

**Чернявская А.П.
Гречин Б.С.**

**СОВРЕМЕННЫЕ СРЕДСТВА
ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ
ОБУЧЕНИЯ**

**Ярославль
2008**

Печатается по разрешению
редакционно-издательского совета
ЯГПУ им. К.Д. Ушинского

Рецензент: Тишко А.П., ст. преп. кафедры теории и истории педагогики

Чернявская, А.П., Гречин, Б.С.

Современные средства оценивания результатов обучения [Текст] : учебно-методическое пособие. – Ярославль : Изд-во ЯГПУ. – 2008. – 98 с.

В пособии представлены 11 конспектов лекций по современным средствам оценивания результатов обучения, контрольный тест и задания для самостоятельной работы.

Пособие предназначено для студентов и преподавателей педагогических вузов, изучающих курс «Современные средства оценивания результатов обучения».

А.П.Чернявской подготовлены темы 1, 2, 5, 6, 7, 11.

Б.С.Гречиным – темы 3, 4, 8, 9, 10, контрольный тест по курсу.

Тема 1
ПОНЯТИЕ О КАЧЕСТВЕ ОБРАЗОВАНИЯ. ОЦЕНКА
РЕЗУЛЬТАТОВ ОБУЧЕНИЯ КАК ЭЛЕМЕНТ
УПРАВЛЕНИЯ КАЧЕСТВОМ

1. Понятие «качество образования»

- 1.1. Что входит в понятие «качество образования».
- 1.2. Стандарты измерения качества образования.
- 1.3. Развитие качества образования.

1.1. **Качество образования** – это общественный продукт, зависящий от позиции и комплексной организации усилий всего образовательного сообщества (педагогов, управленцев образования, обучающихся и их родителей, региональных и федеральных систем образования).

Требования к качеству образования отражены в таких документах, как Закон РФ «Об образовании», «Концепция модернизации российского образования на период до 2010», «Стратегия модернизации содержания общего образования», «Государственный образовательный стандарт» и др.

Концепция модернизации российского образования на период до 2010 года под новым качеством образования понимает [7]:

- соответствие целей и результатов общего образования современным социальным требованиям, связанным с переходом к открытому демократическому обществу с рыночной экономикой, что требует от людей принятия самостоятельных и сознательных решений на основе освоения социального опыта, умения жить в условиях трудовой и социальной мобильности, повышения уровня толерантности;
- соответствие содержания образования его целям и познавательным возможностям всех учащихся; повышение уровня доступности и востребованности школьного образования; развитие навыков, необходимых для приобретения знаний в течение всей жизни, и информационной грамотности;
- соответствие условий образовательной деятельности требованиям сохранения здоровья учащихся и обеспечения психологического комфорта всех участников образовательного процесса.

М.М. Поташник, Е.А. Ямбург, Д.Ш. Матрос так определяют качество образования [10.]:

1. Образование включает в себя обучение, воспитание и развитие.
2. Качество образования – это соотношение цели и результата образования. При этом цели заданы только операционально и спроектированы в зоне ближайшего развития ученика. Результат также формулируется операционально.
3. Качество образования определяется не только соответствием количества и качества знаний учащихся ГОСу, но и качеством личности, духовного и гражданского развития детей. В этом главная общественная ценность образования.

С.Е. Шишов и В.А. Кальней [18] рассматривают качество образования как социальную категорию, которая определяет состояние и результативность процесса образования в обществе, его соответствие потребностям и ожиданиям различных социальных групп и общества в целом в развитии гражданских, бытовых и профессиональных компетенций личности.

1.2. В области обучения и образования действует международный стандарт оценки качества ИСО 9001:2000. В нем используется процессный подход. Если рассматривать образование как процесс, то становится очевидной его непрерывность, что предъявляет особые требования не только к организации самой образовательной деятельности, но и к

оценке ее качества. Возникает необходимость: - организации регулярной и систематической оценки, то есть внедрение мониторинга;

- оценки только конечного результата без анализа факторов, которые на него повлияли (используемых методов и технологий, системы взаимодействия и др.);
- оценки с учетом движущих сил данного процесса – заказчика (общества и различных социальных институтов) и потребителей (обучающихся и их родителей).

Государственный образовательный стандарт (ГОС) - социальная норма образования и образованности, разрабатываемая и вводимая в соответствии с Конституцией РФ. Он включает в себя федеральный и национально-региональный компоненты. Федеральный компонент ГОС определяет обязательный минимум содержания основных образовательных программ, максимальный объем учебной нагрузки обучающихся, требования к уровню подготовки выпускников.

Соответствие требованиям ГОС и другим стандартам качества образовательного учреждения оценивается в процессе лицензирования, аттестации и аккредитации. Лицензирование дает право образовательному учреждению на проведение образовательной деятельности. Аттестация говорит о реализации образовательного процесса в соответствии с лицензионными нормативами, требованиями ГОС и дает право на прохождение аккредитации. Аккредитация образовательного учреждения дает ему право на выдачу документа об образовании государственного образца.

1.3. Показатели качества обучения можно разделить на две основные группы:

- показатели, характеризующие качество учебного процесса;
- показатели, характеризующие уровень подготовки обучающихся.

Цели оценки качества образования:

- определить уровень учебных достижений;
- выявить конкретные сильные и слабые стороны в знаниях и навыках, которыми владеют учащиеся;
- выяснить, имеются ли проблемы с учебными достижениями у тех или иных групп обучающихся;
- выявить факторы, связанные с учебными достижениями;
- отслеживать динамику учебных достижений.

Возможны два механизма совершенствования системы качества образования:

- один из них реализуется в педагогической системе; он включает выявление несоответствий и проведение корректирующих или предупреждающих действий преподавателем при реализации педагогической технологии;
- второй механизм заключается в критическом анализе системы в целом в процессе различных ее рассмотрений, прежде всего, в ходе анализа со стороны руководства различного ранга и доработок системы на их основе.

2. Современное понимание «знания».

Компетентностный подход в образовании

2.1. Современное понимание «знания».

2.2. Компетентностный подход в образовании.

2.1. Традиционная практика обучения видит свою важнейшую задачу в том, чтобы приобщить обучаемых к обобщенному и систематизированному опыту социальной практики общества. Отсюда - утверждение ведущей роли теоретических знаний, ориентация на усвоение основ наук. Широко распространен взгляд на образование как на способ и процесс передачи знаний, умений и навыков. Учебная деятельность обучающихся оказывается как бы вырезанной из контекста реальной жизни - им навязываются цели усвоения накопленной информации. Этим, прежде всего, объясняется падение интереса к учению и профессии.

Информация и знания – это не одно и то же. *Учебная информация* - это знаковая система, учебный текст, звуки произносимых педагогом слов, которые должен воспринять и усвоить обучающийся.

Знания - проверенный практикой результат познания действительности, ее верное отражение человеком, ставшее ориентировочной основой компетентного действия. Знание - это информация, перекодированная на язык мозговых структур. Чтобы получить статус знания, осознанного отражения действительности, информация с самого начала должна «примериваться» к действию и поступку, усваиваться в их контексте. Мышление рождается в проблемной ситуации, а не в ходе решения задачи по заранее известному образцу.

В настоящее время в сфере образования все больше распространяются субъект-субъектный, личностно-ориентированный и деятельностный подходы.

Основные особенности субъект-субъектного образования [17].

1. Ситуация, когда учащийся стремится сам приобрести знания, а учитель создает для этого условия;
2. Использование методов, которые дают возможность совместного «сотворения» или изобретения знания в процессе обучения;
3. Вариативность общего содержания обучения с учетом индивидуальных особенностей учащихся при сохранении инвариантности базового содержания, в том числе вариативность сроков освоения материала при инвариантности общего периода обучения;
4. Наличие общих правил учебной деятельности и взаимодействия, которым подчиняются все субъекты образовательного процесса и вариативность их применения;
5. Наличие особой системы взаимодействия субъектов образовательного процесса и диалог как отражение этой системы;
6. Оценивание не только стандартного набора знаний учащегося, но и его «индивидуализированной» карты знания и процесса развития в соотнесении с его возможностями;
7. Равное внимание к процессам получения знаний и целостного развития личности ребенка.

Суть личностного подхода состоит в том, что образование рассматривается не как способ передачи знаний, умений и навыков, а как способ развития личности. Личность при этом понимается как субъект свободной творческой деятельности. Это предполагает отношение человека к собственному развитию как к ценности. Процесс формирования человека как субъекта творческой деятельности составляет главную линию его развития. С позиции деятельностного подхода, качество образования можно определить как меру соответствия результатов развития личности обучающихся в конце какого-либо возрастного периода возможностям для развития, содержащимся в культуре общества. В процессе образования на соответствующих возрастных этапах у человека должны формироваться определенные целостные виды деятельности и соответствующие им способности. Этим результаты образования, ориентированного на развитие личности, отличаются от результатов обучения и воспитания, ориентированных на передачу знаний, умений и навыков.

Рефлексивные умения выражаются в способности формировать представления о себе, о своей деятельности, прогнозировать ее и планировать, а также выяснять и анализировать мнения других людей о своей деятельности и о себе. Рефлексивные умения являются важнейшим компонентом в структуре инновационной деятельности специалиста, поскольку специалист с развитыми рефлексивными умениями имеет большие возможности в реализации творческого потенциала.

К общим рефлексивным умениям относятся:

- умение ставить реальные цели как основу для оценки результативности своих действий;

- адаптировать, дополнять или изменять план, программу, формы и методы работы соответственно конкретным условиям, с целью достижения оптимальных результатов;
- прогнозировать последствия своих действий, осуществлять контроль и самоконтроль, осознавать выполненную деятельность, самокритично относиться к ней;
- оценивать и обобщать опыт своей практической работы и применять в своей практике опыт других людей.

2.2. По Дж. Равену, «компетентность – это специфическая способность, необходимая для эффективного выполнения определенного действия в конкретной предметной области и включающая узкоспециальные знания, особого рода предметные навыки, способы мышления, а также понимание ответственности за свои действия» [12. С.6]. Природу компетентности можно понять на основе следующей схемы (Схема 1) [17].

Компетентность формируется в той области, где все три параметра накладываются друг на друга.

Схема 1

Природа компетентности

Остается до конца не решенной, в силу сложности и многогранности понятия, проблема классификации компетентностей. Чаще всего проводится классификация по содержанию, в которой выделяются ключевые компетентности. К примеру, Т.В. Иванова выделяет 4 компетенции: социальную, коммуникативную, информационную, учебно-познавательную. Другое основание – по сферам проявления. Так разработчики «Стратегии модернизации содержания общего образования» указывают, что в структуре ключевых компетентностей должны быть представлены компетентности в сферах деятельности: самостоятельной познавательной, гражданско-общественной, социально-трудовой, культурно–досуговой.

Совет Европы определил пять ключевых компетенций:

- политические и социальные,
- компетенции, касающиеся жизни в многокультурном обществе,
- компетенции, касающиеся владения устным и письменным языком,
- компетенции, связанные с возникновением общества информации,
- способность учиться всю жизнь.

С.Е. Шишов приводит совокупность ключевых компетенций, сформулированных в рамках проекта «Среднее образование в Европе» [18]: изучать, искать, думать, сотрудничать, приниматься за дело, адаптироваться.

Тема 2

МОНИТОРИНГ КАЧЕСТВА ОБРАЗОВАНИЯ

1. Средства накопительной оценки

1.1. Необходимость внедрения накопительной оценки в образовательную практику.

1.2. Критерии эффективной оценки.

1.1. Адаптивность системы образования требует определения соответствия деятельности конкретной педагогической системы возможностям и образовательным потребностям конкретного учащегося. Учение в условиях адаптивной системы обучения становится преимущественно активной самостоятельной деятельностью, управляемой посредством использования контроля и диагностики.

Средства контроля и диагностики в новых условиях изменяются. Отметочной системы, измеряющей только единичный конкретный результат, становится недостаточно. Для отслеживания процесса достижения образовательных целей необходимы средства,

дающие возможность проследить и оценить динамику процесса достижения целей. Таким образом, возникает необходимость введения накопительной системы оценок, к которой относятся известные в отечественной системе обучения мониторинг, рейтинговое оценивание, портфолио. К накопительной оценке относятся, кроме того, используемые для оценивания интервью, деловые игры, дневники самооценивания, метод заключения соглашения и другие методы, используемые в западной дидактике.

Накопительные оценки позволяют обучающимся формировать положительное отношение к учебе, так как дают им возможность продемонстрировать то, как много они знают и умеют, а не их недостатки, что характерно для традиционных методов оценки. Они делают процесс обучения более эффективным, особенно при правильно организованной и конструктивной обратной связи. Новые методы оценки, такие как моделирование, практика, ролевые игры, позволяют студенту понять, как применить приобретенные умения и навыки внутри и за пределами образовательной среды. Появляется возможность оценить более разнообразный спектр умений учащихся в большем количестве ситуаций. При этом оценивать могут не только педагоги, но и родители, и, что самое важное, сам учащийся.

1.2. К основным характеристикам эффективной оценки относят то, что она фокусируется на процессе и на продукте; оценивается не только то, чему ученика учат, но и то, что от него ожидается; в процесс оценки активно вовлекаются и педагоги и учащиеся; оценка основывается на разнообразных и вариативных средствах; оценка проходит на всех этапах и уровнях обучения и обеспечивает участников оценки необходимой информацией для совершенствования процесса обучения посредством обратной связи. Накопительная оценка, при ее правильном использовании, реализует все эти требования.

2. Понятие педагогического мониторинга

2.1. Определения мониторинга.

2.2. Виды мониторинга.

2.1. Мониторинг – специально организованное, систематическое наблюдение за состоянием объектов, явлений, процессов с целью их оценки, контроля или прогноза. Педагогический мониторинг – это форма организации, сбора, обработки, хранения и распространения информации о деятельности педагогического коллектива, позволяющая непрерывно отслеживать состояние и прогнозировать его деятельность.

В процессе мониторинга выявляются тенденции в развитии системы образования, соотнесенные во времени, а также последствия принимаемых решений. В рамках мониторинга проводится выявление и оценивание проведенных педагогических действий. При этом обеспечивается обратная связь, осведомляющая о соответствии фактических результатов деятельности педагогической системы ее конечным целям.

Мониторинг затрагивает различные аспекты жизнедеятельности образовательного учреждения:

- анализ целесообразности постановки задач образовательного процесса, планов учебной и воспитательной работы;
- работу с кадрами и создание условий для творческой работы педагогов;
- организацию учебного процесса;
- сочетание контроля с оказанием практической помощи;
- здоровьесберегающий подход в обучении и воспитании обучающихся.

Главное отличие мониторинга качества обучения от контроля, прежде всего в том, что задача мониторинга - установление причин и величины несоответствия результата целям. Кроме того, мониторинг отличается систематичностью и протяженностью во времени, применяемыми критериями и показателями.

К основным функциям мониторинга относятся:

- диагностическая – сканирование состояния системы образования и происходящих в ней изменений, что позволяет дать оценку данным явлениям;

- экспертная – в рамках мониторинга возможно осуществление экспертизы состояния, концепции, форм и методов развития системы образования, ее компонентов и подсистем;

- информационная – мониторинг является способом регулярного получения сопоставимой информации о состоянии и развитии системы, необходимой для анализа и прогноза состояния и развития системы;

- интегративная – мониторинг является одним из системообразующих факторов, обеспечивающих комплексную характеристику процессов.

Выделяют общие особенности мониторинговой деятельности [5]:

а) объекты мониторинга динамичны, подвержены влиянию внешних воздействий, которые могут вызвать различные изменения в состоянии объекта;

б) реализация мониторинга предполагает организацию постоянного слежения за объектом, изучение и оценку его состояния;

в) организация слежения предусматривает отбор обоснованных критериев и показателей, по которым ведется измерение и описание параметров объекта;

г) каждая конкретная система мониторинга ориентирована на конкретного потребителя, которым может быть как отдельное учреждение, так и государство в целом.

2.2. Можно выделить основные виды мониторинга:

1. По содержанию:

- дидактический мониторинг, предметом которого являются новообразования учебного процесса (получение знаний, умений, навыков, соответствие их уровня требованиям ГОС и т. д.);

- воспитательный мониторинг, который учитывает изменения в создании условий для воспитания и самовоспитания обучающихся, «приращение» их воспитательного уровня;

- социально-психологический, показывающий уровень социально-психологической адаптации личности обучающегося;

- управленческой деятельности, показывающий изменения в различных управленческих подсистемах.

2. По характеру используемых методов и методик – статистический и нестатистический мониторинг.

3. По направленности:

- мониторинг процесса – представляет картину факторов, влияющих на реализацию конечной цели;

- мониторинг условий организации деятельности – выявляет отклонения от запланированной нормы деятельности, уровень рациональности деятельности, необходимые ресурсы;

- мониторинг результатов – выясняет, что было сделано из запланированного, какие результаты достигнуты.

3. Условия организации мониторинга в образовании

3.1. Задачи организации мониторинга.

3.2. Методы мониторинга.

3.3. Этапы мониторинга.

3.1. При организации мониторинга важно выполнить следующие задачи:

1. Определить критерии качества реализации мониторинга, выработать комплекс показателей, обеспечивающих целостное представление о состоянии системы, качественных и количественных изменениях в ней.

2. Отобрать средства диагностики.

3. Установить уровень соответствия реального состояния объекта ожидаемым результатам.

4. Систематизировать информацию о состоянии и развитии системы.

5. Обеспечить регулярное и наглядное представление информации о происходящих процессах.

6. Организовать информационное обеспечение анализа и прогнозирования состояния и развития системы образования, выработки управленческих решений.

Информация, собираемая в процессе мониторинга, должна соответствовать требованиям объективности, точности, полноты и достаточности.

3.2. Традиционный мониторинг в форме контрольных работ, экзаменов, инспекторских проверок недостаточно эффективен. Прежде всего, потому:

- контроль состояния обучения носит нерегулярный, эпизодический характер, не вскрывается динамика изменений;
- контролируя итоги обучения, оставляют без внимания сам процесс обучения;
- используются достаточно субъективные балльные отметки и интегральные оценки выполнения проверочных заданий в целом, что не позволяет выяснить, какие конкретно и в какой мере элементы содержания не усвоены;
- по существу не используются диагностические методики, позволяющие вскрыть причины тех или иных ошибок учащихся, недочетов в работе учителя, выявить факторы, влияющие на успеваемость.

Для проведения мониторинга могут применяться общие методы психолого-педагогических исследований – наблюдение, опрос, анкетирование, тестирование, эксперимент. Используются и специфические методы - анализ продуктов деятельности (например, документов), методы изучения состояния воспитательной работы, игровые методы, творческие отчеты, методы экспертных оценок, аналитико-оценочные методы (самооценка, анализ занятия, шкалирование и др.). Для обработки результатов мониторинга используют математико-статистические методы.

3.3. Мониторинг проводится по следующим этапам:

1. Подготовительный этап:

- формирование заказа на мониторинг,
- выделение объекта мониторинга,
- методическое обеспечение мониторинга,
- определение критериев и показателей,
- создание рабочего проекта или программы,
- инструктаж или подготовка кадров, проводящих мониторинг.

2. Этап проведения мониторинга:

- проведение диагностики системы с использованием выбранных методов в соответствии с рабочей программой,
- сбор и анализ, хранение результатов.

3. Этап обработки данных и принятия решений:

- обработка данных, в том числе математико-статистическая,
- анализ, обобщение и систематизация полученных данных,
- подготовка заключительного документа,
- принятие решений,
- комплекс мероприятий, активизирующих использование данных, в том числе информационное обеспечение мониторинга.

В настоящее время предлагаются детализированные технологии педагогического мониторинга [10], описание одной из которых мы приводим ниже.

Таблица 1

**Технология педагогического мониторинга в ОУ
(по В.В. Петрухину)**

Предмет мониторинга (объекты мониторинга)	Цели мониторинга	Периодичность	Ответственные (субъекты мониторинга)	Результат
---	------------------	---------------	--------------------------------------	-----------

Общие показатели и материально-техническая база образовательного учреждения	Сбор статистических данных об оснащенности учебного процесса, об уровне квалификации педагогических работников и динамике изменений качественного и количественного состава обучающихся	2 раза в год	Руководители методических объединений, зам. дир. по УВР	Аналитические записки, диаграммы
Учебно-методические обеспеченности учебного процесса	Сбор информации об обеспеченности учебного процесса учебной и методической литературой, средствами наглядности и т. д.	1 раз в год (май)	Руководители методических объединений	Сводные таблицы, диаграммы
Методическая работа школы	Создание банка данных о методической деятельности методического объединения, создание условий для систематического анализа методической работы	2 раза в год (сентябрь, май)	Руководители методических объединений, зам. директора по УВР	Аналитический отчет, включающий графики и диаграммы
Деятельность учителя	Сбор информации об уровне профессионализма учителей, о направлениях их методического развития и совершенствования, составление рейтинга учителей	Ежемесячно	Зам. дир. по УВР	Таблицы, справки, диаграммы
Деятельность классного руководителя	Определение уровня профессиональной компетентности классного руководителя, создание классным руководителем банка данных учащихся, информации об их уровне обученности, воспитанности, уровне развития классного коллектива и сформированности у учащихся основных компетенций	Ежемесячно	Зам. дир. по УВР, социальный педагог	Справки, таблицы, диаграммы
Нормативно-планирующая документация	Анализ качества составленных рабочих программ и календарно-тематического планирования	2 раза в год (сентябрь, апрель)	Руководители методических объединений	Протоколы
Деятельность учащихся	Определение уровня (качества) обученности учащихся	Ежемесячно	Классные руководители, зам. дир. по УВР	Аналитические отчеты, диаграммы, таблицы

ВИДЫ, ФОРМЫ И ОРГАНИЗАЦИЯ КОНТРОЛЯ КАЧЕСТВА ОБУЧЕНИЯ

Контроль в широком смысле – проверка чего-либо, установление обратной связи. Контроль учебной деятельности учащихся обеспечивает получение информации о результате их учебной деятельности, способствует установлению внешней обратной связи (контроль, выполняемый учителем) и внутренней обратной связи (самоконтроль учащегося).

Функции контроля:

- 1) *диагностическая функция* (определение качества усвоения пройденного материала, успехов и пробелов в ЗУНах);
- 2) *образовательная функция* (приведение ЗУНов в систему и содействие их прочному усвоению);
- 3) *развивающая функция* (внесение корректив в учебный процесс на основе изучения возможностей и способностей учащихся);
- 4) *воспитательная функция* (стимулирование систематических занятий по усвоению ЗУНов и ответственности за их усвоение).

Требования к контролю:

- 1) *индивидуальный характер*;
- 2) *систематичность, регулярность* проведения контроля на всех этапах;
- 3) *разнообразие методов и форм проведения*;
- 4) *всесторонность* (охват всех разделов учебной программы, проверка теоретических знаний, практических умений и навыков);
- 5) *объективность контроля* (отказ от преднамеренных, субъективных, ошибочных суждений).

Виды контроля:

- 1) *предварительный контроль* (перед началом изучения учебного материала для определения исходного уровня ЗУНов);
- 2) *текущий контроль* (проверка ЗУНов, приобретенных в ходе изучения нового материала, его повторения, закрепления и практического применения);
- 3) *тематический* (после изучения темы, раздела для систематизации знаний обучаемых);
- 4) *периодический* (контроль по целому разделу учебного курса);
- 5) *итоговый* (контроль в конце учебного года с учетом результатов периодического контроля).

Данная классификация иногда дополняется еще двумя видами контроля [3]:

- б) *повторный* (проверка знаний параллельно с изучением нового материала способствует прочности и системности знаний учеников);
- 7) *отсроченный* (контроль остаточных знаний и умений спустя какое-то время после изучения темы, раздела, курса).

Формы контроля:

- 1) индивидуальный;
- 2) групповой;
- 3) фронтальный;
- 4) парный.

Иногда в качестве форм контроля выделяют: *зачет, экзамен, собеседование, контрольную работу, общественный смотр знаний* (проводится в присутствии не только одного учителя, но и других учителей, родителей) и т. п.

Оценка и ее функции

Результаты контроля учебно-познавательной деятельности учащихся выражаются в ее оценке. *Оценить* – значит установить уровень, степень или качество чего-либо.

Оценка – качественный показатель (например, «Ты - молодец!»).

Отметка – количественный показатель (пяти или десятибалльная шкала, проценты).

Функции оценки:

- 1) *мотивационная* (поощряет, стимулирует учебную деятельность);
- 2) *диагностическая* (указывает на причины тех или иных образовательных результатов);
- 3) *воспитательная* (формирует самосознание и адекватную самооценку);
- 4) *информационная* (свидетельствует о степени успешности ученика).

Этапы развития пятибалльной шкалы оценок:

- 1) май 1918 года – постановление А.В. Луначарского «Об отмене отметок»;
- 2) сентябрь 1935 года – введено пять словесных (вербальных) оценок: «очень плохо», «плохо», «посредственно», «хорошо», «отлично»;
- 3) январь 1944 года – возврат к цифровой «пятибалльной» системе оценки успеваемости.

Недостатки традиционного пятибалльного оценивания:

- низкая дифференцирующая способность;
- травмирующий характер для ребёнка;
- малая информативность для родителей учащихся;
- невозможность фиксации уровня личных усилий ученика для достижения результата;
- малая пригодность для оценивания личностных качеств и достижений;
- создание формальной мотивации, отсутствия активной позиции по отношению к обучению и личной заинтересованности в его результатах.

Основные ошибки при выставлении отметки:

- 1) *инерция отметок* (выставление их по установившейся традиции, деление на слабых и сильных);
- 2) *искусственный либерализм и жесткий авторитаризм* (завышение и занижение отметок);
- 3) *средний балл успеваемости* (носит формальный характер);
- 4) *двойка как средство борьбы с недисциплинированностью* [4].

Тема 4

СИСТЕМА РЕЙТИНГ-КОНТРОЛЯ КАК СОСТАВНАЯ ЧАСТЬ УЧЕБНОГО ПРОЦЕССА

Рейтинг учащегося (англ. rating, от to rate – оценивать, ранжировать) - это индивидуальный числовой показатель интегральной оценки достижений в учебе, образуемый путем сложения рейтинговых баллов, полученных в результате оценки отдельных учебных действий, по возможности с учетом коэффициента значимости («весового» коэффициента) этих действий в достижении образовательных целей.

Основной принцип рейтинга – ранжирование.

Принципы рейтинговой технологии (по Л.Г. Устиновой)

Ориентационный принцип (чёткое определение для учащихся целей и задач работы, способов их достижения).

Принцип комплексного подхода (единство учебной, научной и воспитательной работы).

Принцип мотивации (формирование у учащихся потребностей в познании, в достижении цели, учении, труде).

Принцип объективности рейтинговой технологии обучения (независимость результатов от личности преподавателя).

Объекты рейтинговой системы контроля:

- совокупность знаний и умений по всем дисциплинам учебного плана;
- социальная активность учащегося.

Конечная цель рейтинговой технологии – становление учащегося как субъекта учебной, научной деятельности, то есть достижения такого уровня развития обучаемых, когда они оказываются способными поставить цель своей деятельности; когда учащиеся могут планировать, корректировать свои действия, соотнося результат с поставленной целью. Учащийся должен осознать себя субъектом учебного процесса. Поэтому в рамках рейтинговой технологии между преподавателем и учащимся возникают субъект-субъектные отношения.

Задачи рейтинга:

- построение системы комплексного мониторинга эффективности образовательной программы;
- повышение уровня мотивации учащихся в разных сферах деятельности
- построение индивидуальных образовательных программ;
- целевая подготовка лучших абитуриентов, отвечающих требованиям высших учебных заведений.

Функции рейтинговой технологии:

- креативная (формирование творческого потенциала учащихся);
- рефлексия (самооценка учебных и личностных достижений);
- стимулирующая (создание условий для развития активности и самостоятельности);
- диагностическая (получение информации об особенностях учебного процесса);
- коррекция (корректировка влияний нововведений на развитие личности учащихся) [3].

Условия организации рейтинговой системы оценивания:

- модульная организация учебного процесса,
- постоянное отслеживание уровня знаний,
- многобалльное оценивание обученности учащихся.

Виды рейтинга

(по М.В. Калужской, О.С. Уколовой, И.Г. Каменских):

- ***академическая успеваемость*** (отражает уровень учебных достижений),
- ***олимпийский*** (отражает участие в любых интеллектуальных, творческих и спортивных состязаниях),
- ***лидерский*** (отражает уровень социальной компетентности, общественную активность учащихся),
- ***званий и сертификатов*** (фиксирует личные достижения в области иностранных языков, искусства, науки и спорта, подтвержденные сертификатами) [6].

Преимущества рейтинговой системы оценивания

(по сравнению с традиционной технологией контроля):

- стимулирование систематической самостоятельной работы;
- снижение роли случайных факторов при получении итоговой оценки по дисциплине;
- повышение роли состязательности в учебе;
- чёткая дифференциация учащихся в соответствии с их успехами;
- количественная характеристика качества учебной работы;
- фиксация результатов учебной деятельности на всех этапах;
- учёт качества и сроков выполнения индивидуальных домашних заданий;
- повышение учебной мотивации.

Эффект от внедрения рейтинговой технологии выражается:

- в росте социальной активности учащихся;
- в признании равноценности различных образовательных траекторий;
- в ориентации на успех;
- в повышении адекватности самооценки;

- в укреплении командного духа;
- в осознанности жизненного выбора.

Управление качеством подготовки учащихся с помощью рейтинга

1. Сравнительный анализ рейтинг-листов и эталонной рейтинг-шкалы.
Поиск информации:
 - о количестве учащихся, достигших максимального предела эталонной шкалы;
 - о количестве учащихся, освоивших 75% содержания образования;
 - о количестве учащихся, освоивших менее 50% содержания образования.
2. Сравнение результатов с рейтинг-листами учащихся предыдущих лет обучения.
3. Характеристика (на основе сравнения и анализа):
 - уровня обучаемости учащихся;
 - уровня образовательной системы в учебном заведении ;
 - качества учебного процесса;
 - эффективности РСК.
4. Решения коррекции учебного процесса, технологии рейтинга.

Этапы проектирования рейтинговой системы контроля по дисциплине:

1. Формулировка целей обучения по учебному предмету (внутренних и внешних, диктуемых теми дисциплинами, которые опираются на данную).
2. Отбор и подготовка учебного материала.
3. Определение содержания контроля, то есть выделение основополагающей части учебного материала.
4. Разработка структуры системы контроля.
5. Назначение средств измерения (распределение заданий-измерителей по учебным темам программы курса).
6. Разработка системы оценивания (оценивание каждого задания - измеряется в баллах).

Основные понятия рейтинговой системы оценивания

Кумулятивный (эталонный) балльный показатель – обобщенный показатель качества обучения, характеризуемый максимально возможной суммой баллов, которую может накопить учащийся в результате изучения всех дисциплин учебного плана. Разрабатывается на базе эталонных показателей по всем изучаемым дисциплинам. При сравнении его с **нормативным (эталонным) показателем** характеризует степень полноты освоения содержания образования.

Рейтинг-шкала – унифицированная шкала рейтинговой оценки, принципы формирования которой являются общими для всех дисциплин. Полученный учащимся рейтинговый показатель в процентах от максимально возможной величины баллов по дисциплине вносится в интегральную **рейтинг-шкалу** учащегося.

Шкала содержит информацию о ее предельных значениях (максимальном и минимальном), о диапазонах рейтингового показателя, отражающих удовлетворительный, хороший, отличный и неудовлетворительный результаты учения и соответствующих привычным пятибалльным оценкам, а также его минимальных пределах, при которых учащийся допускается к экзамену, зачету и т. д.

Повышающие коэффициенты дисциплин. Так как дисциплины учебного плана неравноценны по значимости использования, существует система определения «весовых» значений дисциплин. В дальнейшем «весовой» показатель дисциплины позволяет устанавливать ее балльный показатель в соответствии с ее значимостью по эталонной шкале качества. Сумма рейтинговых баллов должна определяться с применением соответствующих «весовых» коэффициентов для наиболее важных дисциплин.

Критерии дифференциации учебных предметов (с целью определения повышающих коэффициентов):

1. Степень усваиваемости учебной информации (субъективное деление на «сложные» и «простые» предметы);
2. Дисциплины по их значимости в предстоящей учебной и профессиональной деятельности выпускников.

Контрольно-оценочные циклы. Рейтинговая система контроля по дисциплине формируется из системы последовательных блоков, так называемых *контрольно-оценочных циклов*.

Каждый блок организуется для контроля «порции» учебного материала или «порции» знаний и умений, включает в себя акты контроля, проверки, качественной и количественной оценки.

Преподаватель для контроля «порции» содержания обучения выполняет следующие процедуры:

- обеспечивает собственно контроль различными методами (тестирование, устный опрос, письменная работа и т. д.);
- проверяет выполнение работы, анализирует их и измеряет по заранее установленным критериям фактический результат;
- оценивает, то есть сравнивает полученный результат с эталоном (показателями) и определяет меру соответствия результата эталону;
- выставляет отметку.

Знание эталона соответствует норме баллов, ранее установленной для «порции» материала. Отметка может соответствовать норме или быть ниже ее. Полученные в результате баллы суммируются в кумулятивный балльный показатель учащегося. По мере накопления баллов определяется текущий рейтинг учащегося - место учащегося в группе по уровню усвоения им учебного материала. По рейтингу устанавливается уровень достижений обучаемого.

В течение полугодия идет непрерывное повышение кумулятивного показателя учащегося и изменение его рейтинга. Именно рейтинг и является критерием для организации самокоррекции учения.

В конце изучения дисциплины формируется кумулятивный балльный показатель учащегося по дисциплине и его итоговый рейтинг, определяемый в процентах к максимально возможной оценке. Тем самым преподаватель приводит кумулятивный показатель к унифицированной шкале.

В качестве заданий - измерителей обученности могут использоваться тесты, контрольные работы, коллоквиумы и т. п. [3]

Технология рейтинга (по М.В. Калужской, О.С. Уколовой, И.Г. Каменских)

Ранжирование учащихся происходит на основе кумулятивного балльного показателя по всем видам рейтинга (академический, олимпийский, лидерский, званий и сертификатов).

При формировании *рейтинга академических достижений* складываются средние баллы учащегося по всем предметам (а также баллы за поведение и прилежание) за определённый период, умноженные на весовой коэффициент дисциплины. Для простоты подсчёта во многих учебных учреждениях профилирующим предметам (в их число входят, например, русский язык, литература, алгебра, английский язык, история) присваивается весовой коэффициент, равный 1,5. Перечень предметов, которым присваивается повышающий коэффициент, должен определяться профилем учебного учреждения.

При формировании *олимпийского рейтинга* складываются баллы учащегося за участие в конкурсах и соревнованиях, назначаемые из расчёта: 3 место – 1 балл, 2 место – 2 балла, 1 место – 3 балла. Баллы за соревнования различных уровней умножаются на весовые коэффициенты из расчёта: городской уровень – 2, областной уровень – 3, общероссийский уровень - 4.

При формировании *лидерского рейтинга* складываются баллы, назначаемые классным руководителем или консилиумом педагогов за регулярное, систематическое участие:

- в социальных акциях (сбор вещей для детского дома и др., шефство над младшими школьниками);
- в органах классного и школьного самоуправления (совет класса, совет школы);
- в праздничных школьных спектаклях;
- в воспитательных мероприятиях (КВН, редколлегия стенгазеты и др.).

документами». Благодаря ему можно судить об учебных, творческих, коммуникативных способностях школьника.

Портфолио является способом фиксирования, накопления и оценки индивидуальных достижений школьника в определенный период его обучения. Оно выступает важным элементом практико-ориентированного подхода к образованию. Это своеобразный отчет по процессу обучения учащегося, позволяющий увидеть картину конкретных образовательных результатов, обеспечить отслеживание индивидуального прогресса в широком образовательном контексте, продемонстрировать его способности практически применять приобретенные знания и умения.

В зарубежной образовательной практике портфолио активно применяется уже достаточно долгое время.

Портфолио предназначено для того, чтобы систематизировать накапливаемый опыт, знания, четче определить направления своего развития (например, в будущей профессии), облегчить помощь или консультирование – со стороны преподавателей или более квалифицированных специалистов в данной сфере, а также сделать более объективной оценку своего уровня (учебного или профессионального). Оно может собираться с различными целями, поэтому единого определения данной технологии не существует.

1.2. Определения портфолио (по материалам американских исследований).

1. Коллекция работ учащегося, которая демонстрирует усилия, прогресс и достижения в определенной области.

2. Антология работ учащегося, предполагающая его непосредственное участие в выборе работ, представляемых на оценку, а также их самооценку и самоанализ.

3. Форма целенаправленной, систематической и непрерывной оценки и самооценки учебных результатов учащегося.

4. Выставка учебных достижений учащегося по данному предмету за данный период обучения.

5. Систематический и специально организованный сбор доказательств, используемый учителем и учащимися для мониторинга знаний, навыков и отношений обучаемых.

Функции портфолио.

- Диагностическая – фиксирует изменения и рост за определенный период времени.
- Целеполагания – поддерживает учебные цели.
- Мотивационная – поощряет результаты учащихся, преподавателей и родителей.
- Содержательная – раскрывает весь спектр выполняемых работ.
- Развивающая – обеспечивает непрерывность процесса обучения от года к году.
- Рейтинговая – показывает диапазон навыков и умений.

Философия учебного портфолио предполагает:

- смещение акцента с того, что учащийся не знает и не умеет, на то, что он знает и умеет по данной теме и данному предмету;
- интеграцию количественной и качественной оценок;
- перенос педагогического удара с оценки педагога на самооценку школьника [3].

Основной смысл портфолио - дать возможность каждому школьнику показать все, на что он способен, создать для ученика стимул роста. Достоинства портфолио несомненны. Это прекрасное средство углубления и оформления познавательных интересов, развития интеллектуальных рефлексивных способностей учащихся, комплексной проверки уровня усвоения учебного материала, индивидуализации и дифференциации обучения, формирования мотивации достижения, а, следовательно, и создания ситуации успеха.

Цель портфолио для предпрофильной подготовки - выполнять роль индивидуальной накопительной оценки и наряду с результатами экзаменов определять рейтинг выпускников основной школы.

2. Типы портфолио, его структура

2.1. Типы портфолио.

2.2. Структура портфолио.

2.1. В зависимости от целей создания портфолио бывают разных типов.

Первый тип портфолио - это «папка достижений», направленная на повышение собственной значимости ученика, и отражающая его успехи (похвальные грамоты за учебу, достижения в спорте, музыке, шахматах и т. д.; благодарственные письма родителям, табели успеваемости, значки, медали и т. п.).

Второй тип - рефлексивное портфолио, раскрывающее динамику личностного развития ученика, помогающее отследить результативность его деятельности, как в количественном, так и качественном плане. В эту папку собираются все контрольные и творческие работы ученика, которые делались в течение определенного срока, например, года (сочинения, изложения, эссе, рисунки, поделки, зачетные работы, видеокассеты, результаты медицинских и психологических обследований и т. д.).

Третий тип портфолио - проблемно-исследовательский, связанный с написанием реферата, научной работы, подготовкой к выступлению на конференции. Он представляет собой набор материалов по определенным рубрикам (варианты реферата, доклада или статьи, список литературы для изучения, проблемные области, план исследования, дискуссионные точки зрения, статистика, цитаты, афоризмы, интеграция с другими предметными областями, результаты исследования, прогнозы и перспективы и др.). Сбор и систематизация материала в таких портфолио помогает учащимся не только достойно написать какую-либо научную работу, но и развивает их познавательные интересы.

Четвертый тип портфолио - тематический, создаваемый в процессе изучения какой-либо большой темы, раздела, учебного курса. Работа над ним строится следующим образом: учитель сообщает вначале название изучаемой темы, а также форму контроля по ней - защиту своего портфолио, собранного по результатам работы над данной темой.

Учащимся в самом начале предъявляются задания разного уровня сложности, отражающие различные уровни мышления и познания.

Для предпрофильной подготовки разработана другая классификация типов портфолио.

Предполагается, что портфолио достижений ученика будет складываться из трех разделов: «портфолио документов», «портфолио работ», «портфолио отзывов».

В портфолио документов входят сертифицированные (документированные) индивидуальные образовательные достижения - документы об участии в олимпиадах, конкурсах и других мероприятиях (грамоты, дипломы, сертификаты, свидетельства, вкладыш в аттестат и т. д.). Это дает возможность как количественной, так и качественной оценки материалов портфолио.

Портфолио работ - это собрание творческих, исследовательских и проектных работ ученика, описание основных форм и направлений его учебной и творческой активности.

Портфолио работ оформляется в виде творческой книжки с приложением самих работ (текстов, бумажных или электронных документов, видеозаписей и т.д.). Эта часть портфолио ученика дает качественную оценку по заданным параметрам (полнота, разнообразие, убедительность материалов, ориентированность на выбранный профиль обучения, динамика учебной и творческой активности, направленность интересов, характер предпрофильной подготовки).

Портфолио отзывов - это характеристики отношения школьника к различным видам деятельности, представленные учителями, родителями и др., а также письменный анализ отношения самого школьника к своей деятельности и ее результатам (тексты заключений, рецензий, отзывы, резюме, рекомендательные письма и прочее). Эта часть портфолио дает возможность включить механизмы самооценки ученика, что повышает степень осознанности процессов, связанных с обучением и выбором профильного направления.

2.2. Содержание портфолио зависит от его вида и цели, для достижения которой оно создается. Четкого списка материалов (называемых артефактами), необходимых для

включения в портфолио, не существует. Каждый из людей, использующих данную технологию, самостоятельно отбирает способы показа своих работ и достижений. Если портфолио создается всеми учащимися, то педагогу совместно с ними необходимо определить и зафиксировать перечень материалов, обязательных для представления в портфолио, с целью получения той или иной оценки. Например, к артефактам относятся копии заданий, контракты, результаты оценки компетентности на разных уровнях и этапах учебы, выполненные тесты и вопросники, примеры регулярной работы обучающегося (рефлексивные дневники). Можно использовать электронные портфолио, разработанные по установленным критериям. При этом приветствуется любая инициатива, расширяющая данный список или предлагающая равноценную замену перечисленным материалам.

Все материалы рекомендуется распределять по четырем разделам: портрет, коллектор, рабочие материалы и достижения.

3. Работа учителя с учащимися по составлению портфолио

3.1. Особенности работы педагога при использовании технологии портфолио.

3.2. Оценка портфолио.

3.1. Педагогу необходимо помнить, что портфолио создается обучающимися для оценки своего прогресса. Учитель лишь помогает им организовать этот процесс и оценить свои достижения. Следовательно, мотивация обучающихся в начале создания и в процессе работы над портфолио должна быть высокой. Для повышения уровня мотивации необходимо, чтобы и педагог и учащиеся творчески подошли к процессу. Выскажем несколько советов, помогающих повысить уровень мотивации.

А). Педагог сам должен понимать важность создания портфолио обучающимися для их собственного развития (а не только как внедрение новой, модной технологии).

Б). Необходимо провести урок, в ходе которого педагог не только объяснит, каким образом будет проходить работа учащихся и выскажет свои требования, но и обсудит с ними их взгляды на собственное развитие и то, как они видят место преподаваемого им предмета в данном процессе.

В). На вводном уроке очень важно рассказать содержание предмета на ближайшую четверть (семестр, полугодие, год). Только зная основное содержание предмета, сроки и формы контроля, учащиеся по-настоящему смогут поставить цели в разделе «Портрет». Содержание изучаемого материала, сроки и формы контроля, список основной и дополнительной литературы целесообразно представить визуально (в форме краткого текста или схемы) для того, чтобы обучающиеся в любой момент могли обратиться к этому разделу. При постановке учащимися целей необходимо проконтролировать, чтобы цели соответствовали критериям их эффективной постановки. Только в этом случае учащиеся смогут затем определить уровень реализации поставленной цели. При необходимости цель может быть разбита на частные задачи.

Г). Каждый обучающийся сам может выбрать форму создания портрета – эссе, рисунок, коллаж, схема – всего существует более десяти форм портрета. Самостоятельный выбор делает рефлексию более глубокой и повышает уровень творчества.

Д). В раздел «Коллектор» входят материалы, которыми пользуется учащийся при работе по предмету – учебники, список дополнительной литературы, образцы работ других учащихся, основные теоретические положения, глоссарий, статистические данные, другие источники информации. С одной стороны, нет необходимости перегружать данный раздел. С другой стороны, учитель должен помогать детям при составлении данного раздела, обучать их основам поиска информации, ее классификации, фиксации, конспектированию и т. д.

Е). Раздел «Рабочие материалы» включает обязательные и дополнительные работы самого обучающегося и именно он в основном подвергается анализу и оценке. При его создании необходимо соблюдать перечисленные ниже правила оценки. Если учащиеся выполняли

коллективную работу, например, проект, необходимо договориться о форме его краткого отражения в портфолио каждого из выполнявших.

Ж). Раздел «Достижения» чаще всего бывает самым сложным, так как педагоги не располагают разнообразием методов оценки достижений обучающихся. Кроме дневника и контрольной тетради в этот раздел включаются грамоты, сертификаты, похвальные листы, рефлексивный анализ, листы рейтингового контроля и др. Можно ввести систему взаимной оценки обучающихся, фиксировать ее и вносить в данный раздел, использовать характеристики (со стороны одноклассников, учителей, родителей) и другие формы и методы оценки достижений.

Оценка портфолио проводится учащимися совместно с учителем примерно один раз в четверть. Для того чтобы в ходе оценки у обучающихся развивались навыки самооценки и рефлексии, а также для экономии времени педагога, перед уроком по оценке портфолио каждый учащийся должен проделать подготовительную работу. В ходе ее ученик еще раз просматривает поставленные цели и задачи и подбирает материал, свидетельствующий о достижении поставленных задач или о работе в данном направлении. Можно разложить материал в соответствующем порядке или сделать закладки в тетради. Самый эффективный способ – письменный анализ достижений в соответствии с поставленными целями, с приложением рабочих материалов, необходимо предлагать именно его для работы обучающихся.

3.2. Оценка портфолио может проходить:

1. В индивидуальной встрече педагога и обучающихся. В этом случае выделяется время для индивидуальных встреч, или встречи организуются на уроке во время выполнения творческих заданий.

2. Учащимися в малых группах. Учитель заранее дает и разъясняет план работы микрогрупп, назначает лидера. Группа должна заслушать и обсудить каждого ученика и дать каждому рекомендации в течение 5-8 минут. Оценки и рекомендации фиксируются и помещаются в портфолио. Учитель в процессе такого урока перемещается из группы в группу. Если портфолио оценивается в баллах, то после такого урока учитель выставляет отметку на основе рекомендаций группы и содержания портфолио.

3. На основе выступлений учащихся на конференции по портфолио. Конференцию целесообразно организовывать в том случае, если портфолио собирались по различным темам (например, в предпрофильной подготовке или портфолио познавательного интереса). Каждый ученик имеет не более двух-трех минут на выступление и столько же - на ответы по вопросам. В ходе выступления необходимо отразить цели работы, их реализацию, основные достижения. Заранее организованная комиссия оценивает содержание портфолио, выступления и ответы на вопросы (по заранее определенным критериям).

Результатом обсуждения может быть программа дальнейшего углубления познавательного интереса ученика или, напротив, смена темы (проблемы), а соответственно и познавательного интереса. Педагог предлагает разные варианты и траектории личностного развития, а ученик выбирает.

Приведем пример критериев оценки портфолио учащихся, разработанный Е.В. Лебедевым [5].

1. Внутренняя оценка (самооценка) портфолио включает группы критериев:

- самоорганизация (ответственно подходить к делу, доводить начатое до конца, выбирать оптимальную структуру и содержание для представления себя),
- профессиональная ориентация,
- обучение (помогают ли материалы портфолио в процессе обучения, полезность навыков, приобретенных в процессе обучения, самооценка полученных знаний и умений),
- самообучение,
- самореализация.

2. Внешняя оценка включает следующие группы критериев:

- оценка качества оформления портфолио,
 - оценка структуры портфолио,
 - оценка уровня образовательной деятельности по материалам портфолио.
- Оценку портфолио можно заносить в специальный лист оценки.

Тема 6

ИСТОРИЯ РАЗВИТИЯ СИСТЕМЫ ТЕСТИРОВАНИЯ В РОССИИ И ЗА РУБЕЖОМ

1. Возникновение тестирования

1.1. Зарождение тестологии.

1.2. Тесты Дж. Кеттелла, А. Бине, Т. Симоны и др.

1.1. Время зарождения тестологии относят ко второй половине XIX века, когда психологи стали исследовать индивидуальные различия физических, физиологических и психических особенностей человека.

В середине XIX века особое внимание исследователей уделялось изучению умственной отсталости, которую в этот период впервые стали рассматривать как болезнь.

Французский врач Э. Сеген разработал собственную методику и основал первую школу для обучения умственно отсталых. Впоследствии многие из разработанных им приемов вошли в тесты на выявление уровня интеллекта.

Одним из первых, кто начал использовать тестовые технологии для измерения индивидуальных особенностей, стал английский биолог Френсис Гальтон. Он занимался вопросом наследственности, при этом разработал ряд методик для определения зрительной, слуховой и тактильной чувствительности, а также для определения мускульной силы, скорости реакции и др. В ходе сбора эмпирических данных Гальтон обследовал множество учреждений в системе образования с целью получить систематические измерения антропометрических характеристик учащихся. В 1884 году он организовал антропометрическую лабораторию на Всемирной выставке в Лондоне, где каждый желающий за небольшую плату мог измерить свои физические возможности по 17 показателям: рост, вес, сила кисти, сила удара, различения цвета, острота зрения и др. Таким образом, накапливались первые систематические данные по индивидуальным особенностям простых психофизических функций. По мнению Ф. Гальтона, тесты сенсорного различения можно использовать как средство оценки интеллекта человека. Ф. Гальтон был первым, кто применил оценочные шкалы, анкетирование и технику свободных ассоциаций.

Он сформулировал три принципа тестирования, эти выводы сохраняют актуальность до сих пор:

- 1) применение серии одинаковых испытаний к большому количеству испытуемых;
- 2) необходимость накопления и обработки статистических результатов;
- 3) установление эталонов оценки.

Важнейшим вкладом Ф. Гальтона в развитие тестологии является разработка и использование методов математической статистики для обработки полученных данных по индивидуальным различиям. Он вводит метод сравнения двух рядов переменных, для оценки этого соотношения используется особая величина – индекс-коэффициент корреляции. Он также изучает взаимосвязь переменных, для чего впервые использует построение линий регрессии одной переменной на другую.

1.2. Особенно заметный вклад в развитие тестирования внесли работы Джеймса Кеттелла (1860 - 1944). Американский психолог разработал около пятидесяти наборов заданий, которые назывались «умственными тестами», практически же это были тесты сенсорного различения и скорости реакции, которые, по мнению Дж. Кеттелла, можно было использовать для измерения интеллекта. Дж. Кеттелл рассматривал тестирование как

научный метод и выдвигал ряд требований к проведению тестирования, чтобы результаты были приближены к объективным.

Тесты Дж. Кеттелла были типичными среди большого количества серий тестов, разработанных в последнее десятилетие XIX века. Такие серии применялись для школьников, студентов и взрослых людей в Америке повсеместно. В основном они были направлены на измерение простых сенсорных и моторных процессов, хотя их авторы заявляли, что тесты предназначены для измерения интеллекта. Первая проверка этих тестов показала их слабую внутреннюю согласованность и то, что их результаты не соответствовали оценке интеллекта испытуемых независимыми экспертами.

В Европе в это время тесты, например, тесты Э. Крепелина и Г. Эббингауза, были более сложными и объективными.

Дж. Кеттелл активно пропагандировал новые методы измерения. В течение 1895 - 1896 гг. в Америке создаются два национальных комитета, с целью организации тестологов в области теоретических исследований и практики создания качественных тестов и их применения.

Новый этап в развитии тестологии связан с деятельностью французского психолога Альфреда Бине (1857 - 1911). Он разрабатывает оригинальные методики измерения интеллекта. А. Бине не устраивали попытки оценить уровень интеллекта через измерение элементарных психических процессов. Он идет по пути измерения сложных интеллектуальных функций. Работа в Комиссии по изучению методов обучения умственно отсталых детей с 1904 года дала возможность А. Бине реализовать свои идеи на практике. Совместно с Теодором Симоном А. Бине создает тестовые задания, направленные на дифференциацию детей способных к обучению, но ленивых или отстающих, и на детей умственно отсталых.

Шкала Бине-Симона (шкала 1905 года) состояла из 30 заданий, которые были расположены по возрастанию трудности. Уровень трудности был определен по результатам обследования 50 нормальных детей в возрасте от 3 до 11 лет и небольшого числа слабоумных детей. Тесты были рассчитаны на оценку способности к суждению, пониманию и рассуждению, которые, по мнению А. Бине, и являются основными компонентами интеллекта. Вероятность успешного выполнения определялась возрастающей трудностью заданий теста и увеличивалась в зависимости от возраста испытуемого.

В 1908 году появляется новый переработанный вариант шкалы: увеличивается количество заданий, убираются неудачные задания, расширяется выборка стандартизации. Бине и Симон заявляют и новую цель шкалы: теперь это не только дифференциация детей на нормальных и умственно отсталых, но и выделение среди нормальных детей различных возрастных уровней интеллектуального развития. Группировка тестов по возрастным уровням позволила определить нормы для детей разных возрастных категорий. Дальнейшая модификация шкалы была сосредоточена на ее стандартизации и определении ее валидности.

Тесты Бине-Симона привлекли внимание психологов разных стран. Они активно переводились и адаптировались. В Америке появилось много переработанных вариантов тестов Бине-Симона. Одним из удачных вариантов признан тест, разработанный Левисом Мадисоном Терменом (Странфордская шкала интеллекта Бине). В этом варианте был впервые использован коэффициент интеллектуальности - показатель умственного развития (IQ).

Активная работа по созданию, совершенствованию и применению тестов сопровождалась разработкой статистических методов обработки полученных результатов (К. Пирсон, Ч. Спирмен).

В начале возникновения и становления тест как измерительный инструмент использовался только в рамках эксперимента, был предназначен исключительно для индивидуального измерения. Групповые тесты впервые появились в США в 1917 году,

когда Америка вступила в Первую мировую войну, и возникла необходимость быстро определить интеллектуальный уровень полутора миллионов вступающих в ряды армии. С этой целью использовались тесты интеллекта Артура Синтона Отиса. Тесты Отиса включали два вида: альфа-вербальный тест для владеющих английским языком и бета-тест – невербальный, рассчитанный на неграмотных и новобранцев иностранного происхождения. По окончании Первой мировой войны после некоторой переработки эти тесты стали повсеместно использоваться как групповые тесты интеллекта в образовательных учреждениях, среди взрослого населения.

В 1915 г. американец Р.М. Йеркс предлагает новую систему подсчета результатов испытуемых. Он вводит балльную систему (за правильное решенное задание теста испытуемый получает определенное количество баллов) вместо возрастных долей, которые использовались А.Бине. Полученное количество баллов затем переводилось в коэффициент одаренности или успешности в соответствии с разработанными стандартами [3].

2. Психологические и педагогические тесты

В начале XX века также возникает идея использования тестирования для измерения уровня учебных достижений. Американский психолог В.А.Макколл предлагает разделить тесты на психологические (определение уровня умственного развития) и педагогические (измерение успешности учащихся по предметам за определенный период обучения). Целью педагогического тестирования, по мысли Макколла, должно быть выделение и объединение учащихся с близкими показателями уровня обученности.

Основоположником педагогических измерений считается американский психолог Эдуард Ли Торндайк, который создал первый педагогический тест (ability tests). Первые тесты учебных достижений представляли собой тесты на решение арифметических задач, правописание, оценку почерка, рассуждение. Свои выводы по использованию тестовых методов в педагогике Торндайк обобщает в книге «Введение в теорию психологии и социальных измерений» (1904).

3. Тестология в России

В России интерес к разработке тестов и практике их использования формируется в 20-х г. XX века. Видные российские психологи и педагоги занимались этой проблемой, среди них С.Г. Геллерштейн, П.П. Блонский, А.П. Болтунов, М.С. Бернштейн, Л.С. Выготский, Г.И. Залкинд и др. В книге С.М.Василейского «Введение в теорию и технику психологического, педологического и психотехнического исследования» (1927) изложены основные теоретические и практические принципы создания и использования тестов. В работе также были освещены вопросы статистической обработки результатов, проведения тестирования, методы изучения различных профессий. В 1925 году при педагогическом отделе Института методов школьной работы была создана особая Тестовая комиссия, которая способствовала организации деятельности педагогов и психологов по разработке тестов для советской школы. На основе американских были разработаны тесты по счету, решению задач, на понимание чтения, на правописание, на знание географической карты. Тесты сопровождались инструкциями по его проведению и выполнению заданий, а также личными карточками учета роста учебных достижений учащегося.

Под руководством П.П.Блонского осуществлялась проверка шкалы Бине-Симона, создание школьных тестов и построение стандартов. В Ленинградском научно-педагогическом институте им. А.И.Герцена и Детском обследовательском институте им. А.С. Грибоедова разрабатывались тесты диагностики одаренности и тесты школьной успешности для массовых обследований детей нормальных школ.

Постановление ЦКВКП(б) «О педологических извращениях в системе Наркомпроса» (1936) запрещает использование тестов в педагогической практике советской школы.

Тестология как наука и тестирование как метод измерения также прекращает свое существование в нашей стране на долгие годы.

Только в 60-е годы в дидактику возвращаются тесты (через систему программированного обучения). Однако следует отметить, что и в 40-60 годы тесты использовались, но в виде «контрольных заданий», «самостоятельных работ», упражнений и т. п.

В 1970-1980 гг. резко критиковалась традиционная система контроля знаний. В бывшем Советском Союзе развитие дидактической тестологии связано с работами Н.Ф. Талызиной по программированному обучению и В.П. Беспалько по проблемам педагогической технологии.

Отмечая, что учебные методы контроля пригодны для непосредственного общения учителя со школьниками на уроке по конкретным изучаемым на занятии вопросам, В.П. Беспалько пишет, что для оценки качества знаний эти методы контроля не годятся, так как необходимые диагностичность, точность и воспроизводимость результатов в них не заложены. Аналогичными недостатками страдают и обычные письменные контрольные работы. Поэтому необходима специальная разработка материалов тестов, ориентированных на уровни решения тех или иных учебных задач. Сейчас такой подход разделяется многими педагогами.

В настоящее время в нашей стране появилось несколько центров, в которых профессионально занимаются работой с тестовыми технологиями. Среди наиболее активных следует отметить Центр тестирования Министерства образования Российской Федерации, занимающийся тестированием выпускников общеобразовательных учреждений, Центр тестирования профессионального образования, Центр психологического и профессионального тестирования МГУ и целый ряд других.

4. Современное развитие тестологии

В 10-20 гг. XX века тестирование выходит за рамки научного эксперимента и превращается в массовое явление. Для упорядочения работы по разработке и использованию тестов создаются специальные организации и службы государственного уровня. В 1926 г. Совет колледжей США принял тест SAT и утвердил наборы заданий для профессиональной оценки деятельности педагогов.

В 1947 г. в США создается Служба тестирования в образовании (Educational Testing Service), которая является и научно-исследовательским центром, и организацией, контролирующей проведение тестирования на всех уровнях образования в стране.

К 1961 году только в США были созданы 2126 стандартизированных тестов. Здесь существует множество тестовых издательств, общая продукция которых представляет несколько тысяч различных тестов.

Возникло новое направление педагогического тестирования - международные сравнительные исследования, с целью оценки качества образования различных стран. Проведение международных исследований предоставляет странам возможность всесторонней оценки эффективности своих образовательных программ и сравнения подготовки своих учащихся с международными стандартами. Лидирующая роль в проведении подобных исследований принадлежит Международной ассоциации по оценке учебных достижений (IEA - International Association for the Educational Achievement) - независимой организации, объединяющей научные центры более 50 стран мира, и проводящей сравнительные исследования по оценке эффективности образовательных систем стран мира, осуществляющих различную политику в области образования. Первые исследования IEA проводились совместно с ЮНЕСКО. За 35 лет существования, начиная с 1959 года, IEA проведено более 15 исследований. Во многих странах мира результаты исследований IEA дали толчок или явились основой для проведения реформ в области образования.

Россия стала членом IEA в 1991 г. и с этого времени принимает активное участие во всех ее исследованиях (TIMSS - третьем международном исследовании по естественно-

математическому образованию, LES - исследовании по иностранным языкам и CIVIC - исследовании по социальным наукам).

Во второй половине XX века разрабатывается более совершенная модель анализа полученных результатов - современная теория тестов (IRT). В этот период теория тестов из описательной науки начинает превращаться в науку об отношениях, о законах функционирования и оценивания поведения испытуемых. Параллельно разрабатывались компьютерные программы для обработки тестовых результатов с помощью моделей IRT.

В середине 70-х гг. XX века разработаны новые критерии, предъявляемые к тесту как способу оценки качества:

- адаптивность систем тестирования к качеству выводов испытуемых;
- объективность педагогических измерений;
- экономия ресурсов и времени проверок;
- ограничение места для записи тестируемых заключений;
- возможность применения различных форм предъявления тестовых заданий. В

изложении темы использованы материалы курса лекций Ю.А. Воронина и др. [5].

Тема 7

ПСИХОЛОГИЧЕСКИЕ ТЕСТЫ: ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ. ПСИХОЛОГО-ПЕДАГОГИЧЕСКИЕ АСПЕКТЫ ТЕСТИРОВАНИЯ

1. Понятийный аппарат тестологии

1.1. Понятие теста.

1.2. Надёжность и валидность теста.

1.3. Психологические аспекты тестирования.

1.1. Отечественная педагогическая тестология - очень молодая наука. В настоящий момент в различных изданиях по теории и практике тестирования можно встретить различное толкование и определение одних и тех же понятий, для обозначения одного явления может использоваться несколько синонимических терминов, основная группа терминов имеет иноязычное происхождение.

Термин «тест» (от англ. test - испытание, исследование) имеет в русском языке несколько значений (см. словарь иностранных слов). Но нас интересует только значение, которое используется в тестологии.

Существуют различные точки зрения при определении понятия «тест».

1) **Тест** - это инструмент, состоящий из квалитетрически выверенной системы тестовых заданий, стандартизированной процедуры проведения и заранее спроектированной технологии и анализа результатов для измерения качеств и свойств личности, учебных достижений, изменение которых возможно в процессе систематического обучения [8].

2) В.С. Аванесов предлагает различие теста как метода и теста как инструмента.

Тест как метод предполагает технологию измерения, которая включает в себя разработку системы тестовых заданий с заданными качественными и количественными характеристиками для объективного и надежного оценивания учебных достижений испытуемых, стандартизированную процедуру проведения тестирования, методы статистической обработки, анализа и интерпретации полученных результатов.

Тест как инструмент измерения определяется как система заданий (в большинстве случаев возрастающей трудности) специфической формы, позволяющая качественно оценить структуру и эффективно измерить уровень знаний, умений и навыков учащихся [1].

Слово «тест» в первом значении имеет научный аспект и используется учеными-тестологами, которые знают обо всех этапах разработки и применения теста, а также о теоретическом обосновании выбранных способов проверки и обработки полученных

результатов в соответствии с классической теорией тестов или современной теорией тестов.

Второе значение чаще возникает в практической области: в сознании тех, кто его использует или испытуемых, которые выполняют тест. Чаще используется именно второе значение. Но при этом надо помнить, что тест как измерительный инструмент является частью научного метода измерения (тестирования, теста) и соответственно должен отвечать ряду требований по его разработке, использованию и оцениванию результатов, которые определяются научными подходами.

1.2. Отличие тестов от других форм контроля.

1. Содержание теста подвергается четкому планированию. На стадии разработки теста происходит отбор содержания, которое будет подвергаться проверке, планируется форма заданий, их количество и расположение.

2. Форма заданий. В тестах форма заданий стандартизирована - по форме предъявления и по форме записи ответов.

3. Наличие статистических характеристик у тестовых заданий. Заранее известно, какова трудность предлагаемого задания, будет ли оно одинаково выполняться слабыми и сильными испытуемыми или нет (дифференцирующая способность) и др.

4. Наличие специальных шкал, которые соотнесены со стандартизированными нормами для подведения результатов тестирования.

5. Наличие оценок точности измерения (ошибки измерения). С помощью статистических методов мы можем оценить ошибку измерения, а по результатам оценки принять или не принять результаты тестирования.

Отличительные особенности теста определяют преимущества теста перед традиционными формами контроля учебных достижений: объективность, надежность, точность, экономичность измерений.

1.3. Тест, предъявляемый испытуемому, состоит из инструкции и тестовых заданий. В инструкции даются указания, из скольких частей состоит тест, какое количество времени дается на его выполнение, какой стратегии должен придерживаться испытуемый (например, если не знаете ответ на задание, приступайте к выполнению следующего), что надо сделать, чтобы записать правильный ответ. Если тест включает различные формы заданий, то при смене форм, перед каждым субтестом дается дополнительная инструкция по выполнению данной формы задания.

Далее, после инструкции располагаются пронумерованные тестовые задания. Задания в тесте (субтесте), в соответствии с теорией тестирования, должны располагаться по нарастающей трудности, то есть в начало теста включаются легкие задания, потом более сложные.

Тест всегда сопровождается подробной инструкцией для организаторов и наблюдателей тестирования. Данная инструкция должна обеспечить равные условия для всех, кто выполняет тест. Поэтому в этой инструкции четко оговариваются условия тестирования (требования, предъявляемые к помещению, где будет проходить тестирование, количество испытуемых в группе, время тестирования), форма заполнения тестовых бланков (данные испытуемого: имя, возраст, пол, школа и др.; как заносить правильные ответы и т. д.).

2. Надёжность и валидность теста

Основными показателями качества теста являются надежность и валидность теста. А так как тест является инструментом педагогического измерения, то эти показатели имеют численное значение и определяются с помощью различных методов как в классической теории тестов, так и в современной теории тестов.

Надежность - это характеристика теста, отражающая точность педагогического измерения, а также устойчивость результатов тестирования к воздействию посторонних (случайных) факторов.

Для оценки надежности нормативно-ориентированного теста используются две группы методов.

1. Двукратное тестирование:

- ретестовый метод;
- метод параллельных форм.

2. Однократное тестирование (метод расщепления теста).

Оценка надежности чаще всего строится на подсчете корреляции двух наборов результатов выполнения одного и того же теста или двух его параллельных форм. Чем выше корреляция, тем надежнее тест.

Ретестовый метод.

Данный метод оценки надежности предполагает двукратное проведение одного и того же теста в одной группе испытуемых. Повторное тестирование проводится примерно через 2-3 недели, чтобы учащиеся не слишком сильно забыли проверяемый материал и недалеко продвинулись в изучении нового, то есть уровень их учебных достижений остался примерно прежним.

Ретестовый метод основан на подсчете корреляции индивидуальных баллов испытуемых по результатам выполнения первого и второго тестирования.

Коэффициент надежности (коэффициент корреляции Пирсона):

-

Хорошим коэффициентом надежности теста считается, когда показатель колеблется в пределах $0,8 < K < 1$.

Данный метод оценки надежности не очень удобен в педагогических измерениях. Временной фактор играет большую роль: временной промежуток не может быть большим - изменится уровень учебных достижений учащихся, но и не должен быть слишком коротким - испытуемые могут помнить задания теста и ответы на них. И в том и в другом случае результаты измерения нельзя использовать для определения надежности теста.

Метод параллельных форм.

Метод параллельных форм предполагает двукратное тестирование одной и той же группы испытуемых тестами, которые идентичны по содержанию, структуре, включают задания тождественные по трудности, дифференцирующей способности и др., то есть параллельными формами теста.

Коэффициент надежности:

x_i – индивидуальный балл i -ого студента при ответе на тест X

y_i – индивидуальный балл i -ого студента при ответе на тест Y

N – число студентов

Главная трудность применения этого метода в том, что прежде чем использовать параллельную форму теста, разработчику надо создать новый тест, тождественный первому, и доказать идентичность, а это очень трудоемкая процедура. Поэтому этот метод, с практической точки зрения, малоэффективен.

Однократное тестирование (метод расщепления теста).

Этот метод удобен в практическом применении, так как ограничивается однократным тестированием. Метод расщепления теста основан на допущении параллельности двух половин теста и предполагает деление результатов тестирования на две части: данные по нечетным заданиям теста (X) и по четным (Y). Корреляция двух половин тестов возрастает по мере роста однородности (гомогенности) теста. Поэтому метод расщепления теста называют методом оценки внутренней согласованности и используют для определения гомогенности теста.

Коэффициент надежности вычисляется по формуле:

, где результат первого тестирования – это результаты по нечетным заданиям, а второго – по четным заданиям.

Так как подсчет надежности (K) ведется по расщепленному тесту, который в два раза короче, то оценка надежности исходного корректируется по формуле Спирмена-Брауна

В практике педагогических измерений часто используется способ оценки надежности с применением формулы Кьюдера-Ричардсона. Она может применяться только в том случае, когда выполнение задания оценивается дихотомически (1 балл - правильно; 0 баллов - неправильно).

Рассмотренные выше методы определения надежности подходят только для оценки нормативно-ориентированных тестов. По мнению многих тестологов, эти методы нежелательно использовать для вычисления надежности критериально-ориентированного теста, так как дисперсия тестовых баллов в критериально-ориентированном тесте небольшая (здесь не нужна большая дифференциация баллов испытуемых), а соответственно и корреляционная оценка надежности будет низкой.

Поэтому предлагается другая методика оценки надежности. Надежность теста, ориентированного на область содержания, может быть измерена как постоянство результатов тестирования для испытуемых, которым предлагается два набора тестовых заданий, которые соответствуют одному и тому же содержанию.

Практически это выглядит следующим образом. Группа испытуемых выполняет два идентичных теста или два теста объединяются в один, при этом вопросы перемежаются: (1 вариант – 1 вопрос, 2 вариант – 1 вопрос, 1 вариант – 2 вопроса, 2 вариант – 2 вопроса и т. д.). По итогам тестирования испытуемые делятся на две группы: получившие зачет и не получившие зачет по двум тестам (по частям расщепленного теста). Далее проводится сравнение результатов тестирования по двум вариантам или, во втором случае, по вариантам расщепленного теста. По результатам теста составляется таблица сопряжения. Численный показатель коэффициента надежности критериально-ориентированного теста вычисляется с помощью γ -коэффициента корреляции.

Вычисляется также γ -коэффициент как показатель критериально-ориентированного теста по формуле:

$$\gamma = (P - P_c) / (1 - P_c),$$

где $P = A + D$ - доля испытуемых, получивших «зачет» и «незачет» и в первой, и во второй форме теста; P_c - вероятность случайного принятия согласованного решения.

P_c вычисляется по формуле:

$$P_c = (c + d)(d + b) + (a + b)(c + a)$$

Каппа-коэффициент может принимать значения от -1 до 1. Если создается тест для аттестации по итогам определенной ступени обучения, то рекомендуется вычислять γ - и γ -коэффициенты. Их величины должны быть больше 0,8 (0,85) и приблизительно равны между собой, что свидетельствует о хорошей надежности теста.

Валидность (от англ. Validity - значимость, обоснованность, пригодность) - это характеристика способности теста служить поставленной цели измерения, то есть способность теста измерять то, для чего он предназначен.

Выделяют несколько видов валидности, которые отражают различные аспекты этого критерия качества теста (А. Анастаси; А.Н. Майоров): содержательную, критериальную и конструктивную (концептуальную) валидность [2].

1. Содержательная валидность (content validity) - соответствие теста как измерительного инструмента той области содержания, знания и умения, которую проверяют данным тестом. В зависимости от выбранного подхода при создании теста предъявляются различные требования к содержательной валидности. Для критериально-ориентированных тестов одно из требований содержательной валидности - полнота отображения, а в некоторых случаях – детализация проверяемой области содержания. При нормативно-ориентированном подходе тест можно охарактеризовать как валидный по содержанию, если он будет хорошо дифференцировать испытуемых и отражать самые важные характеристики этого теста.

Валидность по содержанию закладывается в тест уже на этапе работы со спецификацией теста и отбора его содержания. Для педагогических тестов содержательная валидность очень важна и оценивается экспертным путем.

2. Критериальная валидность - это характеристика теста, отражающая обоснованность, значимость его результатов, по сравнению с некоторой внешней переменной (степень соответствия между результатами тестирования и внешним критерием).

Критериальную валидность определяют как корреляцию результатов тестирования с некоторым внешним критерием. В качестве внешнего критерия могут браться результаты испытуемых по другому тесту, отметки, выставленные по методике традиционного оценивания. Критериальная валидность, в зависимости от выбранного критерия, может быть текущей (внешний критерий - текущие оценки) или прогностической (критерий - некоторый будущий результат).

Для оценки критериальной валидности необходимо составить таблицу с результатами тестирования и список тех же студентов с экспертной оценкой преподавателя (-ей). В качестве критерия берем оценки преподавателей при традиционной проверке знаний без использования тестов.

Формула:

s_x^2 – дисперсия индивидуальных баллов,

N – число студентов.

Критериальная валидность измеряется от -1 до 1, валидность более 0,3 считается удовлетворительной.

3. Конструктивная (концептуальная) валидность.

Об этом виде валидности говорится в том случае, если представление об измеряемом феномене существует только в проекте и требует доказательства. Разработчик теста только предполагает о том, как будет проявляться объект измерения. Чаще всего это относится к психологическим тестам. Для педагогических тестов понятие конструктивной валидности применимо в процессе разработки теста, когда его качественные и количественные характеристики еще не имеют определенных характеристик. Для выявления валидности на этапе создания теста используются методы корреляционного и факторного анализа.

3. Психологические аспекты тестирования

К человеку, проводящему тестирование, предъявляются определенные требования.

1. Отсутствие влияния экспериментатора (педагога) на результаты тестирования.

Повлиять на результаты можно различными способами. Часто педагоги даже не подозревают об этом влиянии. Например, повлиять может интонация, с которой читаются вопросы или варианты ответов (когда неосознанно выделяются голосом или произносятся в ином темпе правильные варианты), невербальные проявления (кивки головой, жесты и др.). Для того чтобы избежать влияния экспериментатора необходимо соблюдать следующие требования:

- тексты заданий или вариантов ответов не читаются вслух, а даются испытуемым в печатном виде,
- шрифт всех заданий и вариантов ответов должен быть одинаков,
- экспериментатор не отвечает на вопросы, не объясняет содержания заданий, не комментирует варианты ответов и задания.

2. Соблюдение одинаковых для всех условий тестирования.

Все тестируемые должны проходить тестирование в одинаковых условиях. Для этого необходимо четко соблюдать инструкцию для испытуемых и инструкцию для экспериментатора. Соответственно, в этих инструкциях должны быть четко указаны все необходимые для тестирования условия (сроки проведения, требования к помещению и оборудованию, требования к уровню квалификации экспериментатора и др.).

3. Объяснение результатов тестирования испытуемым.

Испытуемым необходимо предоставить информацию о полученных результатах, при этом рассказать о результатах необходимо, с одной стороны, объективно, с другой стороны, таким образом, чтобы испытуемый полностью понял всю информацию.

4. Соблюдение этических требований.

Существует несколько профессиональных этических принципов, обязательных для представителей всех профессий, работающих с людьми. Например, международный этический стандарт психологов включает в себя принципы ответственности (за свою работу и принимаемые решения), компетентности, этической и юридической правомочности, квалифицированной пропаганды профессии психолога, конфиденциальности (сохранения профессиональной тайны), благополучия клиента, информирования клиента о целях обследования, морально-позитивного эффекта исследования, гражданственности и патриотизма.

Тема 8

ВИДЫ ТЕСТОВ И ФОРМЫ ТЕСТОВЫХ ЗАДАНИЙ

1. Классификация педагогических тестов

По целям использования выделяют (классификация предложена Гронлундом, ее же придерживаются М.Б. Челышкова, А.Н. Майоров):

- 1) *входное тестирование* (обеспечивает проверку знаний и умений в начале обучения);
- 2) *формирующее и диагностическое тестирование* (предполагает контроль формирования новых знаний и умений в процессе обучения, обеспечивает текущий контроль);
- 3) *тематическое, итоговое, рубежное тестирование* обеспечивает определение итоговых достижений [16].

Два подхода к разработке тестов

- 1) **нормативно-ориентированный,**
- 2) **критериально-ориентированный.**

Соответственно различают два *вида тестов*:

- *нормативно-ориентированный* (позволяет сравнивать уровень учебных достижений испытуемых друг с другом);
- *критериально-ориентированный* (служит для определения уровня освоения учащимся содержания какого-либо предмета, дисциплины, раздела, темы) [5].

Таблица 4

Сравнительный анализ нормативно-ориентированных и критериально-ориентированных тестов

	НОПТ	КОПТ
Цель тестирования	Сравнение уровня достижений испытуемых друг с другом	Оценка уровня достижений в соответствии со степенью освоения знаний и умений определенной области содержания
Критерий оценки	Стандартизированные нормы	Достаточная доля (%) проверяемого материала
Область	Задания на проверку	Для более полного охвата

содержания	наиболее значимых элементов содержания	содержания составляется детальная спецификация теста (план), отражающая все элементы содержания
Сфера применения	Конкурсные отборы (вступительные экзамены в вузы)	Оценка итогового уровня обученности (начальное, среднее, высшее образование); оценка эффективности программы обучения
Характеристика тестовых заданий	Задания должны быть среднего уровня трудности (от 0,3 до 0,7) и высокой дифференцирующей способности (больше 0,3)	Задания должны соответствовать спецификации и элементам проверяемого содержания. Задания могут быть различного уровня трудности.

Схема 2

Классификация видов тестов

По средствам предъявления педагогические тесты делятся на:

- 1) *бланочные* (испытуемые отмечают или вписывают правильные ответы на бланке);
- 2) *компьютерные* (задания высвечиваются на мониторе компьютера).

По степени однородности задач:

- 1) *гомогенные* (контроль знаний и умений по одному предмету, дисциплине);
- 2) *гетерогенные* (измерение уровня подготовленности по нескольким учебным предметам).
 - а) *полидисциплинарный* (проверка знаний по нескольким дисциплинам)
 - б) *междисциплинарный* (проверка знаний и умений в смежных дисциплинах (математика-физика, химия-биология, история-культурология)).

Типы тестов

I. Открытые (испытуемый сам записывает правильный ответ)

Формы тестовых заданий открытого типа

- а) **задания на дополнение** (краткий ответ: формула, числовое выражение, слово и др.).
- б) **задания со свободно конструируемыми ответами** (ответ в свободной, развернутой форме: решение задачи с пояснениями, небольшое сочинение. Должна оговариваться полнота ответа для получения максимального балла).

II. Закрытые (предполагают выбор правильного ответа из предложенных вариантов. Оптимальное количество альтернативных ответов - 4-5. Неправильные, но правдоподобные ответы, то есть похожие на правильные, называются **дистракторами** (от англ. distract - отвлекать).

Формы тестовых заданий закрытого типа

- а) **задания альтернативных ответов** (задания с двумя ответами: да - нет, правильно - неправильно). Имеют очень высокий процент угадывания (50%), их одиночное использование в тесте считается неэффективным;
- б) **задания с выбором нескольких правильных ответов;**
- в) **задания на установление соответствия;**

г) задания на установление правильной последовательности.

Требования к заданиям на дополнение.

1. Нацеленность каждого задания лишь на одно дополнение, место дополнения обозначается прочерком.
2. Постановка дополнения в конце задания.
3. Равная длина всех прочерков в тексте.
4. Указание единиц измерения после прочерка (при необходимости).
5. Исключение повторов и двойного отрицания в тексте задания.

Требования к заданиям с выбором ответов.

1. Краткая и четкая формулировка задания и вариантов ответа, исключение двусмысленности.
2. Наличие основной, полной информации в задании к тесту, избегание повторяющихся слов и выражений в вариантах ответа.
3. Соблюдение правил грамматики и сочетаемости при формулировании текста задания и ответов.
4. Исключение двойного отрицания.
5. Наличие только одного правильного ответа.
6. Избегание слов: «все», «ни одного», «никогда», «всегда» и выражений «все перечисленные», «ни один из перечисленных» (могут способствовать угадыванию правильного ответа).
7. Одинаковая привлекательность дистракторов, их связь с содержанием тестируемого раздела.
8. Формулирование задания и вариантов ответа исключающих наличие ключа для других заданий.
9. Исключение заданий, выполнение которых основывается на субъективном мнении испытуемого.

Требования к заданиям на установление соответствий.

1. Выбор элементов двух множеств по одному основанию.
2. Размещение в тексте задания названий каждого столбца (обобщающее определение множества).
3. Наличие нескольких правдоподобных дистракторов в правом списке.
4. Указание в инструкции к заданию о наличии дистракторов и о количестве возможных использований элементов второго списка.
5. Число элементов одного списка не должно превышать 10.
6. Расположение всех заданий на одной странице, недопущение переноса отдельных элементов. В изложении темы использованы материалы курса лекций Ю.А. Воронина и др. [3].

Тема 9 КОМПЬЮТЕРНОЕ ТЕСТИРОВАНИЕ

Компьютерное тестирование – разновидность тестирования с использованием современных технических средств, имеющее ряд преимуществ по сравнению с традиционным бланчным тестированием (получение мгновенного результата, исключение предвзятости, нормирование трудности и объёма тестовых заданий, массовость, лёгкость обработки результатов, возможность тестирующих программ работать в режиме обучения).

Функции компьютерного теста

- обучение (в качестве дидактического обучающего средства);
- контроль ЗУНов учащихся;
- самообучение (тренажёр) и самоконтроль;
- дистанционное обучение;
- корректировка учебного процесса.

Инструментальные тестовые оболочки

Для создания тестов по предметной области разрабатываются специальные инструментальные **программы-оболочки**, позволяющие создавать компьютерные тесты путем формирования базы данных из набора тестовых заданий.

Классы программ-оболочек

1. Универсальные (содержат тестовую оболочку как составную часть. Примеры: «Адонис» (Москва), «Linkway» (Microsoft), «Фея» (Томск), «Радуга» (Москва) и т. п.
2. Специализированные (предназначены лишь для формирования тестов. Примеры: «Аист» (Москва), «I_пов» (Иркутск), «Тест» (Красноярск) и др.

Для того чтобы разработать компьютерный вариант теста с помощью одной из названных выше программ, необходимо уяснить, какие формы тестовых заданий они допускают.

Показатели качественного компьютерного теста

- Восприимчивость к угадыванию тестируемым;
- восприимчивость к невнимательности и ошибочным действиям тестируемого;
- положительное влияние на тестируемого и педагога, который использует тест.

Компьютерные формы представления тестовых заданий

1. Альтернативные вопросы (требуют ответа да – нет).
Закрытая форма вопроса: номер правильного ответа.
2. Задание на соответствие: требует соотнести элементы двух множеств.
3. Задание на восстановление правильной последовательности: требует расположить элементы множества в определённом порядке.
4. Вопросы с шаблоном ответа (предполагают числовой или текстовой ввод правильного ответа).
5. Вопросы с выбором нескольких вариантов (ключевых слов, изображений, обозначений).
6. Конструирование ответа (шаблонный и бесшаблонный варианты): ответ формируется путем последовательного выбора его элементов из инструментария по типу меню.
7. Задание на конструирование изображений: с помощью графредактора, меню изображений (аналогично предыдущему примеру).
8. Задание на демонстрацию с движущимися объектами. Ответ в виде действия тестируемого (определенный набор клавиш). (Пример: клавиатурный тренажер на время.)

Выбор формы представления задания определяется

- особенностями инструментальных тестовых программ (тестовыми оболочками);
- особенностями предметной области;
- опытом и мастерством экспертов.

Рекомендации по разработке вопросов к компьютерному тесту

1. При выборе того или иного метода ввода ответов следует исходить в первую очередь из специфики учебного материала.
2. Выборочный метод (вопросы с выбором ответов) целесообразно применять только в тех случаях, когда характер учебного материала не позволяет применять другие методы ввода ответов.
3. Наибольшую объективность проверки знаний обеспечивает результативный метод ввода ответов. Его желательно применять всегда, когда позволяет характер материала.
4. Методы с обозначенным порядком ввода ответов относительно сложны, и их целесообразно использовать только для проверки знаний подготовленных соответствующим образом проверяемых.
5. Вопросы лучше формулировать в форме заданий.
6. При постановке вопросов и формулировании вариантов ответов необходимо избегать двусмысленностей, двояких толкований, вопросов в отрицательной форме, формулировок, несущих в себе элемент подсказки, а также явно абсурдных.

Готовность к компьютерному тестированию

Под *общей готовностью* к компьютерному тестированию понимается интегративное личностное качество. Становление общей готовности осуществляется в процессе целенаправленной, специально организованной деятельности по подготовке.

Содержание готовности к компьютерному тестированию:

- мотивационный компонент,
- когнитивный компонент,
- процессуальный компонент.

Мотивационный компонент готовности к компьютерному тестированию включает:

- положительное отношение к компьютерному тестированию;
- четкое и полное осознание значения компьютерного тестирования для системы образования и себя лично;
- понимание роли компьютерного тестирования на современном этапе развития системы образования;
- осознание необходимости и важности овладения основами теории и практики компьютерного тестирования;
- интерес к компьютерному тестированию, его освоению.

Когнитивный компонент представляет собой знание:

- значений понятий: «компьютерное тестирование», «педагогический тест», «компьютерный тест», «тестовое задание»;
- возможностей и ограничений компьютерного тестирования;
- форм тестовых заданий;
- этапов и форм компьютерного тестирования.

Процессуальный компонент включает следующие умения:

- адекватно воспринимать интерфейс компьютерного теста;
- адекватно относиться к ситуации компьютерного тестирования;
- отвечать на тестовые задания различных форм [15].

Технология проектирования компьютерных тестов предметной области

Проектирование модульной модели знаний

Метод нисходящего проектирования модели знаний состоит в следующем: вначале строится генеральное содержание предметной области с разбивкой на укрупненные модули (разделы). Затем проводится детализация модулей на элементарные подмодули, которые, в свою очередь, наполняются педагогическим содержанием .

Каждый модуль предполагает входящую информацию, состоящую из набора необходимых понятий из других модулей и предметных областей, а на выходе создает совокупность новых понятий, знаний, описанных в данном модуле, рис. 1.

Рисунок 1.

Модуль может содержать подмодули. Элементарный подмодуль - неделимый элемент знания - может быть представлен в виде базы данных, базы знаний, информационной модели. Понятия и отношения между ними представляют семантический граф (рис. 2).

Рисунок 2.

Пример элемента модуля знаний по теме «Исследование графиков функций», рис. 3.

Рисунок 3.

Модульное представление знаний помогает:

- организовывать четкую систему контроля с помощью компьютерного тестирования, поскольку допускает промежуточный контроль (тестирование) каждого модуля, итоговый контроль по всем модулям и их взаимосвязям;
- осуществлять наполнение каждого модуля педагогическим содержанием;
- выявлять и учитывать семантические связи модулей и их отношения с другими предметными областями [9].

Основные понятия компьютерного тестирования

Тестирование - процесс оценки соответствия личностной модели знаний ученика экспертной модели знаний. Главная цель тестирования - обнаружение несоответствия этих моделей (а не измерение уровня знаний), оценка уровня их несоответствия. Тестирование проводится с помощью специальных тестов, состоящих из заданного набора тестовых заданий.

Тестовое задание - это четкое и ясное задание по предметной области, требующее однозначного ответа или выполнения определенного алгоритма действий.

Тест - набор взаимосвязанных тестовых заданий, позволяющих оценить соответствие знаний ученика экспертной модели знаний предметной области.

Тестовое пространство - множество тестовых заданий по всем модулям экспертной модели знаний.

Класс эквивалентности - множество тестовых заданий, таких, что выполнение учеником одного из них гарантирует выполнение других.

Полный тест - подмножество тестового пространства, обеспечивающее объективную оценку соответствия между личностной моделью и экспертной моделью знаний.

Эффективный тест - оптимальный по объему полный тест.

Эффективный компьютерный тест оценивает уровень соответствия или несоответствия *личностной модели знаний ученика и экспертной модели*.

Множество тестовых заданий (*тестовое пространство*), согласно принципу исчерпывающего тестирования, может быть бесконечным. Однако в каждом реальном случае существует конечное подмножество тестовых заданий, использование которых позволяет с большой вероятностной точностью оценить соответствие знаний ученика заданным критериям по экспертной модели знаний (*полный тест*).

Из полного теста можно выделить *эффективный тест* (оптимальный по объему набор тестовых заданий, гарантирующий оценку личностной модели ученика заданным критериям). Выбор эффективного теста зависит от удачного разбиения тестового пространства на классы эквивалентности, пограничные условия, создание тестов на покрытие путей и логических связей между понятиями и модулями.

В дальнейшем необходим тестовый эксперимент на группе учащихся, который позволит провести корректировку и доводку теста до вида эксплуатации.

Способы составления тестовых заданий (по степени сложности)

- 1) формирование вопросов к понятиям, составляющим узлы семантического графа (рис. 2), разработка упражнений, требующих для их выполнения знания свойств выбранного понятия;
- 2) разработка тестовых заданий, определяющих отношения между понятиями;
- 3) конструирование заданий, выявляющих связь понятий между отдельными модулями.

Этапы построения компьютерных тестов:

- 1) формализация экспертной целевой модели знаний;
- 2) проектирование тестового пространства;
- 3) формирование и наполнение тестовых заданий;
- 4) формирование полного компьютерного теста;
- 5) тестовый эксперимент;
- 6) выбор эффективного теста;
- 7) анализ, корректировка и доводка теста до вида эксплуатации [9].

Тема 10 СБОР И СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ

Статистическая обработка полученных данных тестирования (результатов испытуемых) позволяет определить количественные характеристики тестовых заданий и всего теста, что даёт возможность использовать тест как научно обоснованный, объективный инструмент педагогических измерений. Статистическая обработка результатов тестирования проводится с помощью программного обеспечения.

1. Матрица тестовых результатов

Для удобства обработки тестовых данных результаты заносятся в матрицу тестовых результатов. Ниже приведён пример такой матрицы (один балл за правильный ответ, ноль баллов - за неправильный ответ или пропуск задания) [5].

Таблица 5

Матрица тестовых результатов

№ испытуемого i	№ задания j					Индивидуальный балл X_i
	1	2	3	4	5	
1	0	0	0	0	0	0
2	1	0	1	0	1	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	1	1	1	0	4
Число правильных ответов Y_j	4	3	3	1	1	12

Для наглядности и удобства дальнейшего использования необходимо упорядочить матрицу тестовых результатов. Для этого строки таблицы располагают таким образом, чтобы индивидуальные баллы располагались в порядке возрастания. Затем столбцы матрицы располагают в порядке убывания.

Таблица 6

Упорядоченная матрица тестовых результатов

№ испытуемого i	№ задания j					Индивидуальный балл X_i
	1	2	3	4	5	
1	0	0	0	0	0	0
4	1	1	0	0	0	2
2	1	0	1	0	1	3
3	1	1	1	0	0	3
5	1	1	1	1	0	4
Число правильных ответов Y_j	4	3	3	1	1	12

2. Трудность тестовых заданий

По результатам апробационного тестирования определяются характеристики тестовых заданий - трудность и дискриминативность.

Трудность задания – соотношение количества испытуемых, справившихся с данным заданием, и общего количества испытуемых. Трудность задания вычисляется по формуле:

где p_j - доля правильных ответов на j -ое задание; Y_j - количество испытуемых, выполнивших j -ое задание верно, N - число испытуемых в группе, j - номер задания. Чем выше показатель трудности, тем задание легче, чем меньше показатель трудности задания, тем задание сложнее.

Показатель трудности важен для определения характеристики тестового задания и помогает проранжировать задания, входящие в тест по степени сложности. Благодаря этому можно определить место задания в тесте.

3. Дискриминативность тестового задания

Дискриминативность (дифференцирующая способность, различающая способность) задания - это способность задания дифференцировать испытуемых по уровню достижений, на сильных и слабых. Задание, которое одинаково выполняется и слабыми, и сильными, имеет низкую дискриминативность. Задание с высокой дискриминативностью выполняют только сильные учащиеся.

Индекс дискриминативности определяется как разность долей правильных ответов сильной и слабой групп по формуле:

$$(r_{disc})_j = (p_1)_j - (p_0)_j,$$

где r - индекс дискриминативности, p_1 - доля правильных ответов в сильной подгруппе (27 % от всего количества), p_0 - доля правильных ответов в слабой группе (27 %).

Значение индекса дискриминативности находится в интервале [- 1; 1]. Если индекс дискриминативности выше 0,3, это свидетельствует о том, что задание обладает хорошим дифференцирующим эффектом. Если $r = 0$, то это значит, что и слабые, и сильные испытуемые выполняют задание одинаково. Отрицательный показатель дискриминативности появляется в том случае, когда слабые учащиеся выполняют задание правильно, а сильные - неправильно, что свидетельствует о некачественном (невалидном) задании. Тестовые задания с показателями $r_{disc} \geq 0$ и $r_{disc} < 0$ необходимо удалять.

Высокая дискриминативность, которая свидетельствует о сильном дифференцирующем эффекте тестового задания, характерна для заданий со средним показателем трудности (0,5).

4. Дистракторный анализ

Одно из важнейших требований, которое предъявляется к заданиям закрытой формы, - это правдоподобность дистракторов (неправильных ответов). Дистракторный анализ (оценка качества дистрактора) предполагает подсчет долей испытуемых, выбравших каждый дистрактор. В идеальном варианте каждый дистрактор должен выбираться в равной доле от всех неправильных ответов. Ниже представлены таблица 7, с качественными и таблица 8, с некачественными дистракторами.

Таблица 7

Распределение ответов при удачных дистракторах

№ задания	Количество учащихся, выбравших ответ			
	1 ответ	2 ответ (правильный)	3 ответ	4 ответ
j	20%	40%	20%	20%

Таблица 8

Распределение ответов при неудачных дистракторах

№ задания	Количество учащихся, выбравших ответ			
	1 ответ	2 ответ (правильный)	3 ответ	4 ответ
j	45%	40%	5%	10%

В результате дистракторного анализа выявляются задания с неработающими дистракторами (вариантами ответов в закрытом задании, которые испытуемые выбирают мало или вообще не выбирают). Если в задании есть неработающие дистракторы, то вероятность угадывания правильного ответа повышается, снижается его трудность. Такие

задания требуют переработки (замены неработающих дистракторов) или удаления их из теста [3].

5. График (кривая частотного распределения)

Для интерпретации распределения результатов выполнения теста необходимо построить гистограмму или график (кривую частотного распределения), в котором по оси X откладывается количество выполненных заданий, по оси Y – количество учащихся, верно выполнивших эти задания (в абсолютном значении или в процентах).

Схема 2

6. Анализ кривой частотного распределения

Асимметрия. Наличие асимметрии легко установить визуально, анализируя график или гистограмму.

Схема 3

Виды асимметрии

Положительная асимметрия распределения характерна для излишне легких тестов, эффект отрицательной асимметрии встречается в излишне трудных тестах. В хорошо сбалансированном по трудности тесте асимметрия нулевая.

Эксцесс.

С помощью эксцесса можно получить представление о том, является ли гистограмма островершинной или плоской. Островершинная кривая имеет явно выраженный положительный эксцесс, средневершинная имеет нулевой эксцесс, характерный для нормальной кривой, плосковершинная имеет эксцесс меньше нуля.

Схема 4

Виды эксцессов

Понятие «эксцесс» применимо лишь к унимодальным распределениям.

Для нормального распределения характерна кривая с нулевой симметрией, нулевым эксцессом, одной вершиной [5].

7. Меры центральной тенденции

Меры центральной тенденции предназначены для выявления «центрального положения», вокруг которого в основном группируется множество значений такого распределения данных.

Мода (Mo) – это такое значение, которое встречается наиболее часто среди результатов выполнения теста.

В том случае, если два значения встречаются одинаково часто, соответствующее распределение называется бимодальным. Один из наиболее важных выводов в случае бимодального распределения – корректировка трудности заданий теста.

В том случае, когда все значения баллов учеников встречаются одинаково часто, принято считать, что моды у распределения нет.

Среднее выборочное (M или \bar{x}), или среднее арифметическое определяется суммированием всех значений совокупности и последующим делением на их число. Для совокупности индивидуальных баллов x_1, x_2, \dots, x_N группы испытуемых числом N среднее значение будет равно:

Медиана (Me) – среднее (центральное) значение упорядоченного числового ряда.

Например, для ряда 3, 4, 5, 6, 7, 9, 11 медианой будет 6, так как это значение имеет равное число значений справа и слева от себя в числовом ряду. Медиана чётного ряда находится как полусумма двух центральных значений.

Интерпретация мер центральной тенденции.

Меры центральной тенденции помогают при оценке качества теста в том случае, когда она проводится на репрезентативной выборке учеников. Хороший нормативно-ориентированный тест обеспечивает нормальное распределение индивидуальных баллов учеников, когда среднее значение баллов совпадает с модой и находится в центре распределения, около 68% концентрируются вокруг среднего по нормальному закону, а остальные сходят на нет к краям распределения:

Схема 5

График нормального распределения

Нормальное распределение унимодально и симметрично, мода, медиана и среднее значение равны.

На малых выборках любые показатели теряют свою стабильность.

Смещение среднего значения влево или вправо говорит о слишком трудной или слишком легкой подборке заданий теста [12].

8. Меры изменчивости

Размах – измеряет на шкале расстояние, в пределах которого изменяются все значения показателя в распределении. Из индивидуального балла самого сильного ученика вычитают индивидуальный балл самого слабого.

Дисперсия

Подсчет дисперсии основан на вычислении отклонений индивидуального балла каждого ученика от среднего арифметического в распределении: или $x_i - M$ ($i=1,2,\dots,N$). Знак отклонения указывает место результата ученика по отношению к среднему арифметическому по тесту. Для ученика с индивидуальным баллом выше среднего значение разности будет положительно, а для тех, у кого результат ниже среднего арифметического, отклонение отрицательно.

Мера изменчивости, называемая дисперсией обозначается S^2_x (или σ^2_x) и вычисляется по формуле

,
где x_1 – индивидуальный балл первого ученика, x_2 – индивидуальный балл второго ученика и т. п., \bar{x} – среднее выборочное (среднее арифметическое) для данной группы, \sum – сумма квадратов разностей, вычисленных для каждого ученика, N – количество учащихся в данной группе.

Низкая дисперсия индивидуальных баллов говорит о слабой дифференциации испытуемых по уровню подготовленности в группе, что противоречит основной цели нормативно-ориентированных тестов.

9. Определение достоверности различий между сравниваемыми группами

Существует множество статистических критериев и способов для определения того, является ли различие в уровне знаний, умений, качеств двух групп случайным или

статистически достоверным. Один из таких способов – расчёт t-критерия Стьюдента (для независимых выборок) по формуле

где \bar{x}_1 и \bar{x}_2 – среднее выборочное (среднее арифметическое) для первой и второй группы, s_1^2 и s_2^2 – дисперсия для данной группы, n_1 и n_2 – количество учащихся в данной группе. Полученный результат сравнивается с табличным результатом для данной степени свободы ν , которое определяется по формуле:

$$\nu = n_1 + n_2 - 2,$$

где n_1 – количество учащихся в первой группе, n_2 – количество учащихся во второй группе. Например, для двух групп численностью каждая 10 человек $\nu = 10 + 10 - 2 = 18$, $t_{\text{табл}} = 2,10$.

В том случае, если результат превосходит табличное значение, можно говорить о достоверном статистическом различии в уровне знаний, формирований умений или личностных качеств между первой и второй группой.

Таблица 9

Критические значения t-критерия Стьюдента при уровне значимости 0,05

ν	t-критерий
4	2,78
5	2,58
6	2,45
7	2,37
8	2,31
9	2,26
10	2,23
11	2,20
12	2,18
13	2,16
14	2,14
15	2,13
16	2,12
17	2,11
18	2,10
19	2,09
20	2,09
25	2,06
30	2,04
40	2,02
50	2,01
60	2,00
80	1,99
100	1,98

Тема 11

ЕДИНЫЙ ГОСУДАРСТВЕННЫЙ ЭКЗАМЕН

С 2001 года в нашей стране проводится эксперимент по введению единого государственного экзамена (ЕГЭ) для выпускников школ и поступающих в высшие учебные заведения.

Причины введения ЕГЭ

В результате реформ конца 80-90-х годов российское образование перестало носить унифицированный характер и перешло к использованию вариативных программ, учебников и пособий, что, с одной стороны, увеличило возможности системы образования: реализация идей развивающего обучения, профессиональной ориентации, возможности индивидуального подхода, но, с другой стороны, усложнило контроль за качеством получаемого образования. «В условиях резко возросшего разнообразия средств, методов и содержания обучения школьников со всей остротой встала задача обеспечения единого уровня и содержания базового ядра общего и среднего образования для всех выпускников школ» [7]. ЕГЭ может стать эффективным надежным и объективным средством контроля за качеством знаний и умений учащихся, в случае успешного завершения эксперимента. Таким образом, в данный момент ЕГЭ рассматривается как одно из средств модернизации современного образования в России.

Задачи единого экзамена:

- расширение доступности высшего образования благодаря введению однотипных экзаменов для всех выпускников общеобразовательных школ и их параллельного участия в заочном конкурсном отборе сразу в несколько вузов;
- совершенствование системы и практики финансирования высших учебных заведений на основе рыночной состязательности между ними за прием лучших студентов;
- снижение психологической нагрузки на выпускников общеобразовательных учреждений за счет упразднения вступительных экзаменов в вузы;
- объективизация и унификация требований к общеобразовательной подготовке поступающих в вузы;
- стимулирование деятельности педагогических коллективов общеобразовательных учреждений по улучшению качества учебного процесса за счет объективной и независимой сравнительной оценки результатов общеобразовательной подготовки выпускников школ.

Преимущества ЕГЭ перед другими формами контроля

1. Объективность.

Использование ЕГЭ в качестве выпускного экзамена в средней школе и его результатов для поступления в высшие учебные заведения имеет ряд преимуществ перед традиционными устными и письменными экзаменами. Прежде всего, это объективность оценивания. В системе ЕГЭ отсутствует преподаватель, который проверяет знания выпускника, то есть исключается субъективный момент (неприязнь к учащемуся, заинтересованность в хороших показателях, плохое настроение, самочувствие преподавателя и др.) при выставлении оценки. Благодаря стандартизации - единой форме предъявления контрольно-измерительных материалов (КИМ) и единого метода обработки полученных результатов достигается высокий уровень объективности оценивания учебных достижений выпускников.

2. Надежность.

Разработка тестов и анализ результатов тестирования в соответствии с принципами классической или современной теорий тестов позволяют обеспечить точность и надежность оценивания уровня учебных достижений. Чтобы это преимущество ЕГЭ могло быть реализовано, КИМы должны включать тестовые задания, которые прошли экспертную оценку и были апробированы на репрезентативной выборке испытуемых.

3. Достоверность.

Тестовые технологии могут обеспечить достоверные результаты, свободные от фальсификации и искажения. Нельзя не согласиться, что без обеспечения высокой

достоверности единого экзамена, без гарантий его информационной безопасности невозможно завоевать доверие к результатам ЕГЭ, что, безусловно, резко снизит эффективность этого нововведения.

Выделяют ряд возможностей, которые могут способствовать снижению достоверности. Это рассекречивание, подставка, подсказка, подтасовка.

Чтобы не произошло рассекречивания КИМов, существует система информационной безопасности, которая защищает тестовые материалы от преждевременного доступа. Один из возможных способов защиты - это создание большого банка калиброванных тестовых заданий и обеспечение свободного доступа к этому банку (например, через интернет или печатные издания). Знакомство учащихся с множеством заданий банка позволит им лучше подготовиться к сдаче теста. Для самого экзамена формируются многочисленные варианты теста в компьютерном режиме из существующего банка заданий (калиброванных) индивидуально для каждого испытуемого.

Структура КИМов ЕГЭ

«КИМы - это стандартизированная экзаменационная работа, создаваемая в соответствии с требованиями теории педагогических измерений, позволяющая с достаточной объективностью и надежностью провести независимую государственную аттестацию выпускников общеобразовательных учреждений и отбор абитуриентов вузов» [3].

При составлении КИМов определялось эффективное соотношение использования различных форм заданий в тесте для итоговой аттестации выпускников. При этом учитывался многолетний опыт зарубежных коллег. В США сформировалась традиция применения тестов, состоящих из заданий закрытого типа, то есть заданий с выбором правильного ответа (так устроен, например, самый популярный тест академических способностей SAT). Но задания этого типа не могут оценивать способности учащегося рассуждать и анализировать, делать самостоятельные выводы, создавать письменный текст, высказывать собственное мнение и т. д. Включение в КИМы заданий только закрытого типа было бы недостаточным, так как снижало бы возможности контроля знаний выпускников. Для тестологической практики Великобритании характерно использование как закрытых, так и открытых заданий. Так, в тестах на получение общего сертификата о среднем образовании (general certificate of secondary education - GCSE), которые разрабатываются Кембриджским экзаменационным синдикатом, доля открытых заданий выше, чем доля заданий с выбором правильного ответа.

Чтобы обеспечить эффективный контроль знаний и умений выпускников отечественной школы, было решено включить в КИМы различные формы заданий: задания с выбором правильного ответа, задания на дополнение и задания со свободным развернутым ответом (ответ в виде эссе, рецензии, анализа текста).

КИМы по различным предметам включают три части - А, В, С. Каждая часть состоит из заданий, сгруппированных по форме. Общее количество заданий колеблется от 25 до 70. Часть А предлагает задания с выбором ответа (задания закрытого типа). Эти задания достаточно легкие и направлены на проверку знаний фактического материала, правил, формул, определений и др. Доля таких заданий в зависимости от предмета колеблется. Они могут преобладать в количественном отношении над другими видами заданий, но весовой коэффициент таких заданий меньше, то есть в суммарном взвешенном балле за решение одного задания типа А дается меньше очков, чем за решение заданий другого типа.

Часть В состоит из заданий открытого типа - из заданий на дополнение в виде числа или одного слова. В этих заданиях испытуемый сам конструирует правильный ответ, в отличие от заданий закрытого типа, где ответ уже дан и надо только определить правильный. Как правило эти задания сложнее, чем задания части А, поэтому их весовой коэффициент выше.

Задания частей А и В сконструированы таким образом, чтобы проверка проходила в компьютерном режиме.

В части С дается задание, которое предполагает свободный развернутый ответ. Это может быть полное решение математической задачи или написание текста. Задания части С проверяют умения выпускников размышлять, рассуждать на заданную тему, формулировать и грамотно выражать свои мысли письменно. Эта часть теста проверяется независимыми экспертами. Это требует дополнительных материальных расходов и в какой-то степени дает возможность повлиять на выставление баллов по части С, но, тем не менее, включение этой части в КИМ увеличивает возможность проверки широкого спектра знаний и умений выпускников [3].

КОНТРОЛЬНЫЙ ТЕСТ ПО КУРСУ «СОВРЕМЕННЫЕ СРЕДСТВА ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ»

I. Мониторинг качества образования – это ...

1. механизм контроля и слежения за качеством
2. совокупность условий и средств, обеспечивающих непрерывное наблюдение за процессом обучения
3. система органов, контролирующая качество образования
4. обязательный этап аккредитации ОУ

II. Личностно-ориентированный подход рассматривает образование как ...

1. ориентацию на проблемы личности
2. способ решения поставленных человеком профессиональных задач
3. способ развития личности
4. деятельность по согласованию интересов личности и общества.
5. предпосылку социализации

III. Права на выдачу документа об образовании государственного образца дает образовательному учреждению ...

1. лицензирование
2. аттестация
3. аккредитация
4. модернизация качества образования
5. соответствие государственному образовательному стандарту

IV. В системе опережающего обучения, по сравнению с системой поддерживающего обучения, роль традиционных знаний ...

1. возрастает
2. снижается
3. остается прежней
4. изменяется в зависимости от специфики обучения

V. Критериями качества знания в современных педагогических технологиях выступают ...

1. знания, умения, навыки
2. развитие творческих способностей
3. опыт эмоционально-ценностных отношений
4. способности к самореализации в трудовой или учебной деятельности

VI. Мониторинг качества образования контролирует...

1. выполнение образовательным учреждением требований государственного стандарта

2. уровень знаний учащихся
3. способность учащихся к самореализации в учебной деятельности
4. социализированность личности учащихся к моменту окончания учебного учреждения

VII. Мониторинг информирует о соответствии ...

1. подготовки учащихся – требованиям общества и рынка труда
2. ожиданий родителей учащихся – содержанию образования
3. фактических результатов деятельности педагогической системы – ее конечным целям
4. обязанностей учителей и школьных работников – их квалификации

VIII. Информацию о реальных результатах обучения можно получить с помощью ...

1. экзаменов, контрольных и срезовых работ
2. массовых обследований
3. изучения школьной документации
4. психодиагностических методик

IX. Эффективный мониторинг предполагает разработку ...

1. педагогических оценочных материалов
2. системы требований к ответам ученика
3. критериев оценивания традиционных письменных работ
4. способов оценки развития творческих способностей учащихся

X. «Пятёрка» – это ...

1. оценка
2. отметка
3. ранжирование
4. количественный способ выражения знаний
5. эталон

XI. Из нижеперечисленного к видам контроля относятся ...

1. тестирование, портфолио, рейтинг, мониторинг
2. собеседование, контрольная работа, опрос, зачёт, экзамен
3. предварительный, текущий, периодический, итоговый
4. устный, письменный, практический

XII. Диагностика обучения – это ...

1. проверка ЗУНов учащихся
2. контроль сформированности их компетенций и творческих способностей
3. определение результатов, тенденций и динамики педагогического процесса
4. механизм слежения за качеством образования

XIII. К современным средствам оценивания относятся ...

1. тестирование, портфолио, рейтинг, мониторинг
2. собеседование, контрольная работа, опрос, зачёт, экзамен
3. предварительный, текущий, периодический, итоговый
4. устный, письменный, практический

XIV. Рейтинг ...

1. является разновидностью накопительной оценки
2. никак не связан с накопительной оценкой
3. может частично быть накопительным
4. накапливает ЗУНЫ учащихся

XV. Ключевым принципом рейтинга является ...

1. администрирование
2. накопление
3. ранжирование
4. мониторинг

XVI. В рейтинговой системе оценивания показателем качества обучения служит (служат)

...

1. кумулятивный балльный показатель
2. нормативный (эталонный) показатель
3. итоги финальных срезových работ
4. итоги независимого педагогического тестирования

XVII. «Академический рейтинг» – это показатель ...

1. достижений в сфере высшего образования
2. учебных достижений
3. способности к принятию решения
4. сформированности учебных компетенций

XVIII. К средствам накопительной оценки относят в первую очередь ...

1. тестирование
2. портфолио
3. рейтинг
4. мониторинг

XIX. Портфолио можно перевести как ...

1. «портфель»
2. «папка специалиста»
3. «накопитель»
4. «досье»

XX. Основная задача портфолио – ...

1. контролировать домашнюю работу учащегося
2. обеспечить материальное воплощение ЗУНов
3. дать возможность для независимого контроля уровня ЗУНов учащихся
4. создать для учащихся «стимул роста»

XXI. Российская концепция портфолио предполагает следующие разделы портфолио:

1. достижений, рефлексивный, проблемно-исследовательский, тематический
2. репродуктивный, творческий
3. документов, работ, отзывов
4. академический, олимпийский, достижений

XXII. Портфолио позволяет при оценивании делать акцент на ...

1. индивидуальных достижений ученика
2. соответствии ЗУНов ученика государственному образовательному стандарту
3. практических умениях
4. теоретических знаниях

XXIII. Основоположником тестологии считают ...

1. Френсиса Гальтона
2. Джеймса Кеттела

3. Эдуарда Торндайка
4. Бернбаума

XXIV. Шкала Альфреда Бине и Теодора Симона дифференцировала детей по уровню ...

1. творческих способностей
2. умственного развития
3. учебных достижений
4. волевого потенциала
5. нравственного развития

XXV. Служба тестирования образования в США была создана с целью ...

1. проверки результатов итогового тестирования в вузах
2. проверки педагогических и психологических тестов на надежность и валидность
3. независимой оценки результатов психологических и педагогических тестов
4. контроля проведения тестирования на всех уровнях образования в стране

XXVI. В начале прошлого века тестологию в России активно развивал ...

1. А.С. Макаренко
2. В.А. Сухомлинский
3. П.П. Блонский
4. С.Т. Шацкий

XXVII. Надежность теста – это...

1. устойчивость результатов к воздействию случайных факторов
2. способность теста измерять то, для чего он предназначен
3. способность теста давать разные результаты в зависимости от уровня тестируемых
4. эталон педагогического измерения

XXVIII. Валидность теста – это...

1. устойчивость результатов к воздействию случайных факторов
2. способность теста измерять то, для чего он предназначен
3. способность теста давать разные результаты в зависимости от уровня тестируемых
4. эталон педагогического измерения

XXIX. Таксономия Блума – это ...

1. классификация учебных целей
2. иерархия видов учебных заданий
3. структура содержания обучения
4. типология предметных областей

XXX. GCSE - это ...

1. европейский институт тестологии
2. американская национальная служба тестологии
3. аналог ЕГЭ в Англии
4. обозначение уровней владения учебным материалом

XXXI. Диагностическое тестирование, в отличие от формирующего, позволяет ...

1. осуществлять функцию обратной связи
2. определить реальный уровень ЗУНов учащихся
3. учащимся самостоятельно контролировать свои знания и умения
4. установить причины ошибок

XXXII. Нормативно-ориентированные тесты привязаны к ...

1. среднему уровню знаний/умений/качеств учащихся
2. учебной программе и стандартам
3. мерам центральной тенденции
4. нормам выполнения учебных заданий, установленным для каждого предмета

XXXIII. Критериально-ориентированные тесты ориентируются на:

1. средний уровень знаний/сформированность умений/качеств учащихся
2. учебную программу и образовательные стандарты
3. меры центральной тенденции
4. критерии сформированности учебных компетенций

XXXIV. Все предыдущие задания являлись заданиями ...

1. на сопоставление
2. на дополнение
3. множественного выбора
4. альтернативных ответов

XXXV. Оптимальное количество дистракторов в тесте - ...

1. один
2. три-четыре
3. пять-шесть
4. более шести

XXXVI. Один из важнейших аспектов массового компьютерного тестирования – это ...

1. новейшее программное обеспечение
2. владение тестируемыми компьютером на высоком уровне
3. психологическая готовность учащихся к тестированию
4. меры безопасности

XXXVII. Репрезентативность выборки – это её способность ...

1. представлять качества всей популяции
2. соответствовать мерам центральной тенденции
3. интегративное качество, связанное с валидностью и надежностью теста
4. усредненность полученных результатов, отсутствие больших отклонений
5. все вышеперечисленное

XXXVIII. Дистрактор – это ...

1. элемент вычисления дисперсии
2. качество, характеризующее дискриминативность теста
3. качество, характеризующее дифференцирующую способность
4. коэффициент, необходимый для пересчета относительных баллов в абсолютные
5. неправильный ответ

XXXIX. Кривая распределения в виде симметричного колокола показывает на...

1. отсутствие эксцессов
2. репрезентативность выборки
3. валидность теста
4. высокую дискриминативность

XL. Дискриминативность - это ...

1. свойство нормального распределения

2. эффективность конкретного дистрактора
3. способность теста отделять слабых от сильных
4. удельный вес тестового задания

ХLI. Т-критерий Стьюдента проверяет ...

1. зависимость уровня сформированности качества от внешних факторов
2. случайность или закономерность различия между двумя группами тестируемых
3. динамику развития учебных и личностных качеств
4. меры центральной тенденции

ХLII. Задания ЕГЭ ...

1. целиком состоят из тестов различных видов
2. исключают использование тестов
3. предполагают совмещение тестов и заданий на свободное изложение
4. в своих видах целиком определяются спецификой дисциплины

ХLIII. Аббревиатура КИМ расшифровывается как ...

1. контрольно-измерительные методы
2. контрольно-измерительные материалы
3. кривая изменений мет центральной тенденции
4. квалиметрия измерительных материалов

ХLIV. Количество типов заданий при прохождении учащимися ЕГЭ равняется ...

1. одному
2. двум
3. трём
4. четырём

ХLV. Информационная система ЕГЭ ...

1. предназначена для открытого доступа учащихся
2. предполагает возможность открытого доступа учащихся
3. доступна только руководителям образовательных учреждений
4. полностью исключает открытый доступ

ХLVI. ЕГЭ водится с целью ...

1. унификации образования
2. индивидуализации образовательных траекторий
3. обеспечения качественной профильной подготовки
4. эффективного контроля качества образования

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Разделы и темы рабочей программы	Перечень домашних заданий и других вопросов самостоятельного изучения
(1) Понятие о качестве образования.	Охарактеризовать нормативные документы, определяющие качество современного образования.
(3) Виды, формы и организация контроля	Заполнить таблицу «Сравнительные преимущества и недостатки различных

качества обучения.	способов балльного оценивания» (традиционная 5-балльная шкала, 10-балльная шкала, безотметочное обучение).
(5) «Портфолио» как одно из средств накопительной оценки.	Составить портфолио по курсу «Современные средства оценивания результатов обучения» (включает практические задания).
(7) Понятие теста. Психолого-педагогические аспекты тестирования.	Представить результаты тестирования в статистическом виде (графы, диаграммы), оценить свойства распределения.
(8) Виды тестов и формы тестовых заданий. Содержание и структура тестовых заданий по конкретному предмету.	1. Составить задания в тестовой форме с выбором одного правильного ответа (в соответствии с профилем специальности). 2. Составить задания в тестовой форме с выбором нескольких правильных ответов (в соответствии с профилем специальности). 3. Составить тестовые задания открытой формы (в соответствии с профилем специальности). 4. Составить задания в тестовой форме на установление правильной последовательности.

(9) Компьютерное тестирование и обработка результатов.	Разработать документ в формате Microsoft Excel, обеспечивающий компьютерное тестирование школьников.
(10) Единый государственный экзамен. Организационно-технологическое обеспечение ЕГЭ.	Проанализировать задания ЕГЭ по изучаемой специальности (типы тестовых заданий, критерии оценивания, знания и умения учащихся, подлежащие контролю).

ПЕРЕЧЕНЬ ВОПРОСОВ К ЗАЧЕТУ (ЭКЗАМЕНУ)

1. Расскажите о понятии «качество образования». Охарактеризуйте оценку как элемент управления качеством.
2. Сравните традиционные и новые средства оценки результатов обучения. Опишите их достоинства и недостатки.
3. Расскажите историю возникновения тестирования в России.
4. Назовите функции контроля в современном учебном процессе.
5. Назовите традиционные формы контроля. Укажите их достоинства и недостатки.
6. Какие современные средства контроля выделяются в учебном процессе? В чем их преимущество по сравнению с традиционными формами контроля?
7. Укажите место психологических и педагогических измерений в современном образовании.
8. Расскажите об основных подходах к качеству знаний.
9. Охарактеризуйте мониторинг как средство оценки результатов обучения.
10. Назовите основные свойства мониторинга качества образования.

11. Охарактеризуйте виды мониторинга.
12. Назовите методы педагогического мониторинга.
13. Опишите достоинства и недостатки «портфолио» как средства оценки результатов обучения.
14. Охарактеризуйте цели и задачи педагогического и психологического тестирования.
15. Сходство и различие педагогических и психологических тестов в учебном процессе.
16. Назовите психологические тесты, применимые в учебном процессе.
17. Охарактеризуйте основные подходы к структуре учебных достижений.
18. Дайте определение следующим понятиям: тест, тестовое задание, валидность теста, надежность теста.
19. Расскажите о видах тестов.
20. Опишите виды педагогического контроля (текущий, тематический, рубежный, итоговый контроль).
21. Дайте классификацию тестов по разным основаниям.
22. Раскройте понятие гомогенных и гетерогенных тестов.
23. Расскажите о применении компьютерного тестирования.
24. Назовите основные виды заданий в тестовой форме.
25. Как определяются цели тестов?
26. Расскажите о структуре тестового задания.
27. На какие принципы необходимо опираться при отборе содержания тестового задания?
28. Как производится экспертиза качества содержания теста?
29. Расскажите о задачах ЕГЭ. Каковы преимущества и недостатки ЕГЭ перед другими формами контроля?
30. Проанализируйте нормативные документы, регулирующие проведение ЕГЭ (Положение о проведении ЕГЭ, утвержденной приказом Министерства образования РФ от 09.04.2002. № 1306; Положение о государственной экзаменационной комиссии субъекта РФ (ГЭК); Положение о конфликтной комиссии; Положение о системе общественного наблюдения).
31. Опишите организацию проведения ЕГЭ.
32. Опишите структуру заданий ЕГЭ.
33. Расскажите о порядке создания контрольно-измерительных материалов для ЕГЭ.
34. Каков порядок проверки тестовых заданий ЕГЭ?
35. Расскажите о требованиях к пунктам проведения ЕГЭ.
36. Охарактеризуйте особенности тестовых заданий и проведения ЕГЭ по вашей специальности.
37. Охарактеризуйте понятие «рейтинговая оценка», расскажите о ее преимуществах и недостатках, причинах использования рейтинговой оценки в практике преподавания.
38. Расскажите о правилах рейтингового оценивания, видах рейтинга, методах методической поддержки рейтинга

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аванесов, В.С. Композиция тестовых заданий [Текст]. - М., 2002.
2. Анастаси, А., Урбина, С. Психологическое тестирование [Текст]. – СПб., 2002.
3. Воронин, Ю.А., Трубина Л.А., Васильева, Е.В., Козлова, О.В. Курс лекций «Современные средства оценивания результатов обучения» [Текст] : учебное пособие. – Воронеж: ВГПУ. - 2004.
4. Гаврилова, Т. Н., Кузнецов, А. Н. Теория обучения [Текст] : учебное пособие. – Ярославль, Изд-во ЯГПУ им. К.Д. Ушинского, 2003.
5. Золотарёва, А.В. Мониторинг результатов деятельности учреждений дополнительного образования детей [Текст]. – Ярославль, Изд-во ЯГПУ им. К.Д.

- Ушинского, 2006.
6. Калужская, М.В., Уколова, О.С., Каменских, И.Г. Рейтинговая система оценивания. Как? Зачем? Почему? [Текст] – М. : Чистые пруды, 2006
 7. Концепция модернизации российского образования на период до 2010 года. Приложение к приказу Минобразования России от 11.02.2002 N 393 [Текст] // Российская газета. – 2002. – 15 ноября.
 8. Майоров, А.Н. Теория и практика создания тестов для системы образования [Текст]. – М., 2000.
 9. Математика и информатика [Текст] : курс лекций. - Уссурийск : Изд-во Уссурийского государственного педагогического института, 2005
 10. Петрухин, В. В. Проблема организации и апробации системы педагогического мониторинга в образовательном учреждении [Электронный ресурс] // Интернет-журнал "Эйдос". - 2007. - 15 января. <http://www.eidos.ru/journal/2007/0115-7.htm>.
 11. Поташник, М.М., Моисеев, А.М. Управление современной школой (в вопросах и ответах) [Текст] : пособие для руководителей образовательных учреждений и органов образования. – М., 2000
 12. Равен, Д. Педагогическое тестирование: Проблемы, заблуждения, перспективы [Текст]. – М. : Когито-Центр, 2001
 13. Статистический анализ данных и способы представления результатов исследования [Текст] : учебно-методическое пособие к лекционно-практическому курсу «Психодиагностика» / сост. Е. Г. Заверткина, Н. К. Рукавишников. – Ярославль, Изд-во ЯГПУ, 2000
 14. Устинова, Л. Г. Развитие творческого потенциала студентов в условиях рейтинговой технологии обучения [Текст]. – Волгоград, 2000.
 15. Циммерман, Н.В. Условия подготовки студентов педагогических вузов к компьютерному тестированию // Региональная образовательная информационная среда [Текст] : сборник научно-методических статей. - М. : Прометей, 2006
 16. Челышкова, М. Б. Разработка педагогических тестов на основе современных математических моделей [Текст]. – М., 1995.
 17. Чернявская, А.П. Становление партнерской позиции педагога [Текст]. – Ярославль : Изд-во ЯГПУ, 2007.
 18. Шишов, С.Е., Кальней, В.А. Мониторинг качества образования в школе [Текст]. - М., 1998

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная

1. Глас, Дж., Стенли, Дж. Статистические методы в педагогике и психологии [Текст]. - М., 1976.
2. Годфруа, Ж. Что такое психология [Текст]. Т. 2. – М. : Мир, 1999
3. Гузев, В.В. Планирование результатов образования и образовательная технология [Текст]. – М. : Народное образование, 2001.
4. Дьяченко, В.К. Новая дидактика [Текст]. – М. : Народное образование, 2001
5. ЕГЭ. Сборник нормативных документов [Текст]. – М., 2002.
6. Загвязинский, В.И., Атаханов, Р. Методология и методы психолого-педагогического исследования [Текст]. – М., 2003
7. Зверева, В.И. Образовательная программа школы: структура, содержание, технология разработки [Текст]. – М. : Образовательный центр «Педагогический поиск», 1998.
8. Майоров, А.Н. Теория и практика создания тестов для системы образования [Текст]. – М., 2000.
9. Мельникова, М.Б. Теория и практика конструирования педагогических тестов [Текст].

- М., 2002.
10. Сергеева, В.П., Каскулова, Ф. В., Гричненко, И.С.. Современные средства оценивания результатов обучения [Текст] : учебно-методическое пособие. – М. : АПКППРО, 2005.
 11. Чельшкова, М.Б. Теория и практика конструирования педагогических тестов [Текст]. - М., 2002.
 12. Управление качеством образования [Текст] : практикоориентированная монография и методическое пособие / под ред. М.М. Поташника. – М. : Педагогическое общество России, 2000.
 13. Энциклопедия современного учителя [Текст]. – М. : Издательство «Астрель», «Олимп», фирма «Издательство АСТ», 2000.

Дополнительная

1. Валеев, Г. Х. Методология и методы психолого-педагогического исследования [Текст]. - Стерлитамак, 2002.
2. Гребенкина, Л.К., Анциперова, Н.С. Технология управленческой деятельности заместителя директора школы [Текст]. – М : Центр «Педагогический поиск», 2000.
3. Калинин, С.И. Компьютерная обработка данных для психологов [Текст]. – СПб, 2002.
4. Клайн, П. Введение в психометрическое программирование Справочное руководство по конструированию тестов [Текст]. - Киев, 1994.
5. Ксензова, Г.Ю. Перспективные школьные технологии [Текст]. – М. : Педагогическое общество России, 2000.
6. Майоров, А. Н. Тесты школьных достижений: конструирование, проведение, использование [Текст]. – СПб., 1996.
7. Лизинский, В.М. О методической работе в школе [Текст]. - М. : Центр «Педагогический поиск», 2001.
8. Нардюжев, В.И., Нардюжев, И.В. Модели и алгоритмы информационно-вычислительной системы компьютерного тестирования [Текст]. – М., 2000.
9. Поташник, М.М., Моисеев, А.М. Управление современной школой (в вопросах и ответах) [Текст] : пособие для руководителей образовательных учреждений и органов образования. – М., 2000
10. Равен, Д. Педагогическое тестирование: Проблемы, заблуждения, перспективы [Текст]. – М. : Когито-Центр, 2001.
11. Родионов, Б.У., Татур, А.О. Стандарты и тесты в образовании [Текст]. – М., 1995.
12. Селевко, Г.К. Современные образовательные технологии [Текст] : Уч. пособие. – М., 1998.
13. Симонов, В. П. Урок: Планирование, организация и оценка эффективности [Текст] : Учебное пособие. – М., 2005
14. Симонов, В. П., Черненко, Е.Г. Десятибалльные шкалы оценок обученности по предмету [Текст] : Учебно-справочное пособие. – М., 2002
15. Скок, Г. Б. Как проанализировать собственную педагогическую деятельность [Текст]. – М., Педагогическое общество России, 2000.
16. Устинова, Л. Г. Развитие творческого потенциала студентов в условиях рейтинговой технологии обучения [Текст]. – Волгоград, 2000.
17. Федоров, В. А. Теория развития профессионально-педагогического образования в современных условиях [Текст]. – Екатеринбург, 2002
18. Чельшкова, М. Б., Хлебников, В. А. Основные подходы к оценке качества подготовки обучаемых [Текст] // Проблемы качества, его нормирования и стандартов в образовании. Сб. науч. ст. – М., 1999.
19. Чельшкова, М. Б. Разработка педагогических тестов на основе современных математических моделей [Текст]. – М., 1995.

20. Чельшкова, М. Б., Савельев, Б. А. Методические рекомендации по разработке педагогических тестов для комплексной оценки подготовленности студентов в вузе [Текст]. - М., 1995.
21. В помощь организаторам единого государственного экзамена: Учебно-методическое пособие [Текст]. В 2-х ч. / МО РФ; Рос. ун-т дружбы народов. – М., 2002.
22. Единый государственный экзамен [Текст] : Сборник нормативных документов / Министерство образования Российской Федерации. – М. : Интеллект – центр, 2003.
23. Инструктивные материалы для организаторов ЕГЭ в 2003 году [Текст]. – М. : Центр тестирования Минобрнауки России, 2003.
24. Ковалева, Г.С. Подходы к разработке контрольно-измерительных материалов для ЕГЭ [Текст] // Школьные технологии. – 2003. - №4.
25. Макарова, Т.Н. Рекомендации заместителям директоров образовательных учреждений по учебно-воспитательной работе по подготовке школы и учащихся к проведению единого государственного экзамена [Текст] // Завуч. – 2003. – №4.
26. Мендель, В.В. О критериях оценки заданий с развернутыми ответами при проверке экзаменационных работ по математике при проведении ЕГЭ [Текст] // МИФ. – 2003. – №1.
27. Об утверждении Положения и состава Совета по организации подготовки и проведения эксперимента по введению ЕГЭ: Приказ МО РФ ОТ 7.10.03Г. №3804 [Текст] // Вестник образования. – 2003. – №22.
28. О проведении эксперимента по единому государственному экзамену: Приказ Минобрнауки РФ от 05.01.2001г. №50 [Текст] // «ПС» Управление школой. – 2001. – №6. – С. 2.
29. Положение о проведении единого государственного экзамена: Прил. к приказу Минобрнауки России от 27.02. 2001 №645 [Текст] // Образовательное право: Прил. к газете «Учительская газета». – 2001. – №16. // «ПС» Документы. – 2001. – №73. // Образование в современной школе. – 2001. - №4. // Вестник образования. – 2001. – №9. // Учитель. – 2001. - №3. // Народное образование. – 2001. – №5. // Школьное обозрение. – 2001.- №3-4.
30. Правила участия выпускников общеобразовательных учреждений в едином государственном экзамене в 2001 году; Инструкция по обеспечению информационной безопасности процедуры проведения ЕГЭ и обработки его результатов [Текст] // Учительская газета. – 2001. – №20.

СОДЕРЖАНИЕ

Тема 1. Понятие о качестве образования. Оценка результатов обучения как элемент управления качеством	3
Тема 2. Мониторинг качества образования	9
Тема 3. Виды, формы и организация контроля качества обучения	17
Тема 4. Система рейтинг-контроля как составная часть учебного процесса	21
Тема 5. Портфолио как одно из средств накопительной оценки	29
Тема 6. История развития системы тестирования в России и за рубежом	37
Тема 7. Психологические тесты: термины и определения. Психолого-педагогические аспекты тестирования	44
Тема 8. Виды тестов и формы тестовых заданий	
Тема 9. Компьютерное тестирование	59
Тема 10. Сбор и статистическая обработка результатов тестирования	67
Тема 11. Единый государственный экзамен	76
Контрольный тест по курсу «Современные средства оценивания результатов обучения	81
Задания для самостоятельной работы студентов	90
Перечень вопросов к зачёту (экзамену)	91
Библиографический список	39
Рекомендуемая литература	94

**Анна Павловна Чернявская
Борис Сергеевич Гречин**

Современные средства оценивания результатов обучения

Учебно-методическое пособие

Редактор С.А. Викторова

Подписано в печать 16.02.2009.

Формат 62х90/16

6,5 п. л.; 5 уч.-изд. л. Тираж _____ экз. Заказ № _____

Издательство

ГОУ ВПО «Ярославский государственный
педагогический университет им. К.Д. Ушинского»

150000, Ярославль, Республиканская ул., 108.

Используются технологии [uCoz](#)