

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	2
1. ИСТОРИЯ И СОВРЕМЕННОЕ СОСТОЯНИЕ КОМПЬЮТЕРНЫХ СИСТЕМ ТЕСТИРОВАНИЯ	2
1.1. История тестирования и основные понятия тестологии.....	2
1.2. Анализ компьютерных систем контроля знаний.....	8
1.3. Средства тестирования и тестирующие программы	9
1.4. Тестирование в ПАК «Мобитест».....	12
2. ПОНЯТИЕ МОНИТОРИНГОВОЙ РАБОТЫ В ПАК «МОБИТЕСТ».....	13
2.1. Ролевая модель пользователей программной части ПАК «Мобитест».....	13
2.2. Модуль администрирования.....	14
2.3. Понятие мониторинговой работы, ее формирование и параметры	14
2.4. История мониторинговых работ	15
3. ПРОВЕДЕНИЕ МОНИТОРИНГОВОЙ РАБОТЫ	15
3.1. Организационные условия проведения мониторинговой работы и подготовка к ней 15	
3.2. Проведение мониторинговой работы	18
3.3. Проведение работы в компьютерном варианте	18
3.4. Проведение работы в бескомпьютерном (бумажном) варианте	19
3.5. Диагностические и статистические отчеты.....	20
4. ЧТО НУЖНО ЗНАТЬ О СВОЙСТВАХ ЗАДАНИЙ И ПЕДАГОГИЧЕСКОЙ СТАТИСТИКЕ?	21
4.1. Классическая тестовая теория	21
4.2. Трудность задания	23
4.3. Дифференцирующая способность заданий	23
4.4. Рейтинговый балл	24
4.5. Валидность заданий.....	24
4.6. Надежность теста.....	25
ЗАКЛЮЧЕНИЕ.....	26
ЛИТЕРАТУРА	27

ВВЕДЕНИЕ

Программная часть программно-аппаратного комплекса (ПАК) «Мобитест» (далее Система) создана с целью обеспечения мониторинга знаний учащихся. Система работает на стационарных и мобильных компьютерах, позволяет получать данные о результатах выполнения заданий, анализировать проведенные работы и оптимизировать составление блоков тестовых заданий при дальнейшем использовании, поддерживая оптимальные педагогические характеристики.

Программная часть ПАК «Мобитест» позволяет учителю создавать и редактировать задания, формировать по своему усмотрению или использовать уже сформированные варианты КИМ, получать полную информацию о результатах тестирования, в том числе анализ результатов на уровне формирования отчетов первичной статистики.

1. ИСТОРИЯ И СОВРЕМЕННОЕ СОСТОЯНИЕ КОМПЬЮТЕРНЫХ СИСТЕМ ТЕСТИРОВАНИЯ

1.1. История тестирования и основные понятия тестологии

Большинство методов оценивания знаний учащихся, применяемых в школьной практике (опрос у доски, развернутый ответ в самостоятельной работе, сочинение и т. п.), в большей или меньшей степени субъективно. Причин у этого несколько: это и низкая дифференцированность шкалы отметок (от «двух» до «пяти»), и слабая формализованность процесса проверки ответа. В результате процесс оценивания представляет собой для учащихся «тёмное» действие, не свободное от субъективизма, что является причиной значительной части конфликтов и стрессов в образовательной среде. Назовем несколько таких типичных ошибок, весьма частых в работе педагогов при оценивании ответов учащихся:

- ошибки **великодушия**, или **снисходительности**: выставление завышенных отметок;
- ошибки **ореола**, или **шлейфа**: завышение отметок тем, к кому относятся положительно, и занижать – к кому отрицательно;
- ошибки **центральной тенденции**: стремление избежать крайних отметок;
- ошибки **близости**: сложно после очень плохой отметки сразу поставить очень хорошую;
- ошибки **логики**: соотнесение отметок с разным психологическим свойствам и характеристикам, которые *кажутся* логически связанными (например, разные баллы для активного и усидчивого ученика).

Анализ приведенных ошибок наглядно показывает, что личностные характеристики как учащихся, так и учителя могут существенно влиять на ход и результаты оценивания.

Массовая школа, созданная в конце XIX в., достаточно быстро пришла к осознанию необходимости объективных средств контроля знаний учащихся. Одним из таких инструментов является тестирование.

Тестирование, как самостоятельное направление, сформировалось, благодаря деятельности Френсиса Гальтона, двоюродного брата Чарльза Дарвина. Гальтоном разработано множество простых тестов, многие из которых знакомы нам в модифицированном виде («линейка Гальтона» для зрительного различения длины, «свисток Гальтона» для определения верхнего частотного порога слуховых ощущений). Труды Гальтона позволили определить три основных принципа, лежащих в основе тестирования:

- применение серии одинаковых испытаний к большому количеству испытуемых;
- статистическая обработка результатов;
- выделение эталонов оценки.

Ф.Гальтон доказал целесообразность применения математической статистики в экспериментальной практике. Он отобрал и адаптировал большое количество математических приемов, позволяющих количественно оценить результаты экспериментальной работы. Гальтоном была высказана идея сравнения результатов тестирования с каким-либо внешним, независимым от эксперимента критерием, что позволяло узнать, какой тест является более информативным. Им были сформулированы принципы определения валидности и надежности тестов, которые сегодня широко используются в педагогической статистике.

Английские тестологи Фред и Элеонора Шоннел [6] определяют критерии, которым должен соответствовать тест: пригодность (валидность), надежность, объективность, экономичность, простота обработки результатов. Авторы считают, что результаты теста должны использоваться:

- для распределения учащихся по группам;
- для диагностирования трудностей, которые встретились учащимся;
- для изменения в содержании программы на основе сведений, полученных из тестов;

- для изменения в методах преподавания, причем особое внимание следует уделить «remedial method» – методу «исправления», который нужно использовать с отстающими учащимися, нуждающимися в помощи учителя.

Книгу Фреда и Элеоноры Шоннел можно назвать классическим пособием по тестированию: в ней дана классификация тестов, указаны требования, предъявляемые к тестам, определены своего рода нравственные нормы тестирования.

И сегодня сохраняют актуальность рекомендации составителям тестов, предложенные Р. Торндайком и Е. Хаген [7]:

1. Задания должны легко читаться, т.е. по трудности изложения и по словарному составу задания должны быть максимально доступны.
2. Если это тест успеваемости, то задания не должны воспроизводить формулировки, данные в учебнике.
3. Знания одних заданий не должны зависеть от знаний других заданий.
4. В тестах с выборочными ответами правильные ответы по заданиям должны распределяться в случайном порядке.
5. Формулировки заданий не должны содержать двусмысленностей, а тем более ловушек.
6. В тестах альтернативной формы следует избегать слов-подсказок. Задания, начинающиеся словами «все», «всегда», «никогда», часто бывают неправильными, а начинающиеся словами «обычно», «иногда», как правило, бывают верными.
7. В тестах, где требуется заполнение пропусков в заданиях, следует избегать большого количества пропусков в каждом отдельном задании. Пропущенные слова должны быть ключевыми для правильного понимания задания.
8. Для выборочного теста сама суть должна содержаться в головном вопросе. Его не следует перегружать второстепенными деталями. Среди предполагаемых ответов не должно быть явно нелепых, равно как и ловушек.

Оценка качества образования приобретает в последнее время все большую значимость. В связи с этим актуальными стали международные сравнительные исследования в области школьной успеваемости. Эти исследования дают возможность странам всесторонне оценить эффективность своих образовательных систем. Примерами таких исследований являются Международная программа по оценке образовательных достижений учащихся (Programme for International Student Assessment, PISA), Международное мониторинговое исследование качества школьного математического и естественнонаучного образования TIMSS (Trends in Mathematics and Science Study) и др.

Педагогическое тестирование – это форма измерения знаний учащихся, основанная на применении педагогических тестов. Оно включает в себя подготовку качественных тестов, собственно проведение тестирования и последующую обработку результатов, которая даёт оценку обученности тестируемых. **Педагогический тест** – это инструмент оценивания обученности учащихся, состоящий из системы тестовых заданий, стандартизированной процедуры проведения, обработки и анализа результатов. А. Н. Майоров в книге «Теория и практика создания тестов для системы образования (как выбирать, создавать и использовать тесты для целей в образовании)» [5] отмечает, что «в узком смысле тестирование в педагогике означает использование стандартизированных педагогических тестов для измерения и оценки результатов обучения. В широком же смысле тестирование – это любое испытание с целью измерения достижения обучаемого».

В. С. Аванесов считает, что понятие теста сводится к педагогической деятельности по созданию заданий, которые могут применяться как для контроля, так и для обучения. Аванесов [1] предлагает следующую классификацию тестов по различным признакам:

- *по целям* – информационные, диагностические, обучающие, мотивационные, аттестационные;
- *по процедуре создания* – стандартизированные, нестандартные;
- *по технологии проведения* – бумажные, в том числе бумажные с использованием оптического распознавания, натурные, с использованием специальной аппаратуры, компьютерные;
- *по форме заданий* – закрытого типа, открытого типа;
- *по наличию обратной связи* – традиционные и адаптивные.

В педагогическом тестировании выделяются три основные взаимосвязанные функции:

- *диагностическая функция* (основная и самая очевидная) – выявление уровня знаний, умений, навыков обучающегося; по объективности, широте и скорости диагностирования тестирование превосходит все остальные формы педагогического контроля;
- *обучающая функция* – мотивирование обучающегося к активизации деятельности по усвоению учебного материала;
- *воспитательная функция* – периодичность и неизбежность контроля, позволяющая дисциплинировать, организовывать и направлять деятельность обучающихся, способствовать выявлению и устранению пробелов в знаниях, формированию и развитию способностей.

В.П.Беспалько [2] определил следующие требования к созданию педагогически корректного теста:

1. Задание теста должно быть содержательно валидным, т. е. построено на содержании, которое учащемуся должно быть известно из предшествующего обучения.
2. Задание теста должно быть функционально валидным, т. е. тест должен проверять то, для чего его используют.
3. Задание теста должно быть простым, т. е. содержать одну задачу данного уровня.

Применение тестов имеет как свои достоинства, так и недостатки ([3], табл. 1).

Таблица 1. Преимущества и недостатки системы тестирования

<i>Преимущества</i>	<i>Недостатки</i>
<ul style="list-style-type: none"> – Более качественный и объективный способ оценивания (объективность достигается путем стандартизации процедуры проведения, проверки показателей качества заданий и тестов целиком). – Более справедливый метод, поскольку ставит всех учащихся в равные условия, как в процессе контроля, так и в процессе оценки, практически исключая субъективизм преподавателя. – Более объёмный инструмент, поскольку тестирование включает в себя задания по всем темам курса, в то время как на устный экзамен обычно выносятся 2–4 темы, а на письменный – 3–5 (это позволяет выявить знания учащегося по всему курсу, исключив элемент случайности при вытаскивании билета). – Более точный инструмент, поскольку шкала оценивания теста из 20 вопросов состоит из 20 делений (обычная шкала оценки знаний – только из четырёх). – Экономическая эффективность: основные затраты при тестировании приходятся на разработку качественного инструментария, т. е. имеют разовый характер; затраты на проведение теста значительно ниже, чем при письменном или устном контроле (проведение тестирования и контроль результатов в группе из 30 человек занимает до часа, устный или письменный экзамен – не менее четырёх часов); затраты на проверку также на порядок ниже. – Более мягкий инструмент, поскольку 	<ul style="list-style-type: none"> – Разработка качественного тестового инструментария – длительный, трудоемкий и дорогостоящий процесс, поэтому стандартные наборы тестов для большинства дисциплин ещё не разработаны, а разработанные обычно имеют недостаточное качество. – Данные, получаемые преподавателем в результате тестирования, хотя и включают в себя информацию о пробелах в знаниях по конкретным разделам, но не позволяют судить о причинах этих пробелов. – Тест не позволяет проверять и оценивать высокие, продуктивные уровни знаний, связанные с творчеством, а также методологические знания. – Жесткий лимит времени: при тестировании, в отличие от устного или письменного экзамена, учащийся не имеет достаточно времени для сколько-нибудь глубокого анализа темы. – Обеспечение объективности и справедливости теста требует принятия специальных мер по обеспечению конфиденциальности тестовых заданий (при повторном применении теста желательно внесение в задания изменений). – В тестировании всегда присутствует элемент случайности (например, случайная ошибка или угадывание); это искажает результаты теста и приводит к необходимости учета вероятностной составляющей при их анализе.

<i>Преимущества</i>	<i>Недостатки</i>
ставит всех учащихся в равные условия и использует единую процедуру и единые критерии оценки; это приводит к снижению предэкзаменационных нервных напряжений	

В систематизированном виде пути коррекции недостатков закрытых тестовых заданий (заданий, в которых правильным является выбор элементов, их сочетаний или последовательностей), приведены в табл. 2 (по [4]).

Таблица 2. Недостатки заданий закрытой формы и меры по их устранению

<i>Недостаток</i>	<i>Пути коррекции</i>
Правильные ответы на разные вопросы оцениваются одинаково	<p>А. Ранжирование вопросов по сложности и введение весовых коэффициентов для каждого вопроса. Для тестов применяется редко – в них вопросы подбираются (по крайней мере, в рамках одного блока) равносложными</p> <p>Б. Присвоение весовых коэффициентов (по сложности) группам вопросов</p>
Два неправильных ответа на один и тот же вопрос оцениваются одинаково	<p>А. Ранжирование вариантов ответов по степени неправильности: «в корне неправильно»; «скорее неправильно, чем правильно»; «не совсем правильно» и т. д. с разными баллами за каждый вариант</p> <p>Б. Использование отрицательной (минусовой) шкалы, снятие баллов или штрафные баллы</p> <p>В. Абсурдный ответ, который оценивается в минусовой шкале</p>
Неполные ответы оцениваются одинаково	<p>Использование дифференцированных оценок за полноту ответа, системы призов и штрафов (разные варианты для вопроса с выбором нескольких правильных ответов)</p> <p>В открытых тестах – добавление призовых баллов, например, за тщательность: ответ на каждый вопрос оценивается каким-то количеством баллов (не выше максимального); при проверке подсчитывается сумма «сырых» баллов и % от максимального количества баллов за те вопросы, на которые даны ответы; рейтинговый балл складывается из суммы «сырых» баллов и призовых баллов за качество, численно равных 0,1 от процента выполнения заданий)</p>
Угадывание ответов	<p>А. Увеличение числа вариантов ответов (оптимально 5–7 вариантов). Впрочем, при большом количестве вариантов (больше трех) подобрать равно правдоподобные формулировки сложно, и один-два дистрактора в любом случае будут иметь устойчиво низкую выбираемость. Распространенная ошибка – правильным является самый длинный или использующий наиболее на наукообразную лексику ответ.</p> <p>Б. Введение для каждого вопроса теста вопроса-дублера, который в завуалированной форме повторяет содержание основного вопроса (оцениваются одним баллом, только если оба ответа правильные).</p> <p>В. Ввод дополнительного «коэффициента уверенности»: на каждый вопрос теста даются собственно варианты ответов и варианты уверенности в ответе; итоговый результат – сумма произведений баллов за ответ на коэффициент уверенности. Имеет смысл в этом</p>

Недостаток	Пути коррекции
	случае использовать отрицательную шкалу: правильный ответ оценивать как «+1», а неправильный – как «-1»
Списывание	Компьютерное тестирование с генерированием последовательности ответов (или создание батарей тестов в текстовом процессоре с распечаткой большого числа вариантов)

Суммируя вышесказанное, тестология представляет собой важный раздел на стыке психологии и педагогики, которые на основе использования корпуса требований к составлению тестов, проведению тестирования и анализу результатов, позволяет обеспечить высокий уровень точности при анализе уровня знаний учащихся.

Тестовый контроль – эффективное средство определения базового уровня подготовленности, проверки остаточных знаний, благодаря простоте и объективности позитивно воспринимаемый учащимися (по сравнению с традиционными формами опроса). Он дает учителю возможность быстрой проверки большого количества работ и экспресс-анализа ошибок, коррекции проблемных вопросов в знаниях учащихся. Тестовый контроль основан на применении стандартизированных методов, что гарантирует надежность результатов массовых исследований образования (например, государственных экзаменов). Для массовых форм тестового контроля разработаны тренажеры, в которых размещены соответствующие блоки тестовых заданий (готовый образовательный контент). К недостаткам тестовой формы контроля знаний можно отнести низкую корректность при проверке свободного владения материалом по дисциплине. Массовые процедуры тестирования в их традиционной форме подвергаются критике, так как не гарантируют высокую точность измерений (они предназначены для оценки учащихся с широким диапазоном уровня подготовленности, что снижает возможности внутренней дифференциации).

1.2. Анализ компьютерных систем контроля знаний

История информатизации образования неразрывно связана с созданием компьютерных классов различной конфигурации. Классический компьютерный класс предполагает наличие компьютера, выполняющего роль локального сервера (это может быть компьютер учителя или специально выделенная машина), и рабочих мест учеников. Количество последних определяется наполняемостью классов в учебном заведении, требованиями санитарных правил и норм к учебным помещениям и финансовыми

возможностями учреждений образования. Стандартный компьютерный класс включает 12–16 рабочих станций обучаемых.

Существенным этапом развития компьютерного наполнения образовательных учреждений стало появление мультимедийных проекторов, а вскоре – и интерактивных досок, позволявших учителю управлять содержимым компьютера, непосредственно стоя у доски. Интерактивные доски существенно повысили оперативность отображения информации. А так как интерактивные комплексы ряда производителей включали в себя пульты для оперативного контроля знаний («голосовалки»), стало возможным провести быстрый опрос или тестирование и сразу же представить его результаты (например, распределение выборов по вариантам ответа) графически на интерактивной доске.

Новым этапом стало развитие мобильных устройств, в первую очередь – планшетных компьютеров, в меньшей степени, смартфонов, что позволило создать компьютерные классы, в которых в роли «тонкого» клиента учащихся выступал планшетный компьютер. В этом случае возможна полнофункциональная реализация идеологии «один ученик : один компьютер» с полнофункциональной поддержкой индивидуализации обучения средствами как человек-компьютерных (ручное или полуавтоматическое назначение учителем заданий отдельным учащимся), так и полностью автоматизированных компьютерных сервисов (адаптивное тестирование).

Одной из тенденций развития ИКТ в образовании последних лет стало активное использование в ходе занятия личных устройств обучающихся разных типов (смартфонов, ноутбуков, планшетов и т. д.); этот принцип получил название BYOD (*англ.* Bring your own devices – «Принесите свои собственные устройства»). Личные устройства учащихся могут быть использованы, например, в конфигурации мобильного тестирующего комплекса. В этом случае онлайн-сервис организации тестирования доступен учащимся с использованием беспроводных технологий или через веб-интерфейс, или через специальный клиент/плеер контента, устанавливаемый на устройство.

Именно с учетом этих современных тенденций организации информационной среды учащихся в школе и должна строиться идеология современного сервиса контроля и мониторинга знаний учащихся.

1.3. Средства тестирования и тестирующие программы

Исторически первыми компьютерными сервисами для организации и проведения тестирования, включая обработку результатов, было специализированное программное обеспечение (ПО) для подготовки тестовых материалов и проведения тестирования. Свойства такого ПО (в первую очередь, автономность – для их работы не требовалось наличие сетевого соединения) в сочетании с удобством создания и редактирования

тестовых материалов обеспечило им достаточную популярность в конце 1990-х – начале 2000-х годов. Использование специализированного ПО при проведении тестирования также позволяло минимизировать объем передаваемых данных и давало возможность реализовать некоторые сложные типы интерактивных тестовых заданий. Недостатками таких систем являются:

- необходимость предварительной установки и настройки дополнительного ПО, а иногда – привязка к конкретной операционной системе;
- возможные ограничения по форматам представления информации;
- необходимость установки специального клиентского приложения тестирования;
- проблема обновления версий клиентов при совершенствовании ПО.

Отличительной особенностью и общим недостатком автономных тестирующих систем («HotPotatoes», «TestCommander», «Конструктор тестов», «WebQuiz XP» и др.) является отсутствие или ограниченность возможностей по обмену информацией внешними информационными системами (как импорт заданий из офисных форматов, так и экспорт результатов), отсутствие средств адаптивного тестирования и возможностей персональной настройки учителем свойств заданий в траекториях, а также отсутствие поддержки международного стандарта IMS QTI.

Современными решениями являются онлайн-тестирующие системы, создаваемые разработчиками в рамках более крупных проектов (например, Google Формы); онлайн-тестирующие системы, разрабатываемые специально для образовательных целей (ClassMarker, LearningApps, Quizlet, «Тесториум» и др.). Оба типа продуктов ориентированы только на создание оболочки для тестирования, но не на сам контент. С другой стороны, контент-ориентированные системы тестирования в ряде онлайн-систем, ориентированных на проведение текущего тестирования или подготовку к ЕГЭ/ОГЭ (<http://ege.yandex.ru/>, <http://www.school-tests.ru/>, <http://4ege.ru/>, <http://reshuege.ru/> и др.) ориентированы на воспроизведение готового контента, создание вопросов в них редуцировано. Задания размещены в системе в порядке, предусмотренном разработчиками; не редактируются и не переставляются, набор типов заданий ограничен, взаимодействие с системой максимально простое – прохождение теста и сбор общей статистики о выполнении заданий. К тому же большинство из них рассчитаны только на выпускников (учащихся 11-х и 9-х классов), остальные классы, в которых выполнение учебных заданий и тестирование выполняет не только контролирующую, но и обучающую функцию, не охвачены.

Системы организации дистанционного обучения достаточно сложны в использовании для обычной общеобразовательной организации, так как создавались для

потребностей высшего профессионального образования и дистанционного обучения взрослых (корпоративного обучения). Поэтому значительный объем функций является избыточным для работы школьного учителя, тогда как основные повседневные рабочие процессы (например, быстрое распределение учащихся по вариантам, работа в бескомпьютерном, «бумажном» режиме проведения тестов) в таких системах не поддерживаются.

Особый интерес представляют современные «облачные» решения, сочетающие в себе оптимальный для использования в образовательном процессе учреждений общего образования функционал с достаточной «лёгкостью» системы: невысокими аппаратными требованиями к конечным устройствам потребителей, кроссплатформенностью и ориентированностью именно на школьные процессы. Примером такой системы является интерактивный мультимедийный учебно-методический комплекс (ИМУМК) «**Облако знаний**» издательства «ФИЗИКОН» (www.облако-знаний.рф), в котором воплотились лучшие идеи и направления развития образовательных программных продуктов и решений, созданных этой компанией. ИМУМК «Облако знаний» – это информационная система, призванная на основе современных облачных технологий поддержать образовательный процесс в школе, муниципалитете, регионе, обеспечивая учащихся качественным образовательным контентом по разным предметам, жанрам и уровням образования, а учителей и руководителей – удобными сервисами, включая развитую систему тестирования.

Рассмотрим более подробно принятую в ИМУМК систему тестирования. Она включает большой набор интерактивных заданий, порядок заполнения формы для ответа существенно зависит от типа задания. Разработчики выделяют три группы заданий, различающихся по форме ввода ответа:

- указательные – ответ вводится путем указания элемента (группы элементов),
- манипулятивные – ответ вводится путем перемещения элемента (группы элементов),
- клавиатурные – ответ вводится путем формирования нового контента (например, при вводе символов или чисел с клавиатуры, построения формулы, графика или чертежа).

Указательные и манипулятивные типы заданий являются закрытыми (обладают сравнительно небольшим количеством различных вариантов ответа – не более нескольких десятков). Клавиатурные типы заданий являются открытыми – число вариантов ответа в них очень велико.

К указательным типам относятся задания:

- на выбор одного варианта ответа,
- на выбор нескольких вариантов ответа,
- на выбор (указание) элемента на рисунке,
- на выбор элемента из выпадающего списка (в т. ч. при заполнении таблицы),
- на разбор слова,
- на разбор предложения.

К манипулятивным относятся задания:

- на установление соответствия,
- на упорядочивание (сортировку) объектов,
- на распределение по группам (классификацию) объектов,
- на установление графических связей,
- на перемещение объектов на рисунке (составление композиций);
- работа с картой.

К клавиатурным относятся задания:

- на ввод числа (в т. ч. с контролем точности),
- на ввод строки,
- на ввод математической формулы,
- на ввод химической формулы.

К смешанным типам относятся задания:

- на работу с эмуляцией программного обеспечения,
- комбинированный тип заданий.

Такой выбор заданий в сочетании с рядом других важных характеристик (развитой системой аутентификации, поддержкой многобалльных шкал, возможностью сбора статистических показателей тестов и др.) делают «Облако знаний» оптимальной системой в качестве основы для организации компьютерного тестирования и мониторинга знаний учащихся.

1.4. Тестирование в ПАК «Мобитест»

Программная часть ПАК «Мобитест» предназначена для организации тестирования в классах с разным уровнем обеспеченности компьютерной техникой, в том числе с использованием мобильных устройств и личных устройств учащихся (BYOD, см. выше) и в классах без компьютеров. С помощью «Мобитеста» могут быть сформированы и проведены самостоятельные, контрольные и мониторинговые работы разных уровней сложности. В системе реализована возможность быстро и беспристрастно проверить результаты выполнения и проанализировать использованные тесты методами педагогической статистики.

Существенной особенностью Системы является наличие обширной базы заданий по различным предметам для различных уровней обучения с поддержкой возможности для учителей формирования собственной базы заданий.

2. ПОНЯТИЕ МОНИТОРИНГОВОЙ РАБОТЫ В ПАК «МОБИТЕСТ»

2.1. Ролевая модель пользователей программной части ПАК «Мобитест»

Все пользователи работают с Системой в единой ролевой модели. Авторизованным пользователям доступны функциональные возможности согласно их ролям в Системе. Один пользователь может выступать в нескольких ролях. Роль школьного администратора назначается службой технической поддержки Системы. Все остальные роли в системе, а также регистрацию пользователей Системы осуществляет школьный администратор.

В системе реализованы роли школьного администратора, учителя, учащегося, директора. Роль эксперта для проверки заданий с развернутым ответом назначается учителем, создающим мониторинговую работу. Эту роль может выполнять сам учитель или другой учитель/учителя, авторизованные в виртуальной школе.

Распределение функций по ролям пользователей приведено в Таблице 3.

Таблица 3. Функции пользователей различных ролей при работе с мониторинговыми работами на уровне класса

Функции («←» – функция не поддерживается, «+» – функция поддерживается)	Учащийся	Администратор	Учитель	Директор
Регистрация пользователей в Системе в различных ролях	–	+	–	–
Управление свойствами и структурой виртуальной школы	–	+	–	–
Создание и управление вариантами КИМ и спецификациями	–	–	+	+
Создание и управление свойствами мониторинговой работы	–	–	+	+
Назначение мониторинговой работы группе учащихся	–	–	+	+
Назначение экспертов для проверки работ на уровне класса	–	–	+	+
Печать комплектов работ	–	–	+	+
Пробное прохождение мониторинговой работы	–	–	+	–
Выполнение мониторинговой работы	+	–	–	–
Проверка заданий с развернутыми ответами, работа с веером ответов	–	–	+	+
Ввод результатов работ с бланков	–	–	+	–

Функции («←» – функция не поддерживается, «+» – функция поддерживается)	Учащийся	Администратор	Учитель	Директор
Подтверждение результатов работы	–	–	+	+
Просмотр диагностических отчетов по мониторинговой работе	–	–	+	+
Просмотр статистических отчетов по мониторинговой работе	–	–	+	+
Просмотр отчетов по IRT-статистике мониторинговой работы	–	–	+	+

2.2. Модуль администрирования

Виртуальная школа – это модель реальной образовательной организации общего образования (школы, лицея, колледжа) в части электронного обучения. Виртуальная школа состоит из нескольких классов, называющихся, как и в реальной школе, по названию параллели с добавлением литеры (например, 11 «А»).

Имеющие отношение к школе пользователи «привязываются» к ней посредством своих учетных записей. Далее мы будем называть таких пользователей привязанными к школе.

Администрирование виртуальной школы осуществляется пользователем с ролью администратора. Ему доступны следующие разделы меню:

- Школа – в этом разделе пользователь может управлять параметрами виртуальной школы;
- Классы – в этом разделе пользователь может управлять параметрами классов, добавлять и удалять классы;
- Пользователи – в этом разделе администратор может просматривать и изменять параметры пользователей, привязанных к виртуальной школе, создавать новых пользователей, привязывать и удалять пользователей;
- Настройки – в этом разделе администратор может редактировать параметры учетной записи и менять параметры текущей виртуальной школы.

2.3. Понятие мониторинговой работы, ее формирование и параметры

Мониторинговая работа представляет собой совокупность однотипных контрольно-измерительных материалов, являющихся вариантами друг друга. В соответствии с названием в мониторинговую работу могут входить траектории типа «Экзамен».

Мониторинговая работа может быть назначена одному или нескольким учащимся на период времени для ее прохождения (это делается для того, чтобы все учащиеся проходили работу в одно и то же время).

Технологически мониторинговая работа – это отдельный объект системы. Она содержит:

- мета-данные работы,
- равноправные варианты КИМ, составляющие работу,
- свойства назначения работы.

Контент для мониторинговой работы подбирается учителем. Контент может подбираться и из отдельных заданий (в т. ч. готовящихся учителем непосредственно перед сборкой мониторинговой работы).

2.4. История мониторинговых работ

После проведения мониторинговых работ, проверки заданий с развернутым ответом назначенным экспертом или самим учителем, работы перемещаются в раздел «Завершенные мониторинговые работы», и становятся доступными результаты работы и отчет по работе.

Таблица результатов работы содержит список учеников, выполнявших работу, номер варианта работы и тестовый балл, набранный учащимися.

Раздел «Отчеты» содержит данные мониторинга по выбранной работе: регламент, протокол, набранные учащимися баллы, распределение результатов работы по выполняемым вариантам и по трудности заданий, таблицу корреляции и валидности заданий.

Отчеты по проведенным работам доступны учителю, директору образовательного учреждения в разделе «Работы».

3. ПРОВЕДЕНИЕ МОНИТОРИНГОВОЙ РАБОТЫ

3.1. Организационные условия проведения мониторинговой работы и подготовка к ней

Мониторинговая работа создается пользователем в роли «Учитель» (или «Директор») из раздела **Работы** своего личного кабинета. Для этого он использует кнопку **Создать** и на сцене создания работы заполняет поля общего описания работы:

- название работы,

- предмет,
- класс,
- длительность работы,
- инструкция,
- ссылка на шкалу,
- уровень работы (школа, класс).

Помимо заполнения параметров работы, учитель наполняет работу вариантами КИМ. Чтобы добавить вариант, учитель должен сохранить мониторинговую работу и щелкнуть **Добавить вариант** и в появившемся окне выбрать вариант КИМ. Выбор производится с помощью таблицы, в которой показываются все доступные варианты, с указанием названия курса, названия траектории в составе курса, количества позиций, на которые задания подбираются по случайному алгоритму, контактного времени, количества использований данного варианта. У учителя есть возможность добавить сразу несколько вариантов (путем их множественного выделения). После выбора вариантов остается нажать кнопку **Добавить**, чтобы они прикрепились к мониторинговой работе.

Учителю также предоставляется возможность:

- пройти выбранный вариант;
- удалить выбранные варианты;
- сгенерировать дополнительные варианты (указав их количество).

В списке мониторинговых работ указываются:

- название работы,
- предмет,
- класс,
- уровень,
- дата проведения работы,
- количество учащихся, которым доступна работа.

Учителю доступен перечень всех мониторинговых работ, находящихся на его уровне в пределах его уровня доступа.

Статус работы определяется по следующему алгоритму:

- «В разработке» – для работ, у которых не заполнены какие-либо поля;
- «Готово» – для работ, у которых заполнены какие-либо поля, но которые еще не назначены учащимся;
- «Назначено» – для работ, назначенных учащимся, которые в настоящий момент нельзя пройти;

- «На прохождении» – для работ, которые учащимся, которым они назначены, можно пройти в настоящий момент времени;
- «На проверке» – для работ, прохождение которых завершено, но результаты еще не подтверждены;
- «Завершено» – для работ, прохождение которых завершено и результаты подтверждены.

При добавлении варианта КИМ в мониторинговую работу учителю предоставляется список всех доступных ему КИМ.

Назначить мониторинговую работу учащимся имеют возможность пользователи в роли «Директор» или «Учитель» – далее будем их называть организаторами работы. Для этого используются работы, которые имеют статус «Готово».

Назначение работы производится на отдельной вкладке по кнопке **Назначить** напротив записи работы или после нажатия кнопки **Перейти к назначению работы**. При назначении работы указываются:

- время самого раннего начала работы (в формате дата – часы – минуты),
- время самого позднего окончания работы (в формате дата – часы – минуты),
- перечень классов, которым назначена работа,
- назначенные эксперты,
- срок проверки ответов экспертами (в календарных днях с момента завершения мониторинговой работы).

Перечень пользователей, которым может быть назначена работа, формируется:

- для руководителя на уровне класса (учителя) – из числа учащихся класса,
- для руководителя на уровне школы (директора школы) – из числа учащихся школы.

В свойствах назначения мониторинговой работы перечень учащихся показывается с группировкой по классам и школам.

Организатор работы имеет возможность назначить на проверку ответов учащихся одному или нескольким экспертам. Эксперты выбираются учителем из списка доступных экспертов. Учитель может назначить экспертом самого себя.

Распределение вариантов производится автоматически; номера распределенных вариантов указываются напротив фамилий учащихся.

В перечень пользователей система включает только учащихся, класс которых соответствует классу мониторинговой работы. Организатору предлагается список классов, подтверждение выбора осуществляется кнопкой **Назначить** диалогового окна.

Контент, используемый при проведении мониторинговых работ, на период назначенных работ не доступен для обновления.

3.2. Проведение мониторинговой работы

Возможность прохождения мониторинговой работы в штатном режиме предоставляется учащимся в период времени между указанными в свойствах мониторинговой работы началом и концом работы. Прохождение осуществляется учащимся в личном кабинете на любом доступном устройстве (а также в «бумажном» варианте – см. следующий раздел).

Учащемуся не сообщается назначенный ему вариант работы.

3.3. Проведение работы в компьютерном варианте

Для выполнения работы пользователь в роли «Учащийся» должен войти в систему под своими регистрационными данными.

Перечень доступных работ демонстрируется учащемуся во вкладке Работы. Для мониторинговых работ указываются:

- наименование работы,
- графическое изображение работы (зависит от предмета и класса),
- дата и время начала работы,
- индикатор доступности работы,
- кнопка **Приступить**.

Там же ниже приведены мониторинговые работы, которые учащийся уже выполнил. По каждой работе указываются:

- наименование работы,
- графическое изображение работы (зависит от предмета и класса),
- дата завершения работы,
- количество оставшихся на проверке у учителя заданий с открытым ответом,
- набранный тестовый балл (в том случае, если проверка закончена) в виде полосы в нижней части блока.

Результаты завершенных мониторинговых работ доступны учащемуся по кнопке **Просмотреть**.

После запуска мониторинговой работы предьявляется инструкция. С инструкции осуществляется переход к самой работе с помощью кнопки **Начать**.

Учащемуся при выполнении работы не демонстрируются решения и ответы, и не сообщается, какие задания выполнены правильно, а какие – нет.

Мониторинговая работа состоит из интерактивных мультимедийных сцен. Каждая сцена размещается на отдельном экране и просматривается целиком или с помощью полос прокрутки. На сценах размещаются медиа-элементы (текст, формулы, изображения, видеофрагменты и анимация, звук, трехмерная графика), элементы оформления и элементы управления сценой. Другое название сцены – *интерактивное задание*.

Учащийся имеет возможность свободно переходить между заданиями работы, выполняя их в любом порядке. Также учащийся может в любой момент просмотреть свою работу в формате, эквивалентном бланкам ОГЭ/ЕГЭ.

Результаты выполнения мониторинговой работы (включая тестовый балл) становятся доступными для пользователя после того, как будут открыты результаты работы, и демонстрируются в дневнике пользователя.

При выполнении мониторинговой работы в Систему передаются ее результаты (включая неверные ответы пользователей).

В случае, если мониторинговая работа включает задания с развернутым ответом, ответы учащихся на эти задания поступают на проверку экспертам, назначенным на данную мониторинговую работу.

Перечень заданий пользователей, которые должен проверить эксперт, доступны ему в разделе **Проверка** его личного кабинета. Каждое задание каждого пользователя является отдельной единицей перечня. Проверка осуществляется по эталонному решению и критериям, предложенным для оценки авторами задания. После завершения проверки задание становится недоступным эксперту и исчезает из перечня назначенных на проверку заданий.

Если к мониторинговой работе прикреплено несколько экспертов, они будут иметь общий список ответов на проверку. В тот момент, когда эксперт берет очередной ответ на проверку, этот ответ исчезает из списка доступных для проверки ответов. В случае, если эксперт откажется от проверки, ответ возвращается в пул доступных ответов.

3.4. Проведение работы в бескомпьютерном (бумажном) варианте

Одним из направлений развития систем мониторинга является поддержка формы проведения мониторинговых исследований как электронное или бумажное («аналоговое») тестирование. Организация электронного выполнения мониторинговых работ описана выше. Для организации бумажного тестирования в «Мобитест» используется специальный программный модуль для подготовки и распечатки бланков работ для

учащихся. Генерация вариантов, распечатка заданий и бланков может осуществляться заблаговременно для быстрой раздачи и сбора в ходе урока.

Обработка результатов тестов производится с помощью компьютера: электронная проверка работ, выполненных в электронной форме, и обработка бумажных бланков: учитель может вводить ответы учащихся в бланк работы. Проверка открытых заданий проводится по бумажным или оцифрованным работам учащихся экспертами аналогично проверке открытых заданий с вводом развернутого ответа в электронном виде.

3.5. Диагностические и статистические отчеты

После того, как результаты мониторинговой работы подтверждены ее инициатором, появляется возможность формирования *диагностических отчетов*.

Диагностические отчеты доступны:

- для мониторинговой работы на классном уровне – учителю, проводившему работу в классе,
- для мониторинговой работы на школьном уровне – учителям классов, в которых проводилась работа, и директору школы.

Отчет по мониторинговой работе включает в себя ряд разделов.

Регламент работы – общая информация о проведении работы: название, предмет, уровень, дата проведения, количество вариантов, количество учащихся, принявших участие в мониторинговой работе, процент от общего количества учащихся, которым работа была назначена.

Протокол результатов мониторинговой работы представляет собой матрицу, в которой столбцы соответствуют заданиям, а строки – учащимся класса. Каждая запись (строка) представляет собой протокол результатов одного учащегося в данной мониторинговой работе. В него включаются фамилия и инициалы учащегося, номер варианта, первичные баллы за каждое из заданий, суммарный первичный балл за работу в целом и суммарный тестовый балл. Полученная информация может быть сохранена в файлы форматов PDF и XLSX, выведена на экран или на печать.

Распределение по трудности вариантов – отчет, в котором в табличном виде представляется средний балл прохождения работы по всем вариантам, представленным в работе. Отчет по трудности вариантов предназначен для просмотра и сравнения результатов выполнения отдельных вариантов работы.

Распределение по трудности заданий – отчет, в котором в табличном виде представляются данные по всем заданиям, включенным в мониторинговую работу: номер задания, уровень, перечень проверяемых тем по спецификации, средний балл по работе.

Отчет по трудности заданий предназначен для просмотра и сравнения результатов выполнения отдельных заданий работы.

В сводных отчетах представлены также таблицы корреляции и уровень валидности заданий теста.

На основании диагностических отчетов по отдельным работам собираются и анализируются **статистические отчеты**, в которых представляется усредненная информация по результатам выполнения мониторинговой работы какой-то из категорий учащихся.

Статистический отчет о проведении мониторинговой работы формируется в виде файла в формат PDF. Общий отчет должен последовательно включать титульный лист, включающий логотип системы, заголовок отчета в формате «Отчет о мониторинговой работе по [название предмета], место (административный центр проведения работы) и год составления отчета», а также следующие виды отчетов.

Регламент проведения мониторинговой работы. Общее описание мониторинговой работы с указанием предмета, уровня работы, даты проведения, количества вариантов, количества учащихся, принявших участие в работе, доли (процента) учащихся, принявших участие в работе, от общего количества учащихся, которым была назначена работа.

Распределение по набранным тестовым баллам. Отчет, в котором в табличном и в графическом виде представляется количество (доля от общего количества) учащихся, набравших тот или иной тестовый балл. В графическом виде результаты должны представлять собой столбчатую диаграмму. По оси *x* должны быть указаны тестовые баллы (в порядке возрастания), по оси *y* – доля учащихся, набравших данный балл.

4. ЧТО НУЖНО ЗНАТЬ О СВОЙСТВАХ ЗАДАНИЙ И ПЕДАГОГИЧЕСКОЙ СТАТИСТИКЕ?

4.1. Классическая тестовая теория

Современная теория конструирования педагогических тестов, широко известная за рубежом как ***Item response theory*** (IRT), – это семейство математических моделей и статистических процедур для анализа и функционального описания процесса выполнения теста.

Технология тестового контроля имеет длинную историю, однако строгую теоретическую и статистическую основу приобрела только в середине прошлого столетия и получила название **классической тестовой теории** (КТТ). В рамках этой теории были введены концепции параллельных тестовых форм, истинного тестового балла, надежности

теста. Эти концепции позволяют создателям и пользователям педагогических тестов определить, насколько они пригодны и точны. Испытуемые в КТТ также представляются в виде совокупности, как члены выборки или группы (статистической популяции), а не как индивидуумы.

Модели IRT используют одну, две или три характеристики задания (параметра задания), которые отражают различия среди заданий в их статистических признаках:

- a_j – параметр дискриминативности (дифференцирующей способности) задания;
- b_j – уровень трудности задания;
- c_j – параметр «угадывания».

Трехпараметрическая модель, которая учитывает все три параметра, называется также моделью Бирнбаума (Рис. 1); a на данном графике отражает крутизну кривой вероятности.



Рис. 1. Статистические характеристики тестового задания

Двухпараметрическая модель использует две характеристики (a и b). При этом исключается коррекция угадывания, которая представляет собой штраф, пропорциональный количеству дистракторов в задании (так как предполагается, что все учащиеся в той или иной степени угадывают ответы на вопросы). Данный вариант математической модели педагогического теста является наиболее распространенным и интересным как для авторов тестов, так и для их пользователей (в первую очередь – учителей).

Однопараметрическая модель (модель Раша) исключает и характеристику a . Постулируется, что дифференцирующая способность у всех заданий одинакова, только отличается сложность, что сдвигает кривую вправо-влево, но не меняет ее крутизны.

Рассмотрим основные характеристики тестовых заданий более подробно.

4.2. Трудность задания

b_j – уровень трудности задания

Доля учащихся, правильно выполнивших задание. Строго говоря, величина обратно пропорциональная: чем больше доля выполнивших, тем легче задание. Диапазон характеристик заданий приведен в таблице 6.

Таблица 6. Распределение заданий по трудности в зависимости от доли ответивших на задание

Вывод	Значение b_j
Угадывание	$b_j < 0,2$
Сложные задания	$0,2 < b_j < 0,36$
Умеренно сложные задания	$0,36 < b_j < 0,84$
Легкие задания	$b_j > 0,84$

При составлении контрольных работ надо стремиться к заданиям средней трудности: именно среди заданий средней трудности встречаются задания с высокой дифференцирующей способностью.

4.3. Дифференцирующая способность заданий

Показатель, позволяющий оценить, насколько хорошо то или иное задание способствует уточнению оценок (иными словами – разделению учащихся на группы по уровню знаний).

a_j – параметр дифференцирующей способности (дискриминативности) задания.

Значения, принятые в оценке этой характеристики, приведены в таблице 7.

Таблица 7. Распределение заданий по дифференцирующей способности в зависимости от значения показателя

Обозначение	Дифференцирующая способность	Значение a_j
G	1) отсутствует	0
F	2) очень низкая	0,01 – 0,34
E	3) низкая	0,35 – 0,64
D	4) средняя	0,65 – 1,34
C	5) высокая	1,35 – 1,69
B	6) очень высокая	> 1,70
A	5) отличная	> 2,0

Отдельные исследователи дополнительно разбивают класс D на два подкласса: D_1 со значением a_j от 0,65 до 0,99 и D_2 со значением a_j от 1,00 до 1,34.

На основе полученных данных авторы теста анализируют задания, выясняя, в каких именно элементах задания (основа задания, ключи, дистракторы) были допущены неточности, и как могут трактоваться те или иные выборы учащихся (какие рекомендации учителям могут быть сделаны по результатам анализа заданий).

4.4. Рейтинговый балл

Самый важный практический вывод из оценки дифференцирующей способности – это возможность получения более адекватно отражающего знания учащихся рейтингового балла. В ряде случаев может наблюдаться ситуация, при которой у некоторых учащихся рейтинговый балл, рассчитанный с учетом сложности заданий, может быть существенно выше, чем сырой балл (если располагать задания в порядке повышения трудности, такого бы не наблюдалось, но в реальных работах так редко бывает).

4.5. Валидность заданий

Валидность задания позволяет судить о том, насколько задание пригодно для работы в соответствии с общей целью создания теста. Иными словами, она определяет вклад отдельного задания в достижение успеха по тесту в целом (выражаемого через количество набранных сырых баллов).

Измеряют валидность заданий путем расчета *коэффициента бисериальной корреляции*. Это показатель того, как ответ на конкретный вопрос способствует максимальному сырому баллу по тесту в целом. Разница между дискриминативностью и валидностью (Рис. 2) в том, что валидность считает хорошим вопрос, у которого зависимость между вероятностью правильно ответить на вопрос и успехом выполнения теста в целом растет прямо пропорционально, а у дискриминативности – не прямо, а скачком. Поэтому валидность для контрольного теста – показатель менее надежный и менее значимый, чем дискриминативность.

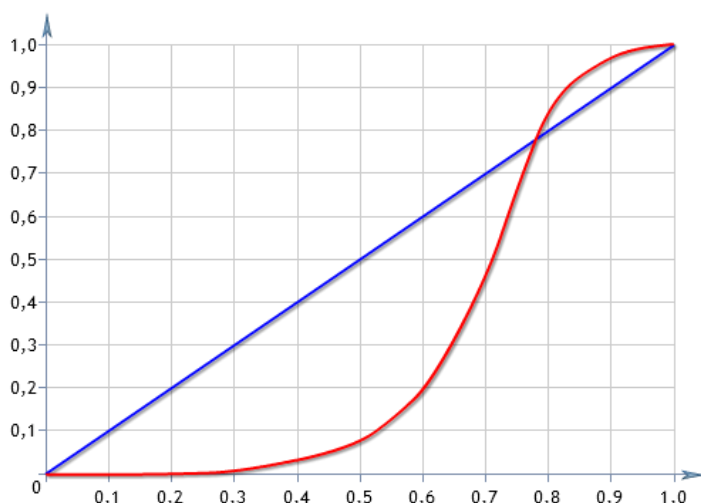


Рис. 2. Распределение ответов учащихся на два задания: с хорошей валидностью (синяя линия) и с хорошей дифференцирующей способностью (красная линия)

4.6. Надежность теста

Эта характеристика измеряется для теста в целом. Для однократного теста (без повторного тестирования, без распределения на варианты и последовательного выполнения сначала одного, а потом другого варианта) основным методом вычисления надежности является *расщепление* – сравнение выполнения испытуемым двух равноценных частей теста. Для оценки надежности этим методом выбираются две эквивалентные по характеру и степени трудности группы задач (например, четные или нечетные вопросы теста). Но данная математическая модель работает только в случае гомогенного теста (построенного из одинаковых по типу и сложности вопросов).

В тех случаях, когда кроме сырого тестового балла имеется в наличии независимая экспертная оценка (например, оценка учителя по результатам устного экзамена), которая постулируется как правильная, можно рассчитать надежность теста специальной методике (альфа Кронбаха); это также весьма популярный метод оценки надежности тестов.

При проведении мониторинговых работ осуществляется накопление данных и формирование отчетов по трудности вариантов и по трудности заданий, а также по их дифференцирующей способности. Отчеты по трудности доступны на всех уровнях (учитель, директор школы). Учителям и экспертам доступны подробные отчеты по всем статистическим характеристикам тестов, рассчитываемых системой. Они также могут быть представлены в табличном, графическом и картографическом вариантах. Так, отчет по трудности заданий в табличном варианте представляет собой матрицу, строками в которой являются задания мониторинговой работы, а столбцами – номер задания (слота) в работе, средний процент выполнения на данном уровне (в школе в целом, в классе),

средний процент выполнения по работе в целом. В графическом виде результаты представляют собой столбчатую диаграмму, в которой по оси x указаны варианты или задания, по оси y – средний процент их выполнения.

ЗАКЛЮЧЕНИЕ

Программная часть ПАК «Мобитест» есть комплексное решение автоматизации мониторинга и контроля знаний учащихся.

Система обеспечивает учителя доступным, легким в освоении инструментом формирования КИМов, возможностью менять содержимое вопросов и спецификаций, дополнять базу вопросов по своему предмету. Руководители и администрация получают возможность проводить срезы знаний и получать полные статистические отчеты по ним. Учащимся система позволяет отработать навыки тестовой формы контроля знаний, в том числе в форматах ЕГЭ и ОГЭ.

ЛИТЕРАТУРА

1. Аванесов В. С. Композиция тестовых заданий. – М., Центр тестирования, 2002.
2. Беспалько В.П. Образование и обучение с участием компьютеров (педагогика третьего тысячелетия) / В.П.Беспалько. – М.: Изд-во МПСИ, – 2008. – 352 с.;режим доступа: http://www.eusi.ru/lib/bespalko_obrasovanie/.
3. Гераськин А. С. Роль и место тестового контроля в современном образовании / А. С. Гераськин, С. В. Синаторов // Образование в современном мире: Сб. научных статей. – Саратов: Издательство Саратовского университета, 2009. Выпуск 4. – С. 260–263.
4. Козленко А.Г. Информационная культура и/или компьютер на уроке биологии / Лекции 5–8. – М.: Педагогический университет «Первое сентября», 2009; режим доступа: http://www.kozlenkoa.narod.ru/docs/game/ic_02.pdf, лекция 8.
5. Майоров А. Н. Теория и практика создания тестов для системы образования (как выбирать, создавать и использовать тесты для целей в образовании). – М., «Интеллект-центр», 2001 – 296 с.
6. Schonell F. J., Schonell, F. E. Diagnostic and Attainment Testing. Including a Manual of Tests, their Nature, Use, Recording and Interpretation. Edinburg London, Oliver and Boyd, 1956.
7. Thorndike R. L., Hagen, E. Measurement and Evaluation in Psychology and Education. New-York, Wiley; London, Chapman and Hall cop. 1955.