K$  – надёжность теста в целом,

$r$  – надёжности теста по двум половинам. Результаты расчётов представлены в таблице 8.

Таблица 8

### Оценка надёжности тестов по ОВОС

| Показатели | Номера тестирований |       |       |       |       |       |
|------------|---------------------|-------|-------|-------|-------|-------|
|            | 1                   | 2     | 3     | 4     | 5     | 6     |
| $r_{xy}$   | 0,922               | 0,920 | 0,836 | 0,919 | 0,913 | 0,906 |
| $K$        | 0,959               | 0,958 | 0,911 | 0,958 | 0,955 | 0,951 |

Как следует из таблицы 8, полученные коэффициенты надёжности ( $K$ ) несколько выше коэффициентов корреляции Пирсона ( $r_{xy}$ ). Авторами работы [11: 68] предлагается 4 уровня надёжности тестов от неудовлетворительного до отличного. Коэффициенты надёжности тестов по ОВОС, представленные в таблице 8, соответствуют высшему уровню надёжности по предложенной шкале от 0,9 до 0,99, т. е. оценка надёжности всех тестов по ОВОС является «отличной».

### Заключение

Тестирование проводилось по всему материалу сборника [9, 10]. Экологи-природопользователи 3-го курса успешно справились со сложностью и объёмом компьютерного тестирования по ОВОС. Подтвердилась также правильность шкалы, составленной для оценки уровня знаний по данному предмету. Надёжность тестов сборника по ОВОС [9, 10] определена как высокая.

Результаты текущего контроля знаний с помощью выполнения тестовых заданий учитывались при определении общей оценки по предмету на экзамене (зачёте) наряду с оценками за решение практических задач и оценкой за курсовую работу. При успешном выполнении всех шести тестирований зачёт по предмету выставлялся автоматически. Устный опрос по теоретической части дисциплины (ОВОС) не практиковался.

Печатная версия сборника [9, 10] использовалась студентами в качестве учебного пособия при самостоятельной подготовке к тестированию. Студенты предпочитали готовиться к компьютерному тестированию по учебному пособию, а не в компьютерном классе, т. к. РС всегда отбирает только часть тестовых заданий, а нужно знать все, к тому же при каждой загрузке программы задания перемешиваются. По опросным данным наиболее успешные студенты при самостоятельной подготовке к компьютерному тестированию выполняли нужные тематические тесты в бланковой форме по 2–3 раза (до получения отличного результата). Всегда худший результат имели студенты, ограничивавшиеся только пониманием сути учебного материала. Исключение составляли студенты, обладавшие фотографической памятью, но таких были единицы.

Из недостатков бланкового тестирования следует отметить два: 1 – невозможность полного предотвращения взаимной помощи студентов во время тестирования (подсказки), а также списывание и использование шпаргалок; 2 – большие затраты времени при проверке результатов тестирования преподавателем: проверка одного теста из 60-ти тестовых заданий группы из 25-ти студентов занимала около 4-х часов рабочего времени. Компьютерное тестирование почти полностью лишено этих недостатков.

При компьютерном тестировании недопустимое поведение студентов (подсказки, использование шпаргалок) может быть полностью исключено. Этому способствует и оснащение компьютерного класса камерами видеонаблюдения, и обязательное присутствие в классе заведующего кабинетом и тестирующего преподавателя. Результат тестирования становится известным сразу же по его окончании. Преподавателю остаётся только занести количество набранных баллов в журнал и выставить оценку. Трудоёмкость проверки выполненных тестов при компьютерном тестировании полностью снимается. Существенные преимущества компьютерного тестирования в сравнении с бланковым достаточно полно перечислены в работе [14: 43].

Из преимуществ тестовой формы контроля знаний студентов перед традиционной (устный опрос на зачёте/экзамене, написание рефератов и курсовых работ) на первое место следует поставить фундаментальный (целостный, холистический) подход в изучении предмета. При каждом тестировании студенту предъявлялось для ответа 60 тестовых заданий. При 6-тикратном тестировании по ОВОС в течение семестра студент должен ответить на 360 вопросов по содержанию предмета, что составляет примерно третью часть всего банка ТЗ по предмету. При подготовке же к тестированию необходимо изучить содержание всего банка тестовых заданий (1131 ТЗ по ОВОС), а не надеяться на знание 2–3-х вопросов в экзаменационном билете при традиционной форме контроля знаний. Необходимость систематической работы в течение семестра, а не штурм предмета перед сессией, является гарантией более прочного усвоения знаний в сравнении с традиционной формой обучения и контроля знаний.

Из других преимуществ тестовой формы контроля знаний, в частности, компьютерного тестирования, следует указать его объективность [2: 9]. Поскольку всем студентам задание даётся одинаковой степени трудности, всем приходится отвечать на одни и те же вопросы, то результат тестирования зависит только от способностей студента и степени его подготовки по данной теме. Претензий к преподавателю по результатам тестирования практически не бывает. Результат компьютерного тестирования совершенно прозрачен. Преподавателя здесь нельзя обвинить в предвзятом отношении к конкретному студенту, в необоснованных придирках и намеренном занижении оценки. При компьютерном тестировании студент получает оценку по мере своих способностей. Обиженных студентов здесь не бывает. В то же время при традиционной форме контроля знаний – зачёте и, особенно, экзамене – субъективизм преподавателя (или студента) является причиной обязательного появления после каждой сессии 1–2-х студентов из группы, не согласных с его проведением и недовольных результатами оценки их знаний. Неизбежный конфликт обуславливает необходимость пересдачи предмета на комиссии.

Регулярный (непрерывный) контроль знаний в течение семестра посредством компьютерного тестирования в итоге даёт значительную экономию учебного времени за счёт сокращения сроков экзаменационной сессии [4: 49], так как итоговая оценка за предмет может быть выставлена автоматически по результатам текущего тестирования. Практика текущего контроля знаний предотвращает неизбежный для студентов стресс в период экзаменационной сессии, в то время как при непрерывном контроле знаний к концу семестра они приходят уже с готовой оценкой знаний по предмету.

Применение компьютерного тестирования предъявляет ряд требований к преподавателю, решившемуся на применение данной формы контроля знаний студентов по читаемым дисциплинам.

Во-первых, необходимо наличие учебного пособия или достаточно полного курса лекций, на основании которых можно было бы составить сборник тестов и которые можно предложить студентам для углубленной самостоятельной подготовки по предмету. Также преподавателю необходимо быть готовым к периодическому обновлению содержания тестовых заданий в связи с появлением новых знаний по изучаемому предмету.

Во-вторых, составление тестовых заданий требует достаточно высокого уровня эрудиции и наличия широкого культурного кругозора у преподавателя. Притом не только в рамках своей узкой специальности, а в максимально широкой области знаний. Это необходимо для разработки вариантов ответов в тестовых заданиях в альтернативной форме, хорошо отличающихся по содержанию друг от друга. Также разработчик тестов должен уметь выделять главное и второстепенное в изучаемом материале для расстановки акцентов важности и формулировать чёткие вопросы по содержанию. По утверждению [2: 248], создание качественных тестовых материалов – крайне сложный творческий процесс, требующий от разработчика воображения, фантазии, глубоких знаний предметной области, правил и стандартов по конструированию ТЗ.

В-третьих, преподаватель должен сам составлять тесты по читаемым дисциплинам. Только при таком подходе он сможет выступать не только в качестве лаборанта-регистратора в процессе оценки знаний, но и реализовать творческий подход к обучению студентов, выступать в качестве консультанта, учитывать индивидуальные особенности студентов. Способность к самостоятельной разработке тестов – это показатель хорошего владения предметом и свидетельство высокого уровня квалификации преподавателя.

Компьютерное тестирование – это современно, технологично, научно, объективно, а в дополнение к этому ещё и очень интересно, но трудно; дисциплинирует студентов и преподавателя; поднимает отношение к предмету у студентов на более высокий уровень (на порядок), т. к. отношение студентов к процедуре компьютерного тестирования гораздо более серьёзное, чем к бланковому тестированию (как к публичному выступлению!); требует от студентов не только умения владения компьютером, но и хорошего знания русского языка, особенно при выполнении тестовых заданий открытой формы, требующей правильного введения ключевого слова с соответствующим смыслу и падежу окончанием; позволяет задействовать современные компьютерные технологии в передаче информации; облегчает труд преподавателя (после того, как тесты составлены).

Компьютерное тестирование – это современная эффективная форма контроля знаний студентов, позволяющая использовать компьютерные технологии не только для обучения, но и для оценки его результатов. Применение тестовой формы позволяет не только рационализировать процедуру контроля знаний, но и оптимизировать весь процесс изучения конкретных дисциплин. Тесты могут и должны применяться как для обучения, так и для контроля знаний студентов в высшей школе.

### Литература

1. Биология. Тестовые задания [Текст] / И. М. Прищепа, Г. А. Захарова, М. А. Щербакова [и др.]. – Минск : Новое знание, 2005. – 760 с.
2. Васильев, В. И. Основы культуры адаптивного тестирования [Текст] / В. И. Васильев, Т. Н. Тягунова. – Москва : ИКАР, 2003. – 584 с.
3. Васильев, В. И. Требования к программно-дидактическим тестовым материалам и технологиям компьютерного тестирования [Текст] / В. И. Васильев, А. А. Киринок, Т. Н. Тягунова. – М. : Изд-во МГУП, 2005. – 27 с.
4. Галочкин, А. И. Болонский процесс, качество образования и эволюция образовательных технологий [Текст] / А. И. Галочкин, Л. В. Петроченко, И. В. Ананьина // Вестник Югорского государственного университета, 2006. – № 3. – С. 32–52.
5. Ивантер, Э. В. Основы практической биометрии. Введение в статистический анализ биологических явлений [Текст] / Э. В. Ивантер. – Петрозаводск, «Карелия», 1979. – 96 с.
6. Кузнецов, А. А. Информатика. Тестовые задания [Текст] / А. А. Кузнецов, В. И. Пугач, Т. В. Добутько, Н. В. Матвеева. – 2-е изд., испр. – Москва : БИНОМ. Лаборатория знаний, 2003. – 232 с.
7. Лакин, Г. Ф. Биометрия [Текст] : учеб. пособие для биол. спец. Вузов / Г. Ф. Лакин. – 4-е изд. – Москва : Высш. Школа, 1990. – 352 с.

8. Методические рекомендации по проектированию оценочных средств для реализации многоуровневых образовательных программ ВПО при компетентностном подходе [Текст] / В. А. Богословский, Е. В. Караваева, Е. Н. Ковтун [и др.]. – Москва : Изд-во МГУ, 2007. – 148 с.
9. Новиков, В. П. Оценка воздействия намечаемой хозяйственной и иной деятельности на окружающую среду (ОВОС): В двух частях. Ч. 1. Введение в ОВОС: Тестовые задания [Текст] / В. П. Новиков. – Ханты-Мансийск : РИЦ ЮГУ, 2007. – 168 с.
10. Новиков, В. П. Оценка воздействия намечаемой хозяйственной и иной деятельности на окружающую среду (ОВОС): В двух частях. Ч. 2. ОВОС: Тестовые задания [Текст] / В. П. Новиков. – Ханты-Мансийск : РИЦ ЮГУ, 2007. – 149 с.
11. Попов, Д. И. Экспертиза качества тестовых заданий [Текст] : учеб. пособие / Д. И. Попов, Е. Д. Попова ; Моск. гос. ун-т печати. – Москва : МГУП, 2008. – 84 с.
12. Пасечник, М. В. Технология тестирования как инструмент оценки знаний учащихся [Текст] / М.В. Пасечник // Актуальные проблемы подготовки профессионалов XXI века в условиях гуманизации образования : материалы второй городской научно-практ. конф. / Югор. гос. ун-т, Бюджет. учреждение доп. проф. образования Ханты-Манс. авт. окр. – Югры, Ин-т развития образования. – Ханты-Мансийск : РИО Ин-та развития образования, 2009. – Т. 1, Ч. 2. – С. 157–162.
13. Чельшкова, М. Б. Теория и практика конструирования педагогических тестов [Текст] : учебное пособие / М. Б. Чельшкова. – М. : Логос, 2002. – 432 с.
14. Семёнов, С. П. Сравнение результатов бланкового и компьютерного тестирования [Текст] / С. П. Семёнов, В. Д. Серегин, П. Б. Татаринцев // Вестник Югорского государственного университета. – 2011. – № 3 (22). – С. 31–44.
15. Скорик, А. В. Экология. Тесты [Текст] : учебное пособие для вузов / А. В. Скорик, О. В. Ларина. – Москва : Экзамен, 2005. – 224 с.
16. Экологическая экспертиза [Текст] : учебное пособие для студ. высш. учеб. заведений / В. К. Данченко, В. М. Питулько, В. В. Растоскуев [и др.]; Под ред. В. М. Питулько. – Москва : Академия, 2004. – 480 с.