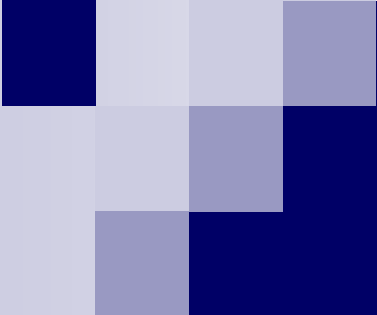


Министерство социальной защиты населения Свердловской области  
Областное государственное образовательное учреждение  
начального профессионального образования  
«Камышловское профессиональное училище – интернат для инвалидов»



## **КОМПЬЮТЕРНЫЙ КОНТРОЛЬ: ЕГО ПЛЮСЫ И МИНУСЫ**

Григорьева Елена Николаевна  
мастер производственного обучения,  
преподаватель теоретических дисциплин

КАМЫШЛОВ - 2009



## СОДЕРЖАНИЕ

Введение .....	4
Глава 1. Система компьютерного контроля как результат новых информационных технологий в обучении .....	5
Глава 2. Особенности оценивания качества обучения .....	8
§ 1. Терминология .....	8
§ 2. Сложности и проблематичность оценивания знаний .....	10
§ 3. Предпосылки использования ИТО в процедурах оценивания .....	13
§ 4. Автоматизированное тестирование .....	16
§ 5. Дополнительные возможности ИТО в процессе оценивания качества обучения и развития .....	21
Глава 3. Актуальные проблемы компьютерного контроля знаний .....	30
§ 6. Место тестирования в учебном процессе .....	30
§ 7. Построение компьютерных систем контроля знаний .....	33
§ 8. Подготовка материалов для компьютерного контроля знаний .....	34
§ 9. Обработка результатов тестирования .....	37
§ 10. Применение средств компьютерного контроля знаний для проведения междисциплинарного экзамена .....	39
Глава 4. Методика проведения компьютерного контроля знаний .....	41
Заключение .....	45
Библиографический список .....	46

## ВВЕДЕНИЕ

В данный момент учебные заведения находятся в периоде интенсивного проникновения ЭВМ в учебный процесс. Использование компьютеров как технического средства обучения имеет достаточный опыт и технологии создания обучающих программ достаточно разнообразны. Однако в настоящее время возникает немало трудностей, связанных с их внедрением в учебный процесс. Причина заключается в том, что практическая компьютеризация процесса обучения во многих учебных заведениях проходит без опоры на надлежащее теоретическое и методическое обоснование.

Необходимо отметить, что первые концепции компьютеризации обучения возникли более 30 лет назад под лозунгом “техническое перевооружение педагога, механизация его труда”, “обучающей машины”, имитирующей процесс индивидуальной работы преподавателя с обучаемым. Со временем росло понимание их ограниченности. Определенное распространение получили попытки использовать в учебном процессе системы искусственного интеллекта, орудийный подход к компьютерам, по которому ЭВМ рассматривается прежде всего как чисто техническое средство деятельности. Это заведомо обедняет возможности компьютера в учебном процессе, ограничивает число функций, которые он мог бы выполнять. Но решающая роль принадлежит новым информационным технологиям, в которых компьютер рассматривается как средство учебной деятельности, изменяющее цели и содержание образования на базе информатизации общества.

Разработка новой концепции компьютерного обучения, учитывающей специфику как индивидуальных, так и коллективных форм познания и мотивации, приобретает статус глобальной проблемы. Традиционная система обучения оказывается совершенно непригодной: чтобы развить и оценить творческие способности каждого обучаемого, преподаватель должен переработать огромные объемы информации. В условиях массовой подготовки специалистов привычные методы использования компьютеров не позволяют раскрыть потенциальные возможности студентов, найти рейтинг каждого и предложить индивидуальную траекторию обучения. Применение компьютера, непосредственно встроенного в управление процессом обучения, требует принципиальной перестройки содержания обучения, приближения его к реальным потребностям жизни.

## **ГЛАВА 1. СИСТЕМА КОМПЬЮТЕРНОГО КОНТРОЛЯ КАК РЕЗУЛЬТАТ НОВЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ОБУЧЕНИИ**

Одной из важных составляющих учебного процесса является контроль знаний и умений обучающихся, опыта их творческой деятельности, эмоционально-ценностного отношения к изучаемому материалу. Значимость контроля трудно переоценить: при эффективной организации он выполняет все три функции процесса обучения: образовательную, развивающую, воспитывающую.

Существуют различные виды и формы контроля, но нас интересует итоговый контроль знаний в форме компьютерного тестирования. Компьютерный контроль знаний в последние десятилетия активно внедряется в систему отечественного образования. Важнейшей его задачей является получение объективной оценки знаний. Примером такого контроля может служить Единый Государственный Экзамен (ЕГЭ), контроль по учебным курсам вузов. Возникли и успешно развиваются целые фирмы, предоставляющие такой контроль учебным заведениям в форме услуги, в том числе, благодаря Интернету, и в интерактивном дистанционном режиме, например АСТ.

В настоящее время определяются следующие направления использования компьютерной технологии обучения в учебном процессе:

- индивидуализация обучения;
- выработка практических навыков и тренировка;
- контроль и оценка знаний.

Остановимся более подробно на третьем направлении – контроль и оценка знаний. Система обучения представляет собой многогранный процесс, состоящий из многих взаимосвязанных элементов. Среди них важное место занимает контроль знаний, навыков, умений. Контроль сопровождает все виды учебной деятельности, без него нельзя проводить управление любым процессом или системой. Осуществление контроля знаний, умений обучаемых, в первую очередь, включает решение проблемы определения совокупности требуемых качеств знаний, без которых критерии оценки знаний и способы определения уровня их усвоения выявить нельзя. Следовательно, задачи, возникающие при разработке средств контроля — это конкретизировать цели, установить адекватность средств и результатов обучения, определить значимость

результатов.

Осуществление контроля знаний обучаемых, в свою очередь, логично приводит к решению проблемы создания надежных методов диагностики качества знаний, способствующих оперативному управлению процессом их усвоения. Актуальность задачи автоматизации процедуры контроля уровня обучаемости определяется целым рядом факторов:

1. освобождения преподавателя от выполнения трудоёмкой и рутинной работы, предоставление времени для творческого совершенствования разных аспектов его профессиональной деятельности;
2. обеспечение всесторонней и полной проверки;
3. повышение объективности контроля и обеспечение его стандартизации;
4. оперативность статистической обработки результатов контроля.

Наиболее эффективно использование компьютера в процессе контроля знаний и умений при наличии самой системы компьютерного контроля (СКК).

В ходе педагогического эксперимента были определены основные требования, предъявляемые к СКК:

1. расчет на различные предметные области и способы деятельности;
2. охват всех периодов обучения и изменяемость в зависимости от них;
3. содержание различных форм, приёмов и способов;
4. зависимость уровня проблемности контроля от уровня проблемности содержания и индивидуального уровня обученности;
5. контроль за степенью достижения каждым обучаемым целевого уровня подготовки;
6. поддержка всех этапов обучения: от целеполагания до оценочно-результативного этапа.

Как и любая система, СКК состоит из различных элементов. Основными в ней являются следующие:

- определение цели контроля;
- отбор учебного материала;
- определение содержания контроля;
- обеспечение процесса контроля;
- сбор, обработка и анализ результатов.

Компьютер, непосредственно встроенный в информационную технологию обучения, становится обязательным элементом образовательной системы, обеспечивает переход к реальной перестройке обучения на основе индивидуального развития личности, повышения уровня профессионализма и культуры через интеллектуальные информационные системы. Внедрение в процедуру контроля СКК – это качественно новое средство управления учебной деятельностью, которое должно опираться на комплексную диагностику при осуществлении единства контролирующей, обучающей, воспитывающей и развивающей функций контроля.

## ГЛАВА 2. ОСОБЕННОСТИ ОЦЕНИВАНИЯ КАЧЕСТВА ОБУЧЕНИЯ

Рост числа студентов, охваченных новыми формами высшего и среднего профессионального образования, внедрение системы тестирования в общеобразовательных учебных заведениях, наконец, широкое распространение многочисленных сертификационных циклов в системе профессиональной подготовки - это основные факторы, которые способствовали росту интереса к возможностям информационных технологий по оцениванию качества обучения. Но в качестве неотъемлемого требования, предъявляемого к процедуре оценивания, необходима гарантия того, что использующиеся методы оценки адекватно отражают уровень достижения целей изучения учебного курса, приобретения знаний, умений и навыков, развития личностных качеств обучаемых.

Особенности качественных и количественных подходов к оцениванию, методика измерения его результатов, наконец, использование различных технических средств довольно долго находятся в центре внимания отечественной педагогической науки. В исследованиях и нормативных документах последних лет представлены четкие требования к используемым заданиям, критериям оценивания, разработаны методики контроля качества обучения и соответствующие технологические процедуры.

### ***§ 1. ТЕРМИНОЛОГИЯ***

Предварительно необходимо разъяснить суть тех терминов, которые используются в описании процедур оценивания качества обучения и развития.

**Оценивание**, по сути дела, сводится к изучению какого-либо образца поведения человека, полученного в некоторый фиксированный момент времени. Основной вопрос: насколько представительным является данный образец поведения оцениваемого человека, поскольку именно на его основе делаются выводы о его статусе — достижениях, потенциале и способностях, интеллекте, мотивации. Для проведения оценивания в первую очередь необходимо поймать, с какой целью выполняется оценивание и кто оценивается. Это очень важно, поскольку на одном и том же материале в ходе тестирования можно оценивать обученность или обучаемость, реакцию учащихся, поведение в сложной обстановке, эффективность использующихся методических приемов, наконец,



значимость самого теста. В частном случае, для процесса обучения, оценивание ставит своей целью получение *оценки*, содержащей как качественные, так и количественные показатели работы обучаемого.

**Критериально-ориентированное оценивание** опирается на явно сформулированные цели и задачи изучения учебного курса и позволяет определить, в какой степени испытуемые овладели идентифицированными компонентами. В этом типе оценивания критерии усвоения задаются заранее.

**Нормативно-ориентированное оценивание** используется в тех случаях, когда необходимо ранжировать достижения в выбранной группе обучаемых, определяя количество достигнувших определенного балла в течение некоторого времени. Соответствующие тесты применяются не только для оценивания хода и результатов обучения, но и для оценки интеллектуальных способностей. Естественно, что соответствующие показатели для ранжирования могут устанавливаться, исходя из конкретного контекста процедуры оценивания: одно и то же значение показателя интеллектуального развития может характеризовать высокий результат для одной группы и средний или даже низкий – для другой.

Для оценивания результатов учебно-познавательной деятельности используются различные функциональные подходы, которым можно придать следующий смысл: **диагностический** — для идентификации сильных и слабых сторон; **обобщающий** — для получения итоговой оценки в конце работы с единицей изучения. Однако на практике в процедуре оценивания первый подход зачастую выступает в качестве подчиненного. Например, результаты, полученные в ходе диагностического оценивания (контрольная работа) могут учитываться в обобщающем и т.д. Представления о сущности диагностики обучения связывают с ней не только проверку знаний, умений и навыков обучаемых, но и возможность рассмотрения полученных результатов в связи со способами их достижения. Анализ данных диагностирования позволяет выявлять тенденции, прогнозировать дальнейший ход учебно-воспитательного процесса и эффективно управлять им. В контексте применения в процедурах оценивания ИТО основной акцент делается на **педагогическое тестирование** — совокупность методических и организационных мероприятий, обеспечивающих разработку педагогических тестов, подготовку и проведение стандартизированной процедуры измерения уровня подготовленности испытуемых, а также обработку и анализ результатов. Разработка педагогических тестов ведется с методическим обоснованием их применения и обработки результатов тестирования, с учетом основных психолого-педагогических

принципов обучения. Задания в тестах должны быть подобраны таким образом, чтобы можно было проверить основные уровни усвоения обучаемыми знаний. Напомним, что к их числу относят:

1. знание основных понятий и определений изучаемой темы;
2. понимание и умение применять полученные знания при решении типовых задач;
3. умение анализировать различные ситуации, находить решения нестандартных задач;
4. умение обобщать изученный материал, устанавливать связи с ранее изученными темами.

Этим уровням соответствуют определенные разновидности тестовых заданий. Тесты с **заданиями закрытого типа**, содержание которых сопровождается несколькими занумерованными вариантами ответа, а испытуемому предлагается выбрать номер (номера) правильного ответа, могут использоваться для проверки разных уровней усвоения:

- выбор одного правильного ответа из нескольких предложенных соответствует уровням 1 и 2;
- выбор нескольких правильных ответов из предложенного списка — уровням 2 и 3.

Задания открытого типа, в которых испытуемому предлагается самостоятельно указать правильный ответ без указания возможных вариантов ответа, пригодны для проверки всех уровней усвоения знаний. Если учащимся младших классов, изучающим правописание гласных после шипящих, предлагается вписать пропущенную букву, то речь идет о 1-м уровне. Но в тесте может быть представлена нестандартная задача, вопрос или задание (без указания возможных вариантов ответа), позволяющие оценить усвоение изученного материала на 3-м и 4-м уровнях. В естественно-научных областях это может быть и расчетная, и качественная задача, ответ которой должен быть однозначным. Более сложно подготавливать тестовые задания 3-, 4-го уровней для гуманитарных дисциплин. Но и это вполне возможно. Вспомним каверзные вопросы из телевизионной передачи «Умники и умницы» на сопоставление различных исторических фактов, также предполагающие вполне конкретные ответы.

## **§ 2. СЛОЖНОСТИ И ПРОБЛЕМАТИЧНОСТЬ ОЦЕНИВАНИЯ ЗНАНИЙ**

Каждому опытному преподавателю известно, что даже простой опрос ученика или студента с целью объективно оценить его знания,

далеко не так прост. Первая сложность заключается в индивидуальных психолого-физиологических особенностях, как ученика, так и преподавателя. Давая оценку знаний учащемуся, преподаватель дает оценку и себе, т.е. такая оценка носит субъективный характер. Субъективность оценки сохраняется, а возможно и усугубляется и при компьютерном контроле. Возникает задача из оценки знаний ученика выделить и удалить долю субъективности оценки экзаменатора. Один из традиционных подходов к решению этой задачи заключается в опросе и оценке знаний одного ученика разными преподавателями с последующим расчетом коэффициента конкордации и вычислением объективной оценки. Прием важных экзаменов комиссией по сути (без явного вычисления конкордации) решает эту задачу. Следовательно, для серьезного компьютерного тестирования по одному предмету, как минимум, нужно подготовить и провести несколько тестов (несколько разных тестов, а не несколько тестирований по одному тесту) и затем обеспечить проверку согласованности и выставление окончательной оценки знаний. На практике это нигде не реализуется. Обычно, главный тезис в защиту компьютерного тестирования – его объективность, т.е. безразличие компьютера к испытуемому. К сожалению, это не достоинство, а недостаток. И убежденность приверженцев такого метода в объективности компьютера ошибочна и опасна. Поясним на примере. Обычный экзаменатор, сформулировав вопрос ученику, сам не осознавая того, оценивает, понял ли испытуемый вопрос и правильно ли понял. Иногда это проявляется в явной форме: «Вопрос понятен?». Если непонятен, формулирует тот же вопрос другими словами. То же самое - в процессе ответа ученика: экзаменатор уточняет правильность ответа дополнительным вопросом в адрес испытуемого. Некоторые компьютерные тесты достаточно сложны и содержат элементы, направленные на повышение объективности оценки. Так в конечной оценке знаний индивидуума одна известная компьютерная система тестирования опирается на средний уровень оценок группы испытуемых (проверяет конкордацию оценок группы испытуемых).

Теперь рассмотрим трудности и особенности тестирования отдельных предметов. Надо отметить, что, чем менее формализован предмет, чем он более гуманитарен, тем проблематичнее его компьютерное тестирование. Главным отрицательным фактором компьютерного тестирования является требование **однозначности** ответа на вопрос теста. В компьютерных тестах вопрос должен быть так сформулирован, что бы он не допускал двусмысленного его толкования и, соответственно, имел единственный бесспорный правильный ответ.

Такому условию отвечает проверка знания таблицы умножения: «Сколько будет  $2 \times 2$ ?» Ответ единственный – «4», а не «3», и не «5». К сожалению, нет уверенности, что среди испытуемых не найдется вундеркинд, увлекающийся алгеброй, который поймет вопрос как «Сколько будет 2 «икс» 2» и надолго задумается. Результат тестирования слишком умного ученика будет отрицательным.

Таким образом, компьютерный контроль знаний даже для очень конкретных дисциплин весьма опасен, что же говорить о многих естественнонаучных (например: ботанике, биологии и т.п.) и, тем более гуманитарных (история, философия и т.д.) дисциплинах. В таких тестах можно проверить только знание некоторых фактографических сведений: «В каком году...?», «Как звали ...?», «Кто сказал...?» и т.п. Ответ на такие тесты чаще всего является выбором из нескольких заданных вариантов ответов. И эффективность такого теста не слишком повышается, если авторы заложили в предложенные варианты ответов возможность двух правильных ответов или другие варианты ответа и т. д.

Особенно неэффективен такой вариант компьютерного тестирования, когда ответом на тест является ответ в произвольной текстовой форме (вопросы с открытыми ответами). Внешне такое тестирование очень похоже на естественный письменный экзамен с человеком-преподавателем. По своей сути, такой тест работает плохо по отношению к думающим и знающим ученикам и студентам, так как компьютер пока не способен к интеллектуальному анализу ответов. Ответ контролируется формально, например, по таблице возможных правильных ответов или правил образования правильных ответов. Яркий пример такого контроля – контроль грамматики в текстовом редакторе MS WORD – далеко не всякое красное или зеленое подчеркивание указывает на ошибку. Насколько мягко MS WORD указывает на ошибку и всегда допускает, что это не ошибка вовсе (см. команду «ПРОПУСТИТЬ»). Пример из практики. В одном из тестов по педагогике студентам предлагалось продолжить предложение: «Школа для крестьянских детей в Ясной поляне была открыта ...». Зная, что Яснополянская школа была открыта Л.Н. Толстым, студентам пришлось пробовать разные варианты, чтобы ответ был оценен правильно: Л.Н.Толстым, графом Толстым, Толстым Л.Н. и т.д. – правильным оказался один единственный «Л.Н. Толстым», т.е. с **обязательным** пробелом между второй точкой инициалов и фамилией, все остальные считались ошибками. Однако в компьютерных тестах такие виды тестирования встречаются и по ним выносятся

однозначные оценки знания испытуемым. А если вспомнить, что авторы правильных ответов тоже не все знают, то опасность такого компьютерного контроля знаний налицо.

На педагогическом факультете МГПУ проводилась экспериментальная проверка различных систем компьютерного тестирования для круга предметов естественнонаучного и гуманитарного профиля. Параллельно с компьютерным контролем знаний студентов использовались традиционные методы контроля: письменные контрольные, собеседование, зачеты, экзамены, разнообразные творческие работы, в том числе с использованием компьютерных технологий. Анализ полученных результатов позволяет утверждать, что средние и слабые студенты на компьютере часто получали завышенные оценки, а знающие и творчески думающие студенты — заниженные. В порядке эксперимента тестированию подверглись и преподаватели: профессора, доценты, старшие преподаватели. Тестирование преподавателей-предметников сопровождал инженер или преподаватель, хорошо владеющий компьютером. Результат ошеломляющий: по большинству тестов обоснованная критика, причем, чем выше уровень преподавателя, тем больше обоснованных нареканий к тестам, как способу объективной оценки знаний студентов по предмету.

В ряде тестов, несмотря на утверждения фирм-разработчиков систем компьютерного тестирования о том, что к подготовке вопросов и ответов в тестах были привлечены «ведущие ученые и преподаватели страны», правильные ответы были неправильными, либо вообще отсутствовали. Таким образом, компьютерное тестирование знаний на педагогическом факультете валидизацию не прошло.

### ***§ 3. ПРЕДПОСЫЛКИ ИСПОЛЬЗОВАНИЯ ИТО В ПРОЦЕДУРАХ ОЦЕНИВАНИЯ***

Можно и нужно ли привлекать средства информационных технологий в практику проведения процедур оценивания? Да, и вот некоторые причины.

Информационные технологии могут использоваться в процедурах оценивания на различных уровнях - от средства управления информацией об оценках, полученных традиционными способами, до полностью автоматизированных систем контроля качества знаний, включающих в себя диагностический, обобщающий и коррекционный модули. В первом случае можно обеспечить представление имеющейся информации в

соответствии с потребностями самых различных лиц - педагогов, обучаемых, администрации учебного заведения и пр. При этом с помощью современных программных средств (электронные таблицы, системы управления базами данных, пакеты статистической обработки) можно улучшить не только внешние характеристики, но и содержательную часть имеющихся данных. Если для самих обучаемых важно узнать набранные баллы или полученные отметки, то для администрации учебного заведения наиболее подходящей формой представления будут диаграммы и графики с иллюстрацией тенденций изучения определенной дисциплины, сравнением результатов, полученных в разных учебных группах, и т.д. Компьютеризация тестирования по сравнению с использованием бумажных технологий открывает ряд интересных возможностей. Создание и развитие базы данных с вопросами позволяет постоянно совершенствовать контроль знаний (но здесь имеются свои проблемы, связанные с идентификацией сложности вопросов и валидности, т.е. значимости получаемых тестов). Такой электронный «классный журнал» благодаря возможностям обобщения и анализа информации помогает преподавателю осуществлять оперативное управление учебным процессом.

Применение ИТО позволяет педагогу привлечь новые и улучшить традиционные методы оценивания, гарантирует качественное выполнение процедуры оценивания. (Педагоги ощущают недостатки традиционной системы контроля, которые вполне очевидны и являются во многом продолжением ее достоинств.)

Устный опрос достаточно универсален и его реализация, на первый взгляд, не представляет проблем. Однако он может быть только выборочным и не давать педагогу возможности проверить знания всех обучаемых, а в случае слабых ответов иногда превращается в своеобразное выяснение отношений. Тестовый контроль в такой ситуации имеет преимущества, позволяя без особых затрат времени опросить всех обучаемых по всем темам изучаемой дисциплины и способствуя повышению рентабельности образования за счет экономии времени преподавателей. Такая форма контроля как экзамен в силу субъективности педагогов часто не позволяет обоснованно оценить уровень знаний обучаемых, а тестовый контроль достаточно универсален и может использоваться как в средней школе, так и в высшей. Преимущество тестового контроля состоит в том, что он является научно обоснованным методом эмпирического исследования и позволяет преодолеть субъективные оценки знаний обучаемых.



Нельзя отрицать тот факт, что в точных науках, где и учебный материал, и требования к качеству обучения структурируются и формализуются естественным образом, составлять тестовые задания легче. Но преподавание всех учебных дисциплин основано на требованиях к базовым знаниям. Например, для такой учебной дисциплины, как история, важно «знание событий, дат, имен, определений основных понятий и многого другого. Проверка базовых знаний средствами тестового контроля позволяет преподавателю в оставшееся время уделить больше внимания общению с обучаемыми на уровне концепций и выводов, проверить традиционными формами не столько знание, сколько понимание проблематики той или иной учебной дисциплины. Следует подчеркнуть, что именно проверка базовых знаний является наиболее доступной сферой для применения тестового контроля». Современные системы тестирования отличаются определенной гибкостью, когда обучаемым можно выбрать индивидуальный график прохождения контрольных точек, а при тестировании — конкретный режим: попытаться ответить на большее число вопросов за большее время или, наоборот, ограничить количество вопросов, но получить меньшее время; выбрать меньшее число трудных опросов или большее число простых и т.п. Системы тестирования часто предлагают испытуемым работу в режиме самоконтроля с заданиями, аналогичными тем, что будут предложены им впоследствии в качестве контрольных. В этой связи выделяют **репетиционные тесты**, позволяющие проверить степень готовности испытуемого к педагогическому тестированию, знакомящие с порядком работы, объемом и сложностью заданий, иногда даже предлагающие справочные материалы, подсказки и т.п. Проведение таких тестов одинаково важно и для педагогов, и для обучаемых, поскольку по их результатам можно судить о качестве полученных знаний и самих тестовых заданий.

Опытные преподаватели знают, как разнообразит учебный процесс и повысит познавательную мотивацию обучаемых использование каких-то новых элементов в преподавании. Автоматизированные тесты привлекают своей необычностью по сравнению с традиционными формами контроля, возможностью проведения быстрого и объективного оценивания качества знаний. Педагогические тесты при регулярном использовании побуждают к систематическим занятиям по предмету, что способствует формированию дополнительной мотивации к обучению. Оперативность обработки тестов обеспечивает эффективную обратную связь, а в условиях, когда обучаемые могут проходить испытания так часто, как им это потребуется, педагог может добиться гарантированного усвоения

базовых знаний, умений и навыков. Для решения вопросов о конструкции контролирующей системы, стратегии оценивания, использующихся для этого методов необходим анализ того, какой тип обучения будет реализовываться в каждом конкретном случае, например, для воспроизводящего типа обучения сами принципы контроля должны в корне отличаться от тех, что могут быть использованы для поискового типа. Существуют проблемы, связанные с чисто измерительными аспектами в процедуре оценивания. Это и выбор валидных материалов для тестирования, и определение подходящих единиц измерения, а также обеспечение того, чтобы процедура оценивания измеряла именно то, что должно быть измерено, надежность оценок и адекватность использующихся для их обработки статистических методов.

#### § 4. АВТОМАТИЗИРОВАННОЕ ТЕСТИРОВАНИЕ

**Предметные тесты.** Говоря об эффективности информационных технологий для организации процедуры оценивания на основе педагогических тестов по определенным предметам, или предметных тестов, ведут речь об использовании вопросов с набором вариантов возможного единственного ответа — тесты типа *MCQ* (англ. *Multi Choice Question*, т.е. вопрос с множественным выбором). Но может потребоваться применение и других вариантов опроса, например, необходимость допускать многократные ответы, ввод текста, чисел в фиксированном или в свободном формате, когда сравнение введенного ответа с эталоном ведется по так называемым ключевым словам. Автоматизированные тесты типа *MCQ* представляют достаточно эффективный метод массовой проверки уровня фактических знаний за относительно короткое время. Однако педагоги должны понимать ограниченность возможностей подобного тестирования. Это касается использования тестов *MCQ* для оценки знаний, умений и навыков, связанных с продуктивной деятельностью обучаемого, поскольку подобное тестирование способно выполнить оценку лишь на уровне узнавания или воспроизведения изученных ранее объектов. Но это не отрицает возможности тестов для объективной проверки качества обучения в ходе диагностического и обобщающего тестирования, разумеется, при условии использования большого количества продуманных вопросов с таким количеством вариантов, которое бы поставило в затруднительное положение тех обучаемых, которые рассчитывали угадать ответ. Современные контролирующие программы обычно не знают ограничений, связанных с необходимостью использования формул (математика, химия и др.): в этих



случаях используются или специальные символы, или в текст вопроса включаются графические объекты.

**Структурирование вопросов и адаптивные тесты.** Автоматизированное тестирование может основываться на различных алгоритмах предъявления вопросов испытуемому, в том числе и на изменяющих ход опроса в зависимости от успешности ответов. Современные контролирующие системы способны адаптироваться и к неверным ответам, предъявляя в этом случае так называемые наводящие вопросы или даже вопросы, содержащие подсказку. Такие функции превращают подобную систему в обучающую. Естественно, что алгоритм экзамена или обучения должен быть запрограммирован заранее так, чтобы с помощью одной и той же программы на одной и той же базе данных с вопросами можно было реализовать и контроль, и обучение. Для наполнения таких систем от преподавателя требуется большая работа по структурированию вопросов: сложный вопрос в случае неверного ответа должен предъявляться в несколько приемов, чтобы даже более длинным путем, но подвести обучаемого к правильному ответу. Экзамен или зачет, проводящийся с помощью системы автоматизированного тестирования, состоит в том, что экзаменуемому задается определенное количество вопросов независимо от того, насколько хорошо или плохо он на них отвечает. Количество набранных баллов при использовании теста такого рода зависит от количества правильных ответов. При этом делается предположение - чем выше качество знаний, тем на большее количество вопросов экзаменуемый отвечает правильно. Такая форма тестирования распространена и используется весьма успешно, однако в применении к конкретному испытуемому количество заданных вопросов может оказаться больше или меньше, чем необходимо для получения адекватной оценки качества его знаний. На практике сложно подготовить тест с вопросами одинаковой сложности: в тесте фиксированной длины могут быть вопросы, которые для определенного человека окажутся слишком легкими, и вопросы слишком трудные для него. В этом случае верные ответы на легкие вопросы и неправильные ответы на трудные вопросы не придадут такому тесту должной степени валидности.

Гораздо лучше, если бы **нелинейная тестирующая система** могла определять тот уровень сложности вопросов, на котором у экзаменуемого начинают возникать проблемы. Этот уровень мог бы как определить оценку (для экзаменатора), так и выявить сложные места (для экзаменуемого). Целесообразность подобного контроля, адаптирующегося к возможностям обучаемого, следует также из необходимости оптимизировать традиционное тестирование. Для каждого педагога

очевиден тот факт, что для обучаемых с хорошей подготовкой легкие задания просто неинтересны, и, наоборот, трудные задания снижают мотивацию к обучению у имеющих относительно слабую подготовку.

Новым шагом в этом направлении стал *CAT* (англ. *computer adaptive test* — компьютерный адаптивный тест). В этом тесте заложена приспособляемость к возможностям экзаменуемого. Принцип тестирования с использованием *CAT* состоит в следующем: при выполнении одного и того же адаптивного теста экзаменуемые с высоким и низким уровнями подготовки получают совершенно разные наборы вопросов: первому будут предложены сложные вопросы, а второму — легкие. Если в итоге доли правильных ответов у обоих даже совпадут, то первый наберет большее количество баллов, так как он отвечал на более сложные вопросы. Фирма Microsoft, разрабатывающая и широко использующая такую форму тестов, предлагает для иллюстрации их особенностей сопоставление с соревнованиями по прыжкам в высоту. Прыгун, независимо от его способностей, быстро достигает такого уровня планки, на котором он имеет примерно равные шансы, как взять высоту, так и сбить планку. «Баллом» для прыгуна является последняя взятая высота. Для получения высокого балла прыгун не должен брать каждую возможную более низкую высоту, также он не должен пытаться брать более высокий уровень планки. Пример из области образования будет более наглядным. В ходе устного экзамена учитель обычно сначала задает вопрос средней сложности, и если ученик отвечает правильно, то ему предлагается более сложное задание. В случае же первого неправильного ответа в качестве следующего задается более легкий вопрос. Этот процесс продолжается, и в течение короткого периода времени у учителя постепенно складывается правильное представление о качестве знаний ученика. При этом ему не надо задавать каждому испытуемому слишком легких или слишком сложных вопросов, а достаточно отталкиваться от того уровня сложности вопросов, на которые ученик дал правильные ответы. Так же должен быть организован нелинейный тест типа *CAT*, обеспечивающий проведение контроля качества обучения на уровне квалифицированного устного экзаменатора. В таком тесте первоначально задается вопрос средней сложности, и полученный ответ немедленно влияет на постепенно формируемую будущую общую оценку. Если ответ правильный, то предполагаемая оценка возможностей экзаменуемого повышается на определенную величину. Затем выбирается и задается более сложный вопрос. Если ответ на него дан неправильно, то предполагаемая оценка возможностей экзаменуемого снижается, а в

качестве следующего вопроса снова предлагается более легкий. По мере того как задаются новые вопросы, более точной становится оценка уровня знаний экзаменуемого. Тест заканчивается, когда точность оценки достигает статистически приемлемого уровня (или когда будет задано максимальное количество вопросов). Так как точно неизвестно, когда адаптивный тест закончится, то обычно он состоит из переменного количества вопросов, причем минимальное и максимальное значения для количества вопросов устанавливаются заранее.

При прохождении адаптивного теста, возможно, что к завершению испытания менее подготовленный человек может ответить на такое же количество вопросов, что и более подготовленный. Сравнение вопросов, на которые даны правильные ответы, покажет, что более подготовленный ответил правильно на более сложные вопросы. Следовательно, получит более высокие баллы. Количество набранных баллов не основано на количестве правильных ответов, а зависит от уровня сложности вопросов, на которые даны правильные ответы.

Главное преимущество адаптивного теста перед традиционным - его эффективность. Адаптивный тест может определить баллы экзаменуемого с помощью меньшего количества вопросов, до уменьшения длины теста на 60%, это - главная причина, по которой следует отдавать предпочтение адаптивным тестам.

**Критериально-ориентированные тесты.** Для объективной оценки достигнутого качества обучения, в том числе и при работе с обучающими программами, особый интерес представляют критериально-ориентированные тесты. Проект отраслевого терминологического стандарта Центра тестирования определяет критериально-ориентированный тест как частный случай теста, предназначенного для абсолютного персонального тестирования, позволяющий оценить, преодолел ли испытуемый определенный порог усвоения учебного материала. При этом результаты тестирования сравниваются с некоторым заранее заданным критерием уровня подготовленности. Таким образом, речь идет не столько о самих тестах, сколько об интерпретации тестовых результатов.

Педагог может получить ответ на вопрос о том, какие элементы содержания учебной дисциплины усвоены конкретным испытуемым, по сути дела, только в виде вероятностной оценки. При подготовке таких тестов на основе содержания учебной дисциплины строится **генеральная совокупность**, т. е. однородное множество заданий для измерения качества полученных знаний, умений, навыков. Затем испытуемому предлагается тест — некая выборка заданий из этой совокупности.

Наконец, на основе ответов делается вероятностный вывод о знаниях учебной дисциплины данным испытуемым. Подобные тесты в оригинале носят название Domain-Referenced Tests, что дословно означает **содержательно-ориентированные тесты**. Понятно, что для надежности результатов требуется основательное определение содержания изучаемой дисциплины и большое число заданий. При этом необходимо, чтобы соблюдались:

- 1). полнота отображения материала образовательной программы при отборе содержания;
- 2). правильность пропорций отдельных разделов и тем предмета (содержательных линий);
- 3). полнота охвата требований государственных образовательных стандартов;
- 4). соответствие содержания заданий знаниям, умениям и навыкам, запланированным для проверки в спецификации тестовых материалов;
- 5). значимость содержания каждого задания для целей проверки.

Подобные тесты можно использовать при проведении экзаменов с точной дифференциацией результатов, поскольку они позволяют получить абсолютную оценку качества обучения. Задания для такого тестирования должны быть ориентированы на диагностику различных уровней усвоения учебного материала: от воспроизведения фактов, понятий, законов и их применения в типовых ситуациях до систематизации и обобщения знаний, позволяющих найти ответ на проблемные вопросы, решить нестандартную задачу и т.д. В таком тесте задания разного уровня имеют разный «вес» — и по тому, как оценивается их выполнение, и по их относительному числу в общей массе заданий. Подобная особенность характеризует, например, тесты, использующиеся при проведении Единого государственного экзамена. Иногда испытуемым может быть предложена и относительно небольшая выборка **однородных** заданий. Такие тесты можно использовать, например, для проверки овладения (на уровне узнавания и воспроизведения) сравнительно ограниченным набором знаний, умений и навыков, выступающих в качестве заданного стандарта или критерия усвоения учебной дисциплины при проведении зачетов. Здесь видно не только объективное измерение разноуровневой подготовки, но и достижение тем или иным испытуемым минимально допустимого уровня. Такой подход удобен для педагогов и организаторов образования в тех случаях, когда необходимо проверить достижение большой группой обучаемых предельно допустимого уровня требований

(при аттестации учебного заведения). В таких случаях **критериально-ориентированная интерпретация** результатов тестирования позволяет сделать вывод: что из заданного стандарта и на каком уровне реально усвоено.

### **§ 5. ДОПОЛНИТЕЛЬНЫЕ ВОЗМОЖНОСТИ ИТО В ПРОЦЕССЕ ОЦЕНИВАНИЯ КАЧЕСТВА ОБУЧЕНИЯ И РАЗВИТИЯ**

Перечислим еще ряд возможностей, открывающихся благодаря использованию ИТО. Помимо непосредственного тестирования имеются и другие направления в процессе оценивания уровня обучения и развития, где информационные технологии могут обеспечить качественно новые результаты.

**Случайный выбор параметров вопроса.** Автоматизация тестирования открывает альтернативный метод создания вопросов (чаще - контрольных заданий) с помощью вариаций случайно подбираемых параметров вопроса. Такие вопросы имеют фиксированный формат, включающий одну или несколько переменных составляющих, которые могут изменяться при составлении вопроса или в ходе тестирования - случайным образом или по некоторой формуле. Например, для расчетных заданий могут задаваться случайным образом значения исходных данных, в тестовых заданиях на знание тех или иных определений словосочетания «необходимое условие» на «достаточное условие» и т.п. В таких тестах сами вопросы носят формальный характер, однако для многих обучаемых подобные тесты оказываются очень полезными, особенно когда необходимо отработать определенный автоматизм реакции на ту или иную ситуацию; тогда речь идет, скорее, о тренировочной, чем контролирующей системе.

**Создание сетевой базы данных для хранения вопросов.** Организация коллективного доступа к базе данных, хранящих тестовые задания и вопросы, актуальна при создании единой образовательной среды для учебных заведений любого уровня. С помощью современных коммуникационных технологий (электронная почта, электронные конференции, дискуссионные группы на образовательных (Web-сайтах) ее пополнением и развитием могут заниматься все заинтересованные педагоги, что значительно улучшает качество тестирования.

**Автоматизированные системы регистрации и анализа результатов оценивания обученности.** Говоря о возможностях информационных технологий для оценивания качества обучения, довольно часто оставляют без внимания ту сферу их применения, которая позволяет добиться быстрых и эффективных результатов. Речь идет о регистрации,



хранении, анализе данных по контролю обученности, а также их использовании для оперативного и долгосрочного управления образовательным процессом. Для этой цели педагоги и администрация учебного заведения могут использовать электронные таблицы, системы управления базами данных, пакеты статистической обработки. Все эти программные средства позволяют вносить имеющиеся данные вручную - в том случае, когда нет возможности автоматически их сгенерировать в электронном виде (например, если соответствующий контроль проводится традиционно - письменная контрольная работа или сочинение, опрос на уроке и т.п.). Практически все современные программные средства, относящиеся к перечисленным категориям, воспринимают данные, подготовленные с помощью других программ (в нашем случае - тестирующих) и имеющие какой-либо из стандартных форматов (текстовый, табличный), что потенциально расширяет их возможности. Подобная автоматизация позволяет сопоставлять и анализировать качество усвоения различных учебных дисциплин, выявлять и прослеживать те или иные тенденции, проявляющиеся при оценке обученности, что обеспечивает высокий уровень обратной связи и управляемости образовательным процессом.

***Итоговые оценки и другая информация - что важнее.*** В ряде случаев, особенно для формирующего тестирования, только лишь информация о том, на какое количество вопросов был дан правильный ответ, явно недостаточна для управления образовательным процессом по данной учебной дисциплине конкретного обучаемого. Здесь на помощь преподавателю может прийти динамическое отслеживание хода тестирования, когда в специальный файл или базу данных записываются все ответы обучаемого, которые впоследствии могут использоваться для более глубокого анализа и диагностики усвоения учебного материала. Также в подобных системах фиксируется полный протокол работы испытуемого: количество попыток, предпринятых для прохождения теста, время, затраченное на ответы на отдельные вопросы и тест в целом. Протоколирование хода тестирования открывает возможность анализировать не только качество усвоения знаний, умений и навыков, но и личностные особенности обучаемых, проявляющиеся в своеобразии прохождения тестирования. Например, протоколирование данных позволяет выделить среди «неудачников» тех, кто стремится к наилучшему результату, затрачивая много времени и делая неоднократные попытки. Противоположной считается категорию лиц, которые ограничиваются более низкими результатами, но тратят на тестирование

значительно меньше времени, чем представители первой группы. Обсуждение с обучаемыми не только результатов прохождения теста, но и использованной стратегии поможет педагогу сориентировать их в нужном направлении: в случае недостаточно высоких результатов при первой попытке желательно направить силы обучаемого на устранение пробелов в подготовке, а затем пройти повторное тестирование. Подобный анализ особенно уместен на этапе промежуточного, диагностического тестирования для улучшения обратной связи. И если на стадии итоговой проверки качества обучения вышеперечисленные параметры (количество попыток, время, затраченное на ответы) играют решающую роль, педагог сможет лучше подготовить обучаемых, уже зная слабые стороны каждого из них и руководствуясь целями проводимого тестирования.

**Вопросы безопасности.** При переходе к автоматизированному тестированию преподавателей волнуют вопросы безопасности, защиты тех материалов, на основе которых проводится тестирование, и данных, представляющих его результаты. При подготовке соответствующего программного обеспечения разработчики предусматривают определенные средства защиты: доступ к базе данных с вопросами теста осуществляется по паролю, который обновляется по прошествии определенного промежутка времени. При наличии достаточного числа компьютеров эффективным оказывается проведение одномоментного тестирования для всех обучаемых или, в крайнем случае, с разбивкой потока на две группы, проходящих тестирование непосредственно друг за другом. Кроме того, при одновременном тестировании группы обучаемых можно использовать один и тот же набор вопросов, но предъявлять их в различном порядке. Еще лучше эта проблема может быть решена при наличии базы данных с вопросами, обеспечивающими проведение сопоставимых, но не идентичных испытаний.

**Информационные технологии в качестве инструмента управления.** Имеется множество примеров интегрированных обучающих систем, включающих полную структуру учебного курса: лекции, задания для практической работы, средства проверки качества усвоенных знаний, дополнительные ресурсы для самостоятельной и творческой работы в виде демонстрационных и моделирующих программ. Однако управление контролирующим модулем даже для таких систем может оказаться более эффективным на основе других, самостоятельных технологических средств. Например, в системе дистанционного обучения обучаемый может получить в свое распоряжение программу, но для более эффективной работы необходимо взаимодействие с педагогом-наставником, который очно или с помощью электронной почты может вовремя напомнить о

необходимости подготовки к очередному тестированию, ответить на имеющиеся вопросы, оптимизируя ход образовательного процесса. Компьютеры представляют собой идеальный инструмент для мониторинга частоты обращения к электронным образовательным ресурсам (образовательный сервер, электронная библиотека, обучающие программы и т.п.) для улавливания тенденций в ходе образовательного процесса как на уровне отдельных обучаемых, так и групп, раннего обнаружения проблем, связанных с успеваемостью. Последовательное занесение на протяжении нескольких лет в электронный классный журнал результатов обучения по отдельным разделам учебной дисциплины дает педагогу и администрации учебного заведения возможность провести анализ и сделать выводы о достоинствах и недостатках используемых учебников и об адекватности методических приемов. Наличие такой информации полезно для начинающих педагогов, способных ориентироваться на объективные данные о результатах работы более опытных коллег.

**Психологическая диагностика обучаемых.** Помимо оценивания обученности педагогу важно иметь представление об индивидуальных особенностях обучаемых, о формировании и развитии их личностных качеств: общих и специальных способностей, интеллекта, памяти, быстроты реакции, коммуникабельности и т.д. Только такое комплексное исследование может обеспечить полное представление о ходе образовательного процесса и его результатах. Речь идет о *психологической диагностике*, которая также может быть проведена с помощью автоматизированного тестирования. Разумеется, учитель не должен заменить собой психолога, но представлять себе возможности, достоинства и недостатки автоматизированной психодиагностики должен каждый педагог.

Нужно различать компьютерные версии «бумажных» тестов и компьютерные тесты, специально разработанные с учетом возможностей современных технологий. Чаше тесты последнего типа в бумажном виде уже непригодны, так как они могут использовать мультимедиа-информацию, адаптироваться к действиям испытуемого и т.д. Автоматизируется психологическая диагностика и в тех случаях, когда «бумажного» прототипа не может существовать в принципе. При диагностике параметров внимания, памяти, реакции, технология мультимедиа позволяет предъявлять испытуемым различные стимулы и в зрительном, и в слуховом вариантах. Еще одна область приложения возможностей мультимедиа-технологии — это *ситуационные тесты*, основанные на



ролевых играх. Ситуация, в которой должен проявить себя испытуемый, задается предельно реалистично с использованием аудио-, видеоэффектов, анимации. Более важно, что аналогичный характер носит и представление вариантов возможных реакций испытуемого. Такие тесты помимо диагностической функции могут носить и обучающий характер, наглядно показывая испытуемому последствия его выбора и подсказывая, как можно исправить совершенную ошибку. Использование информационных и коммуникационных технологий коснулось всех этапов психодиагностического тестирования: упростилось создание тестов благодаря использованию специальных систем-конструкторов, облегчилось проведение группового тестирования, резко повысилась оперативность первичной обработки и интерпретации результатов. Прослеживается также тенденция передачи управления тестированием компьютерным программам: если ранее автоматизировались лишь определенные стадии тестирования, например, предъявление материала, первичная обработка данных, интерпретация результатов, то на современном этапе чаще встречаются программы, выполняющие целиком все исследование вплоть до конечных выводов. На первый взгляд, это сводит необходимость участия психолога к минимуму, однако все не так просто и пользоваться такими программами нужно очень осторожно.

С одной стороны, преимуществами компьютерной психодиагностики являются оперативная и безошибочная обработка данных, обеспечение стандартных и объективных условий тестирования для всех испытуемых, автоматизированный контроль за процедурой тестирования (хронометраж, отслеживание недопустимых или пропущенных ответов). Также можно обеспечить наглядность и занимательность процесса тестирования, поддерживая внимание с помощью цвета, звука, игровых моментов, что особенно важно для учащихся младшего возраста. Специалист-психолог также оценит возможность объединения тестов в «батареи», т.е. комплексы с единой итоговой интерпретацией, возможность проведения массовых исследований через локальные сети или Internet.

С другой же стороны, испытуемых нужно готовить к работе за компьютером, и особенности этой работы будут накладываться на данные тестирования. Часть психодиагностической информации просто теряется без личного контакта психолога с испытуемым. Качество и полнота интерпретации результатов тестирования также могут быть весьма ограниченными. Поэтому если педагогу совместно с психологом предстоит принять важное решение в отношении конкретного испытуемого (зачисление в профильный класс, обучение по особой

программе, выбор профессии и т.п.), нужно использовать разностороннюю информацию, не ограничиваясь компьютерной диагностикой: личное общение с испытуемым, анализ результатов предыдущих тестирований, психологическая диагностика родителей и т.д. Компьютерные психодиагностические программы не заменяют собой психологов, и с ними должны работать профессионалы, умеющие точно определить границы их применимости. Компьютерные тесты, поставляющиеся на компакт-дисках или доступные в сети Internet, реализуют как традиционные и надежные методики, заслужившие доверие среди практических психологов, так и требующие длительной проверки. В последнем случае необходимо перед заказом программы ознакомиться с демонстрационной версией. Разобраться в многочисленных предложениях, исчисляемых сотнями тестов, довольно трудно. Поэтому на практике была бы удобной определенная классификация соответствующих программных продуктов. Можно выделить разновидности тестов по следующим признакам.

**По структуре:** а) аналоги бланковых тестов;

б) собственно компьютерные тесты (КТ).

**По количеству испытуемых:** а) КТ индивидуального тестирования;

б) КТ группового тестирования (компьютеры объединены в локальную сеть, на все компьютеры идет подача материала теста, на сервере локальной сети проводится обработка и создание базы данных).

**По степени автоматизации тестирования:**

а) автоматизирующие один или несколько этапов исследования;

б) автоматизирующие все исследования.

**По решаемой задаче:** а) диагностические КТ;

б) обучающие КТ (тесты-тренажеры, развивающие программы, совмещающие диагностику с возможностью тренировки, обучения).

**По адресату:** а) профессиональные психологические (пользователь — психолог);

б) полупрофессиональные (пользователь — не психолог, например, в помощь педагогу, менеджеру по персоналу), с упрощенной интерпретацией;

в) непрофессиональные (развлекательные).

Широко на российском рынке программного обеспечения представлены в различных вариантах следующие компьютерные психодиагностические тесты:

- СМЛ — многофакторный метод исследования личности

(отечественный вариант *MMPI, Minnesota Multiphasic Personality Inventory* — Миннесотский многофазный личностный опросник) на основе диагностики психического состояния, его динамики под воздействием внешних факторов, широкого спектра типологических особенностей личности и поведения, уровня и качества социальной адаптации, специфики защитных механизмов и эмоциональных реакций в стрессовых ситуациях;

- цветовой тест Люшера (диагностика актуального состояния на основе цветовых предпочтений);
- диагностика межличностных отношений Лири (выделяет восемь типов межличностного поведения, их сочетаний и личностных особенностей, существенных для межличностного взаимодействия, выявляет зоны актуальных личностных конфликтов, уровень и направленность межличностных притязаний, причины нарушения общения в малых группах; позволяет определить степень удовлетворенности собой в межличностных контактах и изменение социально-психологических свойств личности под влиянием различных факторов);
- интеллектуальные тесты Айзенка и Векслера (диагностика уровня интеллекта и преобладающего стиля мышления);
- интеллектуальный тест Кеттелла для измерения уровня «флюидного интеллекта», который является относительно независимым от внешних факторов и в большей степени связан с прирожденными интеллектуальными способностями;
- множество тестов частного характера, использующих методики зарубежных и отечественных психологов (проверяющих невротические нарушения, тревожность, взаимоотношения в семье, самоконтроль, активность и т.д.).

Подобное программное обеспечение разрабатывается и специализированными фирмами, и в учебных заведениях и научно-исследовательских институтах. В любом случае официальная поставка программ предусматривает предоставление методических рекомендаций по их установке и использованию, а при необходимости и обучение работе с ними. Благодаря возможностям информационных технологий удастся совместить диагностику развития и обученности, учитывать личностные особенности испытуемых при проведении автоматизированного предметного тестирования. Исходя из показаний диагностики быстроты реакции обучаемого, его психоэмоционального состояния, можно индивидуально подбирать контрольное время, выделяемое для прохождения предметного тестирования, тип заданий (известно, что у некоторых

испытуемых тестовые задания открытого типа вызывают состояние тревожности, влияющее на показатели). Интересным направлением является создание обучающих программ, настраивающихся на определенного обучаемого, и выводящих итоговые результаты процесса обучения, исходя из комплексной диагностики его личностных качеств и обученности: определение социотипа, креативности, уровня компетентности учащихся в заданной тематике. Почти все разновидности существующих тестов могут найти применение в учебных заведениях: тесты интеллекта, общих и специальных способностей, личностные тесты. Все они постепенно становятся необходимым инструментом школьных психологов, а педагогам и администрации часто приходится действовать в соответствии с результатами их проведения.

**Индивидуальное образовательное пространство.** Наряду с необходимостью оценивания качества знаний по учебным дисциплинам не менее важна **оценка** познавательной деятельности обучаемых, их творческой активности. Здесь тестирование оказывается неуместным. Зеркалом продуктивной, творческой деятельности обучаемых может стать материальный компонент их индивидуального образовательного пространства (ИОП) - представление системы полученных знаний с охватом близких предметных областей с помощью современных информационных технологий. В зависимости от вкусов и возможностей обучаемого это может быть компьютерная программа или база данных, электронный конспект или Web-сайт. Концептуально создание обучаемыми ИОП не является чем-то принципиально новым, поскольку учащиеся и студенты всегда готовили рефераты на основе материалов и дополнительных ресурсов, которые им приходилось отыскивать самостоятельно. Новые информационные технологии позволяют структурировать эти ресурсы, формируя из них базы данных, даже **базы знаний** - основы своего рода экспертных систем, но ориентированных на определенного пользователя - самого обучаемого. Именно поэтому эти разработки могут иметь самую разнообразную структуру, основываться на применении различных информационных технологий - наиболее доступных и удобных для обучаемого. Преподаватели должны поощрять такую продуктивную деятельность, не стараясь особенно ограничивать обучаемых в выборе средств материального воплощения ИОП. Обучаемый может формировать ИОП даже с самых поверхностных знаний (например, на уровне определений), постепенно улучшая качество владения основными понятиями, сформировать структурированные и глубокие представления о сути объекта изучения. Неформальный анализ

педагогом того, как по прошествии времени изменяется Web-страница или база данных, подбираемых обучаемым, поможет оценить не только качество обучения, но и самостоятельность, склонность к исследовательской работе, креативность и ответственность, профессиональные стремления и предпочтения обучаемого.

Материальный компонент индивидуального образовательного пространства в простейшем случае может быть реализован как пакет мультимедиа-информации с простыми и наглядными связями между различными ресурсами (текст, графика, видео, страницы Internet). Создание тематических каталогов взаимосвязанных ресурсов Internet может служить отдельным заданием, технически доступным всем обучаемым. Структура и характер связей созданного комплекса могут служить надежным индикатором и глубины освоения изучаемой дисциплины, и развития системного мышления, уровня поисковой, творческой деятельности. Такая работа может выполняться в рамках коллективных проектов, но от преподавателя потребуются дополнительные и значительные усилия по разграничению степени участия в проекте отдельных членов группы.

Также полезным направлением в создании подобных комплексов может стать подборка справочных и демонстрационных материалов по программным средствам, связанным с профессиональной или профильной подготовкой, анализ и сопоставление возможностей различных пакетов одного профиля (бухгалтерские программы, правовые системы, статистические пакеты и пр.).

Учет преподавателем при выставлении оценки такой творческой, продуктивной деятельности может положительно повлиять на стремление обучаемых более основательно изучить соответствующую дисциплину, проявить самостоятельность и инициативность, продемонстрировать уровень владения современными информационными технологиями. Без адекватной оценки качества обучения и развития обучаемых трудно говорить о сбалансированном учебном курсе. Использование различных информационных технологий расширяет возможности оценивания и создает условия для организации обратной связи в ходе процесса обучения. Однако, рассматривая применение ИТО как части общей стратегии, нельзя ограничивать их место лишь проведением автоматизированных тестирований.

## ГЛАВА 3. АКТУАЛЬНЫЕ ПРОБЛЕМЫ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ

### *§ 6. МЕСТО ТЕСТИРОВАНИЯ В УЧЕБНОМ ПРОЦЕССЕ*

Тесты являются эффективным средством проверки качества знаний, получаемых студентами, и оперативного контроля хода обучения. Информационные образовательные ресурсы (ИОР), содержащие тестовые материалы, можно разбить на две категории:

1. ориентированные на прохождение студентами тестов в письменной форме с дальнейшей проверкой вручную преподавателем (сканирование результатов тестирования с целью их дальнейшей автоматизированной проверки);
2. системы компьютерного тестирования с соответствующим наполнением тестовыми материалами.

Преимущества второй категории тестовых ИОР очевидны. Они позволяют освободить преподавателя от рутинной работы при проведении экзаменов и промежуточной оценке знаний в традиционном учебном процессе, а при обучении с использованием дистанционных технологий становятся основным средством контроля. Необходимо отметить, что если на некоторых уровнях образования существуют объективные препятствия для внедрения компьютерного тестирования, например, многие младшие школьники и даже выпускники школ не умеют работать на компьютере либо имеют крайне ограниченные навыки работы с ним (в большинстве российских школ современные компьютеры появились только в последние два года, причем во многих из них имеются один - два компьютера на всю школу), то студенты вузов обязаны иметь необходимый уровень компьютерной грамотности.

Одними из основных преимуществ использования тестирования являются возможность автоматизации обработки результатов, объективность контроля и быстрая проверка качества подготовки большого числа тестируемых по широкому кругу вопросов. Это позволяет определить разделы, которые представляют наибольшую сложность в изучении, и, возможно, корректировать процесс обучения в зависимости от результатов тестирования. Именно в автоматизированных системах тестирования в наибольшей степени проявляются многие преимущества тестового контроля знаний (оперативность, легкость сбора статистики и пр.).

Естественно, тестирование не может полностью подменить все



формы контроля знаний. Так, для проверки практических навыков решения задач возможность применения средств тестирования ограничена. Например, в тесте можно проконтролировать знание возможных методов решения задачи, предложить учащемуся расположить по порядку основные этапы решения. Для очень простых задач возможно решение их в уме и ввод результата в систему тестирования. Однако во многих случаях желательно дополнять тестовый контроль теоретических знаний решением практических заданий с помощью компьютерных задачников или традиционным образом.

Тестирование является важным элементом не только контроля знаний, но и обучения. При обучающем тестировании пользователю после прохождения теста предоставляются ссылки на те разделы учебного материала, на вопросы по которым он ответил неверно.

Наибольший эффект достигается при совместном использовании различных видов электронных учебных материалов (ЭУМ). Так, встраивание средств тестирования в состав электронных учебников (ЭУ) и электронных учебных пособий (ЭУП) позволяет оперативно оценивать ход изучения теоретического материала студентом, управлять траекторией изучения материала. Входное (до начала работы с ЭУ, ЭУП или их модулем) тестирование позволяет оценить начальный уровень знаний студента, выдать рекомендации по оптимальному для него индивидуальному порядку изучения дисциплины. Выходное тестирование позволяет либо выставить оценку по итогам изучения дисциплины, либо служит для допуска студента к сдаче экзамена (зачета) в традиционной форме. Аналогично, крайне желательно встраивать средства тестирования в состав лабораторных практикумов. Входное тестирование в этом случае проверяет понимание студентами основ теоретического материала по теме лабораторной работы и служит для допуска к ее выполнению. Итоговое тестирование по результатам выполнения лабораторных практикумов позволяет проверить знания, приобретенные студентом в результате выполнения лабораторной работы или блока работ. Таким же образом для проверки навыков решения практических задач целесообразно включать в состав ЭУ и ЭУП вызовы компьютерных задачников (КЗ). Совместное использование систем компьютерного тестирования и КЗ позволяет проверить как теоретические знания (тесты), так и практические навыки (решение типовых задач) при текущей, промежуточной или итоговой аттестации студента. Включение в состав ЭУМ проблемных и исследовательских заданий, интеллектуальных обучающих подсистем позволяют развить культуру мышления и мыслительные способности студента.

Рассматривая средства тестирования, в частности, как традиционные учебные издания, при их разработке необходимо учитывать Санитарные правила и нормативы "Гигиенические требования к изданиям книжным для взрослых. СанПиН 1.2.1253-03", утвержденные Главным государственным санитарным врачом Российской Федерации 30 марта 2003 г. и введенные в действие с 15 июня 2003 года. Раздел II этого документа «Гигиеническая классификация изданий» подразделяет все издания по гигиенической значимости на две категории. Гигиеническая значимость издания определяется реальной зрительной нагрузкой при чтении. Издания первой категории характеризуются большими объемами текста единовременного прочтения, что обуславливает непрерывный длительный процесс чтения и выраженную зрительную нагрузку. Издания второй категории характеризуются небольшими объемами текста единовременного прочтения, что обуславливает эпизодический процесс чтения и незначительную зрительную нагрузку. Далее в СанПиН 1.2.1253-03 к изданиям 1-й категории предъявляются существенно более жесткие гигиенические требования (прежде всего к шрифтовому оформлению) по сравнению с изданиями 2-й категории. Очевидно, что по гигиенической значимости тесты относятся ко второй категории.

Можно выделить следующие основные критерии качества тестов.

- Традиционные требования к тестовым материалам вне зависимости от формы их представления:
  - валидность (содержательная и функциональная);
  - надежность;
  - однозначность (одинаковость оценки качества выполнения теста разными экспертами, корректная постановка вопросов, не допускающая их различное толкование);
  - дифференцирующая способность.
- Наличие средств автоматизированного тестирования.
- Соответствие тестовых материалов международным стандартам (IMS QTI), их интероперабельность.
- Разнообразие типов тестовых заданий.
- Разнообразие способов построения траекторий тестирования (жесткая последовательность, случайная выборка заданий, адаптивное тестирование).
- Разнообразие режимов использования в тестировании (обучающее тестирование, тренировочное тестирование, контрольное тестирование).
- Дифференциация оценки тестовых заданий (возможность назначения



различных баллов разным вопросам, учет частично правильных ответов и пр.).

- Возможность использования в тестах графических иллюстраций, анимации, аудио и видео информации.
- Контроль времени (на тестовое задание или тест в целом), отображение текущего состояния тестирования (количество оставшегося времени и тестовых заданий и пр.).
- Наличие инструментов анализа результатов тестирования и оценки качества тестовых материалов.

## **§ 7. ПОСТРОЕНИЕ КОМПЬЮТЕРНЫХ СИСТЕМ КОНТРОЛЯ ЗНАНИЙ**

Рассматривая компьютерные системы контроля знаний с точки зрения их практической реализации, необходимо отметить, что все они содержат в себе следующие компоненты (иногда не выделенные явно в структуре системы):

1. Подсистему подготовки тестов.
2. Подсистему проведения тестирования.
3. Подсистему анализа результатов тестирования.

Различные разработчики по-разному подходят к выбору программ и технологий создания этих подсистем. Рассмотрим достоинства и недостатки различных подходов в реализации конкретных тестовых систем.

Одним из частых подходов является использование систем клиент-сервер, в которых в качестве клиента выступает браузер, а в качестве сервера – связка из Web – сервера и сервера приложений.

Этот подход при подготовке тестов имеет следующие **недостатки**:

- Составитель тестов во время ввода контрольно-тестовых материалов должен находиться в сети.
- Язык HTML ориентирован на представление данных, а не на их создание, что сильно затрудняет редактирование уже созданных данных. Это является проблемой многих систем основанных на Web – интерфейсе.
- Отсутствует возможность использования при подготовке тестовых заданий технологии OLE, что затрудняет вставку сложных объектов (формул и пр.).

**Достоинствами** такого подхода являются:

- Отсутствует зависимость от операционной системы.
- Отсутствует необходимость в установке и настройке клиентской

части.

Альтернативой является создание специализированного ПО для проведения тестирования и подготовки тестовых материалов. Данный подход при подготовке тестов имеет следующие **преимущества**:

- Автономность (не требуется наличие сетевого соединения).
- Удобство создания и редактирования тестовых материалов.

Его **недостатками** являются необходимость предварительной установки и настройки дополнительного ПО, а иногда и привязка к конкретной операционной системе.

Использование специализированного ПО при тестировании позволяет:

- минимизировать объем передаваемых данных;
- реализовать сложные интерактивных тестовые задания.

**Недостатки** специализированного ПО при проведении тестирования:

- возможны ограничения по форматам представления информации;
- необходимость установки специального клиентского приложения тестирования;
- проблема обновления версий клиентов при совершенствовании ПО.

## **§ 8. ПОДГОТОВКА МАТЕРИАЛОВ ДЛЯ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ**

При разработке тестовых материалов автору необходимо выполнять следующие требования:

1. Обеспечить как можно более полное покрытие материала курса контрольно-тестовыми материалами.
2. Разделить, при необходимости, тестовые задания на используемые для обучения и самоконтроля, с одной стороны, и на используемые для контроля по итогам обучения, с другой. На практике чаще всего для обеих целей используются одни и те же тестовые задания. Для этого подсистема контроля знаний должна предусматривать соответствующие средства построения сценариев тестирования (жесткое задание последовательности вопросов или перемешивание вопросов при случайной выборке).
3. Для обеспечения интероперабельности тестовых материалов необходимо обеспечить их соответствие стандартам тестирования, наиболее распространенным среди которых является IMS QTI. Данный стандарт определяет требования к тестовым материалам,

сценариям тестирования, представлению результатов тестирования. Можно сказать, что спецификация IMS QTI при определении типов тестовых заданий определяет как собственно тип, так и способ его представления пользователю, а также ограничения на типы используемых в вопросе данных. Так, например, классический закрытый вопрос с выбором одного правильного варианта в IMS QTI разбивается на:

- a. Standard True/False (Text) – выбор из двух вариантов ответа: Правда/Ложь (текст);
- b. Standard Multiple Choice (Text) - альтернативный выбор (текст);
- c. Standard Multiple Choice (Images) - альтернативный выбор (изображения);
- d. Multiple Choice with Image Hot Spot Rendering - альтернативный выбор (точки изображения);
- e. Multiple Choice with Slider Rendering - альтернативный выбор (слайдер).

Таким образом, при подготовке тестовых заданий необходимо уделять внимание предполагаемой форме представления вопроса на экране, однако, важнее определить собственно типы тестовых заданий. К наиболее распространенным относятся следующие:

- a. закрытый тип вопроса с одним правильным ответом (альтернативный выбор);
  - b. закрытый тип вопроса с несколькими правильными ответами (многоальтернативный выбор);
  - c. вопрос на упорядочивание;
  - d. свободный ввод краткого числового или текстового ответа;
  - e. вопрос типа «заполнить пробелы» (вопрос на подстановку);
  - f. вопрос на соответствие нескольких частей ответов (напрямую в IMS QTI не определен, частичный аналог - Matrix-based Multiple Response - выбор по одному элементу в каждой колонке матрицы).
4. Соответствие типов тестовых заданий типам, поддерживаемым системой контроля знаний, которую предполагается использовать для контроля знаний по курсу.
  5. Учет объективно различной сложности вопросов и задание дифференцированного уровня сложности вопросов (максимальный балл за вопрос) при поддержке соответствующих возможностей используемой системой контроля знаний.

При разработке тестов надо учитывать, что помимо сложности

вопроса с точки зрения проверяемого материала существует и сложность с точки зрения постановки вопроса, а также комплексности тестового задания, определяемой во многом его типом. Так, с точки зрения проверяемых знаний, один вопрос на упорядочивание или на соответствие может послужить эквивалентом достаточно большого числа альтернативных вопросов.

Для примера рассмотрим вопрос на соответствие следующего вида.

<b>Установите соответствие между писателя-</b>		
<b>Писатель</b>	<b>Год рожде-</b>	<b>Год смерти</b>
А.С.Пушкин	1814	1852
М.Ю.Лермонтов	1809	1873
Л.Н.Толстой	1799	1837
Н.В.Гоголь	1828	1910
	1820	1841

При ответе обучаемый должен правильно указать для каждого писателя год рождения и год смерти. В вопросы занесены заведомо неправильные ответы, не относящиеся ни к кому из представленных писателей (год рождения 1820 и год смерти 1873). С точки зрения проверяемых знаний один такой вопрос эквивалентен восьми классическим закрытым вопросам типа

<b>Укажите год рождения А.С.Пушкина</b>	
<input type="checkbox"/>	<b>1814</b>
<input type="checkbox"/>	<b>1809</b>
<input type="checkbox"/>	<b>1799</b>
<input type="checkbox"/>	<b>1828</b>
<input type="checkbox"/>	<b>1820</b>

При подобной формулировке аналогичные вопросы должны быть заданы и по годам рождения всех остальных писателей, а также годам их смерти.

При использовании комплексных заданий встает вопрос об учете частично правильных ответов на вопросы. В зависимости от целей, которые ставятся при тестировании, можно засчитывать любой неполный ответ как неправильный либо учитывать степень правильности частичного ответа. Так, например, следующий ответ на многоальтернативный вопрос будет неполным (не указан Б.Л.Пастернак).

Если правильный ответ на вопрос оценивать в 1 балл, то за такой ответ может быть засчитано 0 баллов (если считать ответ при неуказании хотя бы одного правильного ответа полностью неправильным), либо, например, 0,75 балла (с учетом того, что правильно указано 3 ответа из 4-х). Вообще, учет частично правильных ответов является нетривиальной задачей, не имеющей однозначного решения. Так, например, даже в многоальтернативном вопросе возникает проблема с учетом указания неправильных вариантов ответов. Еще сложнее учет частичной правильности ответа в вопросах на упорядочивание и на соответствие.

<b>Выберите писателей, родив-</b>	<b>А.С.Пушкин</b> <b>М.Ю.Лермонтов</b> <b>Л.Н.Толстой</b> <b>Н.В.Гоголь</b> <b>Б.Л.Пастернак</b>
<input type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	

Таким образом, использование в тестовых материалах сложных типов вопросов позволяет значительно сократить количество тестовых заданий, необходимых для обеспечения достаточной полноты проверки изученного материала. С другой стороны, в этом случае более трудоемка подготовка тестовых заданий, а также возникают проблемы с учетом частично правильных ответов.

### **§9. ОБРАБОТКА РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ**

Подсистема статистики служит обычно для предоставления преподавателю или администратору результатов тестирования и может выводить краткие или детальные (с указанием всех заданных вопросов и полученных на них ответов, а также баллов за вопрос) протоколы тестирования по каждому из испытуемых с указанием итогового балла, оценки, времени тестирования. Подсистема статистики может предоставлять следующие основные возможности:

- Создание кратких отчетов. Они содержат обобщенную информацию о тестируемом и его результатах.
- Создание протоколов тестирования. В этом случае помимо информации о тестируемом, показывается, как он отвечал на тот или иной вопрос, с указанием набранного балла за вопрос, пропущенных и отложенных вопросов.
- Создание сводной таблицы, в которой для каждого испытуемого указывается состояние теста, набранный балл, оценка, которую поставила система тестирования, время тестирования и количество

пропущенных вопросов.

- Выпуск таблицы с результатами тестирования по группе. В ней указывается количество человек, которые прошли тест с оценкой “5”, “4”, “3”, “2” и их процентное соотношение. Здесь же подсчитывается средний набранный балл и средняя оценка по группе с расчетом соответствующих дисперсий.
- Выпуск таблицы с указанием степени выполнения теста каждым испытуемым. Здесь же подсчитывается средняя величина степени выполнения теста с расчетом дисперсии.
- Сравнение результатов тестирования различных групп. Группы могут формироваться:
  - по фамилиям испытуемых;
  - по названиям их учебных групп;
  - по факультету, на котором они обучаются;
  - по названию теста.
- Создание графиков по результатам тестирования как одной, так и нескольких групп.
- Сохранение результатов тестирования в базе данных.
- Загрузка результатов тестирования из базы данных, которая может осуществляться:
  - по дате сохранения результатов;
  - по дате тестирования;
  - по названию теста, по которому проводилось тестирование;
  - по фамилии тестируемого;
  - по названию учебной группы;
  - по названию факультета;
  - по набранному баллу;
  - по полученной оценке;
  - с использованием комбинированных запросов.

Для того, чтобы тестирование обеспечивало эффективный контроль знаний учащихся, необходимы качественные тестовые материалы. Выше уже были рассмотрены проблемы учета частично правильных ответов и обеспечения различными типами тестовых заданий разного покрытия предметной области. Однако автору тестов часто трудно объективно оценить качество своих разработок и обеспечить валидное и высоконадежное тестирование. Для автоматизации расчета качества тестовых заданий необходимо оценивать их статистические характеристики. Поэтому дальнейшие работы по развитию подсистемы



статистики ведутся в направлении автоматизации проверки качества тестовых заданий и качества теста в целом. Алгоритмы разрабатываются на основе последовательности статистической обработки результатов и формул.

### ***§10. ПРИМЕНЕНИЕ СРЕДСТВ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ ДЛЯ ПРОВЕДЕНИЯ МЕЖДИСЦИПЛИНАРНОГО ЭКЗАМЕНА***

В качестве примера использования в учебном процессе сочетания традиционных и тестовых методов контроля знаний рассмотрим организацию Государственного междисциплинарного экзамена по специальности «Системы автоматизированного проектирования» в Пензенском государственном университете. Процедура сдачи междисциплинарного экзамена состоит из двух частей.

1. Компьютерное тестирование по теоретическим вопросам.
2. Решение практических заданий.

Таким образом, проверка практических умений и навыков студентов выполняется отдельно. В ходе ее студентам необходимо решить практические задачи в соответствии с выбранным билетом. Для повышения объективности оценки письменные работы в момент получения задания шифруются и проверяются преподавателями под шифром. Однако из-за временных ограничений таким способом реально проверить знания только по одной — двум дисциплинам.

Проверка теоретических знаний по всем выносимым на экзамен дисциплинам осуществляется с помощью СКК. Каждому студенту необходимо пройти тест из 48 вопросов по 13 дисциплинам. Преподаватели готовят банк вопросов по каждой из дисциплин междисциплинарного экзамена (от 30 до 70 вопросов на дисциплину). Далее создается сценарий тестирования для междисциплинарного экзамена. Используемая система компьютерного тестирования позволяет реализовывать различные сценарии, но для проведения экзамена используется относительно простой: для теста по каждой дисциплине указывается количество вопросов, которые будут случайным образом выбраны и войдут в сценарий экзаменационного тестирования. «Квота» каждой дисциплины составляет от двух до четырех вопросов. Применяемая система тестирования поддерживает шесть типов тестовых заданий, что дает в распоряжение преподавателя достаточно гибкий инструмент для проверки знаний по своему предмету.

Для подготовки к экзамену до студентов доводится общий список

вопросов по всем предметам (около 600). Разумеется, при этом им не сообщаются ни ответы на вопросы, ни типы вопросов. То есть, студент должен быть готов ответить на вопрос по определенной тематике, но не знает, будет ли ему предложен выбор из готовых вариантов ответа, или ему придется вводить ответ вручную. На тестирование отводится 30 минут. Так как каждый студент должен в сжатые сроки пройти свой индивидуальный сценарий тестирования, возможности «взаимопомощи» в ходе экзамена минимизированы. Правильный ответ на вопрос оценивается в один балл. Частично правильные ответы также учитываются (например, из трех правильных ответов выбраны два – засчитывается 0,67 балла).

Полученные в результате теста баллы (максимум – 48 баллов) суммируются с баллами за практическое задание (максимум – 20 баллов). На основании суммарного балла выставляется итоговая оценка. Многолетний опыт использования СКК знаний для проведения междисциплинарного экзамена показал высокую достоверность получаемых результатов: в большинстве случаев «компьютерная» оценка соответствовала тому, как студент учился в течение пяти лет.

## ГЛАВА 4. МЕТОДИКА ПРОВЕДЕНИЯ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ

Системы компьютерного контроля знаний — это системы тестирования, позволяющие проводить анализ знаний учащихся при помощи современных информационных технологий. Одно из преимуществ автоматизированных систем контроля знаний в том, что они могут использовать сложные методики представления заданий учащимся, называемые стратегиями тестирования.

Однако для того, чтобы использование систем компьютерного тестирования было оправдано, необходимо выполнение нескольких условий. Во-первых, компьютерное тестирование должно предоставлять результаты не хуже, чем его неавтоматизированный аналог. Во-вторых, поскольку любой вид тестирования есть не что иное, как попытка структурировать процесс проведения экзамена или опроса, то результат компьютерного тестирования должен качественно стремиться к результату экзамена. Тогда основным преимуществом автоматизированного тестирования будет заметное сокращение времени, отведенного на проведение экзамена, увеличение пропускной способности самого процесса экзамена, а результат тестирования можно будет считать объективным. В-третьих, для того, чтобы получить дополнительные преимущества перед экзаменом, тестирование должно максимально абстрагироваться от человеческого фактора при проверке работ и при вычислении оценки. К сожалению, этот момент практически не достижим при использовании неавтоматизированного тестирования, поскольку проверяющий почти всегда знает, работу какого ученика он проверяет, а, следовательно, существует вероятность необъективной оценки. Для выхода из сложившейся ситуации необходимо оставлять в тесте только простейшие типы контрольных вопросов, а это, в свою очередь, сильно упрощает тест и снижает его качество.

Таким образом, можно определить основные характеристики компьютерной системы для проведения тестирования и методики, лежащей в ее основе:

1. полученные результаты должны быть не хуже, чем результаты обычного тестирования;
2. полученные результаты должны соответствовать результатам устного экзамена;
3. содержание теста должно быть нетривиальным и разнообразным;
4. временная нагрузка на прохождение теста должна быть минимизирована;

5. результаты тестирования обязаны быть объективными.

Для достижения указанных целей была создана методика и разработана компьютерная система проведения тестирований, которая и описывается далее.

Условно, созданную методику можно разбить на две части:

1. Технология проведения тестирования;
2. Технология вычисления результата;

Сразу необходимо отметить, что предлагаемая методика не дает рекомендаций по составлению тестовых последовательностей, а лишь регулирует их набор, указывая типы контрольных вопросов, которые могут использоваться.

При разработке технологии проведения тестирования была проведена попытка привести систему к адаптивному типу, при этом оставив ее максимально простой в использовании. Для этого поведение системы было построено на событийно-ориентированном принципе. Это означает, что любой процесс тестирования можно описать с помощью понятий: «задание» и «событие», при этом любую последовательность заданий и событий можно представить в виде ориентированного графа. Эти последовательности получили название «схемы тестирования» и являются основой предлагаемой методики. С их помощью можно описать практически любой алгоритм поведения системы, а простой визуальный инструмент для создания таких схем делает эту технологию удобной для широкого круга пользователей.

В методике определен набор контрольных вопросов, которыми оперирует система. Конечно же, он ограничен. В число этих контрольных вопросов вошли:

1. Вопросы типа выбора (*Выбор одного из многих и Выбор многих из многих*).
2. Вопросы типа заполнения.
3. Вопросы типа порядка (*Установление порядка и Установление связей*).
4. Вопросы со свободно-конструируемыми ответами.

Проведенные эксперименты показывают, что с помощью перечисленных типов контрольных вопросов возможно создание тестовых последовательностей, достаточных для суждения об уровне знаний учащихся.

Все типы контрольных вопросов, кроме последнего, обладают свойством автоматической оценки. Это означает, что для определения степени правильности ответа человеческое участие не требуется, а

объективность оценки — абсолютна и определяется только заложенными в работу системы алгоритмами. Эти алгоритмы были созданы в рамках работы над системой и характеризуются следующими свойствами:

1. Оценка за ответ на вопрос зависит от параметров вопроса, таких как его сложность, степень правильности каждого варианта ответа, временных рамок и т.д.
2. Оценка является не дихотомической, а распределенной на интервале от минимальной до максимальной. Причем это распределение не равномерное, а слегка смещено в сторону максимальной оценки, поскольку, исходя из данных экспериментов, преподаватели обычно не ставят оценки, равных минимальной за совсем неправильный ответ.
3. Итоговая оценка вычисляется, опираясь на оценки каждого вопроса.

Ответы на вопросы со свободно-конструируемым ответом невозможно проверить без участия человека. Но в этом случае проверяющий не должен видеть имя ученика, ответ которого он проверяет, чтобы избежать необъективной оценки. Поэтому, в рамках данной методики, вопросы со свободно-конструируемым ответом предлагаются преподавателю анонимным набором. Необходимо заметить, что эта особенность отличает автоматизированную систему контроля знаний от «ручной», где субъективный фактор заведомо велик.

Для проверки описанной методики был проведен ряд экспериментов, а также планируется проведение других. Все эти эксперименты призваны проверить характеристики автоматизированной системы контроля знаний, базирующейся на созданной методике. Было выделено три подвида экспериментов:

1. Ученикам предлагается пройти одинаковые по смыслу, но разные по содержанию тесты в «ручном» и автоматизированном виде. При этом ясно, что такие тесты должны быть максимально просты и не иметь в своем составе вопросов со свободно-конструируемыми ответами. Результаты тестов сравниваются, и анализируется их схожесть. Условия, в которых проводятся тестирования, должны быть идентичными. По результатам такого эксперимента делается вывод о соответствии результатов простого и автоматизированного тестирования. Такой эксперимент необходимо провести в различных возрастных группах и, желательно, по различным областям знаний.
2. Второй подвид эксперимента проверяет соответствие результатов автоматизированного тестирования и устного экзамена. Для этого нескольким группам учащихся предлагается сначала пройти тест на заданную тему, а потом по этой же теме проводится устный зачет

или проверочная работа. После этого анализируется схожесть полученных результатов. Рекомендуется провести эксперименты в разных возрастных группах и по различным темам курса.

3. Третий подвид эксперимента определяет широту предметных областей, в которых может применяться данная методика и разработанная на ее основе система. Необходимо провести любые комбинации первого и второго подвидов экспериментов по разным предметам как школьной, так и вузовской программы обучения. Рекомендуется провести тестирования в различных учебных заведениях и по различным дисциплинам, как математическим и техническим, так и гуманитарным.

Результаты описанных экспериментов должны подтвердить надежность разработанной методики проведения тестирований и показать возможность применения ее в школе и в вузе.

Следуя описанной методике, разработана система проведения компьютерного контроля знаний *TestSys*, позволяющая проводить как локальные, так и удаленные тестирования, которая успешно внедрена в некоторых ВУЗах и школах города Москвы.



## ЗАКЛЮЧЕНИЕ

Следует отметить еще одну важную проблему компьютерного контроля знаний в России. Главным аргументом совершенствования контроля знаний и перехода на компьютерный контроль, в том числе и на ЕГЭ, можно назвать повышение объективности результатов, борьбу с коррупцией среди преподавателей и т.д. На самом деле, как следует из приведенных выше аргументов, реальная объективность такого контроля ниже естественно-человеческого контроля группы преподавателей (комиссии). И не следует преувеличивать полицейские возможности компьютера, да и с коррупцией пусть борются профессионалы.

Хочется напомнить, что объективный контроль знаний в первую очередь нужен самому учащемуся: зная, чего он, ученик или студент не знает, он может восполнить имеющиеся пробелы в своем образовании, ибо без этого невозможно ему учиться дальше. По существу, компьютерный контроль знаний у нас в стране берется на службу бюрократизированной частью системы образования: как хорошо подстричь всех под одну гребенку, а «если, что не так – не наше дело», компьютер виноват.

И речь здесь не идет о запрете компьютерного контроля знаний студентов. Такой контроль хорош в качестве текущего формального контроля усвоения знаний с обязательной интеллектуальной оценкой преподавателем адекватности результатов контроля. Что же касается компьютерных систем контроля, то над ними надо работать, только в этом случае можно ожидать в этой сфере значимого прогресса.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аванесов В.С. Научные проблемы тестового контроля знаний. – М., 1994.
2. Башмаков А.И., Башмаков И.А. Разработка компьютерных учебников и обучающих систем. – М.: Информационно-издательский дом «Филинь», 2003.
3. Захарова И.Г. Информационные технологии в образовании: Учебное пособие для студ. высш. пед. учеб. заведений. – М.: Издательский центр «Academa», 2003.
4. Полат Е.С., Бухаркина М.Ю., Моисеева М.В., Петров А.Е. Новые педагогические и информационные технологии в системе образования: Учебное пособие. – М., 2001.
5. Роберт И.В. Современные информационные технологии в образовании. – М., 1994.
6. Челышкова М.Б. Теория и практика конструирования педагогических тестов. – М.: Информационно-издательский дом «Филинь», 2003.



