

Елабужский государственный педагогический университет

Ф.М. Сабирова

**СОВРЕМЕННЫЕ СРЕДСТВА
ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ
ОБУЧЕНИЯ**

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Елабуга – 2007

Печатается по решению Редакционно-издательского совета Елабужского государственного педагогического университета (прот. № 14 от 22.02.2007 г.).

УДК 731.3
ББК 74.202.8
С 12

Рецензенты:

Ахтариева Р.Ф., канд. пед. наук, доцент каф. педагогики
ЕГПУ

Кадырова Л. Х., канд. пед. наук, доцент НГПИ

Сабирова Ф.М. Современные средства оценивания результатов обучения /Ф. М. Сабирова/ Учебно-методическое пособие для студентов педвузов. – Елабуга: изд-во Елабужского пед. ун-та, 2007. – 79 с.

ВВЕДЕНИЕ

В соответствии с решением Министерства образования Российской Федерации в программу педагогических вузов в 2004 г. введена новая дисциплина «Современные средства оценивания результатов обучения». Целью изучения данной дисциплины является знакомство студентов с методологическими и теоретическими основами тестового контроля знаний, порядком организации и проведения единого государственного экзамена (ЕГЭ)¹. Актуальность введения этой дисциплины очевидна: в практику контроля знаний все шире внедряются тестовые методики, с 1999 г. многие школьники России участвуют в централизованном тестировании, а с 2000 г. Министерством образования проводится эксперимент по введению единого государственного экзамена (ЕГЭ). В связи с этим будущему учителю необходимо быть готовым к подготовке и участию своих учащихся в централизованном тестировании и в ЕГЭ, а также в целом к применению современных средств контроля в учебном процессе.²

Педагогическое тестирование – один из наиболее стандартизованных и объективных методов контроля и оценивания знаний, умений и навыков испытуемого, который лишен таких традиционных недостатков других методов контроля знаний, как неоднородность требований, субъективность экзаменаторов, неопределенность системы оценок и т.п.³ В курсе «Современные средства оценивания результатов обучения» изучаются: история и современное состояние тестирования в России и за рубежом, традиционные и современные подходы к оценке учебных достижений, этико-психологические аспекты тестирования, классификации тестов, формы тестовых заданий, различные методы оценивания результатов тестирования, нормативные документы, регламентирующие проведение ЕГЭ, структура и содержание контрольно-измерительных материалов для ЕГЭ по соответствующему предмету. Студент, изучивший дисциплину, должен уметь давать экспертную оценку заданиям тестового типа, использовать на практике тесты различных видов и форм; проводить тестирование и анализировать полученные данные в рамках классической и современной теории создания тестов.

¹ Примерная программа дисциплины «Современные средства оценивания результатов обучения». – М.: Министерство образования РФ, 2004

² “О создании Совета по организации подготовки и проведения эксперимента по единому государственному экзамену”. Приказ министра образования РФ от 26 октября 2006 г.

³ Отраслевой стандарт Министерства образования РФ. Педагогические тесты. Термины и определения. – М., 2001.

Настоящее пособие предназначено для организации аудиторной и самостоятельной работы студентов педагогического вуза. Оно состоит из двух частей. Первая часть представляет собой конспекты лекций по изучаемой дисциплине, которые содержат необходимый теоретический материал, предусмотренный образовательной программой. Во вторую часть включены материалы для организации семинарских и практических занятий. В ней подробно рассмотрены формы тестовых заданий и их примеры, что должно помочь студентам при составлении различных типов тестовых заданий по соответствующей их основной специальности дисциплине.

ЧАСТЬ 1. КОНСПЕКТЫ ЛЕКЦИЙ ПО КУРСУ **Современные средства оценивания результатов обучения**

Тема 1. ОЦЕНКА И КАЧЕСТВО ОБРАЗОВАНИЯ

1. Понятие о качестве образования

В любой сфере деятельности человек стремится достичь максимально возможного качества этой деятельности. В настоящее время особо остро стоит проблема качества образования. Однако до сих пор не существует однозначного определения понятия «качество образования». Обсуждению этого понятия Международным институтом планирования образования были посвящены две международные конференции в 1969 и 1978 гг. Особые споры вызвало одновременное использование и противопоставление понятий «качество» и «количество».

В широком смысле *под качеством образования понимается соотношение цели и результата, мера достижения целей при том, что цели (результаты) заданы только операционально и спрогнозированы в зоне потенциального развития обучаемого*¹.

Иначе говоря: образование, полученное школьником, признается качественным, если его результаты соответствуют операционально² заданным целям и спрогнозированы в зоне потенциального развития ученика. При этом результаты образования обязательно должны включать в себя и оценки того, какой ценой эти результаты достигнуты.

¹ Управление качеством образования/ Под ред. М. М. Поташника. – Педагогическое общество России, 2000.–С.33.

² Операционально в данном случае означает предельно конкретно и так, чтобы можно было определить, достигнута ли цель.

Из этого определения следует, что необходимо измерять в одинаковых единицах: цель, поставленную перед учебным заведением; результат, достигнутый этим учебным заведением¹. В связи с этим образовательное учреждение, школа должна:

- 1) научиться прогнозировать, проектировать, моделировать те качества подготовки выпускника, которые предполагается получить на «выходе» образовательного процесса;
- 2) обеспечивать достижение требуемого качества образования;
- 3) повышать качество образования, приводя его в соответствие с растущими требованиями внешних заказчиков;
- 4) выявлять и оценивать реальное качество образования, его соответствие стандартам.

Однако для практических целей под качеством образования решили понимать «качественные изменения» в учебном процессе и в среде, окружающей обучаемого, которое можно идентифицировать как улучшение знаний, умений и ценностей, приобретаемых обучаемым по завершению определенного этапа².

Таким образом, понятие «качество образования» можно сузить до понятия «качество подготовки». Обсуждение понятия «качество подготовки» можно вести на трех различных уровнях:

1-й уровень – *планирование обучения*, когда определенные представления о планируемом качестве подготовки закладывается в образовательные программы по каждому предмету;

2-й уровень – *этап реализации образовательных программ* в учебном процессе. В зависимости от обстоятельств, внешнего окружения и качества работы педагога можно говорить об определенном реализуемом качестве подготовки учеников и студентов;

3-й уровень – *качество подготовки обучаемых* рассматривается как *результат осуществления учебного процесса*. Этот уровень рассмотрения качества подготовки обычно в зарубежных исследованиях отождествляется с термином «учебные достижения»³. Отечественные же исследования подвергают сомнению такой подход, так как представление о качестве подготовки может ассоциироваться с внутренним состоянием обучаемого, в то время как результаты обучения проявляются во внешних наблюдаемых признаках и результатах учебного процесса. Однако это противоречие снимается при использовании специальных моделей

¹ Матрос Д.Ш., Полев Д.М., Мельникова Н.Н. Управление качеством образования на основе новых информационных технологий и образовательного мониторинга. – М.: Педагогическое общество России, 2001.–с 8.

² Чельшкова М.Б. Теория и практика конструирования педагогических тестов. М.:Логос, 2002.–с.420.

³ Указ. соч. – с.421.

измерения, соединяющих латентные параметры испытуемых с наблюдаемыми результатами выполнения заданий.

2. Оценка как элемент управления качеством учебного процесса.

Целью создания системы оценки качества является создание условий повышения качества образования через усиление потоков информации о результатах обучения и соответствующих факторах.

Под «качеством» учебного процесса (в соответствии со вторым и третьим уровнями обсуждения понятия «качество подготовки», см. предыдущий параграф) понимается система устойчивых характеристик в:

- деятельности педагога и обучаемых,
- методике преподавания,
- результатах обучения,

выраженных в форме качественных и количественных показателей, которые поддерживаются в ходе учебного процесса на определенном уровне, не ниже принятых норм и требований образовательной программы.

Критерий «качества» определяет специфику достигнутых результатов и указывает ту относительную границу, при которой эти результаты соответствуют (или не соответствуют) нормативным показателям, установленных теоретическим или опытным путем. «Качество» учебного процесса может быть выражено **по показателю обученности**, опосредованно отображая в нем выполнение тех требований, которые предъявляются образовательной программой к деятельности обучаемых и их результатам.

Качественная оценка – это единичная или совокупная определяющая характеристика педагогического объекта, в которой абстрагированы и обобщены его наиболее существенные свойства (качества) и представлены в форме вывода или заключения об этом объекте (иногда с указанием степени качества: больше-меньше, лучше-хуже).

Количественная оценка также имеет дело с качественной характеристикой объекта и отличается тем, что дополняет ее, уточняет и углубляет, что дает возможность ответить на вопрос: «сколько», «насколько больше», «насколько меньше». Количественную оценку нельзя противопоставлять качественной. В обоих понятиях речь идет об одних и тех же свойствах, лишь в последнем случае они выступают в единстве с числовыми значениями. Следовательно, количественная оценка по отношению к качественной всегда является вторичной, производной от первой. Важно понимать, что количественная оценка –

это не число, получаемое в результате измерений и суждений, а приписанное оценочному суждению значение. Количественные манипуляции с оценками недопустимы.

Большинство оценок результатов образования, понимаемого в широком современном толковании этого понятия, такие, как оценки личностного развития, различных сторон воспитанности, готовности поступать тем или иным образом в обычной или экстремальной ситуации, креативности и т.д., осуществляется квалиметрическим (качественным, описательным) путем. Известны и широко применяются следующие квалиметрические (описательные) методики определения результатов и качества образования:

- достижение прогноза, сделанного учителями (как одними из экспертов) в зоне ближайшего развития обучаемого;
- оценка поведения в естественно возникших экстремальных ситуациях, которые всегда представляет жизнь;
- оценка поведения учеников в специально продуманных и организованных педагогами диагностических ситуациях;
- достижение высшей качественной оценки по шкале критериев, разработанных самой школой;
- использование совокупности так называемых сквозных (то есть годных для оценки любых действий, поступков и т.п.) психологических показателей, например:
 - а) знает, как надо (интеллектуальный показатель);
 - б) способен реализовать (волевой показатель);
 - в) хочет реализовать (эмоциональный показатель).

Принцип построения элементарных критериев оценки заключается в выражении соотношения двух или нескольких *эмпирических величин*:

- затраченного времени к планируемому;
- числа правильных ответов по всему множеству заданных вопросов;
- числа правильно решенных задач ко всему перечню предложенных;
- числа допущенных ошибок к возможному среднему их значению;
- совокупность выполненных действий ко всей необходимой численности в ходе достижения цели.

Указанные критерии принадлежат к виду относительных показателей, значения которых выражаются отвлеченным числом или в процентах. Они просты в использовании и удобны для проведения элементарного педагогического анализа (сравнения эффективности отдельных фрагментов однородной деятельности обучаемых), но мало пригодны для оценки сложной учебной работы.

В настоящее время основой психолого-педагогических измерений служит пооперационный или поэлементный анализ, позволяющие предмет исследования расчленять на «равные» законченные части.

Однако при оценке предмета в целом не учитываются их весовые коэффициенты и не принимается во внимание характер и внутренние закономерности между структурообразующими компонентами.

В зарубежной педагогической литературе используется близкое к измерению понятие "*evaluation*", что можно перевести как "*оценивание*", имея в виду не только конечный результат, но и *процесс формирования* оценки. Соответственно выделяются два основных вида оценивания - *формирующее* и *итоговое*. Цель первого – оказывать, посредством оценивания, формирующее влияние на текущий процесс обучения, в смысле его улучшения, за счет установления обратной связи от студента к преподавателю. Цель второго – получить итоговые результаты обучения.

3. Мониторинг качества образования.

Мониторинг качества образования является важнейшим элементом системы качества (системы управления качеством) в учебном заведении, обеспечивая механизм обратной связи в системе управления качеством образования.

Под *мониторингом* понимается отслеживание состояния объекта и фиксация проблем, осуществляемые на основе систематизации и обработки, прежде всего, существующих источников информации, а также специально организованных исследований и измерений; прогнозирование тенденций и выработка рекомендаций по корректировке развития объекта мониторинга

Объектом педагогического мониторинга являются результаты учебно-воспитательного процесса и средства, которые используются для их достижения. Он позволяет определить, каково состояние образования на данном этапе, насколько рациональны педагогические средства, в какой мере они соответствуют заданным целям, насколько эффективны педагогические технологии.

Длительность слежения за состоянием педагогических явлений позволяет выявлять тенденции их изменения, устанавливая зависимость их от определенных факторов и условий. Таким образом, педагогический мониторинг представляет собой системную и лонгитюдную диагностику качества образования.

Можно выделить разные уровни мониторинга.

Макромониторинг организуется в масштабах общегосударственной системы, является основой разработки стратегии развития образования. Формы организации макромониторинга в России: развитие федеральной системы контроля качества образования, создание центра национального

тестирования, функционирование системы лицензирования, аттестации и аккредитации образовательных учреждений.

Микромониторинг организуется на региональном и муниципальном уровнях, а также в пределах конкретного образовательного учреждения и является основой разработки путей совершенствования учебно-воспитательного процесса.

Мониторинг качества образования в учебном заведении организуется для отслеживания и анализа состояния и тенденций развития образовательной системы высшего учебного заведения.

Мониторинг качества образования в учебном заведении заключается:

1) в сборе информации о состоянии основных видов деятельности учебного заведения, связанных с подготовкой выпускников (в случае вуза – специалистов с высшим образованием),

2) анализе этой информации с точки зрения качества образования,

3) выявлении динамики и тенденций развития образовательной системы учебного заведения,

3) подготовке и реализации необходимых управленческих решений, направленных на повышение качества образования.

Мониторинг качества образования в учебном заведении является частью комплексного системного мониторинга качества образования в России, осуществляемого на федеральном, региональном, отраслевом (уровне УМО) уровнях, уровне образовательных областей и отдельного образовательного учреждения (вуза).

Объектами мониторинга в высшем учебном заведении являются:

– *качество образования* под которым понимается качество подготовленности выпускников учебных заведений с точки зрения государственных и общественных потребностей, а также потребностей личности в получении образования;

– *система обеспечения качества образования*, включающая в себя следующие основные структурно-функциональные аспекты деятельности учебного заведения:

а) качество содержания образования;

б) состояние научно-педагогических кадров и условия преподавательской деятельности;

в) состояние организации учебного процесса и использования образовательных технологий;

г) состояние материальной базы, библиотечного, учебно-методического и информационного обеспечения учебного процесса в вузе;

д) состояние системы внутривузовского, внутришкольного контроля качества подготовки студентов и выпускников;

е) состояние системы управления образовательным процессом;

ж) состояние воспитательной работы в вузе.

Для активизации и координации работ по повышению качества образования в каждом вузе в настоящее время создается специальное подразделение, находящиеся в ведении учебного управления – Центр мониторинга качества образования. Основными задачами Центра являются:

- совершенствование методики мониторинга качества образования на основе непрерывного контроля уровня подготовки студентов;
- внедрение сквозного непрерывного мониторинга качества образования на основе специального программного пакета;
- формирование необходимого набора банков тестовых заданий путем приобретения готовых сертифицированных банков, а также за счет разработки собственных банков с последующей их сертификацией;
- подготовка и переподготовка преподавателей и специалистов в области технологий разработки банков тестовых заданий;
- формирование условий, организация работ и методическая поддержка разработки электронных банков тестовых заданий преподавателями и сотрудниками вуза;
- совершенствование методик проведения экспертизы качества программно-дидактических тестовых материалов.

Следует отметить *социальную обусловленность* мониторинга качества образования, поскольку сами процессы общего и профессионального образования, его содержание и формы, структура и организация на всех уровнях (государственном, региональном, местном и т. п.) подвержены постоянному и глубокому влиянию социальных, экономических, политических, духовных процессов. Поэтому при определенной устойчивости основополагающих параметров педагогического мониторинга необходимо постоянно отслеживать и учитывать изменяющиеся условия жизни общества и школы. Иными словами, мониторинг должен быть достаточно динамичен, иначе он может быстро превратиться в некую заданную схему, и его задачи могут свестись лишь к простой регистрации наблюдаемых фактов и явлений.

Благодаря такому подходу результаты мониторинга могут дать богатый материал для анализа и оценки, например, социальной мотивации, устойчивости или изменения поведения, профессиональных устремлений учащейся молодежи, а также для определения факторов, влияющих на формирование личности будущего специалиста.

4. Педагогический контроль. Функции и принципы педагогического контроля.

Управление любым процессом, в том числе и качеством образования, предполагает осуществление контроля, то есть определенной системы проверки его эффективности его функционирования. Крайне необходим он и для успешного протекания педагогического процесса. Благодаря контролю удовлетворяется потребность в получении информации об эффективности функционирования всей системы учебных заведений, осуществляется обратная связь в обучении. Суть проверки результатов обучения состоит в выявлении *уровня освоения* знаний обучаемыми, который должен соответствовать образовательному стандарту по данной программе, предмету.

Основные *задачи* контроля:

- выявление уровня правильности, объема, глубины и действенности усвоенных учениками знаний,
- получение информации о характере познавательной деятельности, об уровне самостоятельности и активности учащихся в учебном процессе,
- определение эффективности методов, форм и способов их учения.

Функции контроля:

- *диагностирующая* (определение объективно существующего уровня владения знаниями, умениями и навыками на конкретном этапе обучения, выявление пробелов и трудностей в усвоении учебного материала);
- *обучающая* (систематизация знаний и умении, их обобщение, их применение в разных познавательных ситуациях);
- *корректирующая* (соотнесение реального уровня сформированности знаний, умений и навыков с прогнозируемым и их совершенствование путем коррекции учебного процесса);
- *стимулирующая* (создание положительных мотивов овладения изучаемым курсом, повышение интереса к ее изучению);
- *развивающая* (развитие умений быстрой концентрации усилий для решения в определенный срок поставленной задачи, сосредоточенности, мобилизации внутренних резервов психики, самостоятельной мыслительной деятельности, логики, умения анализировать, обобщать, систематизировать, классифицировать и т.д.);
- *воспитывающая* (воспитание воли, настойчивости выдержки, самообладания, критического отношения к учебной деятельности,

культуры мышления, культуры общения с педагогом и со сверстниками, культуры организации учебной деятельности).

Важнейшие *принципы* контроля:

- *объективность* (заключается в научно-обоснованном содержании диагностических заданий, тестов, вопросов, диагностических процедур; - равном, дружеском отношении педагога ко всем обучаемым; точном, адекватном установленным критериям оцениванию знаний, умений);
- *систематичность* (означает осуществление контроля на всех этапах дидактического процесса – от начального восприятия знаний до их практического применения; регулярному диагностированию должны подвергаться все обучаемые с первого до последнего дня обучения в учебном заведении; комплексный подход к проведению диагностирования: различные формы, методы и средства контролирования, проверки, оценивания используются в тесной взаимосвязи и единстве, подчиняются одной цели);
- *наглядность (гласность)* (подразумевает проведение открытых испытаний всех обучаемых по одним и тем же критериям; рейтинг каждого учащегося носит наглядный, сравнительный характер; оглашение и мотивация оценок. Оценка – это ориентир, по которому учащиеся судят об эталонах требований к ним и об объективности педагога; объявление результатов диагностических средств, их обсуждение и анализ, составление планов ликвидации пробелов. Педагог, вынося оценку, должен каждый раз обосновывать ее, руководствуясь логикой и существующими критериями).

Выполняя функцию руководства учебно-познавательной деятельностью учащихся, контроль не всегда сопровождается выставлением оценок. Он может выступать как способ подготовки учащихся к восприятию нового материала, выявления готовности обучаемых к усвоению знаний, навыков и умений, их обобщению и систематизации. В процессе контроля формируются и совершенствуются как общие, так и специальные практические навыки и умения учащихся, т.е. контроль имеет важное образовательное и развивающее значение.

Контроль выполняет также большую воспитательную роль в процессе обучения. Он способствует повышению ответственности за выполняемую работу не только учащихся, но и преподавателя, приучению обучаемых к систематическому труду и аккуратности в выполнении учебных заданий, формированию у них положительных нравственных качеств и коллективистских отношений, созданию здорового общественного мнения.

5. Традиционные виды, методы и формы контроля

По *месту* контроля на основных этапах обучения различают предварительный, текущий, тематический и итоговый контроль. По *характеру получения информации* в процессе контроля выделяют устный, письменный и лабораторный контроль. По средствам, используемым в ходе контроля, его подразделяют на машинный и безмашинный. По способу организации контроля все ранее названные виды его можно разделить на программированные и непрограммированные.

На различных этапах обучения используются различные виды контроля. Предварительный контроль направлен на выявление знаний, умений и навыков обучаемых по предмету или разделу, который будет изучаться, то есть он проводится с диагностической целью перед изучением новой темы или в начале учебного года, четверти. Его цель – *ознакомиться с общим уровнем подготовки учащихся по предмету*. На основе полученных результатов преподаватель планирует, если необходимо, повторение (объяснение) материала, учитывает эти результаты в дальнейшей организации учебно-познавательной деятельности обучаемых.

Текущий контроль осуществляется в повседневной учебной работе и выражается в систематических наблюдениях преподавателя за учебно-познавательной деятельностью обучаемых на каждом занятии. Главное его назначение – *оперативное получение объективных данных об уровне знаний обучаемых и качестве учебно-воспитательной работы на занятии*. Полученная информация о том, как обучаемые усваивают учебный материал, как формируются их умения и навыки, помогает преподавателю наметить рациональные методы и приемы учебной работы, правильно дозировать материал, осуществлять постоянное руководство учебной деятельностью обучаемых, активизировать внимание и пробуждать интерес к изучаемому материалу. Можно сказать, что текущий контроль решает задачи руководства учебным процессом.

Тематический (периодический) контроль обеспечивает выявление и оценку знаний и умений обучаемых, усвоенных не на одном занятии, а в целом цикле занятий. Его цель – установить, насколько успешно обучаемые владеют **системой** определенных знаний, каков общий уровень их усвоения, отвечает ли он требованиям образовательного стандарта. Периодический контроль проводится, как правило, после изучения логически завершенной части учебного материала- темы, подтемы, нескольких тем, раздела или полного курса.

Итоговый контроль обеспечивает усвоение знаний и умений обучаемых за более длительный период обучения: за четверть, полугодие, семестр, год или по окончании курса обучения. Цель его – установить систему и структуру знаний обучаемых. Итоговый контроль учитывает результаты всех других видов контроля. Особая форма итогового контроля – зачеты и экзамены.

Для получения информации о подготовленности учащихся, об их конкретных знаниях используются различные методы. **Методы** контроля – это способы диагностической деятельности, позволяющие осуществлять обратную связь в процессе обучения с целью получения данных об успешности обучения, эффективности учебного процесса.

Они должны обеспечивать систематическое, полное, точное и оперативное получение информации об учебном процессе.

Методы *устного контроля* – это беседа, рассказ учащегося, объяснение, чтение текста, технологической карты, схемы, сообщения об опыте и пр. Основу устного контроля составляют монологический ответ учащихся (в итоговом контроле это более полное, системное изложение) и/или вопросно-ответная форма – беседа, в которой преподаватель ставит вопросы и ожидает ответа учащихся. Устный контроль, как текущий, проводится ежеурочно в индивидуальной, фронтальной или комбинированной форме. Индивидуальный опрос учащихся позволяет учителю получить более полные и точные данные об уровне усвоения, однако он оставляет пассивными на уроке других учеников, что вынуждает учителя решать проблему их занятости во время опроса. Фронтальный опрос занимает всех учащихся сразу, но дает более поверхностное представление об усвоении.

Зачет и *устный экзамен* являются наиболее активной и обстоятельной проверкой знаний за определенный период обучения. Экзамены как способ проверки знаний вызывают неоднозначную оценку как педагогов, так и учащихся. Отличаются два главных недостатка экзамена с традиционным выбором экзаменационных билетов учащимися. В “вытягивании” удачного/неудачного билета есть элемент случайности (“экзамен - лотерея”). Второе – экзамен является в известной степени стрессовой ситуацией для экзаменуемого, часто блокирующей его интеллектуальные возможности. Однако человек всю жизнь в процессе своей деятельности подвергается проверке и оценке, поэтому и учащиеся должны быть готовы к контролю и воспринимать ситуацию экзамена как стандартную, что, кстати, и свойственно большинству экзаменуемых. Поэтому экзамен остается эффективной формой проверки знаний, умений, развития мышления, сформированности взглядов, отношений, оценок.

У каждого экзаменатора даже при наличии общих критериев оценки знаний имеется свой подход к экзамену, свое понимание педагогических принципов требовательности и справедливости, свое отношение к оценке. Один особенно ценит точность изложения и основательность знаний, другой – оригинальность; одни экзаменаторы являются воплощением строгости, другие – самой доброжелательностью. Субъективизм и волюнтаризм в оценке устного опроса и экзамена ставит школьников в неравные условия, делает оценки несопоставимыми между собой. Все это говорит о том, что существующую систему методов контроля знаний у учащихся нужно изменить и, прежде всего, стандартизировать как по содержанию, так и по структуре и формам поведения.

Письменный контроль (контрольная работа, изложение, сочинение, диктант, реферат) обеспечивают глубокую и всестороннюю проверку усвоения, поскольку требует комплекса знаний и умений учащегося. В письменной работе нужно показать и теоретические знания и умения применять их для решения конкретных задач, проблем. Письменные работы позволяют документально установить уровень усвоения материала, но требуют больших затрат времени.

Выполнение *практических работ* можно считать эффективным, но малоприменяемым способом проверки результатов обучения. Имеется в виду, прежде всего, проведение учеником лабораторных опытов, создание изделий, монтаж аппарата и пр.

6. Современные методы и формы педагогического контроля (тестовый, модульно-рейтинговый)

В условиях сокращения учебного времени в учебных заведениях одной из проблем становится повышение эффективности контроля знаний, для проведения которого необходимо сочетать традиционные и нетрадиционные его методы и формы.

К традиционным обычно относят, во-первых, субъектно-объектные отношения в системе “учитель-ученик”, во-вторых, экзамены, коллоквиумы, зачеты, контрольные работы и опросы. Им, несмотря на их значимость и важность, присущи определенные недостатки: пассивность учащихся, психологическая боязнь экзаменов, эпизодичность проверки и другие. Как правило, они носят характер интуитивного оценивания, часто недостаточно объективны. Именно педагогический субъективизм является главной причиной, по которой нынешние школьники отдают предпочтение компьютерным и тестовым формам контроля с минимальным участием педагогов.

Поэтому в системе контроля очень важным является использование нетрадиционных методов контроля, которые позволяют преодолеть данные ограничения. К одному из таких методов относится тестовый контроль.

Тестовый контроль – целенаправленное, одинаковое для всех испытуемых обследование, проводимое в строго контролируемых условиях, позволяющее объективно измерять изучаемые характеристики педагогического процесса. Тестирование отличается от других методов обследования точностью, простотой, доступностью, возможностью автоматизации. Тест позволяет получить весьма информативный результат проверки, помогает определить, какие из исследуемых факторов влияют в наибольшей степени на успеваемость, какие – в меньшей или не влияют вовсе.

Контроль с использованием тестов или заданий тестового типа в последнее время всё больше внедряется в учебный процесс. Он находит всё новые и новые области применения, в частности при построении модульно-рейтинговой системы аттестации студентов.

Модульно-рейтинговая система складывается из двух взаимосвязанных и дополняющих одна другую частей (модульной и рейтинговой), которые могут функционировать и по отдельности, но с меньшей эффективностью.

Под модулем понимают логически завершённую часть учебного материала, обязательно сопровождаемую контролем знаний и умений студентов. Модуль совпадает с темой учебной дисциплины или блоком взаимосвязанных тем, но в отличие от темы, в рамках модуля все измеряется, все оценивается: посещение студентом занятий, выполнение им аудиторных и внеаудиторных заданий, исходный, этапный и итоговый уровень учебных достижений студентов.

В структуре модуля должны содержаться следующие элементы: цель, задачи, уровень овладения, содержание учебного материала, формы, виды, методы учебной деятельности студента, направленные на овладение учебным материалом, рекомендации по организации и методике самостоятельной работы студента по данному модулю, задания для самоконтроля и самооценки результатов обучения. Число модулей зависит от сложности и значимости учебного материала, от необходимости и возможной частоты контроля. Модуль содержит теоретическую и практическую части. Первая формирует теоретические знания, вторая – профессионально важные умения и навыки.

Целью создания модуля является достижение заранее запланированного результата обучения. С помощью модулей достигается индивидуализация темпа обучения. Модульная система имеет целью

поставить студента перед необходимостью регулярной учебной работы в течение всего семестра.

Исходя из понятия «модуль», под учебно-тематическим модулем мы понимаем комплекс учебных материалов по теме теоретического и практического характера, подлежащих усвоению на заданном уровне на основе целенаправленно организованной, лично-ориентированной учебно-познавательной деятельности студентов, а также обязательному измерению и оцениванию ее результатов.

Основной целью контроля является определение степени достижения задач обучения, воспитания и развития. Контроль по модулю может быть содержательным, деятельностным или содержательно-деятельностным. Результаты контроля являются основой для оценивания.

Одной из современных организационно-методических форм контроля является рейтинговый контроль. Понятие "рейтинг учебных достижений" является категорией педагогической квалиметрии. В условиях модульно-рейтинговой технологии обучения и контроля используются различные шкалы оценивания: количественная абсолютная, количественная относительная, порядковая. Эталоны оценки учебных достижений разрабатываются на основе образовательных стандартов, таксономии учебных целей.

Максимально возможная сумма баллов предметного рейтинга может устанавливаться по-разному: либо в соответствии с объемом часов, отводимых учебным планом на аудиторную и внеаудиторную работу по данной дисциплине, либо в соответствии с суммарной оценкой всех заданий по учебной дисциплине. При этом величина оценки за выполнение конкретного вида учебно-познавательной деятельности выбирается самим преподавателем. Полученные студентом баллы при прохождении модульных видов контроля суммируются и составляют его рейтинг по данной дисциплине. В соответствии с рекомендациями методики высшей школы и требованиями педагогической квалиметрии минимальный рейтинг должен быть не менее 70 % от максимального. Перевод набранной суммы баллов в 4-балльную шкалу оценки осуществляется с учетом набранной суммы баллов в процентном отношении от максимально возможной суммы баллов.

Деятельность по разработке рейтинга по учебной дисциплине может быть представлена системой последовательных и взаимосвязанных действий: от составления рабочей учебной программы модульного типа до разработки структуры и содержания учебных занятий, предусматривающих самоконтроль и самооценку различных видов учебно-познавательной деятельности студентов.

Рейтинговый контроль создает условия для формирования навыков самоконтроля и самооценки результатов учения, превращает студента в

субъекта контроля и оценивания процесса и результатов своей учебно-познавательной активности. Рейтинговый контроль позволяет повысить объективность и полноту контроля, более эффективно реализовать его организующую, обучающую, воспитывающую и развивающую функции, способствует переводу внешнезаданных целей обучения во внутреннезаданные.

Таким образом, особенности модульно-рейтингового обучения заключаются в следующем: в модульном построении содержания учебного курса и организации его изучения; акценте на самостоятельную работу студентов по модулю; индивидуализации темпа, времени овладения модулем; наличие четких критериев оценки различных видов учебно-познавательной деятельности и ее результатов; широкое использование тестового результатов обучения (при этом контролю подлежали не только знания, умения и навыки, но и такие социально и личностно важные качества, как учебная мотивация, уровень притязаний, способность к самоконтролю и самооценке учебных достижений и т.д.); постоянном самоконтроле и самооценке результатов учебной деятельности студентами; осуществлении мониторинга и управления качеством учебных достижений студентов.

7. Папка индивидуальных учебных достижений «портфолио»

Одним из современных средств оценивания результатов обучения является портфолио.

Термин «портфолио» пришел в педагогику из политики и бизнеса: каждому знакомы понятия министерский портфель, портфель инвестиций. Portfolio – 1) предмет для хранения и переноски письменных работ, документов и т.д. (портфель, «дело», папка); 2) должность министра; 3) серия работ, документов, бумаг и т.д., объединенных общей темой: портфель ценных бумаг – security portfolio, portfolio investments – портфельные инвестиции, папка с рисунками.

Портфолио (в широком смысле слова) – это способ фиксирования, накопления и оценки индивидуальных достижений школьника в определенный период его обучения. Портфолио относится к разряду «аутентичных» (то есть истинных, наиболее приближенных к реальному оцениванию) индивидуализированных оценок, ориентированных не только на процесс оценивания, но и самооценивания.

Портфолио – это рабочая файловая папка, содержащая многообразную информацию, которая документирует приобретенный опыт и достижения учащихся. Портфолио позволяет учитывать результаты, достигнутые учащимся к разнообразным видам деятельности

– учебной, творческой, социальной, коммуникативной и др. и является важнейшим элементом практико-ориентированного подхода к образованию. Можно отметить, что портфолио дополняет традиционные контрольно-оценочные средства. Конечная цель учебного портфолио заключается в доказательстве прогресса обучения по результатам, по приложенным усилиям, по материализованным продуктам учебно-познавательной деятельности.

Основной смысл портфолио – «показать все, на что ты способен».

Педагогическая философия портфолио предполагает:

- смещение акцента с того, что учащийся не знает и не умеет, на то, что он знает и умеет по данной теме, разделу, предмету;
- интеграцию количественной и качественной оценок;
- перенос педагогического удара с оценки на самооценку

Важная *цель* портфолио – представить отчет по процессу образования обучаемого, увидеть «картину» значимых образовательных результатов в целом, обеспечить отслеживание индивидуального прогресса ученика в широком образовательном контексте, продемонстрировать его способности практически применять приобретённые знания и умения.

Оценка тех или иных достижений (результатов), входящих в портфолио, а также всего портфолио в целом (либо за определенный период его формирования) может быть как качественной, так и количественной.

Портфолио не только является современной эффективной формой оценивания, но и помогает решать важные педагогические *задачи*:

- поддерживать высокую учебную мотивацию школьников;
- поощрять их активность и самостоятельность, расширять возможности обучения и самообучения;
- развивать навыки рефлексивной и оценочной (самооценочной) деятельности учащихся;
- формировать умение учиться – ставить цели, планировать и организовывать собственную учебную деятельность.

Состав портфолио зависит от конкретных целей обучения. Именно цели являются свидетельством усилий, достижений и прогресса в обучении учащегося по конкретному предмету, разделу, теме. Каждый элемент портфолио желательно датировать, чтобы можно было проследить динамику прогресса учащегося. Так же желательно, чтобы при оформлении окончательного варианта портфолио включал три элемента:

- описание цели сбора портфолио, его предназначения и краткого описания;

– содержание (или оглавление) портфолио с перечислением его основных элементов;

– перспективы дальнейшего использования, т.е. взгляд в будущее.

«Портфолио документов» – портфель сертифицированных (документированных) индивидуальных образовательных достижений. Этот раздел предполагает возможность как качественной, так и количественной оценки материалов портфолио. Итоговый документ представлен в виде вкладыша в аттестат и содержит итоговый балл, а также перечень и оценки сертификатов, входящих в его состав. Документы или их копии могут быть помещены в приложении к портфолио. В «Портфолио документов» ученик представляет сертификаты официально признанных на международном, федеральном, региональном, муниципальном уровне конкурсов, соревнований, олимпиад, документы об участии и грантах, об окончании музыкальных или художественных школ, сертификаты о прохождении тестирования, например TOEFL и т.п.

«Портфолио работ» представляет собой собрание различных творческих и проектных работ ученика, а также описание основных форм и направлений его учебной и творческой активности: участие в научных конференциях, конкурсах, учебных лагерях,хождение элективных курсов, различного рода практик, спортивных и художественных достижений и др. Данный раздел портфолио предполагает качественную оценку: например, по параметрам полноты, разнообразия и убедительности материалов, качества представленных работ, ориентированности на выбранный профиль обучения и др. Портфолио может оформляться в виде творческой книжки ученика с приложением его работ, представленных в виде текстов, электронных версий, фотографий, видеозаписей.

В «Портфолио работ», учащийся представляет, во-первых, «ведомость» или «зачетную книжку» о прохождении им курсов по выбору. Во-вторых, может представить сами работы, проекты, исследования, которые он выполнил в ходе обучения на курсах по выбору или при обучении в учреждениях дополнительного образования, других учреждениях или выполнил самостоятельно, не получив за это подтверждающего документа. В-третьих, могут быть представлены сертификаты обучения, конкурсов, соревнований и т.п., не имеющих «официального признания» на уровне федерации, региона, муниципалитета. В-четвертых, ученики могут представлять собственные изделия: например модели, картины, стихи, музыкальные произведения собственного сочинения, фотографии, компьютерные программы и т.п.

«Портфолио отзывов» включает в себя характеристики отношения школьника к различным видам деятельности, представленные учителями,

родителями, возможно одноклассниками, работниками системы дополнительного образования и др., а также письменный анализ самого школьника своей конкретной деятельности и ее результата. Портфолио может быть представлен в виде текстов заключений, рецензий, отзывов, резюме, эссе, рекомендательных писем и пр.

В «Портфолио отзывов» ученикам предлагается представлять отзывы на творческие работы, исследовательские и другие проекты, социальные практики, участие в конференциях и самых различных сферах приложения сил. Важной составляющей данного раздела является самооценка ученика, его рефлексия на собственную деятельность. Бывает, что учащийся представляет неплохие результаты по каким-нибудь предметам или в конкурсах, олимпиадах, а выполняет при этом работу с большим нежеланием, о чем часто никто не подозревает. В результате он может получить рекомендации заниматься на профиле, где он будет успешен, но не будет испытывать морального удовлетворения. Помочь школьнику «найти себя», грамотно выстроить профориентацию будет возможно используя этот раздел портфолио, где будет представлена саморефлексия ученика на разнообразную выполняемую им деятельность, начиная с учебной и урочной и заканчивая хобби.

Метод портфолио обеспечивает дифференцированный подход к принятию решений, который повлияет на отношение детей к работе и школе в целом. Это освобождает преподавателя от ограничений, связанных со стандартизацией исследований. Наконец, используя портфолио, преподаватель имеет возможность расширить работу в классе и с каждым ребенком. Кроме того, использование портфолио позволяет сосредоточить свое внимание на ребенке и развивать межличностные отношения с ним.

Это особенно важно преподавателям и учащимся для совместной работы, для того, чтобы выделить критерии, которые будут использованы в качестве основы для оценивания развития учащихся, как полностью (т.е. в процессе всего учебного периода), так и частично (то есть, как часть законченного проекта, деятельности, или оценки для выявления степени, которую необходимо достигнуть по определенным стандартам).

8. Оценка: понятие, функции, критерии.

Понятие «оценка» в философской литературе используется в широком и узком смысле. В широком смысле оценка понимается как сравнение с эталоном изучаемого объекта, категоризация объекта. Такую оценку называют объективной, так как она опирается на достижения науки, подтверждается практикой. В узком смысле оценка отражает

значимость объекта для субъекта, соотношение объекта с целями, потребностями и интересами субъекта. Такую оценку называют субъективной, так как она дает объективно-субъективную информацию о ценностях, а не о сущностях.

Следовательно, педагогическая оценка используется в так называемом «узком смысле» и носит, в сущности, субъективный характер.

Функции оценки: стимулирующая; фиксирующая; ориентирующая; информирующая; воспитательная, контролирующая.

Критерии выставления оценки:

- *нормативный* критерий – результат контроля оценивается с точки зрения установленной нормы, образца проверяемого объекта;
- *сопоставительный* критерий – результат контроля одного обучаемого сравнивается с аналогичным результатом другого;
- *личностный* – результат контроля обучаемого сравнивается с прошлыми результатами этого же ученика и тем самым устанавливается динамика его продвижения в овладении оцениваемыми знаниями, умениями и навыками.

Формы оценок:

- *нормативная* оценка – используется, чтобы оценить насколько знания и умения соответствуют тому уровню, который принят за норматив;
- *критериальная* оценка – используется для определения содержания знаний студента или для выявления его конкретных предметных учебных навыков;
- *процессуальная* оценка – осуществляется для того, чтобы понять, как протекает учебная деятельность студента.

Негативные стороны современной оценочной системы.

1. Различные виды познавательной деятельности, выполняемые студентом и являющиеся неравнозначными, оцениваются по единой оценочной шкале.

2. При выставлении оценки не проводится различие между собственно успеваемостью и личностью студента; не учитывается тот путь, те усилия, которые стоят за каждым из оценочных фактов.

3. Зачастую носят травмирующий характер. Процесс учебы заменяется спорами, как и за что ставятся отметки, о справедливости различных моментов этой процедуры.

В связи с этим при оценивании учащихся рекомендуется придерживаться следующих **принципов**:

- учащиеся должны постоянно иметь в виду, по каким критериям будет оцениваться их деятельность;
- оценка не должна обладать карающими функциями, как это часто встречается в современном российском образовании;

- необходим баланс проверки знаний и умений;
- оценивание должно производиться как за групповую работу, так и за индивидуальную;
- оценка должна осуществляться как на основе индивидуального прогресса, так и на основе абсолютных достижений.

Проблема оценивания чрезвычайно актуальна. Деятельность оценивания имеет наибольшую по времени продолжительность в профессиональной деятельности педагогов и ценностное отношение к миру у обучаемых формируется именно как отражение преподавательского оценивания. К сожалению, господствующим пока является представление, что вузовский преподаватель должен хорошо знать в первую очередь свой предмет, а воспитание переадресовывается к родителям, молодежным организациям, соответствующим структурам вуза и т.п.

Попытки изменения существующей систем оценивания сводятся к:

- изменению содержания, качественного наполнения оценки (например, вербализация отметки; учет суммарного «веса» рассматриваемых параметров; введение рейтинговой системы);
- изменению количественного выражения оценки (вместо пятибалльной предлагается 12, 50, 100-балльная шкала)
- включению аутентичной оценки в систему оценивания результатов обучения школьников.

Понятие аутентичной оценки введено группой американских педагогов (Линн, 1995; Шепард, 1989; Валенсия, 1990), предложивших взгляд на оценку как на естественный многоаспектный процесс. Они высказали необходимость пересмотра самого процесса оценивания и смещения акцента в сторону оценки *реальной деятельности школьников* на занятии, а не ее результата, predeterminedенного заранее взрослыми. При этом может быть оценена любая, не только мыслительная деятельность ученика. Это позволяет ребенку реализовать имеющиеся возможности. Аутентичная оценка, по сути своей, гуманна, поскольку она позволяет каждому ученику быть успешным.

Таким образом, в современных условиях предъявляются высокие требования к методологии и методике оценивания личностных и учебных достижений обучаемых. Для обеспечения объективности и надежности оценки учебных достижений должна вестись планомерная работа по широкому внедрению тестовых методик оценки учебных достижений в педагогическую практику.

Тема II. ИСТОРИЯ РАЗВИТИЯ СИСТЕМЫ КОНТРОЛЯ ЗНАНИЙ И ТЕСТИРОВАНИЯ

9. История развития системы контроля знаний и тестирования за рубежом.

Контролю качества подготовки придавалось большое значение еще в странах древнего мира (Вавилон, Египет, Китай, Греция, Рим), особенно это касалось лиц, претендующих на должности государственных служащих, врачей, военнослужащих. К участию в испытаниях в качестве «экзаменаторов» привлекались ученые, высшие государственные чиновники и даже главы государств. В зависимости от особенностей государства испытания выпускников проводились в разных формах (письменно и устно) и в разной обстановке, причем нередко - в специально создаваемой обстановке стресса.

К концу XIX века в Европе сложились **две традиции в контроле знаний**. В одной (ее условно можно назвать *немецкой*) упор делался на **устные** формы экзамена, в присутствии комиссии, состоявшей, по меньшей мере, из двух человек. На экзамене давались один или два сравнительно объемных вопроса, ответ на которые должен был свидетельствовать об уровне знаний всего курса. Другая традиция, которую можно назвать *английской*, – это **письменная** форма контроля, во время которого испытуемым давалось десять–двенадцать коротких заданий из разных тем. В странах Западной Европы традиционны были устные экзамены, и много лет ушло на то, чтобы прижилась письменная форма контроля знаний.

Существовал и другой подход к оценке уровня подготовки – «бесконтрольный», когда все было отдано на откуп преподавателю (или группе преподавателей), поскольку считалось, что только он мог оценить успехи учащегося должным образом. Ни о каком объективном сопоставлении оценок при этом не приходилось и говорить. Борьба тенденций – «за контроль» и «против контроля» с переменным успехом продолжалась на протяжении столетий и формально закончилась только в 20 веке.

Появление первых тестов относится к концу 19 века. Ф. Гальтон, а позже Дж. Кеттел проводили измерения некоторых физиологических и психомоторных параметров и возможностей испытуемых, результаты измерений подвергали статической обработке и на основании этого делали определенные выводы. В начале 20 века была сделана попытка измерения интеллектуальных способностей детей разного возраста. Для этого А. Бине и Т. Симон применили систему заданий (тест), снабженную инструкцией по проведению тестирования. Пригодность, эффективность,

трудность и дифференцирующая способность заданий, включенных в тест, были предварительно оценены путем их предъявления детям разных возрастных групп и сопоставления результатов тестирования с мнением преподавателей, обучавших их. Первые опыты тестирования показали, что на основании результатов выполнения заданий можно делать выводы о скрытых возможностях и способностях человека.

Применение тестов для определения профпригодности началось во время первой мировой войны, что было продиктовано сложностью военной техники и высокой стоимостью подготовки обслуживающего ее персонала. В послевоенное время интерес к тестам возрастает: делаются попытки их применения при приеме на работу, для профориентации, в социологических и психологических исследованиях и, наконец, для оценки знаний школьников и студентов. Особенно больших успехов в области применения тестов достигли в США.

Популярность тестов росла, но одновременно усиливалась и ужесточалась их критика. Это было вызвано, с одной стороны, несовершенством самих тестов, а с другой - различного рода злоупотреблениями при тестировании, в том числе, этического характера. А поскольку применение тестов часто касалось жизненно важных для людей вопросов, связанных с учебой и работой, то в значительной степени критика была обоснованной, хотя и не всегда объективной.

Тесты, направленные на исследование свойств личности, были по своей сути психологическими. Первые тесты для объективного контроля знаний, умений и навыков появились в начале XX века. Они быстро завоевали популярность среди преподавателей вузов и школ в Англии и США, а позже в России и СССР. Их стали в США называть *педагогическими*.

Были разработаны и внедрены тесты О. Стоуна по арифметике, Б. Зекингема для проверки правописания, Э.Торндайка по диагностике большинства школьных предметов. Т.Келли разработал способ измерения интересов и наклонностей обучаемых (при изучении алгебры), Ч.Спирмен предложил общие основы использования корреляционного анализа.

С 1931 г. начинается современный этап развития школьного тестирования. Поиски специалистов направляются на повышение объективности тестов, создание непрерывной системы школьной диагностики, подчиненной единой идее и общим принципам, более совершенных средств предъявления и обработки тестов, накопления и эффективного использования информации.

Наиболее развитые в тестовом отношении страны – Нидерланды, США, Англия, Япония, Дания, Франция, Израиль, Финляндия, Канада, Австралия, Новая Зеландия и др.

В США тесты применяются в двух основных сферах: в образовании и в сфере профотбора-профориентации. Каждый год знания и способности учащихся проверяются там с помощью 47 миллионов тестовых бланков, выпускаемых примерно 400 тестовыми компаниями. Самая крупная - это ETS (Educational Testing Service). Ежегодной проверкой интеллектуальных способностей и знаний подвергаются практически все студенты высших и учащиеся средних учебных заведений. В США нежелание тестироваться расценивается как слабость и неспособность показать, в честной конкуренции, свою образованность, тестирование стало там нормой жизни, обязательным элементом систем образования.

В области определения системы оценок (отметок) наблюдается большое разнообразие как в принципах, так и в конкретных подходах, выборе способов оценивания и выставлений оценок. В зарубежных школах практикуются различные системы оценивания знаний, умений, навыков, приняты различные шкалы оценок, включая сто-, двенадцати-, десяти-, двухбалльную и др. Во французском лицее, например, при сдаче выпускных экзаменов результаты определяются по 20-балльной шкале. При этом для каждого предмета устанавливается определенный весовой коэффициент, определяющий значимость данного предмета для избранной обучаемым специальности, профиля или отделения лицея. Благодаря этому оценки по профилирующим предметам приобретают большую весомость.

При значительной децентрализации большинства зарубежных школьных систем, при условии свободного выбора школой учебных планов и программ, а обучаемыми — учебных предметов редко проводятся централизованные проверки и оценивание знаний, умений. В Великобритании, например, функцию практического контроля за качеством академической подготовки берут на себя университетские комиссии, к которым на период сдачи выпускных экзаменов прикрепляются все грамматические и привилегированные школы, ориентирующиеся в подготовке своих выпускников на дальнейшее университетское образование.

В целом же нужно признать, контроль и учет успеваемости, диагностирование обученности в зарубежных школах носят характер объективной констатации результатов. Он не предполагает заботу о всеобуче. Общепринятый принцип индивидуализации обучения диктует один подход – каждый идет своим путем и темпом, обучается в меру своих возможностей, потребностей, реальных оценок будущего.

10. Развитие оценивания и тестирования в России.

Россия конца XIX века позаимствовала в основном немецкий вариант проверки знаний с упором на устные формы контроля знаний. До революции в России существовала шестибальная система оценки знаний с баллами от нуля до пяти. В начале 1918 г. оценка «0» была упразднена. Но постепенно и оценка «1» стала использоваться все реже. В школах РСФСР с мая 1918 года были отменены все экзамены – вступительные, переходные и выпускные как «не дающие правильного представления о знаниях и уровне умственного развития учащегося и оказывающего на него разрушительное действие». Была отменена также балльная система оценки. Для поступления в вуз в это время не требовалось никакого документа, который бы удостоверял уровень образования. Порочность «бесконтрольного» подхода к оценке знаний, не позволяющего судить об их реальном уровне, быстро привела не только к его снижению, но и к заметному падению у учащихся мотивации к учебе. Экзамены были снова введены в 1932 году. Начиная с 50-х годов все меньше стала использоваться и оценка «2». Пятибалльная система оценок фактически превратилась в трехбалльную, а для большинства обучаемых, которые не могут учиться на «4» и «5», эта шкала стала двухбалльной. «Ступенька» между тройкой и четверкой стала непреодолима для большинства обучаемых. Однако многие педагоги используют «дополнения» к обычной пятибалльной шкале в виде знаков «плюс», «минус». Реально получается три градации пятерки («пять с плюсом», «пять», «пять с минусом»), три градации четверки (аналогично), три градации тройки и двойки. Получается типичная десятибалльная шкала оценок.

В нашей стране судьба тестирования в образовании в это время связана с прикладным направлением педагогики – педологией. **Педология (наука о комплексном развитии ребенка)** начавшая развиваться в России в начале века, безоговорочно приняла тестовую основу объективного школьного контроля.

В 1936 году это направление, также как и тестирование, было запрещено. Пострадали авторы выпускавшихся в конце 20-х - начале 30-х годов сборников "Тесты: теория и практика". Это М.Я. Басов, М.С. Бернштейн, П.П. Блонский, А.П. Болтунов, С.М. Василейский, психолог, логик И.Челпанов. Всего в 1937 году было репрессировано более 300 ученых - педагогов и педологов. Тесты запретили в 1936 году как "буржуазные и вредные", но истинные причины запрета - расстановка сил в политической борьбе того периода. Эта ситуация продолжалась в течение примерно сорока лет. В течение всего этого времени печатались в основном статьи о вредности тестов, о недопустимости использования последних в отечественной педагогической науке и практике. Типичные

аргументы оппонентов тестового метода сводились, в обобщенном виде, к следующим утверждениям:

- тесты используются в капиталистических странах, где с их помощью решаются вопросы расовой и классовой дифференциации;
- применение тестов унижает достоинство личности, особенно в случаях, когда получаемые баллы оказываются ниже среднего уровня;
- никакие методы измерения не могут заменить преподавателя и его личный опыт;
- в педагогике нет и не может быть точной единицы измерения, и потому не следует терять время, силы и средства на разработку неточных методов.

Педагогика того времени потеряла интерес к тестам и к проведению эмпирических исследований; отчасти поэтому она в течение многих лет оставалась одной из немногих общественных наук, лишенных качественного эмпирического базиса.

Начиная с 30-х годов, в контроле знаний наметились две основные тенденции. Первая связана с применением традиционных форм и длительным, растянувшимся на полвека, периодом активного игнорирования тестов. Эта тенденция привела в наши дни к тотальному субъективизму в выставлении оценок, несопоставимости оценок, к образовательно-технологическому и организационно-педагогическому отставанию школы от общемирового образовательного процесса, к несопоставимым с имеющимся качеством выпускников большим затратам живого труда преподавателей. Не секрет, что исторически субъективность у нас всегда играла большую роль, чем объективность. В сфере образования это проявлялось в повсеместной приверженности к традиционным экзаменам, в то время как во многих странах уже давно тот, кто учит, не берется ставить итоговые оценки учащимся за свою же, по существу, работу. Там это считается нарушением профессиональной этики и потому общественное мнение склоняется в пользу независимого тестового контроля знаний.

Состояние застоя в нашей стране наблюдалось вплоть до начала 80-х годов 20 века, когда интерес к тестированию стал возрождаться, однако ситуация с ними не только не улучшилась, а в некотором отношении даже ухудшилась. Потому что кажущаяся простота создания тестов породила множество некачественных самоделок, дискредитирующих этот перспективный метод научной организации самоконтроля и объективного педагогического контроля знаний. Тестирование становилось своего рода модой, особенно в престижных школах, где тесты начинают применять вместо приемных и выпускных экзаменов. Отчасти это происходило из-за того, что наиболее распространенным и признанным в мире методом объективной оценки знаний считался именно тест.

Одной из форм оценки знаний выпускников стал единый государственный экзамен. Многие страны мира ввели ЕГЭ десятки лет тому назад, и эта система успешно работает. Среди них отметим такие страны, как Англия, Германия, Австралия, Турция, Китай и др. Система Единого государственного экзамена введена в Казахстане.

В России также имеется некоторый опыт проведения подобных экзаменов. На протяжении шести (с 1999 по 2005 гг.) лет Центр тестирования при МО РФ проводил для желающих школьников из разных регионов России централизованное тестирование. Более 100 вузов принимали сертификаты этого Центра при приеме абитуриентов. В течение шести лет независимая аттестация выпускников проводится в ряде школ Москвы Лабораторией аттестационных технологий при Московском институте повышения квалификации работников образования. И наконец, летом с 2001 по 2006 гг. был проведен эксперимент по тестированию выпускников общеобразовательных школ по правилам Единого государственного экзамена. К 2009 году ЕГЭ должен стать обязательным на всей территории России.

Тема III. ПЕДАГОГИЧЕСКИЕ ТЕСТЫ

11. Современные подходы к созданию тестов, их сопоставление

В настоящее время в зарубежной педагогической практике существуют несколько подходов к созданию тестов. Первый подход, базируется на применении стандартных для математической статистики корреляционных и факторных методов анализа. Второй, современный подход зародился в рамках теории латентно-структурного анализа, основная цель которого – оценить скрытые (латентные) параметры испытуемых. Пионерской работой в области второго подхода является труд датского математика Г. Раша¹ вышедший в 1960 г. Именно он дал толчок интенсивному развитию за рубежом серьезной теоретической базы тестирования, вылившейся в современную теорию тестирования – Item Response Theory (IRT). Первые сведения по этой теории на русском языке появились только в 1995 г. в работе Р. Хамблетона².

Основная цель состоит в разработке математической модели процесса тестирования, параметрами которой, подлежащими определению, служат различные характеристики участников

¹ Rash G. Probabilistic Models for Some Intelligence and Attainment Tests, 1960, Copenhagen, Denmark: Danish Institute for Educational Research

² Чельшкова М.Б. Разработка педагогических тестов на основе современных математических моделей, 1995, М.: с31.

тестирования и самого теста. В настоящее время IRT представляет собой довольно обширную и, к сожалению, непростую теорию, активно использующую разнообразный арсенал математических методов¹. Поэтому зачастую целесообразно бывает использовать сочетание двух упомянутых теорий. С помощью статистических методов классической теории проводится первичный анализ качественных характеристик полученного варианта теста, а с помощью IRT, более углублённый анализ характеристик тестовых заданий. Наиболее эффективным использование IRT становится при сформированном банке тестовых заданий, что обеспечивает возможность автоматической генерации тестов, с определёнными наперёд заданными и научно-обоснованными характеристиками точности и надёжности.

Независимо от выбранной концепции, существенным моментом в тестологии является начальный этап конструирования, на котором происходит формирование целей создания тестов. Этап целеполагания является наиболее трудным и вместе с тем наиболее важным: от результатов его выполнения зависит в первую очередь качество содержания теста или содержательная валидность, являющаяся базовым качественным критерием как для критериально-ориентированных так и для нормативно-ориентированных тестов. Оценка тестовых заданий на пригодность к достижению поставленных целей всего теста, реализуется с помощью экспертов.

Таким образом, на данный момент создание тестов реализуется в рамках двух базовых итерационных подходов: классический подход обеспечивает начальный этап создания, а также оценки качества тестов, и современный IRT-подход, позволяющий получить устойчивые оценки латентных параметров испытуемых и трудности тестовых заданий и поддерживающий конструирование тестов на продвинутом этапе, причём применение математического аппарата обеих этих теорий целесообразно только при грамотно сформированной таксономии целей тестового контроля и удовлетворяющих ей тестовых заданиях. Ниже кратко изложены основные возможности создания тестов в рамках классической и современной теории тестирования, см. таблицу.

¹ Handbook of Modern Item Response Theory. Editors Win J. van der Linden, R. K. Hambleton, 1997, Springer-Verlag, New York, p.510.

<i>Классическая теория.</i>	<i>IRT</i>
<p data-bbox="244 204 479 228"><u>Основная задача теста:</u></p> <p data-bbox="132 231 563 395">Получить значение истинного балла (Т) испытуемого исходя из наблюдаемого результата (X), с учётом случайной ошибки измерения (E), откуда вытекает основной постулат классической теории тестирования:</p> $X_i = T_i + E_i$ <p data-bbox="288 426 434 450"><u>Возможности:</u></p> <p data-bbox="132 453 563 531">Извлечь первичную информацию о тесте, на основании матрицы результатов апробационного тестирования, а именно:</p> <ol data-bbox="132 534 563 1225" style="list-style-type: none"> 1) Оценить статистическую сложность заданий 2) Интеркорреляцию между заданиями теста и корреляцию баллов заданий и внешнего критерия (суммы баллов испытуемых), для определения валидности тестовых заданий 3) Быстро оценить качества теста на основе графического вида кривой распределения тестовых баллов испытуемых. 4) Получить оценку надёжности результатов тестирования посредством корреляционного анализа баллов испытуемых по тесту, либо по нескольким его вариантам. 5) Построить доверительный интервал, в пределах которого находится истинный балл испытуемого, либо получить точечную регрессионную оценку. 6) Интерпретировать результаты в терминах Z-шкалы, либо процентильной шкалы, то есть определить место(рейтинг) испытуемого в выборке. 	<p data-bbox="675 204 910 228"><u>Основная задача теста:</u></p> <p data-bbox="596 231 977 367">Получить устойчивую объективную оценку латентного параметра уровня знаний испытуемого по исследуемому предмету, не зависящую от конкретного теста.</p> <p data-bbox="714 426 865 450"><u>Возможности:</u></p> <ol data-bbox="596 453 977 1393" style="list-style-type: none"> 1) Установление связи между латентными параметрами испытуемых и наблюдаемыми результатами выполнения теста. Наблюдаемые результаты выполнения теста порождаются взаимодействием двух множеств латентных параметров теста: уровнем знаний испытуемых и трудности заданий. 2) Параметры уровня знаний испытуемого и трудности заданий теста отображаются в единую шкалу логитов, что позволяет реализовать идею адаптивного тестирования, когда для каждого испытуемого (с конкретным уровнем знаний) отбираются задания определённой сложности. 3) Существует возможность помимо стандартных критериев качества, ввести новый критерий-эффективность теста, путём введения особого класса информационных функций, посредством которых происходит оценка количества информации, обеспечиваемое j-м заданием при уровне знаний θ_i. 4) Реально только в рамках IRT реализована возможность применения не только рейтинговой, но и интервальной шкалы, а это значит уровень подготовки можно оценить количественно.

12. Достоинства и недостатки педагогического тестирования

Тесты дают возможность заметно улучшить образовательный процесс, потому что обладают рядом *преимуществ* перед другими методами контроля знаний:

- снижаются затраты на проверку знаний,
 - более полный охват содержания дисциплины (минимум 30 заданий в тесте против 2 – 5 вопросов в экзаменационном билете),
 - помогают выявить индивидуальный темп обучения,
 - позволяют систематизировать срез знаний и умений учащихся,
 - выявляются пробелы в текущей и итоговой подготовке
- Тестовому методу присущи:
- гибкость, приспособляемость ко многим требованиям, уровням, условиям,
 - равные условия и права всем сдающим тесты,
 - более высокая точность оценки знаний и построения относительных шкал успешности владения предметом,
 - технологичность, особенно при компьютерном тестировании.

В сочетании с персональными ЭВМ, тесты помогают перейти к созданию современных систем адаптивного обучения и контроля - наиболее современных организации учебного процесса.

Кроме того, тесты позволяют наладить самоконтроль – самую полезную для обучения и гуманную форму контроля знаний, а также организовать рейтинг – эффективное средство повышения учебной мотивации.

При традиционных формах контроля оценки заведомо субъективны, поскольку основаны на предъявлении нестандартизированных заданий и зависят, в конечном счете, от квалификации и добросовестности преподавателя или комиссии, выставляющих оценку. В силу этого оценки, полученные учащимися в разных учебных заведениях одного профиля и даже у разных преподавателей в одном учебном заведении, не могут быть сопоставлены. Пятибалльная шкала (а на самом деле она трехбалльная, ведь оценка «1» при аттестации не выставляется вообще никогда, а оценку «2» преподаватели стараются не выставлять по мере сил), принятая при проведении традиционных испытаний позволяет лишь грубо ранжировать учащихся, не предоставляя возможности сравнить уровень учащихся с одинаковыми оценками. Применение такой шкалы имеет далеко идущие последствия, а именно вынужденное дублирование экзаменов за среднюю школу и вступительных при поступлении в вуз.

С помощью тестов, используя специальные шкалы, можно объективно оценить и сопоставить уровень подготовки, можно выявить

структуру полученных знаний (т.е. выявить пробелы в усвоении учебного материала), а также оценить качество преподавания.

Подчеркнем, что указанные преимущества тестов реализуются в результате большой подготовительной работы по их составлению и анализу и затраты значительного времени и средств.

Недостатки тестирования.

1. Результаты теста указывают не на причинную связь, а на главную тенденцию (на результаты теста могут повлиять и побочные обстоятельства—непониманием задачи, случайно неправильный выбор буквы вследствие усталости, нервозности)

2. Тесты выявляют результат, а не ход работы. Возможно, тестируемый рассуждал правильно, а ответ написал неправильно, или случайно правильно ответил.

3. Тесты недостаточно выявляют эмоционально-волевую сферу деятельности, старательность ученика, его интерес к работе.

4. Учебные тесты упрощают задачу, стоящую перед учеником. Вместо того чтобы самостоятельно найти нужный ответ, воспроизвести в памяти учебный материал, или составить рассказ, в котором даётся полный и исчерпывающий ответ на поставленный вопрос, от ученика требуется лишь выбрать правильный ответ из числа предложенных ему готовых вариантов.

Таким образом, следует признать, что тесты не в состоянии охватить все стороны учета успеваемости. Только наряду с другими формами контроля тесты с успехом можно использовать.

13. Тест, педагогический тест, предтестовое задание.

Обычно любой вид проверки или испытаний называют тестированием. В переводе с английского тест означает пробу, испытание.

Тест – совокупность заданий стандартной формы, по результатам выполнения которых можно судить о психофизиологических и личностных характеристиках, а также о знаниях, умениях, навыках испытуемого.

Различают *педагогические* и *психологические* тесты. Предметом психологического тестирования являются психические свойства, а объектом – носитель *психических свойств – человек*, а педагогического – измерение уровня усвоения определенных знаний, умений и навыков, предусмотренных учебной программой. В области психодиагностического обследования системы представлений граница между психологическими и педагогическими тестами размыта.

Педагогический тест – это система заданий определенной формы, расположенных по возрастанию трудности, которая дает возможность измерить уровень подготовки испытуемых и оценить структуру их знаний.

Ключевым в этом определении является слово «измерить», которое приписывает тесту свойства измерительного инструмента. Соответственно, педагогическое тестирование процесс измерения уровня подготовки испытуемых и способ оценки структуры их знаний (в дальнейшем слово «педагогическое» будет опущено).

Создание теста, который бы отвечал требованиям, содержащимся в определении, – задача простая лишь на первый взгляд. Эта обманчивая простота во многом способствовала дискредитации тестирования, породив появление, в том числе и в печати, огромного количества так называемых «тестов», большинство из которых в лучшем случае можно назвать лишь «тестоподобными» контрольными материалами. Результат применения таких продуктов для контроля знаний вызывает недоверие. Причины этого различны, например, при тестировании разные по уровню подготовки испытуемые могут получить одинаковые баллы, велика вероятность угадывания правильного ответа и т.д.

Следует сразу оговориться, что любые виды и формы контроля, помогающие преподавателю в его работе, имеют право на существование, но не следует опрометчиво считать их тестами.

Предтестовое задание – это единица контрольного материала, которая по содержанию и структуре обеспечивает однозначность оценок испытуемых в выбранной шкале. Не каждое предтестовое задание становится тестовым, для этого требуются многочисленные проверки и улучшения содержания. Требования к предтестовому заданию:

- наличие порядкового номера задания по априорной шкале трудности задания;
- наличие стандартной инструкции;
- наличие эталона правильного ответа к заданию и правила для оценки результатов его выполнения.

Предтестовое задание превращается в тестовое после доведения его характеристик до уровня, соответствующего определенным научно обоснованным критериям качества.

14. Основные требования, предъявляемые к тестам. Характеристики тестов.

Тест должен отвечать следующим основным требованиям.

1. Тест должен быть пригоден для измерения того, что предполагается измерять. Пригодность теста в целом и заданий, входящих в него, для заявленной цели измерения называется **валидностью**.

Для определения валидности требуется независимый *внешний критерий*, то есть оценка эксперта (преподавателя). Валидность зависит от качества заданий, их числа, от степени полноты и глубины охвата содержания учебной дисциплины (по темам) в заданиях теста. Кроме того, валидность зависит также от баланса и распределения заданий по трудности, от метода отбора заданий в тест из общего банка заданий, от интерпретации тестовых результатов, от организации сбора данных, от подбора выборочной совокупности испытуемых.

В современной технологии выделяются следующие виды валидности: очевидная валидность; валидность по содержанию (содержательная); конструктивная (концептуальная); валидность по критерию (критериальная или эмпирическая валидность); сравнительная или конкурентная валидность. Для всех тестов учебных достижений должна быть определена содержательная валидность, а для тестов, используемых для аттестации учащихся, содержательная и критериальная валидность.

2. Инструмент должен иметь определенную точность (*ошибку измерения*). Любой тестовый балл можно представить как сумму истинного и ошибочного компонентов измерения:

$$X = T + E,$$

где E – символизирует некоторую ошибку (или точнее, ошибки измерения, проистекающие по различным причинам); измеряемое значение (X) не равно истинному (T). Погрешность может оказаться случайной или систематической. Последнюю можно учесть, в случае необходимости, в виде поправок, и потому для теории тестов они не представляют интереса. Другое дело – **случайные ошибки**, вызванные состоянием испытуемого, случайное изменение некоторых условий проведения теста, различиями в наборах тех или иных заданий, предлагаемых различным испытуемым и многое другое. Взятые вместе, они рассматриваются как случайная ошибка измерения, хотя слово "ошибка" при этом надо трактовать не в привычном смысле допущенной (а потому и легко устранимой) оплошности, а в смысле неизбежной погрешности, определяемой неконтролируемыми факторами.

3. Тест должен быть **надежен**, т.е. результаты, полученные с его помощью, должны быть воспроизводимы, а их значения устойчивы. **Надежность** теста – это характеристика методики, отражающая точность измерений знаний, а также устойчивость результатов теста к действию посторонних случайных факторов. **Надежность теста** ρ тем выше, чем более согласованы результаты одного и того же человека при повторной проверке знаний посредством того же теста или эквивалентной его формы (**параллельного теста**).

Параллельный тест, имея внешне отличающееся содержание, должен, концептуально, измерять то же самое, что и исходный тест, с той же точностью. **Параллельными** называются варианты, которые имеют сходное предметное содержание в пределах укрупненной учебной единицы, равные средние арифметические, равные дисперсии и равные интеркорреляции.

Если измерение валидно, оно должно быть надежно. Однако надежность еще не гарантирует валидности, поскольку надежность зависит от случайных ошибок измерения, а валидность – как от случайных, так и от систематических. Иными словами, надежный тест может быть невалидным, в то время как валидность теста предполагает его надежность.

4. Должна быть **шкала**, обеспечивающая определенный уровень измерений. По степени возрастания уровня измерения шкалы подразделяются на: номинальные, порядковые, интервальные, относительные (характеристики отдельных шкал ниже).

5. Тест должен быть **стандартизирован**, т.е. должны быть выработаны определенные нормы его выполнения (среднее, стандартное отклонение и т.д.) для интерпретации результатов применительно к испытуемым из различных выборок, и сертифицирован.

6. Тест должен быть построен из заданий возрастающей трудности. Трудность определяется по соотношению правильных и неправильных ответов на тестовые вопросы. Трудность задания можно определяться двояко:

1) умозрительно, на основе предполагаемого числа и характера умственных операций, необходимых для успешного выполнения заданий;

2) после эмпирического апробирования заданий, с подсчетом доли неправильных ответов (отношения числа неправильных ответов W_j к числу испытуемых N)

$$q_j = W_j / N .$$

Содержание теста не может быть только легким, средним или трудным. Легкие задания теста создают только видимость наличия знаний у учащихся, потому что ими проверяются минимальные знания. Искажает результаты тестирования и подбор заведомо трудных заданий,

в результате чего у большинства школьников оказываются заниженные баллы.

Вариация и дифференцирующая способность: если на какое-то задание правильно отвечают все тестируемые, то такое задание становится не тестовым. Испытуемые отвечали на него одинаково; между ними нет вариации. Не тестовым надо считать и то задание, на которое нет ни одного правильного ответа. Вариация по нему также равна нулю. Нулевая вариация означает практическую необходимость удаления задания из проектируемого теста. Оно, для данной группы, не тестовое. Возможно, в другой группе это задание заработает, но это будет задание уже другого, а не данного теста, если под тестом понимать метод и результат измерения знаний.

Тестовые задания должны обладать **дискриминационной (дискриминативной) способностью** – способностью дифференцировать обучаемых на худших и лучших.

Применение регулярного тестирования не приводит к повышению уровня тревожности у учащихся. Более того, возможно и его снижение. **Тестовая искущённость** – приобретение обширной практики выполнения тестов.

15. Классификации педагогических тестов.

Педагогические тесты весьма разнообразны.

1. По **содержанию** тесты делятся на гомогенные и гетерогенные.

Гомогенный тест – тест, обладающий дисциплинарной общностью заданий. Гомогенные тесты распространены больше других. В педагогике они создаются для контроля знаний по одной учебной дисциплине или по одному разделу.

Гетерогенный тест – охватывает систему заданий по нескольким учебным дисциплинам. Гетерогенные тесты бывают:

- полидисциплинарными (состоящими из гомогенных субтестов);
- междисциплинарными (для выполнения которых потребуются применение некоторых обобщенных, интегративных знаний и умений).

Обычно гетерогенные тесты используются для комплексной оценки выпускника школ, оценки личности при приеме на работу и для отбора наиболее подготовленных абитуриентов при приеме в вузы.

2. С точки зрения **целей применения**, можно выделить:

- тесты **достижения**;
- **критериально-ориентированные** тесты, позволяющие сопоставить уровень индивидуальных учебных достижений с полным объемом знаний, умений и навыков (в тесте **образовательный стандарт**

выступает в виде **критериального** балла – минимального тестового балла, который должен получить обучаемый, чтобы считаться усвоившим тот или иной материал);

- **нормативно-ориентированные тесты**, сравнивающие испытуемых друг с другом по уровням и учебным достижениям;
- **аттестационные тесты**, определяющие степень обученности;
- тесты **прогнозирования** результатов обучения.

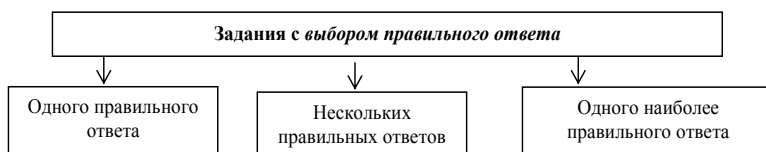
3. В зависимости *от характера деятельности* испытуемого при выполнении педагогического теста различают:

- тесты выбора верных ответов из общего числа предложенных в задании (альтернативные тесты);
- тесты со свободно конструируемыми ответами;
- тесты на дополнение ответов, приведенных в заданиях.
- комбинированные

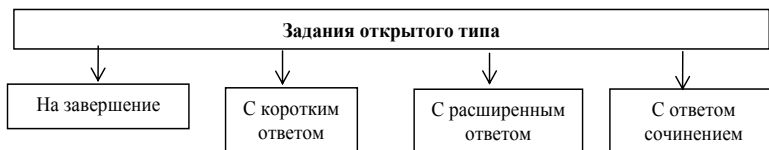
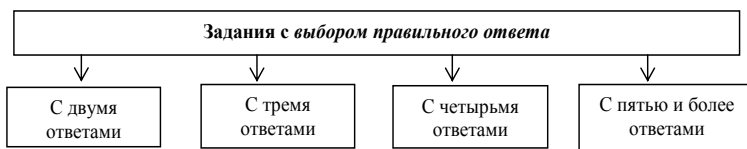
Формы тестовых заданий



Виды заданий с выбором правильного ответа



Варианты заданий с выбором правильного ответа



4. По *характеру задач*, решаемых преподавателем с помощью тестов в процессе обучения, выделяют

1) задачи, стоящие перед преподавателем *на входе* в обучение (решаются с помощью тестов, нацеленных на выявление степени овладения базовых знаний, умений и навыков, а также на определение степени владения *новым* материалом до начала его изучения);

2) *текущие задачи* (решаются с помощью *формирующих* тестов, нацеленных на осуществление дифференцированного подхода к успевающим и неуспевающим ученикам в повседневном учебном процессе, *диагностических тестов*, нацеленных на установление причин пробелов в знаниях учеников и их устранение);

3) задачи, соответствующие *концу* определенного периода учебного процесса (решаются с помощью *тестов достижений* (*суммативных тестов*)). Основная цель итогового тестирования – обеспечение объективной оценки результатов обучения после завершения некоторого курса. Оценка ориентирована, во-первых, на характеристику степени освоения содержания курса, во вторых – на дифференциацию учеников по результатам выполнения теста.

Классификация педагогических тестов в системе образования в соответствии с задачами приведена на рисунке.



5. Различают тесты традиционные и нетрадиционные. Перечисленные выше тесты можно отнести к традиционным. **К нетрадиционным относят адаптивные тесты** – это тесты, где

используются способы регулирования трудности и числа предъявляемых заданий в зависимости от ответа учеников. Каждый учитель понимает, что хорошо подготовленному ученику нет необходимости давать легкие и очень легкие задания. Из-за высокой вероятности неправильного решения нет смысла давать трудные задания слабому ученику. После появления компьютеров эта мера легла в основу методики адаптивного контроля знаний. При успешном ответе следующее задание более трудное, при неуспешном - легкое. Адаптивный тест представляет собой вариант автоматизированной системы тестирования с заранее известными параметрами трудности и дифференцирующей способностью каждого задания. Эта система создана в виде компьютерного банка заданий, упорядоченных в соответствии с интересующими характеристиками заданий. Прежде чем попасть в банк, каждое задание проходит эмпирическую апробацию на достаточно большом числе типичных учащихся интересующего контингента, в результате чего определяется уровень трудности каждого задания. Основной причиной появления адаптивных систем, основанных на принципе индивидуализации обучения, стала недостаточная информированность о реальном уровне знаний учеников и естественные различия в их способностях усвоить предлагаемые знания.

16. Модели педагогического тестирования.

Одним из основных моментов при создании и использовании тестов является правильный *выбор модели педагогического тестирования* – схемы предъявления тестовых заданий и оценивания результатов тестирования. При сравнении моделей тестирования первоочередное значение имеет надежность тестовых результатов — это показатель, зависящий от точности измерения знаний, умений или навыков заданиями теста.

Ниже описываются основные модели тестирования.

1. *Классическая модель.* Из n имеющихся заданий случайным образом выбираются k ($k < n$), на которые и отвечает обучаемый. Результат ответа на каждое задание теста — «правильно» или «неправильно». Результатом выполнения теста является процент правильных ответов обучаемого. Это простейшая модель, и она имеет массу недостатков, основной из которых — задания не разделяются по сложности. У данной модели самый низкий показатель надежности.

2. *Классическая модель с учетом сложности.* Отличается от предыдущей модели тем, что для каждого задания определен уровень сложности. Подсчет результатов ведется с учетом сложности заданий, на которые обучаемый дал правильный ответ. Основной недостаток: заранее

не определено, какие по сложности задания достанутся обучаемому: одному могут достаться k легких, а другому — k сложных. Иначе говоря, используя эту модель невозможно добиться параллельности тестов по сложности (как и в предыдущем модели). Показатель надежности у данной модели также недостаточно высок.

3. *Модель с возрастающей сложностью.* Имеется n заданий. Для каждого задания определен уровень сложности. Имеется m уровней сложности. Обязательным условием при композиции теста является наличие в тесте заданий всех уровней сложности. Из множества заданий выбираются и сортируются в порядке возрастания сложности k заданий. Эти задания предъявляются обучаемому. Важное условие: распределение заданий в тесте по уровням сложности должно подчиняться нормальному закону. Конечный балл прямо пропорционален сложности заданий, на которые были даны правильные ответы. У данной модели надежность выше, чем у предыдущей, поскольку обеспечивается параллельность тестов по сложности.

4. *Модель с разделением заданий по уровням усвоения.* Уровень усвоения — это показатель, характеризующий качество владения учебным материалом. Согласно данной модели, для каждого из уровней усвоения составляются задания. Проводится тестирование обучаемого для каждого уровня в отдельности, начиная с нижнего. Для перехода на каждый последующий уровень необходимо, чтобы степень владения учебным материалом, определенная на текущем уровне, была не ниже заранее определенного минимального для данного уровня. Степень владения материалом определяется как отношение количества правильно выполненных заданий текущего уровня к общему количеству заданий текущего уровня в тесте. Уровень усвоения учебного материала используется для оценивания качества знаний у обучаемого и выставления оценки согласно заранее определенным критериям.

5. *Модель с учетом времени ответа на задание.* В данной модели время ответа обучаемого на задание теста влияет на результат по тесту в целом. В основе данной модели лежит предположение о том, что длительное время размышления над ответом свидетельствует о плохой подготовленности обучаемого. Для каждого задания в тесте задается время, достаточное для того, чтобы обучаемый прочитал задание, осмыслил его и ответил. Это время может задаваться, в частности, как функция от сложности задания, или как функция от индивидуальных характеристик обучаемого, которые определяются заранее. Описанная модель дает хорошие результаты при использовании ее совместно с моделью с учетом сложности заданий.

6. *Модель с ограничением времени на тест.* Из n имеющихся заданий случайным образом выбираются k ($k < n$), на которые и

предлагается ответить обучаемому. Устанавливается максимальное время выполнения для всего теста. Результат тестирования формируется лишь из тех заданий, на которые обучаемым был дан правильный ответ. Сам тест может быть построен по моделям 1-4.

7. *Адаптивная модель.* Данная модель может быть использована только при компьютерном тестировании. Адаптивным называется тест, в котором сложность последующих заданий испытуемого зависит от правильности ответов на предыдущие задания: чем более правильные ответы на предыдущие задания, тем сложнее последующие. Предусмотрена возможность повторного задания вопросов по некоторым темам для более точного определения уровня знаний испытуемого в данной области. Тестирование заканчивается, когда обучаемый выходит на некоторый постоянный уровень сложности, например, отвечает подряд на несколько вопросов одного уровня сложности.

8. *Модель на нечеткой математике.* Нечеткая логика, на которой основана нечеткая математика, ближе к человеческому мышлению, чем традиционная логика. Данная модель тестирования является развитием любой из предыдущих моделей, где вместо четких характеристик тестовых заданий и ответов используются их нечеткие аналоги. Например, уровень сложности задания в этой модели может быть «легким», «средним» и т.д., а процент правильных ответов «маленьким», «средним» и т.д. При использовании данной модели преподавателю легче оценивать сложность, а испытуемому более понятна его оценка.

На основе описанных возможно построение комбинированных моделей: модели с возрастающей сложностью и ограниченным временем тестирования или модели на нечеткой математике с учетом времени ответа на задание и т.п.

17. Этапы создания тестов.

Работа над созданием теста проводится поэтапно.

1 этап. Она должна начинаться с *определения цели* использования теста и отбора содержания дисциплины в соответствии с этой целью и вида теста (он может быть критериально- или нормативно-ориентированным).

2 этап. Отбор учебного материала, подлежащего тестовому контролю, и его спецификация. Определяется круг тем, включаемых в тест, и относительное количество заданий, которым должен быть представлен каждый раздел курса. При планировании содержания теста следует заранее предусмотреть: количество заданий по каждой из контролируемых тем (в соответствии с образовательным стандартом), их оптимальную форму, желательный уровень трудности каждого из них,

количественное соотношение заданий разной трудности и время на их выполнение. Необходимо следить за соблюдением пропорций содержания теста и тем дисциплины; не перегружать тест малозначимым материалом, который не вносит заметного вклада в определение уровня подготовки.

3 этап. Создание заданий в тестовой форме по всему курсу или по проверяемой его части, объединение их в тематические группы, комплектование первичного, пробного, теста. На этом этапе работы автор должен также руководствоваться требованиями государственного образовательного стандарта, а также интуицией, здравым смыслом и опытом преподавательской работы. Очевидно, что задания должны быть ясными, точно и четко сформулированными, оптимально краткими, определенными и нетривиальными. Выбор формы заданий (одной или нескольких) диктуется целью тестирования и особенностями содержания контролируемой дисциплины. Задания одной формы объединяются в блоки и снабжаются стандартной инструкцией выполнения.

Составитель теста определяет правило оценивания заданий, систему подсчета тестового балла, способ предъявления теста испытуемым т.д. Все это, во избежание неопределенностей и двусмысленностей, затрудняющих выполнение теста и интерпретацию его результатов, должно быть отражено в инструкциях для испытуемых и проводящих тестирование преподавателей и другого персонала.

Результатом этого этапа является создание системы так называемых **предтестовых** заданий, т.е. инструмента, которым *предположительно* можно измерять. Назвать тестом его можно пока еще только условно. Прежде, чем им пользоваться, он должен пройти *экспертизу*, при которой априорно оцениваются все шаги указанных выше этапов, и при необходимости коррекцию.

4 этап. Проверка первичного теста на группе испытуемых (студентов или учащихся). Для этого необходимо проведение выборочного исследования (апробации теста на репрезентативной выборке). Создание представительной выборки достаточного объема, в которой основные признаки генеральной совокупности были бы представлены в той же пропорции и с той же частотой, – процесс сложный и дорогостоящий. Независимо от вида теста, формы входящих в него заданий, и способа предъявления его испытуемым результаты тестирования представляют собой **матрицу**, в строках которой записаны числа, кодирующие ответы каждого испытуемого на все задания, а в столбцах – числа, кодирующие ответы всех испытуемых на каждое задание. В простейшем случае при использовании дихотомической оценки матрица состоит из нулей (неверно выполненное задание) и единиц (верно выполненное задание). Сумма единиц в строке равна

тестовому баллу испытуемого, сумма единиц в столбце – количество испытуемых, правильно выполнивших задание. В более сложных случаях возможно применение политомической оценки, введение весовых коэффициентов для заданий, учет вопросных пропусков (если задание пропущено, оно не считается выполненным неверно) и т.д. В любом случае тестовый балл, полученный суммированием баллов за каждое задание, называется наблюдаемым (первичным, сырым).

5 этап. Статистический анализ результатов первичного тестирования, выбраковка и корректировка тестовых заданий. Апробации теста включает в себя статистическую обработку результатов в рамках классической и современной теории тестов с целью определения и последующего уточнения его характеристик: трудности и дифференцирующей способности заданий, их корреляции друг с другом и с тестовым баллом, гомогенности теста и т.д. В процессе анализа выясняют, насколько распределение отличается от нормального, строят характеристические кривые заданий и информационную кривую теста. Здесь указано далеко не все: многое из того, что следует сделать, остается за рамками даже простого перечисления. Регрессионный, корреляционный, факторный, кластерный, дистракторный и другие виды анализа проводят с применением современных программных средств. На основании статистического анализа делают выводы о валидности, надежности, эффективности и информативности теста, а также составляют рекомендации по каждому заданию.

Задания, неудачные по тем или иным характеристикам, подвергаются либо исправлению и доработке, либо удалению из теста и замене другими. Последнее все же нежелательно, поскольку на этапе апробации вновь создаваемого теста заданий с заведомо хорошими характеристиками еще просто не существует. Только после апробации задания, включенные в тест, могут быть названы тестовыми.

6 этап. Формирование из прошедших проверку заданий собственно теста, который должен состоять из заданий в тестовой форме возрастающей трудности с учетом необходимого уровня усвоения знаний и максимально охватывающих всю программу дисциплины. После внесения в тест исправлений и изменений снова проводится тестирование репрезентативной выборки, коррекция и т.д.

7 этап. Эмпирическая проверка теста для уточнения педагогических характеристик как отдельных тестовых заданий, так и всего теста в целом, его валидности, надежности и др. После нескольких итераций и достижения нужного уровня характеристик тест можно считать стандартизированным (строго говоря, только такой продукт и можно считать тестом). Задания, включенные в стандартизированный тест, называются калиброванными и могут быть помещены в банк заданий.

Создание и пополнение банка калиброванных заданий - также весьма важная задача. Большой объем банка позволит создавать тесты с заранее заданными свойствами, вплоть до индивидуальных тестов, которые будут особенно востребованы в связи с развитием дистанционной формы обучения. В наиболее передовых в тестологическом отношении странах (США, Австралии и т.д.) существуют банки заданий по разным дисциплинам и блокам дисциплин, содержащие десятки тысяч единиц. Причем доступ к ним открыт. Попутно отметим, что открытость банка не противоречит соблюдению конфиденциальности самого теста, для обеспечения которой принимаются специальные меры.

18. Определение целей тестирования.

Этап целеполагания является наиболее трудным и вместе с тем наиболее важным: от результатов его выполнения зависит качество содержания теста. В процессе целеполагания преподавателю необходимо решить вопрос о том, какие результаты обучаемых он хочет оценить с помощью теста. Для создания средств измерения необходима предварительная конкретизация целей.

1. Непосредственно для целей создания теста наибольший интерес представляют этап, связанный с конкретизацией результатов обучения. Планируемые результаты обучения в сознании преподавателя-предметника преломляются сквозь призму содержания знаний по преподаваемой им учебной дисциплине. В состав *требований* к планируемым результатам обучения обычно включают:

- систему изучаемых объектов,
- описание видов учебной деятельности,
- описание качества усвоения учебного материала.

К важнейшим элементам системы знаний можно отнести понятия и факты, законы, теории, идеи, знания о способах деятельности, методологические и оценочные знания.

При создании теста ставится задача отобразить в его содержании то главное, что должны знать ученики в результате обучения. **Необходимо структурировать цели и ввести определенную иерархию в их взаимное расположение.** Работа по построению системы целей была проведена Б. С. Блумом¹. В настоящее время классификация целей (или, как ее иначе называют, таксономия целей), является наиболее известной в мировой педагогической литературе. Понятие "таксономия" обозначает

¹ Bloom B.S. et al. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain. N.Y.: David Mckay Co., 1956.

такую классификацию и систематизацию объектов, которая построена на основе их естественной взаимосвязи и используется для описания категорий, расположенных последовательно, по нарастающей сложности. Б. С. Блум выделяет 6 категорий целей обучения: **знания, понимание, применение, анализ, синтез, оценка**. Принципиальная особенность таксономического подхода определяет основную задачу обучения как формирование основы, на которой базируется оценка, путем овладения знаниями на уровне всех категорий таксономии учебных целей. То есть "Оценка", являясь высшей категорией таксономии учебных целей и используя все другие категории, не обязательно становится последним этапом процесса решения задачи, а может быть прелюдией к приобретению новых знаний, новому пониманию или применению, анализу или синтезу.

В последнее время таксономия Б.С. Блума подвергалась значительной критике в связи с недостаточным отражением современных достижений в области психологии обучения. К числу недостатков ими были отнесены излишняя упрощенность моделей, не позволяющая использовать современные теории процесса обучения, избыточное внимание к оценке результата обучения, а не к процессу формирования результата. В нашей стране работ таксономия Б.С. Блума также неоднократно подвергалась критике, в частности, из-за того, что в ней произошло методологически недопустимое смешение конкретных результатов обучения (знания, понимание и т.д.) с операциями, представляющими необходимое условие их достижения (анализ, синтез, оценка). В этой связи в пособии предлагается иная классификация целей. В ее основу положен уровневый системный подход описания достижений учащихся, который позволяет сгруппировать результаты обучения в зависимости от уровней учебной деятельности.

Первый уровень связан с непосредственным воспроизведением по памяти содержания изученного материала и его узнаванием.

Второй уровень предполагает понимание и применение знаний в знакомой ситуации по образцу, выполнение действий с четко обозначенными правилами.

Третий уровень включает применение знаний в измененной или незнакомой ситуации.

2. После постановки целей в общем виде переходят к конкретизации планируемых результатов обучения. Основные категории конкретизации учебных целей, разработанные М. В. Клариным¹ в виде, удобном для задач измерения, приводятся в таблице:

¹ Кларин М. В. Инновационные модели обучения в зарубежных педагогических поисках. М., 1994.

<i>Обобщенные учебные цели, планируемые преподавателем</i>	<i>Примеры конкретизированных учебных целей, достигнутых учеником</i>
Знание на уровне запоминания и воспроизведения	Знает смысл употребляемых терминов, основные понятия и определения, формулы, законы, принципы
Знания на уровне понимания	Понимает и интерпретирует термины, интерпретирует понятия и определения, преобразует словесный материал в математические выражения, интерпретирует словесный материал на схемах и графиках
Умения по применению знаний в известной ситуации	Умеет применять термины, понятия и определения в знакомой ситуации по образцу, а также формулы, законы и принципы в знакомой ситуации
Умения по применению знаний в незнакомой ситуации	Использует законы и принципы в новых ситуациях, осуществляет перенос известных методов на незнакомые ситуации
Анализ	Видит ошибки и упущения в логике рассуждений, корректирует неполные или избыточные постановки задач, выделяет скрытые предположения, проводит различия между фактами и следствиями

3. Очередной этап конкретизации целей связан с *переходом* от планируемых результатов обучения к тому, что можно отобразить в средствах педагогического измерения, например в тесте. На этом этапе операционализация позволяет структурировать, расчленять, а иногда, наоборот, укрупнять, уточнять и детализировать содержание дисциплины для его **отображения в содержании теста**. Именно этот этап обеспечивает **переход** от теоретического уровня анализа целей обучения к эмпирическому и позволяет концептуально правильно подойти к выделению эмпирических индикаторов – заданий теста.

4. Следующий этап конкретизации целей сопряжен с задачами *планирования содержания теста*. Далеко не весь набор целей обучения по разным причинам можно отобразить в содержании теста. Конечно, чем глубже и полнее отображение, тем выше содержательная валидность теста, тем больше оснований доверять тестовым баллам учащихся. Однако на практике приходится думать не только о требованиях тестовой технологии, но и о реальных возможностях школьников или студентов, которые должны справиться с выполнением теста. Поэтому набор целей необходимо структурировать, чтобы в тест попали самые важные цели учебного процесса. Это позволяет уточнить определенные предметные области, разделы, темы, содержание которых необходимо отразить в тесте.

5. Последний этап – оценка степени реализации целей. Знание предметных областей выражается правилами измерения с перечислением

элементов содержания в совокупности с желаемыми, планируемыми при обучении уровнями владения этими элементами.

19. Принципы отбора содержания теста

При разработке тестов необходимо придерживаться определенных *принципов*:

- определение значимости проверяемых знаний в общей системе проверяемых знаний;
- взаимосвязь содержания и формы;
- содержательная достоверность тестовых заданий. Каждое задание теста опирается, как правило, на факт, правило, теорему, норму, закон или на апробированный в практике метод;
- репрезентативность содержания учебной дисциплины в содержании теста, полнота и достаточность числа заданий;
- возрастающая трудность;
- соответствие содержания теста уровню современного состояния науки;
- комплексность и сбалансированность содержания теста (итоговый тест не строится на содержании одной темы, проверяться должны не только теоретические знания, но и практические умения);
- системность содержания (формулирование содержания тестовых заданий, отвечающего требованиям системности);
- вариативность содержания (множество вариантов заданий одного и того же теста, которые примерно равны по трудности, имеют сходные показатели вариации тестовых баллов испытуемых; содержание теста для слабых испытуемых заметно отличается от содержания теста для сильных испытуемых.)

Задания, входящие в тест, подбираются так, чтобы они давали основу для проверки некоторых из таких категорий приобретенных знаний, как названия, имена; формулы; смысл слов, названий и имен; факты; определения; сравнение, сопоставление объектов; противоположности, противоречия, антонимы и т.п.; ассоциации; классификации; причинно-следственные отношения; алгоритмы, процедуры; технологии и технологические понятия; вероятностные понятия; абстрактные понятия; методология предмета¹.

Короткие инструкции, общие для всех испытуемых, обычно помещаются перед заданием или группой заданий и по шрифтовому оформлению отличаются от содержательной основы задания и ответов к

¹ Аванесов В.С. Композиция тестовых заданий. М.: Ассоциация инженеров-педагогов, 1996.

нему. Инструкции адекватны форме и содержанию задания. Если задания представлены одной формой, инструкция пишется один раз для всего теста. Если же тест включает в себя задания различных форм, то перед каждой сменой формы задания пишется новая инструкция.

Задание, чаще всего, формулируется *в утвердительной, а не в вопросительной форме*. Содержательную часть задания не перегружают второстепенными деталями; она включает минимум ключевых слов, необходимых для правильного понимания задания, потому что формулировки заданий должны иметь однозначное толкование.

Немаловажное значение имеет шрифтовое оформление задания. Оно должно быть таким, чтобы суть задания понималась с одного взгляда. Традиционно текст задания пишется прописными буквами, а варианты ответов – строчными.

Правила оценки выполнения каждого задания разрабатываются вместе с тестом, они всегда одинаковы для всех испытуемых. Чаще всего, за верный ответ принято давать один балл, за неверный – ноль. В таком случае сумма всех баллов, полученных студентом, равна числу его правильных ответов. Однако можно использовать и другие шкалы оценок. Сумма баллов ассоциируется с уровнем знаний студента. Совокупность сумм баллов испытуемых используется для установления их рейтингов, то есть порядковых номеров, показывающих сравнительную оценку достижений в рамках данной группы.

20. Экспертиза качества содержания

Оценка качества содержания теста обычно проводится по определенной методике независимыми экспертами, не участвовавшими в разработке теста. Как правило, число экспертов составляет не менее трех человек по каждому тесту. К экспертизе привлекаются наиболее опытные учителя, имеющие большой стаж работы с теми учащимися, для которых в конечном итоге предназначен тест.

Методика экспертизы качества содержания теста обычно включает три раздела, которые выстраиваются сообразно трем направлениям работы экспертов. Перед началом работы каждый эксперт должен ознакомиться со спецификацией рецензируемого теста, содержащей пояснения по его структуре и планируемому к проверке содержанию. Один из вариантов подобных методик для оценки качества содержания аттестационного теста приводится далее.

Работа эксперта по первому направлению заключается в анализе содержания отдельных заданий теста. Во-первых, дается собственная оценка содержанию каждого задания; во-вторых, проводится анализа со-

держания заданий в терминах требований к уровню подготовки выпускников; в-третьих, определяется уровень базовости (какую степень достижения требований проверяет задание: на минимальном уровне, по трудности превышающие минимально достаточный уровень, но находясь в пределах минимума содержания образования поданному предмету или это задания, выходящие за пределы минимума содержания и требований к уровню подготовки выпускников. Эксперт также оценивает уровень значимости содержания каждого задания теста. Указывает предполагаемый процент испытуемых, справившихся с каждым заданием тестом, приводит предполагаемое время выполнения теста. Характеризует качество формулировок, указывает неудачные задания теста.

Второе направление работы эксперта связано с анализом качества содержания всего теста, имеющего несколько параллельных вариантов. Содержание, например, аттестационного теста определяется как оптимальное отображение требований к уровню подготовки выпускников в системе заданий теста. Требование оптимальности выделяет определенные критерии качества отображения.

Первый критерий — полнота охвата требований к уровню подготовки выпускников каждым вариантом теста и всеми вариантами вместе. Эксперт подсчитывает число требований, охваченных по одному варианту, а затем делит это число на общее количество требований и приводит полученный таким образом результат (в процентах) по каждому варианту теста.

Второй критерий качества – правильность пропорций содержания теста. Простой оценки полноты охвата требований недостаточно. Необходима также уверенность в том, что задания теста охватывают все важные аспекты предметной области и в правильной пропорции. Зачастую при разработке теста возможно смещение пропорций, так как тест легко перенасытить теми разделами содержания, по которым легче составить задания. Для оценки правильности пропорций теста эксперт подсчитывает процент заданий в тесте, ориентированных на материал каждого раздела (содержательной линии). Затем приводит свое видение оптимального соотношения разделов и вычисляет разность, характеризующую отклонение мнения разработчиков от своих оценок.

Третий критерий – проверка соответствия содержания системы заданий спецификации теста. Несоответствие может возникать как при уменьшении, так и при увеличении планируемого числа заданий в тесте. Подсчитывается общее число заданий, не соответствующих спецификации теста.

Третье направление работы эксперта рассчитано на подготовку обобщающих выводов и рекомендаций по улучшению содержания теста. Эксперт приводит свое общее впечатление о содержании теста. Здесь должны быть высказаны все сомнения и пожелания эксперта, его

рекомендации по улучшению содержания. Возможна оценка соотношения заданий, проверяющих знание теории предмета и его практики. Желательно выявить задания, предназначенные для проверки системы понятий, а также задания интегрального характера, предназначенные для проверки умений учащихся обобщать знания по различным разделам предмета.

21. Понятие «измерение». Измерительные шкалы.

Эффективность тестовых оценок зависит не только от качества теста, но и от используемых методов сравнения и интерпретации первичных (сырых) баллов. Стремление выявить истинные уровни знаний испытуемых, получить возможность сравнивать эти уровни между собой даже тогда, когда они получены по разным тестам, привели исследователей к разработке новых методов интерпретации результатов тестирования. Одной из сильных сторон тестового метода является не оценивание результатов подготовки, а *измерение уровня знаний* учащихся.

В самом широком смысле **измерение сводится к приписыванию чисел объектам согласно определенным правилам.** *Правила* устанавливают соответствие между некоторыми свойствами объектов и чисел, позволяющее сравнивать между собой эти объекты по состоянию измеряемого свойства. В зависимости от целей измерения выбираются различные правила, отображающие различные свойства объектов. Всякий раз при измерениях нужного свойства выбирают соответствующую измерительную шкалу.

С физической точки зрения измерение – это процесс сравнения измеряемой величины с эталонной. С математической точки зрения измерение – отображение измеряемого объекта в некоторое множество действительных чисел, называемое шкалой.

Шкала – это средство фиксации результатов измерения определенных свойств объектов путем упорядочения их в определенную числовую систему, в которой отношение между отдельными результатами выражено в соответствующих числах. В педагогических измерениях шкалы различаются в зависимости от характера свойств, лежащих в основе их построения.

Педагогическое измерение — это процесс установления соответствия между оцениваемыми характеристиками обучаемых и точками эмпирической шкалы, в которой отношения между различными оценками характеристик выражены свойствами числового ряда.

При педагогическом измерении в качестве характеристик обучаемых обычно выступают знания и умения, освоенные учениками на момент выполнения теста. Роль единицы измерения играют тестовые задания,

объектом оценки являются сами ученики, а результатом измерения — шкала баллов тестируемых учеников. Количество правильно выполненных заданий теста дает основания для присвоения испытуемому определенного места на шкале.

*Операция упорядочения исходных эмпирических данных путем перевода их в шкальные носит название **шкалирования**.*

В образовательной практике можно выделить четыре основных вида шкал, получивших наиболее широкое распространение. Согласно С.С.Стивенсону¹, они получили следующие названия:

- **шкала наименований (номинальная),**
- **шкала порядка (ранговая шкала),**
- **интервальная шкала и**
- **шкала отношений.**

Измерения на первых двух шкалах считаются качественными, а на двух других – количественными. Шкалы качественных измерений называют «дискретными», а количественных – «непрерывными». В каждой из этих шкал определены свойства чисел, приписываемых объектам. Шкалы перечислены в таком порядке, что в каждой последующей, кроме свойств предыдущей шкалы, добавляются и новые. Поэтому, чем больше порядок шкалы, тем больше арифметических действий разрешается проводить над числами, приписанными объектам в этих шкалах.

Номинальная шкала измерений. По этой шкале процесс измерения осуществляется группированием предметов в классы, когда объекты, принадлежащие к одному классу, идентичны в отношении некоторого признака или свойства. Далее классам даются обозначения, вместо обозначений классов могут также принимать и часто принимают для *идентификации* числа, что может служить объяснением заголовка «номинальное измерение». Мы выполнили бы номинальное измерение, если бы присвоили число 1 – англичанам, 2 – немцам, а 3 – французам. Равна ли одному французу сумма одного англичанина и одного немца ($1+2=3$)? Конечно, нет. **Арифметические действия с числами номинальной шкалы не имеют смысла.** Примером является дихотомическая шкала: выполнившие задание получают число 1 (зачет), а не выполнившие – 0 (незачет). Методы обработки таких результатов оценивания знаний называют «статистикой качественных признаков», она позволяет подсчитать количество объектов в каждом классе признаков, найти процентное отношение частоты классов, выделить класс с наибольшим (или меньшим) числом объектов данного признака, т.е.

¹ Экспериментальная психология (глава: Математика, измерения и психофизика) – М., 1960.

разделить на «моды». Данные, соответствующие номинальным шкалам, составляют наблюдаемые значения частот появления каждой из разновидностей изучаемой переменной, эти результаты, как правило, используются при построении матриц результатов.

В **порядковой** шкале вводятся числа и отношение «больше–меньше», поэтому по числу, соответствующему оцениваемому объекту, можно узнать о месте объекта в совокупности. В порядковой шкале измеряются только *качественные* признаки. Объекты оцениваются с точки зрения отношения равенства между ними или отношения «больше–меньше», а расстояния между объектами не имеют никакого смысла. Например, по результатам экзаменов можно приписать ранги каждому ученику, но только в пределах той группы, где проводился экзамен. *Пятибалльная шкала, по которой сейчас выставляются оценки в школе, является частным случаем порядковой.* В ней все ученики делятся на отдельные группы. Например, группа троечников может включать довольно много учеников, объединенных вместе и никак не ранжированных внутри своего объединения. По оценкам в порядковой шкале можно ранжировать учащихся, но делать вывод о том, насколько один лучше другого, нельзя в силу *отсутствия единицы измерения.*

Понятие среднего балла при таком оценивании лишено смысла, оно не отражает объективных закономерностей результатов учебных достижений, так как на такой шкале оцениваются только качественные признаки. (Кроме моды, на такой шкале может быть оценен «медианный объект», который по измеряемому свойству переходит за 50%. Число, приписываемое такому объекту, называется «медианой» и принимается за меру центральной тенденции группы объектов).

В **интервальной** шкале определено расстояние между объектами и предусмотрена общая для всех объектов постоянная единица измерения. Интервальное измерение – это такое присвоение чисел предметам, когда равные разности чисел соответствуют равным разностям значений измеряемого признака или свойства предметов. Интервальная шкала – количественная. В ней возможны все арифметические действия над числами, кроме деления. Таким образом, в интервальной шкале нельзя определить, во сколько раз один объект больше или меньше другого. Например, если ученик ответил правильно на 30 заданий, то это не означает, что он знает вдвое больше ученика, ответившего на 15 заданий теста.

Недостатком этой шкалы является неизвестность абсолютного нуля. Эта шкала задает взаимное положение измеряемых объектов относительно друг друга, но не показывает расположение объектов относительно начала координат. Например, при оценке выполнения учащимся теста нуль верных ответов не означает полного отсутствия

знаний. В интервальной шкале допустимы почти все статистические операции, кроме тех, которые предполагают знание «истинно» нулевой точки шкалы.

Измерение **по шкале отношений** отличается от интервального только тем, что нулевая точка не произвольна, а указывает на полное отсутствие измеряемого свойства. Измеритель может заметить отсутствие свойства и имеет единицу измерения, позволяющую регистрировать различающиеся значения признака. В ней можно выполнять все арифметические и статистические операции. Так же как и интервальная, она позволяет производить количественные измерения.

22. Методы шкалирования при оценке знаний

Традиционные методы тестирования (классические) используют порядковые шкалы, отличающиеся друг от друга длиной, масштабом и значением центрального индекса. Балл тестируемого определяется количеством правильно выполненных заданий «А» из общего числа заданий «К». Тогда отношение A/K можно выразить в процентах и получить 100-балльную шкалу, называемую *процентной*.

Другой подход характеризуется тем, что окончательный балл участников тестирования зависит от его относительных успехов по сравнению с успехами других тестируемых. Такие шкалы называют *процентильными*. как и процентные, они имеют ранговый смысл. Недостатком этих шкал является невозможность сравнения полученных результатов между распределением результатов разных выборок тестируемых.

Наиболее простой метод выявления места результата i -го ученика (X_i) в сравнении с результатами основан на подсчете отклонения балла X_i от среднего значения баллов \bar{X} по группе тестируемых учащихся. Отклонение находят путем вычисления разности ($\bar{X} - X_i$). Если разность больше 0, то результат i -го ученика выше среднего по группе. Отрицательное значение разности указывает на результат ниже среднего значения \bar{X} .

Так среднее арифметическое, полученное по различным группам и тестам, существенно разнятся, возникает проблема сопоставимости отклонений. Специфической мерой вариативности показателей является стандартное отклонение. Чем меньше стандартное отклонение, тем больше показателей группируется около среднего, чем больше стандартное отклонение, тем больше разброс. Стандартное отклонение вычисляется по формуле:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}},$$

где S – стандартное отклонение, X_i – индивидуальный результат обследуемого, \bar{X} – среднее арифметическое в данной группе, n – количество обследуемых. Зная средние арифметические и соответствующие стандартные отклонения, полученные в данной группе обучаемых в других контролях знаний, мы можем сопоставить результаты выполнения заданий (или контрольных работ между собой) между собой.

Отсюда можно выразить индивидуальный показатель как долю от стандартного отклонения. Полученные доли называют Z -оценками, которые могут стать основой любой стандартной шкалы, так как, во-первых, они характеризуют относительное положение результата обследуемого среди всех результатов группы независимо от среднего и стандартного отклонения, во-вторых, они не имеют размерности:

$$Z = \frac{X_i - \bar{X}}{S}$$

С помощью Z -оценок можно проводить сопоставление результатов, полученных различными путями для отдельного индивида. Недостатком Z -оценок является то, что приходится иметь дело с дробными и отрицательными величинами, поэтому Z -шкалы преобразуют в другие виды шкал. Для перевода сырых баллов в одно их стандартных шкал можно воспользоваться следующей формулой:

$$T = M + Z \cdot s$$

где T – балл по новой шкале, M – выбранное среднее значение новой шкалы, s – выбранное стандартное отклонение новой шкалы. Традиционно чаще всего используются следующие шкалы:

- шкала станайнов – $M=5$; $s=2$
- шкала стенов – $M=5,5$; $s=2$
- T -шкала: $M=50$; $s=10$
- IQ -шкала: $M=100$; $s=12, 14, 16, 18$ и т.д.

Эти шкалы взаимосвязаны друг с другом.

Шкала станайнов удобна для интерпретации результатов, полученных на относительно небольшой группе испытуемых.

Современная теория тестов (IRT), предполагает использование **единой интервальной шкалы логитов**, в которой можно объективно сравнить уровни подготовки испытуемых и определить, насколько задания теста эффективны для оценки испытуемых. Единицей измерения метрической шкалы является **логит**; на такой шкале одновременно

измеряются латентные параметры: как трудности тестовых заданий, так и уровень подготовленности тестируемых при ответе на каждое задание.

В современных тестовых технологиях мерой трудности заданий является *логит трудности j-го задания*, определяемый натуральным логарифмом $\ln q_j/p_j$ отношения доли неправильных q_j к доле правильных ответов p_j .

$$\beta_j = \ln \frac{q_j}{p_j} = \ln \frac{1-p_j}{p_j}$$

Симметрично была введена и логарифмическая оценка уровня знаний i-го учащегося – *логит уровня знаний*, определяемый натуральным логарифмом отношения доли правильно выполненных к доле не выполненных заданий:

$$\theta_i = \ln \frac{p_i}{q_i} = \ln \frac{p_i}{1-p_i}$$

Пересчет уровня учебных достижений испытуемых, выраженный в логитах, на любую другую шкалу должен осуществляться только путем линейных преобразований, не нарушающих метричность шкалы. В частности, для пересчета в 100 – бальную шкалу Центр тестирования использует формулу:

$$B_i = 50 + \frac{50}{t\sigma} (\theta_{0i} - \theta_{с.в.})$$

где t – коэффициент, значение которого выбирается так, чтобы возможность получения высших баллов была регламентирована; σ – среднее квадратичное отклонение подготовленности испытуемых нормативной выборки относительно средневзвешенного значения параметра $\theta_{с.в.}$; θ_{0i} – значение подготовленности испытуемого.

Так, например, при коэффициенте 2.5 вероятность получить 100 баллов составляет около 0,01, а при коэффициенте, равном 3, не более 0,003.

Рассмотренные выше особенности шкалирования полностью отвечают сути тестирования как *измерительного процесса* и согласуются с особенностями результатов измерения в любой другой области.

Уровень подготовки каждого участника тестирования и трудность всех заданий теста оцениваются сначала в логитах на логистической шкале в интервале от -5 до $+5$. Начало шкалы логитов “закрепляется” репрезентативной выборкой из 1000 испытуемых, адекватно отражающих структуру генеральной выборки участников тестирования (по регионам, видам образовательных учреждений, городских – сельских и др.). После этого значения оценок в логитах и их средние квадратичные ошибки преобразуются (без деформации метрики) в 100-бальную шкалу для

выдачи результатов учащимся и использования в образовательной статистике.

Окончательный тестовый балл является следствием учета не только качества верно выполненных заданий, но и характеристик трудности этих заданий, их дифференцирующей силы (способности к дискриминации уровня подготовленности). Такой балл имеет основные черты количественной меры, а поэтому позволяет проводить объективное сопоставление результатов, производить математико-статистический анализ, изучать динамику различных образовательных процессов в одних и тех же единицах измерения уровня подготовки учащихся и трудности заданий тестов.

Тема IV. ЕДИНЫЙ ГОСУДАРСТВЕННЫЕ ЭКЗАМЕН

23. Единый государственный экзамен.

Единый государственный экзамен призван объединить итоговую аттестацию учащихся в школе и отбор абитуриентов для поступления в вузы. Эксперимент по введению ЕГЭ начат в России по инициативе МО в 2001 г., продолжается до сих пор. Его целью является создание системы объективной оценки подготовленности выпускников общеобразовательных учреждений и абитуриентов, поступающих в вузы и ссузы. Для создания такой системы потребовалось подготовить:

- нормативные документы,
- разработать и проверить в эксперименте технологию проведения экзаменов,
- подготовить соответствующих специалистов,
- разработать объективный инструментарий для объективной проверки экзаменуемых – контрольно-измерительные материалы.

Сначала к участию в ЕГЭ было привлечено ограниченное число учебных заведений. К настоящему времени список учебных заведений и регионов существенно расширился. В него включился и Татарстан.

Главными доводами являются объективность, демократичность, снижение физической и психологической нагрузки, повышение престижности знаний, наличие мирового и отечественного опыта подобной формы итоговой аттестации выпускников средних учебных заведений.

Правильно составленные контрольные измерительные материалы (КИМ) и правильно проведенный ЕГЭ наиболее объективный способ оценки знаний и умений по сравнению со всеми другими формами проверки этих навыков.

Объективность обуславливается тем, что такой экзамен – письменный и все школьники сдают его по параллельным (эквивалентным) заданиям по предметам. Эти предметы они выбирают сами в соответствии с профилем вуза, в который они хотят поступить. Тесты ЕГЭ составляют специалисты. Работы на проверку поступают не под фамилией ученика, а под его шифром.

В тех регионах России, где проводится ЕГЭ, каждый учащийся пишет тест и свои результаты может послать в любые вузы, причем сразу в несколько. При этом ему не надо ехать в тот город, где находится выбранный им вуз, не надо тратить большие деньги на дорогу и на проживание в этом городе. Важно отметить, что снимается прямой контакт абитуриента и преподавателя вуза, а также снижается вероятность прямых или замаскированных взяток.

В настоящее время выпускники школ России вынуждены сдавать выпускные экзамены в школе и следом вступительные экзамены в вуз, причем некоторые учащиеся сдают экзамены в 2—3 вуза. Это приводит к нервным срывам абитуриентов. О тяжелых последствиях летней экзаменационной лихорадки для школьников и их родителей пишут многие газеты.

Единый государственный экзамен настраивает школьников на получение реальных знаний в школе. В этом случае именно знания, а не деньги родителей или их связи определяют дальнейшую судьбу школьника.

И наконец, летом с 2001 по 2006 гг. был проведен эксперимент по тестированию выпускников общеобразовательных школ по правилам Единого государственного экзамена.

По результатам эксперимента можно сделать следующие выводы;

1. Эксперимент по проведению ЕГЭ усилил внимание общественности к системе образования.

2. Подтверждена возможность объективной оценки подготовленности выпускников и абитуриентов: все учащиеся оказались в одинаковых условиях и за одинаковые знания получили одинаковые оценки.

3. Учащиеся успешно справились с новой формой государственной аттестации, и, более того, эта форма контроля им понравилась.

4. Процент пятерок и четверок по результатам ЕГЭ оказывается несколько ниже, чем соответствующие проценты в аттестате в прошлом году. При двойке на обычном экзамене назначалась переэкзаменовка, а при тестовом экзамене ее нет; в аттестат ставилась текущая оценка – тройка. Таким образом, ЕГЭ высветил реальную картину знаний, хотя школьники при этом не пострадали. Главная цель ЕГЭ — выявить наи-

более узкие места в преподавании физики и других предметов для того, чтобы обратить внимание педагогов на недостатки в их работе.

5. Существенного разрыва между результатами эксперимента в городских и сельских школах не произошло. Шансы на поступление выпускников из удаленных сельских школ возросли.

6. Повысилось доверие общества к идее Единого государственного экзамена. В целом отмечена позитивная реакция родителей абитуриентов, что свидетельствует о том, что выбран правильный путь оценки знаний школьников и правил зачисления в вузы.

7. ЕГЭ в регионах рассматривается как важный инструмент аттестации и аккредитации школ.

24. Основные нормативные и правовые документы по ЕГЭ

Основные нормативные документы, регламентирующие проведение единого государственного экзамена.

- Приказ № 1306 от 09.04.2002 "Об утверждении Положения о проведении единого государственного экзамена"
 - Приказ № 50 от 14.01.2003 Порядок приема в государственные высшие учебные заведения РФ, учрежденные федеральными органами исполнительной власти Утвержден приказом № 50 от 14.01.2003
 - Положение об аттестации выпускников IX и XI(XII) классов общеобразовательных учреждений РФ
 - Инструкция по рассмотрению апелляций к процедуре и результатам единого государственного экзамена в 200_ году
 - Письмо МО РФ органам управления о пробном ЕГЭ в 2007 году (от 25.01.2007 г.)
и др.¹
- Содержание экзаменационной работы соответствует:
- Федеральному компоненту государственного стандарта основного общего и среднего (полного) образования по физике (Приказ МО от 5 марта 2004 г. № 1089).
 - Обязательному минимуму содержания основного общего образования по физике (Приказ МО от 19. 05. 1998 г. № 1236).
 - Обязательному минимуму содержания среднего (полного) общего образования по физике (Приказ МО от 30.06. 1999 г. № 56).

При этом учитываются:

¹ См., например: Единый государственный экзамен. Сборник нормативных документов /Министерство образования Российской Федерации. М.: Интеллект-Центр, 2004.

- Требования к уровню подготовки выпускников основной школы по дисциплине
- Требования к уровню подготовки выпускников средней (полной) школы по дисциплине. Уровень В

ЕГЭ проводится в специально подготовленных пунктах. Список пунктов проведения экзаменов (далее ППЭ) утверждается ОУО субъекта Федерации. Администрация образовательных учреждений должна получить из ОУО субъекта Федерации необходимое количество пропусков для выпускников, участвующих в ЕГЭ (до 30 апреля).

Выпускники до установленного срока (обычно конец апреля) должны заявить в письменном виде администрации своего образовательного учреждения о своем желании участвовать в ЕГЭ по конкретным предметам.

ЕГЭ проводится по расписанию, утвержденному Минобразованием России. Все выпускники, участвующие в ЕГЭ, должны до 15 мая текущего года получить у администрации своего образовательного учреждения пропуск, в котором указаны предметы ЕГЭ, адрес пункта проведения ЕГЭ, даты и время начала экзаменов, коды образовательного учреждения и пункта проведения ЕГЭ, настоящая инструкция и иная информация.

После размещения выпускников в аудиториях организаторы объясняют им: порядок проведения ЕГЭ, порядок получения экзаменационных материалов, порядок сдачи заполненных бланков регистрации бланков ответов, черновики и экзаменационных материалов, виды и порядок подачи апелляций. В каждую аудиторию экзаменационные материалы поступают в запечатанных доставочных пакетах, содержащих как правило, 15 индивидуальных прозрачных полиэтиленовых пакетов с экзаменационными материалами для каждого выпускника и пакеты для упаковки заполненных бланков ответов.

Организаторы раздают выпускникам запечатанные индивидуальные полиэтиленовые пакеты с вложенными в них **КИМами, бланком регистрации, бланками ответов № 1 и № 2, бланком черновика** и проводят инструктаж по работе с экзаменационными материалами и бланками. Бланк регистрации и бланки ответов № 1 и № 2, полученные выпускником в индивидуальном запечатанном пакете, имеют специальную нумерацию (кодировку) и **рассматриваются только во взаимосвязи друг с другом.**

Контрольно-измерительные материалы по каждому предмету содержат три части:

задания типа **А** с нумерацией А1, А2, А3... и т.д.;

задания типа **В** с нумерацией В1, В2, В3... и т.д.;

задания типа С с нумерацией С1, С2, С3... и т.д.;

Задания типа А предлагают выпускникам выбрать правильный ответ из нескольких предлагаемых вариантов. В бланке ответов № 1 для каждого такого задания под заголовком «Номера заданий типа А с выбором ответа из предложенных вариантов» выпускник должен под номером задания пометить знаком «Х» ту клетку, номер которой соответствует номеру выбранного ответа.

Задания типа В предлагают выпускникам дать ответ в виде последовательности символов (не более 17). В бланке ответов № 1 для таких заданий выделены поля с заголовком «Результаты выполнения заданий типа В с ответом в краткой форме». Выпускник аккуратно вписывает свой ответ рядом с номером задания. Ответ записывается буквами или цифрами по образцу из верхней части бланка ответов. Запись формул или математических выражений не допускается. Какие-либо словесные заголовки или комментарии к числу недопустимы.

Задания типа С предлагают выпускникам дать развернутый ответ в виде решения задачи или краткого рассказа (сочинения), которые записываются в бланке ответов № 2.

Результаты ЕГЭ определяются на основании ответов выпускников в заполненных полях бланков ответов № 1 и № 2. Результаты ЕГЭ объявляются выпускникам в виде: баллов в столбальной системе; отметок в пятибалльной системе.

Например, шкала перевода баллов в отметки для результатов по физике в 2006 году была установлена:

- 0 - 34 баллов – отметка "2";
- 35 - 51 баллов – отметка "3";
- 52 - 69 баллов – отметка "4";
- 70 и более баллов – отметка "5".

Результаты ЕГЭ оформляются для каждого выпускника в свидетельствах утвержденной формы (приказ Минобробразования России «Об утверждении формы свидетельства о результатах единого государственного экзамена»)

Выдача выпускникам свидетельств о результатах ЕГЭ (далее – **свидетельство**) осуществляется в соответствии с инструкцией Минобробразования России по организации выдачи свидетельств о результатах ЕГЭ.

Свидетельство о результатах ЕГЭ выдается выпускнику (его законному представителю или доверенному лицу, представляющему его интересы) администрацией образовательного учреждения, в котором обучался выпускник, в обмен на пропуск с подписями ответственных организаторов в аудиториях пункта проведения ЕГЭ и соответствующими печатями по каждому сдаваемому экзамену.

Часть 2. МАТЕРИАЛЫ ДЛЯ ПРАКТИЧЕСКИХ И СЕМИНАРСКИХ ЗАНЯТИЙ

ФОРМЫ ТЕСТОВЫХ ЗАДАНИЙ

1. Задания закрытой формы

(задания с выбором одного или нескольких правильных ответов)

Это простейший вид задания, в котором правильный ответ уже содержится, и задача испытуемого состоит в его узнавании. Рассмотрим основные элементы заданий с выбором правильного ответа. К ним относятся инструкции для испытуемых, содержание заданий, форма, содержание и число ответов, а также оценки за правильность выполнения.

Инструкции. В заданиях с выбором ответа применяют один из двух вариантов инструкций, соответствующих двум вариантам заданий. Инструкция помещается перед заданием, она печатается шрифтом, отличающимся от шрифта самого задания

Содержательная основа задания. Задание представляет собой часть утвердительного предложения, но не вопрос. Предлагаемые варианты ответов дополняют его до полного утвердительного предложения. Содержательная основа задания должна быть такой, чтобы для выбора правильного ответа достаточно было вспомнить и применить лишь то, что звучало на лекциях и было написано в рекомендованных для изучения пособиях. В конце содержательной части задания нет знаков препинания.

Ответы. Ответы должны быть содержательными и, по возможности, короткими. В конце ответов нет знаков препинания. Места для правильных ответов в разных заданиях выбираются случайным образом. Располагать ответы можно в одну, две и три колонки. Минимум средств (слов, символов, рисунков и графиков) должен обеспечивать максимальную ясность смысла задания. При подготовке ответов избегают повторов слов, применения малопонятных, редко употребляемых слов, а также неизвестных студентам или учащимся символов, иностранных слов, затрудняющих восприятие смысла.

Для объективной оценки знаний обучаемым предлагаются так называемые **ответы-дистракторы** (от англ. to distract - отвлекать). *Дистракторы* – это правдоподобные, но неправильные ответы. В хорошо составленном задании правильные и неправильные ответы испытуемыми, плохо знакомыми с предметом, выбираются с равной вероятностью (не должно быть вариантов ответов, вводящих в заблуждение знающих учеников).

Ответы к заданиям могут быть выражены словами, числами, графиками. Они не должны представляться в форме "да" или "нет", "верно" или "неверно". Исключается применение таких вариантов ответа, как "правильного ответа нет", "все ответы правильные" или "все ответы неправильные".

Перечислим ряд требований, предъявляемых к заданиям закрытой формы:

1) в тексте задания должна быть устранена всякая двусмысленность или неясность формулировок;

2) основная часть задания формулируется предельно кратко, как правило, не более одного предложения из семи-восьми слов;

3) задание имеет предельно простую синтаксическую конструкцию, в основной текст задания вводится не более одного придаточного предложения;

4) в основную часть задания следует включать как можно больше слов, оставляя для ответа не более двух-трех наиболее важных, ключевых слов для данной проблемы;

5) все ответы к одному заданию должны быть приблизительно одной длины либо правильный ответ может быть короче других, но не во всех заданиях теста;

6) из текста задания необходимо исключить все вербальные ассоциации, способствующие выбору правильного ответа с помощью догадки;

7) частота выбора одного и того же номера места для правильного ответа в различных заданиях теста должна быть примерно одинакова либо номер места для правильного ответа выбирается в случайном порядке;

8) из ответов обязательно исключаются все повторяющиеся слова путем ввода их в основной текст заданий;

10) в ответах не рекомендуется использовать слова «все», «ни одного», «никогда», «всегда» и т. п., так как в отдельных случаях они способствуют угадыванию правильного ответа;

11) из числа неправильных исключаются ответы, вытекающие один из другого;

12) при формулировке дистракторов не рекомендуется использовать выражения «ни один из перечисленных», «все перечисленные» и т.п., так как они способствуют угадыванию правильного ответа;

13) из числа тестовых исключаются задания, содержащие оценочные суждения и мнения ученика по какому-либо вопросу;

14) все дистракторы к каждому заданию должны быть равновероятно привлекательными для испытуемых, не знающих правильного ответа;

15) ни один из дистракторов не должен являться частично пра-

вильным ответом, превращающимся при определенных дополнительных условиях в правильный ответ;

16) основная часть задания формулируется в форме утверждения, которое обращается в истинное или ложное высказывание после подстановки одного из ответов;

17) ответ на одно задание не должен служить ключом к правильным ответам на другие задания теста, т.е. не следует использовать дистракторы из одного задания в качестве ответов к другим заданиям теста;

18) все ответы должны быть параллельными по конструкции и грамматически согласованными с основной частью задания теста.

Выполнить все эти требования для начинающего разработчика трудно, а иногда и попросту невозможно. Правда, часть требований носит рекомендательный характер, что несколько облегчает ситуацию с анализом отдельных недостатков формы. Например, вряд ли тест не удастся, если будет задействовано не семь-восемь, а десять, двенадцать или даже больше слов в формулировках основной части заданий либо задания будут сформулированы в виде вопроса. Другая часть требований обязательна к выполнению.

2. Принципы конструирования заданий с выбором одного или нескольких правильных ответов.

При конструировании заданий с выбором одного правильного ответа применяются две группы *принципов композиции*.

1. *Первая группа* используется при разработке ответов к заданиям. К ней относятся принцип противоречия, принцип противоположности, принцип однородности, принцип кумуляции, принцип градуирования, принцип удвоенного противопоставления.

2. *Вторая группа*, состоящая из принципа фасетности и принципа импликации, используется при разработке содержания заданий.

Рассмотрим принципы композиции разработки ответов.

1. Принцип противоречия. Если задание содержит **два ответа**, то, согласно *принципу противоречия*, второй ответ образуется из первого простым прибавлением отрицательной частицы “не”, отрицающих предлогов и слов так, что этим ответом отрицается смысл не самого задания, а содержания первого ответа:

Примеры:

Обвести кружком номер правильного ответа:

ПРЯМЫЕ $y=1+3x$ И $y=1-5x$

1) параллельны

2) не параллельны

ОРГАНИЧЕСКОЕ СТЕКЛО – ЭТО ПЛАСТМАССА, СОСТОЯЩАЯ ИЗ ТЕРМОПЛАСТИЧНОЙ СМОЛЫ

- 1) с наполнителями 2) без наполнителей

В этих заданиях отрицается смысл не самого задания, а содержания первого ответа.

2. Принцип противоположности.

Примеры ответов, построенных по *принципу противоположности*.

ПРОИЗВОДНАЯ ЛИНЕЙНОЙ ФУНКЦИИ

1) постоянная 2) переменная
ЯЧЕЙКИ ПРИ ОБРАЗОВАНИИ ПЕНОПЛАСТОВ

- 1) открыты 2) закрыты

ОТЛОЖЕНИЕ ЛИГНИНА В ОБОЛОЧКЕ РАСТИТЕЛЬНОЙ КЛЕТКИ ПРОИСХОДИТ

- 1) снаружи 2) внутри

В заданиях с **тремя ответами** второй ответ может быть противоположен первому, а третий ответ – первому и второму, так что множество ответов полно.

Примеры:

ЕСЛИ $b > 1$, то функция $f(x) = \log_b(x)$

- 1) убывает 2) возрастает 3) немонотонна

Противоположность может быть введена также внутрь самих ответов.

ПРИ ОСТРОМ ИНФАРКТЕ МИОКАРДА АРТЕРИАЛЬНОЕ ДАВЛЕНИЕ

- 1) повышается
2) понижается
3) вначале повышается, затем понижается
4) вначале понижается, затем повышается

3. Принцип однородности состоит в том, что ответы в заданиях должны быть однородными по форме. Не следует использовать в разных ответах одновременно глагол, прилагательное и существительное.

Примеры:

РАЗНОСТЬ ПОЛИНОМОВ – ФУНКЦИЯ

- 1) рациональная 2) полиномиальная 3) трансцендентная

КИСЛАЯ СРЕДА ОБРАЗУЕТСЯ ПРИ ГИДРОЛИЗЕ

- 1) карбоната натрия 3) нитрита свинца
2) сульфида калия 4) ацетата натрия

Усиливает эффективность заданий, сконструированных по принципу **однородности**, использование сходных по написанию и звучанию слов, похожих формул.

Примеры:

$z = x^2 + y^2$ – УРАВНЕНИЕ

- 1) параболы 2) параболоида 3) параболоида

ТИП ПИТАНИЯ ТИОНОВЫХ БАКТЕРИЙ

- | | |
|----------------------|----------------------|
| 1) фотоавтотрофный | 3) хемоавтотрофный |
| 2) фотогетеротрофный | 4) хемогетеротрофный |

4. Принцип кумуляции. Суть принципа в том, что каждый следующий ответ содержит на один элемент больше, чем предыдущий:

Примеры:

ПРИ ЙОДОМЕТРИИ ОПРЕДЕЛЯЮТ СОДЕРЖАНИЕ

- 1) окислителей
- 2) окислителей, восстановителей
- 3) окислителей, восстановителей, кислот

СТЕПЕНЬ ПЛОДородИЯ ПОЧВ ОБУСЛОВЛИВАЕТ

- 1) совокупность элементов
- 2) структура и совокупность элементов
- 3) микроорганизмы, структура и совокупность элементов

5. Принцип сочетания. Применяя *принцип сочетания*, используют соединение двух, трех или четырех слов в каждом ответе. Можно, например, сочетать более или менее однородные и правдоподобные пары ответов. Ответы сочетаются также по правилу цепочки, причем последнее слово в первом ответе становится первым во втором ответе, последнее во втором – первым в третьем и т.д. Понятия сочетаются по два и по три.

Примеры:

ФУНКЦИЯ $y = -2x^2 + 5x - 1$

- 1) непрерывна, дифференцируема
- 2) дифференцируема, монотонна
- 3) монотонна, непрерывна

ВОДА, ОЧИЩЕННАЯ В АПТЕКЕ, ЕЖЕДНЕВНО АНАЛИЗИРУЕТСЯ НА СОДЕРЖАНИЕ

- 1) хлоридов и сульфатов
- 2) сульфатов и нитратов
- 3) нитратов и хлоридов

6. Принцип градуирования. Ответы в задании упорядочиваются по возрастанию какого-то количественного признака.

Пример:

НА РЕАЛИЗАЦИЮ ПРИРОДООХРАННЫХ МЕРОПРИЯТИЙ КРАЕВОГО ЗНАЧЕНИЯ ОТ ВСЕЙ СУММЫ, ПОСТУПИВШЕЙ В ЭКОЛОГИЧЕСКИЙ ФОНД, ВЫДЕЛЯЕТСЯ

- | | | | | | |
|-------|--------|--------|--------|--------|--------|
| 1) 9% | 2) 10% | 3) 27% | 4) 30% | 5) 54% | 6) 60% |
|-------|--------|--------|--------|--------|--------|

ПРИ УВЕЛИЧЕНИИ ВЛАЖНОСТИ СВЫШЕ 30% ПРОЧНОСТЬ ДРЕВЕСИНЫ

- 1) возрастает
- 2) не изменяется
- 3) снижается

7. Принцип удвоенного противопоставления применяется в заданиях с четырьмя ответами, части которых, построенные по принципу противоположности, сочетаются попарно.

Примеры:

СКОРОСТЬ РЕАКЦИИ ЗАВИСИТ ОТ ТЕМПЕРАТУРЫ

- 1) прямо пропорционально, линейно
- 2) прямо пропорционально, экспоненциально
- 3) обратно пропорционально, линейно
- 4) обратно пропорционально, экспоненциально

ЕСЛИ ДЛЯ ЛЮБОЙ ТОЧКИ (x,y) ИЗ НЕКОТОРОЙ ОКРЕСТНОСТИ ТОЧКИ (a,b) ВЫПОЛНЯЕТСЯ УСЛОВИЕ $f(x,y) > f(a,b)$, ТО В ТОЧКЕ (a,b) ФУНКЦИЯ $f(x,y)$ ИМЕЕТ

- 1) относительный максимум
- 2) относительный минимум
- 3) абсолютный максимум
- 4) абсолютный минимум

Вторая группа принципов используется при разработке содержания заданий и состоит из принципа фасетности и принципа импликации.

8. Принцип фасетности содержания задания позволяет использовать фасеты в основном тексте задания. **Фасет – это форма записи нескольких вариантов одного и того же задания. Множество слов и словосочетаний, образующих фасет, помещается в столбик и обрамляется фигурными скобками.**

С помощью фасета создаются параллельные задания.

Фасет содержания задания может иметь большое количество вариантов. Например, список математических выражений может быть сделан сколь угодно длинным.

Пример:

ПЛОСКОСТЬ $\left. \begin{array}{l} 2x + 3y - 9z - 1 = 0 \\ 2x + 3y - 9z = 0 \\ \dots \end{array} \right\}$ ЧЕРЕЗ НАЧАЛО КООРДИНАТ

1) проходит 2) не проходит

ПРИ НАМОКАНИИ $\left. \begin{array}{l} хлопкового \\ вискозного \\ шерстяного \end{array} \right\}$ ВОЛОКНА ЕГО ПРОЧНОСТЬ

- 1) увеличивается
- 2) остается без изменений
- 3) уменьшается

9. Принцип импликации. Когда в содержательной части задания используется оборот “Если..., то” или эквивалентные ему, то имеет место применение **принципа импликации**.

Пример:

41. ЕСЛИ ПРИМЕНИТЬ ПОВЫШЕННЫЕ ДОЗЫ ФОСФОРНЫХ УДОБРЕНИЙ, ТО СНИЗИТСЯ ДОСТУПНОСТЬ

- 1) бора
- 2) меди
- 3) цинка
- 4) молибдена

При конструировании заданий в тестовой форме, как правило, применяются **сочетания принципов**, о которых шла речь выше.

Пример сочетания принципа фасетности и импликации.

ЕСЛИ В ПОЛДЕНЬ СТАТЬ СПИНОЙ К СОЛНЦУ, ТО

}	<i>спереди</i>
	<i>справа</i>
	<i>слева</i>
	<i>справа</i>

БУДЕТ

- 1) юг
- 2) север
- 3) запад
- 4) восток

В заданиях с **выбором нескольких правильных ответов** применимы те же принципы конструирования, что и в заданиях с выбором одного правильного ответа. При этом верным может быть один ответ, несколько ответов или даже все предложенные ответы.

Примеры:

К ПОЛИНОМАМ ОТНОСЯТСЯ ФУНКЦИИ

$$\begin{array}{cccc}
 \frac{ax+b}{cx^3+dx} & & & \\
 1) y= & 2) y=ax+b & 3) y=ax^2+bx+c & 4) y=\frac{1}{ax+b}
 \end{array}$$

Если требуется указать не один верный ответ, а несколько из большого числа вариантов ответов, то без знания учебного материала угадать ответ маловероятно, так что информационная ценность заданий

этого вида, вообще говоря, выше, чем у заданий с выбором одного правильного ответа.

В заданиях с выбором нескольких правильных ответов используется практически только два принципа композиции – **однородность ответов и фасеты** в основной части задания.

Разработчик теста с выбором одного правильного ответа вправе использовать любую обоснованную систему оценки выполнения заданий. Хотя в большинстве случаев за правильно выполненное задание принято давать один балл, а за неправильно выполненное – ноль (дихотомная оценка).

Задание с выбором нескольких правильных ответов считается выполненным правильно, если точно выбраны все без исключения правильные ответы. Как правило, один балл дается за правильно выполненное задание, ноль – за неправильно выполненное. Можно использовать и более широкую шкалу оценок. Например, за полностью выполненное задание давать 3 балла, за одну ошибку (один лишний ответ или один пропущенный) - 2 балла, и так далее, до нуля. Лишь бы правила оценки были определены заранее и известны испытуемым перед началом тестирования.

3. Задания открытой формы (задания на дополнения)

Задание открытой формы конструируется в виде утверждения, рядом с которым готовые ответы с выбором не приводятся. Испытуемый сам дописывает в отведенном для этого месте свой ответ так, чтобы в результате получилось истинное высказывание. Эта форма задания сводит возможность догадки к минимуму. С помощью заданий открытой формы проверяют знание названий, формул, имен, фактов, свойств, признаков, дат, причинно-следственных отношений.

1. Каждое задание должно быть нацелено только на одно дополнение, место для которого обозначается прочерком или точками.

2. Прочерк ставится на месте ключевого элемента, знание которого является наиболее существенным для контролируемого материала.

3. Все прочерки в открытых заданиях для одного теста рекомендуется делать равной длины.

4. Дополнения лучше ставить в конце задания или как можно ближе к концу.

5. После прочерка, если это необходимо, указываются единицы измерения.

6. Текст задания должен обладать предельно простой синтаксической конструкцией и содержать минимальное количество информации, необходимое для правильного выполнения задания.

7. В тексте задания нет повторы и двойное отрицание.

Инструкция к заданиям открытой формы может иметь вид:

ДОПОЛНИТЕ:

Если несколько заданий открытой формы следуют одно за другим, то инструкция пишется один раз перед всей группой таких заданий.

Чем ближе к концу фразы находится место для ответа, тем лучше понимается суть задания. В конце задания открытой формы ставится точка.

Примеры:

НИ ОДНОГО ЭЛЕМЕНТА НЕ СОДЕРЖИТ _____ МНОЖЕСТВО.

При создании заданий открытой формы используются такие *принципы композиции*, как логическая определенность содержания задания, фасетность, параллельность, обратимость, логическая соразмерность объема определяющего понятия объему определяемого, краткость, неотрицательность и импликация. Рассмотрим на примерах применение этих принципов.

Выполняя задание, построенное в соответствии с *принципом логической определенности содержания*, испытуемый, знающий содержание предмета, легко находит ответ. Содержание и форма правильно сконструированного задания ведут его мысль прямо к правильному результату.

Пример:

СОВОКУПНОСТЬ ОБЪЕКТОВ, ОБЪЕДИНЕННЫХ ОБЩИМ ПРИЗНАКОМ ИЛИ СВОЙСТВОМ, НАЗОВЕМ _____.

Использование *принципа фасетности* обогащает возможности создания параллельных заданий.

Примеры:

ГРАФИК $\left. \begin{array}{l} \text{квадратичной} \\ \text{показательной} \\ \text{гиперболической} \end{array} \right\}$ ФУНКЦИИ НАЗЫВАЕТСЯ _____.

Чтобы проиллюстрировать применение *принципа обратимости*, покажем, как одно и то же утверждение может быть преобразовано в несколько заданий открытой формы.

Примеры (из арифметики):

$$6+5=_____.$$

$$6+_____ = 11.$$

$$_____ + 5 = 11.$$

Посредством такого приема создаются параллельные задания для разных вариантов теста.

Еще один принцип конструирования заданий в тестовой форме – *принцип краткости*: чем меньше слов, тем меньше недоразумений:

Примеры:

$$\sum_{i=1}^n (x_i - x) = \underline{\hspace{2cm}}$$

Посредством заданий открытой формы можно проверять знание терминов.

Пример:

КРИВАЯ, К КОТОРОЙ ПРИБЛИЖАЕТСЯ ГРАФИК ДАННОЙ ФУНКЦИИ, НИКОГДА НЕ СЛИВАЯСЬ И НЕ ПЕРЕСЕКАЯ ЕЕ -

Несколько последовательных заданий открытого типа, предназначенных для проверки знания терминов, можно объединить в таблицу под условным названием “КАК НАЗЫВАЕТСЯ”. Ответы в этой таблице вписываются испытуемым на специально предназначенные для этого места во втором столбце.

Пример:

КАК НАЗЫВАЕТСЯ ПРОЦЕСС,
ПРОИСХОДЯЩИЙ

ОТВЕТЫ:

1) ПРИ ПОСТОЯННОМ ДАВЛЕНИИ

2) ПРИ ПОСТОЯННОМ ОБЪЕМЕ

3) ПРИ ПОСТОЯННОЙ ТЕМПЕРАТУРЕ

При оценке правильности выполнения заданий открытой формы, обычно используется дихотомическая шкала {0,1}. Однако в случаях, когда в одном задании проверяется более одного элемента знаний, допустимо использовать и более широкие шкалы.

Задания на дополнение кажутся ученикам более трудными, так как в них исключается догадка. Это свойство делает задания на дополнение исключительно привлекательными для педагогов, особенно для тех, кто привык в своей работе опираться на традиционные средства контроля и не доверяет тестам.

Есть и недостатки. Они связаны с трудностями, возникающими при оценке ответов учеников. Дописывая ответ на месте прочерка, ученик может выбрать синонимы пропущенного запланированного разработчиком слова или изменить порядок следования элементов в

пропущенной формуле, что значительно затрудняет проверку и оценку результатов учеников.

Основной сферой применения заданий на дополнение остается текущий контроль. Реже они используются в итоговом контроле из-за трудностей с проверкой, которая становится крайне трудоемкой, довольно субъективной и требует привлечения большого количества педагогов в случаях массового тестирования учеников

4. Задания на установление соответствия.

Задания, где элементам одного множества требуется поставить в соответствие элементы другого множества, называют заданиями на установление соответствия. **Задания на установление соответствия эффективны при самоконтроле и текущем контроле знаний.** С их помощью проверяются так называемые ассоциативные знания, то есть знания о связи формы и содержания, сущности и явления, о соотношении между различными предметами, свойствами, законами. Испытуемый должен собрать ответ из элементов списка левой и соответствующих им элементов правой колонки.

Инструкция к заданиям этой формы имеет вид:

УСТАНОВИТЕ СООТВЕТСТВИЕ:

Номер задания ставится рядом с заголовком первого столбца. Заголовки пишутся прописными буквами. Заголовок каждого столбца относится ко всем элементам соответствующего множества. Элементы должны соответствовать названиям столбцов и быть короткими. Элементы левой колонки нумеруются цифрами, элементы правой – прописными буквами. Номер и буква отделяются от элементов столбца круглой скобкой и пробелом.

При разработке заданий на соответствие следует руководствоваться следующими правилами:

1) задание формулируется так, чтобы все содержание можно было выразить в виде двух множеств с соответствующими названиями;

2) элементы задающего столбца располагаются слева, а элементы для выбора справа;

3) желательно, чтобы каждый столбец имел определенное название, обобщающее все элементы столбца;

4) необходимо, чтобы правый столбец содержал хотя бы несколько дистракторов. Еще лучше, когда число элементов правого множества примерно в два раза больше числа элементов левого столбца;

5) необходимо, чтобы все дистракторы в одном задании были равновероятно правдоподобны;

6) элементы столбцов должны быть выбраны по одному основанию для

включения только однородного материала в каждое задание теста;

7) в дополнительной инструкции к заданию необходимо сообщить ученику о наличии дистракторов в правом столбце, а также о том, используется ли каждый элемент правого столбца только один раз или более одного раза;

8) все задания необходимо располагать на одной странице, не допуская переноса отдельных элементов.

На каждый элемент слева должен найтись, по крайней мере, один элемент справа, а каждому элементу справа должен соответствовать только один элемент слева. Испытуемый пишет ответы буквами из правого столбца в соответствующих пробелах рядом с номерами.

Пример:

УСТАНОВИТЕ СООТВЕТСТВИЕ

Процесс	Газовый закон
1) изотермический	А) $p/T = \text{const}$
2) изобарный	Б) $V/T = \text{const}$
3) изохорный	В) $pV = \text{const}$

ОТВЕТЫ: 1 ___; 2 ___; 3 ___.

В определенном смысле задания на установление соответствия довольно близки к заданиям в закрытой форме. И в том, и в другом случае при их выполнении ученик совершает одно и то же действие — выбирает из числа предложенных преподавателем правильный ответ. Не случайно задания на соответствие легко переделываются в ряд заданий закрытой формы, а закрытые задания в том случае, когда они имеют к нескольким содержательным основам одно и то же множество ответов, можно трансформировать в два множества для установления соответствия.

Оценивая правильность выполнения задания на установление соответствия, можно использовать дихотомическую шкалу: 1 за безусловно выполненное задание, 0 даже за одну допущенную ошибку. Однако можно выставлять оценки на основе более широкой порядковой шкалы: 0 – “совершенно неверно”, 1 – “есть верные элементы”, 2 – “много верных элементов” или “совершенно верно”, и т.д.

5. Задания на установление правильной последовательности. Сравнительная характеристика всех форм.

В тех случаях, когда требуется установить правильную последовательность действий или слов в определениях, используются задания на установление правильной последовательности. Это более сложный тип задания в тестовой форме, в процессе выполнения которого испытуемый конструирует ответ из предложенной неупорядоченной последовательности слов.

Задания на установление правильной последовательности используются для проверки знаний хода процесса, цепочки событий, действий и операций, а также определений и понятий. Они помогают формировать у учащихся алгоритмическое мышление, знание и умение. Задания этой формы полезны как в качестве средства контроля знаний, так и в качестве средства обучения.

Инструкции к заданиям этой формы имеют вид:

УСТАНОВИТЕ ПРАВИЛЬНУЮ ПОСЛЕДОВАТЕЛЬНОСТЬ:

Пример:

“ПРАВИЛО ТРЕХ ШАГОВ” ОТЫСКАНИЯ ПРОИЗВОДНОЙ

- избавиться от Δx в знаменателе
- найти предел выражения при $\Delta x \rightarrow 0$
- найти отношение приращения функции к приращению аргумента

ФАЗЫ ДЕЛЕНИЯ КЛЕТКИ ПРОТЕКАЮТ В СЛЕДУЮЩЕЙ ПОСЛЕДОВАТЕЛЬНОСТИ

- анафаза
- метафаза
- профаза
- телофаза

Испытуемый должен расставить в отведенных для ответов местах в начале каждой строчки (—) порядковые номера элементов последовательности действий.

Посредством заданий этой формы можно проверять знание последовательности доказательства теорем. Задания на установление правильной последовательности полезны при проверке знаний формулировок утверждений и определений.

Инструкция к заданию может сопровождаться необходимыми пояснениями. Название задания пишется заглавными буквами; оно декларирует, знание чего должен продемонстрировать испытуемый, причем ключевое слово в нем должно иметь именительный падеж.

Ранжируемые элементы в задании ставятся в случайном порядке. Чтобы окончания слов не служили подсказкой, их пишут в именительном падеже. Предлоги и союзы из множества ранжируемых элементов могут исключаться.

Чаще всего используется дихотомическая оценка правильности выполнения задания: 0 или 1. Однако можно применить и другое правило

оценивания. Например, правильно выполненное задание оценивать 3 баллами, ошибка в конце задания снижает оценку на 1 балл, ошибка в середине – на 2 балла, а ошибка в начале – к снижению оценки выполнения всего задания до нуля.

Каждая из рассмотренных выше форм заданий может быть применена для проверки знаний. Выбор зависит от цели применения теста, специфики учебной дисциплины, от технологии проверки, сбора и обработки эмпирических данных, от технического и материального обеспечения процесса применения теста. Задания с выбором ответов наиболее удобны при итоговом контроле в силу ряда технологических преимуществ, повышающих эффективность процессов применения итогового теста и оценки результатов учеников. Для таких заданий легко организовать компьютеризированный сбор и анализ результатов. Задания на дополнение предпочтительны при обучающем контроле. Результаты их выполнения требуют ручной обработки. Как правило, для оценки результатов их выполнения приходится привлекать экспертов, а это требует дополнительных материальных затрат и времени на проверку. Задания на установление соответствия эффективны при самоконтроле и текущем контроле знаний.

В каждом тесте можно использовать несколько форм, но с точки зрения тестологов желательно выбрать одну форму и выдержать ее от первого до последнего задания. Совмещение различных форм в одном тесте значительно затрудняет подведение итоговых результатов тестирования.

ЛИТЕРАТУРА

1. Аванесов В.С. Теоретические основы разработки заданий в тестовой форме: Пособие для профессорско-преподавательского состава высшей школы. – М.: МГТА, 1995.
2. Аванесов В. С. Композиция тестовых заданий. – М.: Ассоциация инженеров-педагогов, 1996.
3. Аванесов В. С. Математические модели педагогического измерения. – М.: Исследовательский центр проблем качества подготовки специалистов, 1994.
4. Аванесов В. С. Научные проблемы тестового контроля знаний. – М.: Исследовательский центр проблем качества подготовки специалистов, 1994.
5. Амонашвили Ш. А. Обучение, оценка, отметка. – М.: Знание, 1980.
6. Амонашвили Ш. А. Воспитательная и образовательная функции оценки учения школьников. – М.: Педагогика, 1984.

7. Анастаси А., Урбина С. Психологическое тестирование. – СПб.: Питер, 2003.
8. Беспалько В. П. Педагогика и прогрессивные технологии обучения. – М., 1995.
9. Беспалько В. П. Программированное обучение: (Дидакт. основы). – М.: Высш. шк., 1970.
10. Битинас Б.П., Катаева Л.И. Педагогическая диагностика: сущность, функции, перспективы. // Педагогика, 1993, №2
11. Вашкевич Ю В, Васильеве С Дидактические возможности тестового контроля.– М., 1974.
12. Все грани образовательного тестирования // Стандарты и мониторинг в образовании, 1999. – №2.
13. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. – М.: Прогресс, 1976.
14. Грабарь М. И. Проблема измерений и проверка гипотез при мониторинге результатов обучения // Стандарты и мониторинг в образовании, 2000. – №3.
15. Дерябо С.Д. Учителю о диагностике эффективности образовательной среды -М.: Молодая гвардия, 1997.
16. Единый государственный экзамен. Сборник нормативных документов /Министерство образования Российской Федерации. М.: Интеллект-Центр, 2004.
17. Ингекамп К. Педагогическая диагностика. Пер. с нем. – М.: Педагогика, 1991.
18. Кабанов А.А. Тестирование студентов: достоинства и недостатки // Педагогика– 1999, №2.
19. Кальней В. А., Шишов С. Е. Технология мониторинга качества обучения в системе «учитель-ученик». – М.: Педагогическое общество России, 1999.
20. Качество общего среднего образования в Российской Федерации: состояние, тенденции, проблемы, прогнозы. Ежегодный аналитический доклад. – М.: Исследовательский центр проблем качества подготовки специалистов. –1999.
21. Кларин М. В. Инновационные модели обучения в зарубежных педагогических поисках. – М.: ИОСО РАО, 1994.
22. Кларин М. В. Педагогическая технология. – М., 1989
23. Занина Л. Проверка знаний или поиск истины? Высшее образование в России № 1, 2000
24. Матрос Д.Ш., Полев Д.М., Мельникова Н.Н. Управление качеством образования на основе новых информационных технологий и образовательного мониторинга. – М.: Педагогическое общество России, 2001.

25. Михеев В. И., Потапова А. Ю. Методы измерения в педагогических исследованиях – М.; Издательский центр АПО, 2002.
26. Мишин В. М. Управление качеством: Учеб. пособие для вузов. – М.: ЮНИТИ-ДАНА, 2000.
27. Панасюк В. П. Системное управление качеством образования в школе. /Под ред. А.И.Субетто. – СПб., М.: Исследовательский центр проблем качества подготовки специалистов. –2000.
28. Педагогика: Учебное пособие для студентов педагогических учебных заведений/В. А. Сластенин, И. Ф. Исаев, А. И. Мищенко, Е.Н.Шиянов. – М: Школа-Пресс, 1997.
29. Поддубная Л. М., Татур А. О., Чельшкова М. Б. Задания в тестовой форме для автоматизированного контроля знаний студентов. –М.: МИФИ, 1995.
30. Поташник М.М. Качество образования: проблемы и технология управления. – М.: Педагогическое общество России, 2002 .
31. Рысс В.Л. Контроль знаний учащихся. - М.: Педагогика, 1982.
32. Субетто А. И. Качество образования в России: состояние, тенденции, перспективы. – М.: Исследовательский центр проблем качества подготовки специалистов, 1998.
33. Талызина Н.Ф. Контроль и его функции в учебном процессе.// Советская педагогика, 1989, №3.
34. Талызина Н.Ф. Теоретические основы контроля в учебном процессе. – М.: Знание, 1983.
35. Теоретические основы содержания общего среднего образования /Под ред. В.В. Краевского, И.Я. Лернера. -М.: Педагогика, 1983.
36. Управление качеством образования/ Под ред. М. М. Поташника. – Педагогическое общество России, 2000.
37. Управление качеством: Учебник для вузов / С. Д. Ильенкова, Н. Д. Ильенкова, В. С. Мхитарян и др.; Под ред. С. Д. Ильенковой. – 2-е изд., перераб. и доп. – М.: ЮНИТИ-ДАНА, 2003.
38. Хубаев Г. О построении шкалы оценок в системах тестирования. Высшее образование в России № 1, 1996.
39. Чельшкова М. Б. Теория и практика конструирования педагогических тестов. – М.: Логос, 2002.
40. Чельшкова М. Б. Адаптивное тестирование в образовании (теория, методология, технология). – М.: Исследовательский центр проблем качества подготовки специалистов, 2001.
41. Чельшкова М. Б. Разработка педагогических тестов на основе современных математических моделей. – М.: Исследовательский центр проблем качества подготовки специалистов, 1995.
42. Чошанов М. А. Гибкая система проблемно-модульного обучения. – М.: Народное образование, 1996.

43. Шишов Е.С., Кальней В.А. Мониторинг качества образования в школе. – М.: Пед. об-во России, 1999.

СОДЕРЖАНИЕ

Введение	3
Часть 1. Конспект лекций по курсу <i>Современные средства оценивания результатов обучения</i>	
Тема 1. ОЦЕНКА И КАЧЕСТВО ОБРАЗОВАНИЯ.	
1. Понятие о качестве образования	4
2. Оценка как элемент управления качеством учебного процесса.	6
3. Мониторинг качества образования	8
4. Педагогический контроль. Функции и принципы педагогического контроля	11
5. Традиционные виды, методы и формы контроля	13
6. Современные методы и формы педагогического контроля: тестовый, модульно-рейтинговый	15
7. Папка индивидуальных учебных достижений «портфолио»	18
8. Оценка: понятие, функции, критерии	20
Тема II. История развития системы контроля знаний и тестирования	24
9. История развития системы контроля знаний и тестирования за рубежом	24
10. Развитие оценивания и тестирования в России	27
Тема III. Педагогические тесты	29
11. Современные подходы к созданию тестов, их сопоставление	29
12. Достоинства и недостатки педагогического тестирования	32
13. Тест, педагогический тест, предтестовое задание	33
14. Основные требования, предъявляемые к тестам, характеристики тестов	35
15. Классификация педагогических тестов	37
16. Модели педагогического тестирования	40
17. Этапы создания тестов	42
18. Определение целей тестирования.	45
19. Принципы отбора содержания теста	48
20. Экспертиза содержания теста	49
21. Понятие «измерение». Измерительные шкалы.	51
22. Методы шкалирования при оценке знаний	54
Тема IV. Единый государственный экзамен и его нормативная база.	57
23. Единый государственный экзамен	57
24. Основные нормативные и правовые документы по ЕГЭ	59
Часть 2. Материалы для практических и семинарских занятий	62
Формы тестовых заданий.	

1. Задания с выбором одного или нескольких правильных ответов.	61
2. Принципы конструирования заданий с выбором одного или нескольких правильных ответов.	63
3. Задания открытой формы (задания на дополнения)	68
4. Задания на установление соответствия.	71
5. Задания на установление правильной последовательности.	72
Сравнительная характеристика всех форм.	72
Литература	74

Сабилова Файруза Мусовна
Современные средства оценивания результатов обучения
Учебно-методическое пособие для студентов педвузов.

Подписано к печати 22.02. 2007

Формат бумаги 60×84/16

Условно печатных листов 5

Тираж 200 экз

Заказ №

Цена договорная

Елабужский государственный
педагогический университет

423604, г.Елабуга, ул.Казанская, 89.