

Лекция №1

«Он-лайн тестирование как форма активизации и контроля ЗУН.»

История возникновения.

История тестов учебных достижений насчитывает, по мнению В.Кадневского, по крайней мере, несколько тысячелетий⁵. В.Аванесов указывает на факты, свидетельствующие о еще более древнем применении тестов.

Древние вавилоняне знали 400 клинописных знаков, использовали шестидесятеричную систему счета, умели вычислять проценты, измерять площадь и объем различных геометрических фигур. Среди изучаемых предметов были те, которые отвечают современному понятию «профессиональная пригодность». За 2200 лет до н.э. в Китае успешно действовала система проверки способностей и отбора персонала для различных государственных должностей. В частности проверялось умение писать, читать, знать порядок проведения придворных ритуалов и церемоний. В течение последующих 2000 лет в систему отбора чиновников были внесены экзамены по гражданскому праву, военному делу, финансам, сельскому хозяйству, географии.

По мнению А.Н.Майорова одним из первых ученых, попытавшихся измерить различия между людьми в области элементарных психических процессов, был англичанин Френсис Гальтон (Galton F.; 1882-1911).

Гальтон ввел в теорию тестирования три фундаментальных принципа, используемых и по сей день:

1. Применение серии одинаковых испытаний к большому количеству испытуемых.
2. Статистическая обработка результатов.
3. Выделение эталонов оценки.

Все современные тесты построены на основе статистической теории измерений, а идея эталона оценки лежит в основе определения теста как стандартизованного инструмента.

Термин «умственные тесты» ввел Дж. Кеттел (Cattell J., 1860-1944). Дж.Кеттел считал тест средством для проведения научного эксперимента с соответствующими требованиями к чистоте эксперимента. Такими требованиями он определял:

1. одинаковость условий для всех испытуемых;
2. ограничение времени тестирования приблизительно одним часом;
3. в лаборатории, где проводится эксперимент, не должно быть зрителей;
4. оборудование должно быть хорошим и располагать людей к тестированию;
5. одинаковые инструкции и четкое понимание испытуемыми, что нужно делать;
6. результаты тестирования подвергаются статистическому анализу, находят минимальный, максимальный и средний результат, рассчитывают среднее арифметическое и среднее отклонение.

Эти идеи, выдвинутые Дж.Кеттелом, составляют основу для современной тестологии. Одинаковость условий для всех испытуемых, одинаковые инструкции и четкое их понимание испытуемыми – фундаментальные принципы, положенные в основу стандартизации процедуры проведения тестирования; ограничение времени, в настоящее время, после дополнительных исследований, устанавливается в зависимости от возраста испытуемых и особенностей применяемого инструментария; идеи статистической обработки результатов реализованы в достаточно сложных методах статистического анализа и моделирования.

Большой вклад в развитие тестов интеллекта внес французский психолог Альфред Бине (Binet A., 1857-1911). Совместно с Теодором Симоном (Simon T., 1873-1961) он разработал тест, позволяющий дифференцировать нормальных и умственно отсталых детей.

В 1911-1912 годах американские психологи Л.Термен и Х.Чальдс дополнили тест Бине - Симона четырьмя новыми:

- 1) «Образец обобщения», то есть пояснения сущности или морали басни.
- 2) Постепенный дополнительный тест по методу Эббингауза.
- 3) Тест для испытания запаса слов (из 100 слов).
- 4) Испытание «Практическое суждение» (тест мяча и поля).

Если первоначально развивалось психологическое тестирование, то в последующем, Маккол В.А. предложил различать тесты психологические - тесты умственного развития (Intelligence Test) и педагогические - тесты учебных достижений (Educational Test).

Основоположником педагогических измерений считается Эдуард Ли Торндайк. Именно Торндайком были созданы первые научно обоснованные педагогические тесты, снабженные нормами.

Большой вклад в развитие теории тестирования внесли Spearman C.E., Gulliksen H., Guttman L. Lord F.M. & Novick M., Kuder G.F. & Richardson M.W. (теория надежности тестов), Crocker Linda & Algina James (современная классическая теория тестов).

В настоящее время за рубежом и в нашей стране широкое применение находит современная теория тестирования - Item Response Theory (IRT). Однопараметрический вариант IRT предложен Георгом Рашем (G.Rasch). Развитие IRT основывалось на появлении двух и трехпараметрических моделей - Birnbaum A. . Обширная деятельность по развитию IRT осуществляется Д.Эндричем (D.Andrich), Б.Райтом (B.Wright)

Фундаментальный труд Анны Анастаси «Психологическое тестирование» представляет собой классическую работу, вобравшую в себя все достижения западной тестологии. В

2006 году вышло 7-е издание совместно с С.Урбиной, куда были добавлены главы по современным методам тестирования.

В своей «Педагогической диагностике» Карлхайнц Ингенкамп утверждает, что традиционные способы оценки, существующие в системе образования, срочно нуждаются в систематическом дополнении объективными методами. При этом необходимо найти научное обоснование методики оценок. Без разумного использования информативных тестов добиться существенного улучшения в оценочной практике учителей невозможно. Это, безусловно верное утверждение, справедливое не только для Германии, но и для России.

Наряду со всеобщим распространением тестов, нарастала и их критика. Дж.Равен указывает на научные и этические аспекты экспансии тестологии в сфере образования. Он называет «безнравственным» пренебрежение ущербом, который практика тестирования наносит судьбам детей и интересам общества. Дж. Равен считает, что традиционные тесты достижений не могут должным образом оценить результаты педагогического процесса, в частности, они не подходят для выявления одаренности учащихся.

Отмечая справедливость его критики, следует все же признать, что развитие тестологии, дающей в руки педагога качественный измерительный инструмент, явление нужное и прогрессивное. Правда, это должен быть не единственный измеритель, определяющий ход учебного процесса.

Развитие тестологии в России.

Среди российских исследователей занимавшихся тестированием, можно назвать П.П.Блонского, Г.И.Залкинда, М.С.Бернштейна и др. К сожалению, в 1936 году вышло постановление ЦК ВКП(б) «О педологических извращениях в системе Наркомпросов». Тестирование было признано противоречащим советской идеологии со всеми вытекающими последствиями.

В послевоенные годы, работы в области тестирования начали возрождаться, а в 70-80-х годах прошлого столетия педагогическое тестирование стало усиленно развиваться в рамках технологии программированного обучения.

Важную роль в становлении отечественной тестологии сыграли работы Беспалько В.П. и Талызиной Н.Ф. Согласно В.П.Беспалько процесс обучения должен быть технологичным и диагностичным. Если нет достоверной диагностики, то нет и учебного процесса. Н.Ф.Талызина, рассматривая вопросы управления процессом учения, анализирует проблемы педагогической оправданности применения тестов различного типа.

Работы отечественных и зарубежных тестологов были в основном изолированы от друг от друга. В СССР фундаментальные труды зарубежных тестологов были практически неизвестны.

В этой связи следует отметить трудно переоценимую деятельность В.С.Аванесова по применению и развитию передовых идей и достижений зарубежной тестологии в отечественной теории и практике педагогического тестирования.

Под руководством В.С.Аванесова в 1985 году на базе Московского института стали и сплавов был организован Исследовательский центр по проблемам управления качеством подготовки специалистов. В этом центре началась планомерная переподготовка руководящих работников и преподавателей в системе высшего образования СССР. В Москву на краткосрочные (1 месяц) курсы съезжались преподаватели, доценты, профессора со всех регионов страны - от Дальнего Востока, до Прибалтики и Средней Азии. Именно этот период, видимо следует считать началом широкомасштабного применения современных научных методов в педагогическом тестировании.

В 1989 году В.С.Аванесовым было выпущено учебное пособие «Основы научной организации педагогического контроля в высшей школе», внесшее большой вклад в развитие теории и практики отечественной тестологии. В.С.Аванесов является приверженцем современных технологий в образовании, большую просветительскую деятельность он осуществляет в глобальной сети Интернет. Его сайт www.testolog.narod.ru содержит массу постоянно обновляемой информации весьма полезной для тестологов. Журнал «Педагогические измерения», главным редактором которого является В.С.Аванесов, является авторитетным изданием, где публикуются работы отечественных и зарубежных тестологов. А.Н.Майоров называет Вадима Сергеевича Аванесова классиком отечественной тестологии и с этим трудно не согласиться.

Вместе с В.С.Аванесовым в Исследовательском центре начала свою деятельность М.Б.Челышкова. К этому моменту она защитила кандидатскую диссертацию и увлеченно читала в Исследовательском центре лекции по Item Response Theory (IRT). Это были очень актуальные лекции, следует отметить, что значимость IRT в тестологии возрастает с каждым годом. Ныне, профессор Марина Борисовна Челышкова широко известна в кругах тестологов, а ее учебное пособие «Теория и практика конструирования педагогических тестов»²⁷, вышедшее в 2002 году пользуется всеобщим признанием.

Очень важные вопросы стандартизации педагогических тестовых материалов подняты в работе Б.У.Родионова, А.О.Татура²⁸. Педагогический тест является измерительным инструментом и это должен быть качественный инструмент, позволяющий получать достоверные результаты. В создании качественных педагогических тестов чрезвычайно велика роль стандартов, которым должны соответствовать педагогические тестовые материалы.

Полный комплекс работ по составлению и использованию тестов школьных достижений представлен в работах А.Н.Майорова. В нашей стране остро стоит проблема подготовки кадров для системы тестирования. А.Н.Майоров отмечает, что существует «миф о том, что для составления тестового инструмента нет необходимости иметь специальные знания. В последние годы, особенно в связи с введением ЕГЭ, появилось множество книг с описанием тестов по любым школьным предметам. Следует понимать, что это не тесты, а некоторые совокупности сырых заготовок, которые следует еще переработать в задания

в тестовой форме, а затем, если удастся – в тестовые задания. Только после этого можно говорить, что создан тест для той или иной предметной области.

Если работы А.Н.Майорова имеют больше практическую направленность, то работы Ю.М.Неймана и В.А.Хлебникова носят больше теоретический характер. Им принадлежит русскоязычная трактовка названия теории Раша (G.Rasch) - «Теория моделирования и параметризации педагогических тестов» (ТМПТТ). Эти работы имеют большое значение для развития теоретических основ отечественной тестологии. Традиционные контрольные процедуры страдают субъективизмом и неопределенностью оценок. В этой связи Ю.М.Нейман и В.А.Хлебников отмечают, что принципиально изменить ситуацию можно лишь в том случае, если подходить к оцениванию знаний как к процессу объективного измерения, а результаты таких измерений обрабатывать стандартными математическими методами и сопровождать стандартными характеристиками точности. Ими указывается, что педагогический тест, в отличие от, например, контрольной работы, можно рассматривать как своеобразный измерительный инструмент определенной разрешающей силы и точности.

Информационные и телекоммуникационные технологии оказывают сильное воздействие, как на организационные формы, так и на обработку результатов тестирования. В работе В.И.Нардюжева и И.В.Нардюжева рассмотрены вопросы построения системы компьютерного тестирования. Программные разработки этих авторов использовались для организации абитуриентского компьютерного тестирования Федеральным центром тестирования Минобрнауки РФ (ЦТ МО РФ). Прикладные программные средства Tester - для проведения тестирования, Operator - для конфиденциальной передачи результатов тестирования в ЦТ МО РФ, StatInfo - для статистической обработки результатов тестирования, показали себя как надежные и удобные программные продукты.

Термин «Дидактическая тестология» вводит Е.А.Михайлычев. Если исходить из того, что дидактика - это теория обучения, а педагогика - теория и обучения и воспитания, то термин, предложенный Е.А.Михайлычевым представляется более точным, нежели термин «педагогическое тестирование». Однако следует отметить, что в научной терминологии уже устоялся термин «педагогическое тестирование». Е.А.Михайлычевым очень обстоятельно описаны проблемы валидации теста и пути их решения.

Из приведенного, очень краткого и неполного обзора следует, что тестирование учебных достижений широко используется за рубежом и довольно высокими темпами развивается в России.

Определение термина «тест» и «тестирование».

Термин «тест» был введен в научный обиход американским психологом Дж. Кеттеллом в 1890 г. для наименования психологических проб, т. е. специально разработанных стандартизированных методик, с помощью которых пытались измерить различия между индивидами или реакциями одного индивида в разных условиях. Однако ни в психологии, ни в других научных областях (социологии, педагогике, медицине, физиологии, вычислительной технике) все еще нет единого определения термина.

Слово «тест» в английском языке означает «испытание», «эксперимент», «проверка», «проба». При такой полисемии терминологические неточности очевидны.

Иногда тест трактуется чрезвычайно широко: под ним понимаются и простые контрольные работы, и контрольно-тренировочные упражнения, и собственно процесс контроля, а также все взятое вместе.

Англичанин Френсис Гальтон определил основные принципы в развитии теории тестов. Применение серии одинаковых испытаний к большому количеству испытуемых. Статистическая обработка результатов. Выделение эталонов оценки.

Французский психолог Альфред Бине считается родоначальником современных тестов, предназначенных для диагностики уровня развития интеллекта.

Проблемой разработки тестов в разное время занимались видные российские ученые: М.С. Бернштейн, П.П. Блонский, А.Я. Болтунов, С.Т. Геллерштейн, Г. Залкинд, К.Н. Шпильрейн, А.М. Шуберт, А.Н. Леонтьева, А.Р. Лурии, А. А. Смирнова и другие. В настоящее время в нашей стране появилось несколько центров, в которых достаточно профессионально занимаются работой с тестовыми методиками. По мнению П. П. Блонского огромная заслуга тестов состоит в том, что они позволяют обычный ответ ученика "Так мне кажется" заменить словами "Я это знаю" или "Я это не знаю".

Термин «тест» — базовый компонент других понятий, связанных с идеей измерений способностей, свойств, достижений личности. Так, теоретико-прикладная наука, исследующая тест с точки зрения его формы, структуры, функционального назначения, показателей качества, изучающая возможности использования тестов для выявления особенностей формирования и деятельности личности в конкретно-исторических условиях, называется **тестологией**. Соответствующий этому наименованию английский термин **testing** многозначен. Он переводится на русский язык и как **тестирование** (testing).

Под **тестированием** в широком смысле имеют в виду совокупность всех процедурных этапов, а именно: планирование, составление, апробация, обработка и интерпретация результатов первичного, предварительного теста — **предтеста**, перепланирование, оформление и подготовка спецификаций, инструкций окончательного теста.

В узком смысле **тестирование** есть форма контроля (и обучения) с помощью проведения и использования теста.

Роль тестирования в педагогическом процессе.

Одной из задач повседневного учительского труда является необходимость осуществлять контроль знаний учащихся. Формы контроля, применяемые учителями, очень разнообразны, но наиболее часто используются письменный или устный опросы. К сожалению, эти формы не лишены недостатков. При проведении устного опроса — это относительно большая затрата времени урока при небольшом количестве выставляемых

оценок, при проведении письменных работ количество оценок возрастает, но много времени уходит на проверку.

Тестирование как эффективный способ проверки знаний находит в школе все большее применение. Одним из основных и несомненных его достоинств является минимум временных затрат на получение надежных итогов контроля. При тестировании используют как бумажные, так и электронные варианты. Последние особенно привлекательны, так как позволяют получить результаты практически сразу по завершении теста.

Функции тестирования.

Тестирование в педагогике выполняет три основные взаимосвязанные функции: диагностическую, обучающую и воспитательную:

Диагностическая функция заключается в выявлении уровня знаний, умений, навыков учащегося. Это основная, и самая очевидная функция тестирования. По объективности, широте и скорости диагностирования, тестирование превосходит все остальные формы педагогического контроля.

Обучающая функция тестирования состоит в мотивировании учащегося к активизации работы по усвоению учебного материала. Для усиления обучающей функции тестирования, могут быть использованы дополнительные меры стимулирования студентов, такие, как раздача преподавателем примерного перечня вопросов для самостоятельной подготовки, наличие в самом тесте наводящих вопросов и подсказок, совместный разбор результатов теста.

Воспитательная функция проявляется в периодичности и неизбежности тестового контроля. Это дисциплинирует, организует и направляет деятельность учащихся, помогает выявить и устранить пробелы в знаниях, формирует стремление развить свои способности.

Тестирование — более справедливый метод, оно ставит всех учащихся в равные условия, как в процессе контроля, так и в процессе оценки, практически исключая субъективизм преподавателя.

Следует отметить, что именно тестирование постепенно становится и основной формой сдачи экзаменов. С 2009 года для всех выпускников школ основной формой итоговой государственной аттестации в школе Российской Федерации является Единый Государственный Экзамен. И реалии таковы, что требуется обязательно вводить тестовые технологии в систему обучения. С их помощью в течение года следует оценить уровень усвоения материала учениками и формировать у них навык работы с тестовыми заданиями. Такие тренировки позволят учащимся при сдаче ЕГЭ реально повысить балл. Также во время таких тренировок развиваются соответствующие психотехнические навыки саморегулирования и самоконтроля. В связи с этим тестирование, как средство измерения и контроля знаний учеников, становится основой обучающего процесса в российской школе.

Тестовые задания могут составляться с использованием разнообразных компьютерных инструментов, начиная от различных редакторов и программ для разработки презентаций и до использования языков программирования и возможностей сети Интернет.

Форма теста признана в системе образования многих стран мира. Актуальность и выбор темы состоит в том, что проблема контроля за уровнем знаний, умений и навыков учащихся - вечная проблема образования. Современный учебный процесс не мыслится без системы тестового контроля. Тест как метод изучения индивидуальных различий возник сравнительно недавно. Ввели его в конце XIX и начале XX веков психологи для определения физических, физиологических и психических особенностей человека.

По определению, тест – стандартизированные, краткие, ограниченные во времени испытания, предназначенные для установления количественных и качественных индивидуальных различий. Тестирование может частично заменить обычные формы проверки знаний. Но основное их достоинство – это простота и скорость, а также то, что тесты воспринимаются большинством учащихся как своеобразная игра.

«Компьютерные» - педагогические тесты имеют свои особенности, которые необходимо учитывать при их разработке. Анализ информации на экране монитора компьютера часто бывает затруднен из-за неправильной подачи материала. Хотя именно компьютеризованное тестирование имеет свои преимущества и перспективы применения в образовании. Особенная роль здесь принадлежит компьютерно-индивидуальному тестированию, когда каждому испытуемому предъявляется уникальный набор заданий. Выявление индивидуальных особенностей ребенка с помощью компьютерных тестов позволяет говорить об их актуальности для современной системы образования, Компьютерные программы, используемые для проведения тестирования, - это инструменты для проведения полного цикла тестирования - начиная с разработки теста и заканчивая учетом и анализом результатов. Одно из главных требований к компьютерным тестам - многовариантность. Многовариантность должна быть двух типов: во-первых, по данным заданий, во-вторых, по порядку их предъявления. Реализация этого требования снимает сразу две организационные проблемы: списывание и запоминание ответов. Тест – это стандартизированное задание, по результатам которого судят о знаниях, умениях и навыках испытуемого. Поэтому к тестам предъявляют определенные требования.

Основные требования, предъявляемые к тестам.

Тесты должны быть:

1. Относительно краткосрочными, т.е. не требовать больших затрат времени для выполнения тестовых заданий.
2. Однозначными, т.е. не допускать произвольного толкования тестовых заданий.
3. Правильными, т.е. исключать возможность формулирования многозначных ответов.
4. Относительно краткими, требующими сжатых ответов.
5. Удобными, т.е. пригодными для практического использования всеми учащимися.
6. Проверяемыми, т.е. соответствовать проектируемым целям обучения.
7. Надежными, т.е. показывать устойчивые результаты в сходных условиях.
8. Валидными, т.е. ценными, имеющими значение для процесса обучения.

9. Объективными, т.е. демонстрирующими в результате измерения именно то, что предусмотрено целью контроля.

Тесты обученности составляются по всему курсу или по отдельной изучаемой теме в основном по узловым вопросам программы и применяются на всех этапах дидактического процесса. Образовательный, воспитательный характер тестов направлен на реализацию принципов гуманизма и толерантности. Чтобы добиться развивающего и выравнивающего эффекта и формулировки учебной задачи тесты удобно использовать на первом ориентировочно-мотивационном этапе. Обычно составляется небольшой тест, например, из 6 вопросов. 5 вопросов на активизацию необходимых знаний, а последний вопрос помогает создать проблемную ситуацию, приводит к формулировке учебной задачи. За минимальный срок я успеваю получить информацию о степени усвоения предыдущей темы, создать мотивацию к усвоению новой темы, оценить знания. Оптимальны тесты, содержащие 10-12 вопросов при 3- 4 выборочных ответах, из которых правильный лишь один. Для обработки результатов тестирования обычно применяю следующую шкалу: 85% — 100% — «5», 65% — 85% — «4», 40% — 65% — «3», 0% — 40% — «2».

При правильном отборе контрольного материала содержание теста может быть использовано не только для контроля, но и для обучения. Таким образом, позволяя испытуемому самостоятельно обнаруживать пробелы в структуре своих знаний и принимать меры для их ликвидации. В таких случаях можно говорить о значительном обучающем потенциале проектируемой системы, использование которого станет одним из эффективных направлений практической реализации принципа взаимосвязи и единства контроля и обучения.